

**FAST TRACK: THE PRACTICE OF DRUG DEVELOPMENT
AND REGULATORY INNOVATION IN THE LATE TWENTIETH
CENTURY U.S.**

Donna A. Messner

Ph.D., Science and Technology Studies

University of Edinburgh

2008

For my Parents

DECLARATION OF ORIGINALITY OF SUBMITTED WORK

In conformance to University regulations, I hereby declare that that:

1. this thesis has been composed solely by me;
2. this thesis is entirely my own work; and
3. this thesis has not been submitted in part or whole for any other degree or professional qualification

Signed

Date

ABSTRACT

This thesis examines the laws and regulations created in the 1980s and 1990s in the U.S. to hasten development, evaluation, and approval of drugs to treat serious and life-threatening diseases, and to allow access of seriously ill patients to investigational drugs on a pre-market approval basis. Using detailed historical exposition in tandem with the social-theoretic tools of the sociology of scientific knowledge (SSK), and particularly Barnes's account of meaning finitism, this thesis examines the social origin, definition, and case-by-case application of conceptual categories in the regulatory oversight of drug development and approval. With this approach, rules and standards for drug approval are shown not to be fossilised machinery for decision-making, but rather living, socially produced and maintained, inherently revisable resources for action. Key conclusions from this study are that: the regulatory actions taken to confront AIDS in the 1980s, often considered to be a radical break with previous practice, had their conceptual origins in the 1960s and 1970s; rule-making is often constitutionally related to a creative process of rule-'breaking'; tacit processes of consensus outside of, and prior to, formal consensus mechanisms for rule-making are often fundamental to the rule-making process, resulting in de facto 'rules' on which later, formal rule-writing can be based; as predicted by finitism, newly created categories of action in drug development and approval require reinterpretation of underlying concepts in related existing categories. The thesis also demonstrates the flexibility and revisable nature of existing conceptual resources for application to current circumstances, consistent with a finitist view of knowledge. While the conclusions of this research are based on only one area of regulation, they are suggestive for more general descriptions of regulatory action. Contemporary theories of regulation are typically designed as economic models or are viewed through traditional categories of law and political science. As a result, they tend to abstract reality, ignoring day-to-day administrative practice, idealizing the nature of rule-following and rule-making, and ignoring tacit processes of consensus. This thesis brings an interdisciplinary perspective to the theory of regulation, suggesting the outlines of a 'social' theory of regulation more fully sensitive to the empirical reality of the social process of rule-making and rule-breaking in contemporary regulation.

TABLE OF CONTENTS

TABLE OF CONTENTS	iv
List of Acronyms	vi
Acknowledgements	viii
1. Introduction.....	1
1.1 Drug Development and Accelerated Approval in Context.....	1
1.2 Literature in the Sociology and Recent History of Therapeutic Drug Regulation and Clinical Evaluation	6
1.3 Finitism and the Theory of Regulation.....	11
1.4 Plan of thesis.....	30
2. Research Design and Methodology	35
2.1 Documentary Sources and Source Selection Methods.....	36
2.2 Interviewing.....	47
2.3 Data Reduction and Analysis.....	49
2.4 Research Ethics.....	53
2.5 Research Topic, Researcher Motivations, and Method Choice.....	54
3. Background: The Roots of Reform	62
3.1 The 1962 Drug Amendments	63
3.2 FDA Growing Pains	66
3.3 Rising Criticism and Efforts for Reform.....	73
3.4 FDA Response to Criticism: <i>Concept Document</i>	89
3.5 IND and NDA rewrite proposals	98
3.6 Two Decades in Review	104
4. AZT: Reflections of the Past and a Model for the Future	107
4.1 AZT as a Model for New Rule-Making: Subpart E.....	111
4.2 Treatment IND and the Collision with AIDS.....	115
4.3 Comments on Subpart E.....	121
4.4 Preliminary Interpretations	131
5. War on Cancer: Three Cases.....	138
5.1 Mitoxantrone: Shades of Future Legislation	139
5.2 Ifosfamide: Constitution of a Single-Study Drug Approval.....	150
5.3 Fludara: Retrospective Data Mining and the Meaning of Clinical Benefit ...	164
5.4 Conclusions: The ODAC as a ‘Microclimate’ of Standards-Making.....	174
6. Accelerated Approval	181
6.1 Surrogate Endpoints and dDI	181
6.2 The Re-Vision of Subpart E and Creation of Accelerated Approval.....	191
6.3 Subpart H and ddC.....	197
6.4 Reactions to Accelerated Approval.....	211
7. Accelerated Approval and Cancer	218
7.1 Dexrazoxane: Misalignment of Rule and Application.....	218
7.2 Liposomal Doxorubicin for AIDS-Related KS	227
7.3 Discussion: Adaptations for Accelerated Approval.....	239
8. Reinvention and Modernization.....	243
8.1 National Politics and New Directions for Reform.....	244
8.2 Congressional Politics and the Footrace to Reform	250
8.3 FDAMA	266
8.4 Evolving Regulatory and Statutory Interpretations.....	273

8.5	Fast Track in Application: One Brief Case.....	279
8.6	Conclusions.....	290
9.	Conclusions: FDA History, Finitism, and a ‘Social’ Theory of Regulation	293
9.1	FDA Decision-Making and Meaning Finitism.....	293
9.2	Toward A Finitist-Collectivist Theory of Regulation	310
9.3	Afterward.....	325
	WORKS CITED	328

LIST OF TABLES

Table 3.1	The Clinical Phases in Selected Regulations and FDA Reform Proposals	99
Table 3.2	Expedited Approval, Phase IV Commitments and Pre-Market Access to Drugs in FDA Reform Proposals	99

LIST OF FIGURES

Figure 5.1	Therapeutic Scheme for Mitoxantrone Clinical Trials.....	141
Figure 9.1	Regulatory Process and Knowledge Praxis.....	313

LIST OF ACRONYMS

ACTG	AIDS Clinical Trial Group
ADAC	Antiviral Drugs Advisory Committee
AIDS	Acquired immunodeficiency syndrome
AZT	Azidothymidine (chemical name); zidovudine (generic name)
BLA	Biologics license application
BMS	Bristol Myers Squibb
CAST	Cardiac Arrhythmia Suppression Trial
CBER	U.S. FDA Center for Biologics Evaluation and Research
CDC	U.S. Centers for Disease Control
CDER	U.S. FDA Center for Drug Evaluation and Research
CFR	U.S. <i>Codes of Federal Regulations</i>
CLL	Chronic lymphocytic leukaemia
CML	Chronic myelogenous leukemia
CRC	Colorectal cancer
CRS	Congressional Research Service
DESI	Drug Efficacy Study Implementation
DOX	Doxorubicin
DZR	Dexrazoxane
E.O.	Executive Order
FAC	5-fluorouracil, adriamycin (doxorubicin), and cyclophosphamide
FDA	U.S. Food and Drug Administration
FDAMA	<i>Food and Drug Administration Modernization Act</i>
FDCA	<i>Federal Food, Drug and Cosmetic Act</i>
FR	U.S. <i>Federal Register</i>
IBA	Industrial Biotechnology Association
IND	Investigational new drug
KS	Kaposi's Sarcoma
LTI	Liposome Technology, Inc.
MIPI	Medicine in the Public Interest
MTD	Maximum tolerated dose
NAC	N-acetyl cysteine
NAS	National Academy of Sciences

ACRONYMS (CONTINUED)

NAS/NRC	National of Academy of Sciences/National Research Council
NIH	National Institutes of Health
NCI	National Cancer Institute
NDA	New drug application
NPR	National Performance Review
ODAC	Oncologic Drugs Advisory Committee
OMB	Office of Management and Budget
PDUFA	Prescription Drug User Fee Act
PhRMA	Pharmaceuticals Research and Manufacturer's of America
PHS	Public Health Service (Act)
PMA	Pharmaceutical Manufacturer's Association
PVB	cisplatin [platinum], vinblastine, and bleomycin
RCT	Randomized, controlled trial
SWOG	Southwest Oncology Cooperative Group
TTP	Time to progression
VIP	VP-16 or vinblastine (Velban) and (cis)platinum
VP-16	Etoposide or VePesid

ACKNOWLEDGEMENTS

It will be difficult to do justice in this short space to all who have helped me in the process of doing this research. My thesis advisors, Steve Sturdy and James Mittra, are a veritable supervisory ‘dream team’, providing a level of responsiveness and insight that many students would envy. Their input has enriched this project immensely — although, of course, any deficiencies are entirely my responsibility. David Bloor, Donald MacKenzie, Harry Marks and Alex Preda were kind enough to read early versions of some of this material. Harry Marks was especially generous with his time in acting as an informal mentor for a summer while I conducted research in Washington D.C. Special thanks go to the staff of the Food and Drug Administration’s (FDA) Dockets Division, especially to Ryan Cleaver for hunting down long lists of archived electronic transcripts for me in the FDA’s antiquated system. While in Washington, Jessica, Joe and Samantha Kaputa welcomed my husband and me like cherished family members, providing not merely lodging, but sustenance for the body and spirit (including, as needed, beer, vodka and barbeque). Vivian Messner and Rob Congedo generously provided for transportation around the region. I am grateful to the School of Policy at the Georgia Institute of Technology for hosting me as a visiting student, and to PRIME for making it possible. Thanks especially to Ann Bostrom for her enthusiasm and geniality as my Georgia Tech advisor. My fondest thanks go the superb faculty and students of the Science Studies Unit, a group with which I am privileged and delighted to be associated. John Henry has been unfailingly supportive of my work and a most welcome source of wit and anecdotes over coffee and tea. I echo countless authors of Unit books and theses in thanking Carole Tansley for both administrative competence and tea time hospitality beyond the call of duty. She has also been a valued source of strength and friendship. Special thanks to Graeme Beale for being an incomparably congenial office mate and for graciously providing an empty sofa and delectably improvised meals on my return visits. I am deeply indebted to my husband, E.J. Messner, for allowing me to uproot our comfortable, settled lives to pursue this goal. His devotion and good humour are treasured gifts. Finally, I dedicate this thesis to my parents, Dean and Neva Hurley, who by example taught me how to live life with diligence, faithfulness, kindness and laughter, and who unwaveringly sustained me with love and acceptance. I remember them daily, and hope to honour their memory through this accomplishment. If only they were here to share it with me.

1. INTRODUCTION

1.1 Drug Development and Accelerated Approval in Context

In May 2007, the *New England Journal of Medicine* published online a meta-analysis of 42 studies in the literature revealing that the leading drug to treat type II diabetes, Avandia, was found to increase significantly the risk of death from heart attack and other cardiovascular problems (Nissen and Wolski, 2007). The finding was especially significant since a majority of diabetic patients ultimately die from cardiovascular diseases. The U.S. media trumpeted that a ‘blockbuster’ diabetes drug was a ‘heart death risk’ (Sternberg 2007). The drug-maker, Glaxo, disputed the study and insisted that the drug was still safe for patients to use. Nevertheless, many observers compared Avandia to Vioxx, Merck’s infamous non-steroidal anti-inflammatory drug approved in 1999 and voluntarily withdrawn from the market in 2004 following revelations of increased risk of heart attack and stroke for users of the pain-relieving product (Marchione 2007).¹

Reporters and politicians were unanimous in asking: how could this happen — *again?* Rep. Henry Waxman called a 6 June hearing of the U.S. House Oversight and Government Reform Committee to examine the Food and Drug Administration’s (FDA) handling of Avandia (U.S. House 2007). Meanwhile, others rendered judgment. In answer to the question ‘How did drugs like Vioxx or Avandia get approved?’, one medical reporter wrote that such drugs ‘typically are tested on small groups of people, and with narrow and sometimes flawed definitions of success. Such studies often are not large or long-lasting enough to reveal rare side effects’ (Marchione 2007)². Secondly, in the case of Avandia, the drug was approved ‘because it led to short-term improvements in certain

¹ See the FDA information page on non-steroidal anti-inflammatory drugs (NSAIDs) at <http://www.fda.gov/cder/drug/infopage/cox2/> (accessed 25 May 2007).

² This Associated Press article appeared widely in online news sources on 23 May 2007. The quotations were originally taken from Washingtonpost.com on that date. However, as of this writing the story is no longer available at that website, so an alternative link is provided in the bibliography.

blood-sugar measurements whose true value as a sign of diabetes control some experts now question. 'The FDA did not insist on evidence of clinical benefit — fewer blood-sugar crises, hospitalizations, etc.' (Ibid.). A similar assessment was made by a medical expert on one of the major television network news programmes, the *ABC Evening News*. Asked by the presenter Charles Gibson whether the Avandia case 'point[s] up a basic flaw in the way we approve drugs in this country', medical expert Dr. Timothy Johnson replied, 'I think it does'. Johnson described the same perceived flaws in clinical trial design as Marchione (2007), concluding, 'I think we really have to follow the old adage in medicine which is to "first do no harm" — and that means longer-term studies that show real benefit or real problems' (ABC News 2007a).

To hear these observers explain the basis of decision-making for drug approval, one would wonder why larger studies are not required by the FDA and why direct evidence of clinical benefit is not sought to establish firmly the risks and benefits of investigational drugs. Part of the answer to this question can be seen in another news story from the same period. In the spring of 2007, the FDA made an unusual decision to contravene an advisory committee vote to approve a prostate cancer drug called Provenge (Stein 2007), a new immunotherapy approach to combating the growth of cancer cells. While one study of the two presented to the committee demonstrated some survival advantage for patients on the drug as compared to a placebo group, the number of patients on which this conclusion was based was sufficiently small that two temporary voting members of the advisory committee were unconvinced that the apparent survival advantage was not a statistical anomaly.³ These two members later convinced the FDA to seek supplemental data from the drug sponsor rather than approve the drug immediately. When the decision became public in May 2007, patient advocacy groups as well as investors in the small biotechnology drug-maker Dendreon were outraged over the delay. Dendreon's stock price plummeted. Patients and their allies initiated letter-writing campaigns and protests. They staged a rally on Capitol Hill on 4 July and obtained a personal meeting with the FDA commissioner (Stein 2007). The two committee members seen as responsible for the FDA's decision received threatening letters and telephone calls. One irate investor created an 'Approve Provenge Now' webpage

³ Press accounts like Stein's (2007) consistently represent the 13-to-4 approval vote as based on two randomized, placebo-controlled, double-blind clinical studies. However, in the advisory committee transcript the drug sponsor admits that the 'primary evidence of clinical efficacy in this application is the results from Study 1' (FDA 2007, 31). Study 2 was discontinued prior to completion of enrolment.

declaring, ‘Hey, Hey, FDA, How Many Dads Did You Kill Today?’ with images of the two advisory committee members exhibited against a background depicting chaotic and crookedly arranged white granite crosses, like an aged and overcrowded virtual graveyard, with Mozart’s *Requiem* playing as theme music (Ibid.).⁴

All of this for a drug which cured no one, and provided at best a median survival benefit over placebo of 4.5 months in the course of a 36-month clinical trial (FDA 2007). In an opening line laden with irony for anyone familiar with FDA history, one *ABC News* report on Provenge began by saying that the ‘Food and Drug Administration is often criticized for being too lax for approving drugs that turn out to be dangerous and must be pulled from the market’ (ABC News 2007b). Another notable bit of irony: a few years earlier the FDA was criticized for not sweeping out older diabetes therapies after the new and improved Avandia hit the shelves. A 2003 law review article (Bean 2003) found it ‘disturbing’ that the older drug Rezulin remained available, despite its high level of liver toxicity, when Avandia had been on the market since 1999 and ‘worked similarly to Rezulin but without the high level of risk’ (908).

The examples of Avandia and Provenge, coincidentally juxtaposed by history in ironic counterpoint, illustrate vividly the tension perpetually negotiated by the FDA between, on the one hand, gathering sufficient information to make a judgment of the safety and effectiveness of new drugs and, on the other hand, making approval decisions for drugs to treat life-threatening diseases as early in the process as possible (with correlatively less information on which to base the decision). These examples also demonstrate the potential consequences arising when that tension is perceived as imbalanced. In this thesis we will witness the development and growth of this tension from the 1960s through the end of the century, examining the circumstances of the regulatory and legislative rule-writing which sought to define its boundaries. Historically, the clinical practices criticized in the Avandia case — the use of smaller clinical trials and reliance on things like blood sugar laboratory tests in lieu of direct measurements of clinical benefit — represent the culmination of decades of FDA reform. This reform came about in response to decades of criticism of the FDA as plodding, over-cautious, and too often favouring the information-gathering side of the tension. To be sure, the most immediate cause of the reformism of the 1980s and 1990s was the urgent need and

⁴ At the time of this writing, the page has added photographs of FDA officials and others. It also features a gruesome image of the Grim Reaper with ‘FDA’ emblazoned on the blade of its scythe. The webpage can be found at http://www.myspace.com/approve_provenge_now (accessed 1 Feb 2008).

political action of AIDS patients — political forebears to the Provenge activists. On the role of AIDS activists in changing the way clinical studies were done, Epstein's (1995; 1996; 1997) work is definitive. However, one result of this thesis will be to see Epstein's work in a larger context. This thesis will demonstrate that many of the regulatory approaches used to address AIDS, and many of the concepts imbedded in AIDS-era rule-making, had been in circulation for at least a decade prior to the advent of AIDS — a conceptual lineage observers often miss when they focus only on the immediate circumstances surrounding the rule-writing of the AIDS era. For this reason, authors such as Orlando (1999) suggest that reforms for expanded (pre-market) access to investigational drugs (another practice which negotiates the tension between information-gathering and early action) 'began in the late 1980s' (543), when in fact these practices began informally as early as the late 1930s, after the passage of the 1938 *Food, Drug and Cosmetic Act*,⁵ but were not formalized in the regulations until 1987.

The type of shorthand used by Orlando may be acceptable for the purposes of law or policy, but for me it will not do because it tends to marginalize the repertoire of pre-existing cultural and historical resources as if history were a discontinuous function. Yet it is these resources which form the basis of human experience for understanding and approaching any new situation, including the AIDS crisis. While we may draw from our experiences in ways that are creatively synthetic and analogous — and of course we may reach well beyond our own personal experience to what we know of past institutions, historical figures and events, hearsay from others, etc.— versions of our past experience form the pattern or design we use to interpret and understand each new event confronting us, and ultimately to produce new knowledge as a result. Hence, my study of AIDS era regulation means recognizing how conceptual tools already in circulation were brought to bear on the new situation.

At the same time, we must recognize that *which* elements of past experience will be applied in any given situation, and *how* they will be applied, ultimately cannot be anticipated in advance. Close examination of the events of this period will reveal that the related processes of knowledge application and knowledge creation necessitate step-wise, case-by-case judgements as situations develop. This is true, even if a pattern or trajectory can be identified in the judgement process on a retrospective basis. I will identify such

⁵ This according to the U.S. Senate report on what became the 1997 *Food and Drug Administration Modernization Act* (U.S. Senate 1996).

patterns in this thesis, but this account makes clear that the next application of existing knowledge to a new situation is always indeterminate. For this reason, we can benefit by approaching this subject with a finitist conception of knowledge (Barnes 1982).⁶ Such a conception leads us to view knowledge as thoroughly conventional and constructed from the 'bottom-up' through practice: 'It is not that knowledge is a system of conventions which determines how we think and act. On the contrary it is our decisions and judgments which determine what counts as conventional, and thus which sustain and develop a conventional framework' (ibid, 30). Such a conception implies that all applications of a term are ultimately determined in contingent, moment-by-moment practice, and are ultimately revisable as they continue to be applied. Consistent with this finitist view of knowledge, we will see that during the 1980s and 1990s, researchers and regulators responded to AIDS by marshalling conceptual and other resources already at hand: existing FDA practice and ongoing approaches to reform; informal FDA procedures already in use for years; a drug that had already been in cancer testing (and failed) and had lain on a shelf for years. We will see that the meaning and application of the resources applied to each problem were themselves modified on a case-by-case basis in praxis, becoming in turn part of the basis for future decision-making and concept application. In this way we will come to see the rules guiding drug development and approval, and even scientific standards used to assess drug applications, as flexible resources adapted to each new situation.

A core supposition of this thesis is that while the AIDS era of regulation was extraordinary in terms of the circumstances driving regulatory action, the mechanisms of human action differed only in degree, not in kind, from regulatory decision-making processes under less urgent conditions. In this sense, my work is akin to the time-tested approach in science studies of examining scientific controversies to expose the social relations which perpetually exist in scientific activity but tend to be obscured under the veneer of normal routine.⁷ Analogously, my contention is that we can use this tumultuous period of regulatory history to improve our understanding of the process of regulatory rule-writing and decision-making generally. In so doing, we will discover that even in the

⁶ On finitism, see also Barnes 1995, section 2.4; Barnes, Bloor and Henry 1996. For more on a 'naturalistic' approach to knowledge see Barnes 1977 and Bloor 1991.

⁷ See, for example, Collins 1975; Collins and Pinch 1982; Pickering 1981 (and all of *Social Studies of Science* vol. 11 no. 1, a special edition on 'Knowledge and Controversy: Studies of Modern Natural Science'); Richards 1988 and 1991; Epstein 1997; as well as many of the chapters in Collins and Pinch 1994 and 2002.

context of regulatory action — an arena often viewed as rigidly rule-guided — rules turn out to be much more flexible and adaptive than often believed.

While the conclusions of this research are based on only one area of regulation, they are suggestive for more general descriptions of regulatory action. Contemporary theories of regulation are typically designed as economic models or are viewed through traditional categories of law and political science. As a result, they tend to abstract reality, ignoring day-to-day administrative practice, idealizing the nature of rule-following and rule-making, and ignoring tacit processes of consensus. This thesis brings an interdisciplinary perspective to the theory of regulation, suggesting the outlines of a ‘social’ theory of regulation more fully sensitive to the empirical reality of the social process of rule-making and rule-breaking in contemporary regulation.

In the sections which follow, I will review the literature in the sociology and recent history of therapeutic drug regulation, showing the contribution my work makes to the field. Then I will build the theoretical apparatus for this work by discussing in more detail the finitist view of knowledge and its implications for social institutions. I will also provide an overview of the primary theoretical basis for contemporary theories of regulation, showing where this theory fails when seen from the perspective of collectivist social theory, and will outline in broad terms the major theories of regulation. In the final section I will provide the reader with a roadmap of the historical narrative to follow, showing in general terms how that narrative builds the case for a finitist view of regulation.

1.2 Literature in the Sociology and Recent History of Therapeutic Drug Regulation and Clinical Evaluation

In the sociological and historical literature, a number of authors have studied the development and practice of randomized, controlled trials (RCTs) for the evaluation of drugs. Lilienfeld (1982) has written on the general history of the clinical trial, tracing the individual concepts of comparing groups, randomization and control, and blinding. While the systematic comparison of groups can be seen in sources as early as the book of Daniel in the Old Testament, Lilienfeld shows that these concepts did not finally come together to form the randomized, controlled trial (RCT) until the 1930s. Oudshoorn (1993) similarly noted that, contrary to what is often believed, a version of the clinical trial was

practiced in the U.S.⁸ as early as the 1920's and 1930's. Marks (1997) has traced the historical development of 'rational therapeutics' in the U.S., where the controlled clinical trial arose to prominence in the twentieth century as the preeminent method for testing the safety and efficacy of new therapies. This development came about through the persistent efforts of 'therapeutic reformers' (Marks 1997)—a disparate group of individuals 'joined by their belief in the power of science to unite both medical researchers and practitioners *despite* obvious differences of training and circumstance' (p. 3, original emphasis). As a result of their efforts, the RCT is now thought of as the 'gold standard' in drug evaluation; as an 'unparalleled technique for measuring the value of novel treatments' (Marks 1988, p. 297).

Notwithstanding this reputation for evaluative excellence, Richards (1988) writes that 'clinical trials, no matter how rigorous their methodology, will inevitably embody the values or commitments of the assessors'; ultimately they 'will not provide definitive answers' (p. 654). The example Richards proffers is that of clinical trials to determine the efficacy of vitamin C in cancer treatment (1988; 1991)—trials the results of which were hotly debated and marked by political interests. Another study (Markle and Peterson 1980) describes the conflict between the medical establishment and a grassroots social movement over another putative cancer cure, Laetrile. Indeed, the interest-driven and value-laden nature of medical research and drug approval has been explored in some detail, especially by Abraham (1995b), who has found that industrial interests can be a major factor influencing the reception of scientific papers in medicine (1995a); that in conditions of scientific uncertainty in assessing drug safety, the regulatory agencies tend to favour industrial interests (1994); and that a range of socio-political factors and institutional interests can be cited to explain differences in risk assessment of the same drug in the US and UK (1993, Abraham and Sheppard 1999). Other studies have been done showing the effect of cultural values on assessments of drug effectiveness and risk (McCrea and Markle 1984); of conceptions of a 'successful' therapeutic study being influenced as much by journalistic accounts as by 'scientific merit' (Oudshoorn 1999); and of widespread acceptance of therapeutic or diagnostic findings for which 'scientific merit' is lacking (Lipton and Hershaft 1985; Casper and Clarke 1998).

⁸ This thesis will be focused on the conduct of clinical trials and drug regulation in the United States. For accounts of the rise of clinical trials for therapeutic evaluation in the British context, see Cox-Maksimov 1998 and Yoshioka 1998

The design of clinical trials can be no less controversial than the results. In his book on the development of the breast cancer drug Herceptin, Bazell's (1998) journalistic account describes in detail the technical difficulties and political conflict surrounding the conduct of large efficacy trials. Marks (1997) describes the unsuccessful attempts to initiate a diet-heart study. A planned major study on the connection between diet and coronary disease became bogged down in political maneuvering and disagreements on design, including the practice of 'blinding', the use of placebos, and role of biostatisticians.⁹ Moreover, in recent years patients and other interested laypeople have sought credibility to influence clinical trial design, desiring to bolster ethical considerations over strictly statistical ones (as in, for example, eliminating the use of placebo controls in trials for drugs to treat life-threatening diseases) and pressing for ways to accelerate the testing and approval process for new drugs to treat life-threatening illnesses (Epstein 1995, 1996), even while debates ensue over the best characteristics of patient condition ('endpoints' or 'markers') to use as a measurable index of drug effectiveness (Epstein 1997). Similarly, Meldrum (1994) has observed that the clinical trial in practice is 'a social exercise in problem solving' involving various groups of actors each having 'a unique definition of the social-medical problem to which the trial is designed to provide a solution' (iv). As a result, she argues that the outcome of clinical trials is 'not value-free data, but a social construction' (iv). In saying this, she does not imply that clinical trial data is meaningless, but rather that the questions to be asked and the approach to answering them are ultimately a matter of social negotiation.

Taking a wider view, Keating and Cambrosio (2007) have looked beyond the conduct of the clinical trial itself to the 'distinctive configuration of institutions, scientific practices, and materials that generates specific ways of identifying and investigating research questions, of producing and assessing results, and of regulating these activities' (199). They have termed this configuration a 'style of practice' and argue that a new style of practice has arisen in clinical oncology, instantiated by the cooperative group clinical trial systems initiated by the National Cancer Institute (NCI) in the U.S., and the European Organization for Research and Treatment of Cancer (EORTC). A key aspect of this style of practice is rooted in larger changes taking place in biomedicine, where according to Cambrosio, Keating, et. al. (2006), a novel form of objectivity has arisen

⁹ On the controversy regarding connections between cholesterol, dietary fat, and heart disease see also Garrety 1997. For another example of controversy over clinical trial results, see Jones 2000.

which they call ‘regulatory objectivity’ (189). Here they do not use ‘regulatory’ in the sense I will be using it, referring to regulatory rules external to the practice of medicine. Rather, they refer to a form of objectivity based on ‘systematic recourse to the collective production of evidence’ (189) and which ‘incorporate[s] unprecedented levels of reflexivity, in the sense that biomedical practitioners in their debates and discussions take into account the conventional dimension of their endeavors’ (190). We will see later in this thesis that while Keating and Cambrosio’s view of the formation of conventional practices is entirely consistent with mine, my concern will be the way in which such practices become formalized and inscribed in the law.

In her (1996) ethnographic study of the development of the drug Interleukin-2, Löwy demonstrates the ‘fragile cooperation between clinicians and immunologists’ (33) working on the drug. In this study, Löwy points to limits in the way constructivist accounts of drug development have been conducted in the past, noting that most of these accounts are based on documentary sources rather than long-term direct observation, and consequently ‘have suffered from the difficulties inherent in such reconstructions’ (20). A more complete response to Löwy’s concerns will be offered in the next chapter. Suffice it to say that while certainly there are limitations to historical studies of drug development, ethnography poses other difficulties — most notably that the social scientist who restricts herself to real-time observation necessarily forecloses on past events. Yet the ensemble of practices associated with clinical trials, the social attitudes and values towards them, and even the participants engaged in negotiation over clinical study have changed over time (e.g., the rise of oncology as a speciality). Hence, historical study is well situated to reveal the construction of clinical trials as an expression of (changing) social priorities as a dynamic process. As I will argue in the next chapter, the choice of method must vary according to the questions to be asked and the subject to be studied.

Related studies in medical sociology have been concerned with diagnostic and therapeutic decision-making. Berg (1992) suggests that physician decision-making is based on habit and routine rather than strictly fitting the stereotype of objective assessment of symptoms progressing linearly to diagnosis and treatment. Moreover, Anspach (1987) has shown that judgments of prognosis are rooted in one’s organizational role. A growing awareness of physicians’ subjectivity in medical decision-making has led

to studies of decision-support tools for ‘evidence-based medicine’ (Berg 1997; Timmermans and Berg 2003).¹⁰

Although much has been written on drug approval practices, a great deal of this literature is geared toward evaluating the impact of regulation on innovation, the pharmaceutical industry, and the drugs market (e.g. Dranove and Meltzer 1994; Lichtenberg and Waldfoegel 2003; Temin 1979, 1980; Thomas 1990). Much has also been written on conflicts—or conflicts of interest—between the medical profession and the pharmaceutical industry (e.g. Angell 2000a, 2000b, 2004; Bodenheimer 2000; Fisher 2003). In a more sociological context, Jasanoff (1995) has written extensively at the interface of science, law, and policy, examining how developments in science and technology become bound-up with changing social values, ending up in the courtroom where issues can range from the politics of expert scientific witnesses, toxic torts, regulation of new technologies, and science-based re-definition of fundamental social terms like ‘family’ or legal categories like ‘life’ and ‘death’.¹¹ More relevant to the current study is Jasanoff’s (1990) work on how government agencies and regulatory bodies use expert advisors and advisory committees in policymaking. Here she describes the rise of a scientific bureaucracy within government for policy decision-making and demonstrates how advisory committees (such as those used by the U.S. Food and Drug Administration) are not so much in the business of truth-finding as of social negotiation, boundary-work (see Gieryn 1983), and consensus-building for the agencies they serve (depending on the agency and the circumstance).

Other studies have looked at social and political aspects of drug approval including the difference in risk perception between medical experts and non-experts (Abraham and Sheppard 1997); peer review in the regulatory process (Jasanoff, 1985); international comparisons of science advising (Brickman and Rip, 1979); and international comparison

¹⁰ For more on physician diagnostic and therapeutic decision-making see Anspach 1993; Berg and Mol 1998; Clark, Potter, and McKinlay 1991; Frohock 1986; Fujimura 1987; Timmermans and Angell 2001. For a foundational work on medical work as related to its social organization, see Strauss et. al. 1985.

¹¹ Jasanoff has also written on how judges’ remarks and rulings crucially establish whose version of ‘facts’ is to be considered authoritative or ‘expert’ (1998). In related work, the court decision establishing standards of admissibility of scientific evidence in court (*Daubert vs. Merrell Dow Pharmaceuticals*) has been widely studied (see especially Edmund and Mercer 1996/97; Solomon and Hackett 1996). Edmund and Mercer have also written at length on tort reform and ‘junk science’ in the courtroom (see, for example, Edmund and Mercer 1998, 2000). See Mercer’s website for a full list of publications: <http://www.uow.edu.au/arts/staff/d Mercer/>. Many authors have examined aspects of expert testimony, including expert witness and DNA typing (Daemmrich 1998); expert witness and fingerprinting (Cole 1998); expert witnesses and linguistic privilege (Stygall 2001).

of drug regulation (Daemmrich 2002, 2004). Mullins (1981) has written on the historical development of the national advisory system in the 20th century U.S.

While all of this work provides valuable insight into the historical development and conduct of clinical trials in the U.S., the evolution of science advising in the U.S., and the social and political negotiations connected with drug approval, only one study in the field of which I am aware attempts to address in specific terms the advent of a relatively recent practice in drug development and approval: accelerated or ‘fast track’ approval for certain drugs. Abraham and Davis (2007) examine these practices from the perspective of outcomes, assessing the recent history of drug regulation in terms of pharmaceutical innovation, drug effectiveness and safety. Consistent with earlier work (Abraham 1995), Abraham and Davis suggest that regulatory agendas have been dominated by corporate interests (they have undergone ‘regulatory capture’),¹² and that standards for drug safety have suffered. Journalistic accounts of accelerated approval likewise ask whether accelerated approval has produced a spate of inappropriate drug approvals and drug recalls (e.g. Borchardt 2000).¹³ What seems to be lacking in any account to date is a discussion of how—in a more dynamic, sociologically well theorized and historically sensitive sense—these new categories of drug approval came about and how they alter the existing system of drug evaluation. What follows is a study of accelerated approval and drug evaluation in just such a sense.

1.3 Finitism and the Theory of Regulation

1.3.1 *Finitism and Related Social Theory*

To pursue a ‘sociologically well theorized’ study, I must define more fully the theoretical tools to be brought to bear on the subject, and in so doing elaborate the specific nature of the research questions to be considered within the overall thematic goals described above. The core of a finitist view of knowledge can be found in Barnes’ (1982) description of concept application, in which humans make case-by-case evaluations of new prospective instances of a term, I_x , based on comparison to salient characteristics of a

¹² Abraham subscribes to the notion of ‘regulatory capture’, which as we will see was elaborated by a group of regulatory theorists who advocate deregulation and allowing market forces to determine many regulatory outcomes. Presumably, Abraham does not embrace the conclusions of these theorists, since Abraham and Courtney (2007) call for enhanced regulatory oversight.

¹³ For the U.S. Food and Drug Administration’s response to these kinds of charges see Friedman et. al.1999.

stock of such instances, $\{I\}$, learned from experience. The salient characteristics to use for comparison and how close of a match is needed are both matters of conventional practice within a given community. Hence, Barnes argues that the proper application of all knowledge categories is ultimately a matter of community consensus.

For commonly used terms, concept application will be routine and seemingly natural. The dynamics are different, however, for relatively small $\{I\}$ such as, say, 'black hole' or 'solar system', where only a small stock of exemplars is available for comparison. In such cases, each new I_2 will tend to provoke conscious evaluation and comparison to salient characteristics of existing exemplars. Each new I_n added to $\{I\}$ changes the working definition of $\{I\}$ to some degree (perceptibly or imperceptibly) and also with time some exemplars previously thought to be applicable might subsequently be ejected, as recently happened when in 2006 the International Astronomical Union declared Pluto not to be a full-fledged planet after all.¹⁴

Categories of knowledge having a relative scarcity of I_n will occupy a substantial portion of this thesis. Each time a new category of drug approval is created, a very small stock of I_n exist as a reference for future decision-making. One purpose of the thesis will be to observe concept application under these circumstances. We should expect early applications of the new category to be variable and non-uniform, altering the content of $\{I\}$ itself. We can also expect that if other related categories exist — ones for which prospective I_n share some defining characteristics in common, and might have been assigned to existing categories before the new one was made — then we should see some definitional jockeying between the existing categories and the new one. If these types of shifts can be identified through new rule-making and the application of those rules to cases of drug approval, then I will have identified an important and ongoing process of creation of meaning in regulatory rule-writing and rule-following.

The example of the 'planet' Pluto illustrates Barnes' argument that concept application is 'open-ended and revisable' (Barnes 1982, 30). 'Nothing in the nature of things, or the nature of language, or the nature of past usage, determines how we employ, or correctly employ, our terms' (30). Therefore, 'nothing external determines the truth or falsity of verbal statements' (30) and 'correct' applications of concepts can be redefined at any time by a community. Every instance of use of a concept 'must, in the last analysis be

¹⁴ A quick Google search on 'Pluto not a planet' will bring up dozens of articles on this decision, such as http://www.usatoday.com/tech/science/2006-08-24-pluto_x.htm (accessed 12 December 2007). The decision was surprisingly controversial.

accounted for separately, by reference to specific, local, contingent determinants' (30).

This is, he tells us, the core precept of finitism.

To be clear, Barnes is *not* claiming that there is no basis for anyone to make judgments of truth and falsity, nor is this position an unintended consequence of his ideas. This is a mistaken interpretation frequently applied to relativist conceptions of knowledge. For an example relevant to the topic of this thesis, Abraham (1995) fell into this trap in his study of 'bias' in the regulation of therapeutic drugs when he erroneously argued that 'for a relativism, which is supposedly impartial with respect to the truth or falsity of statements made by scientists, and believes that truth is no more than a social construction, this [self-reflexivity] is especially problematic because, to be consistent, there can be no criteria for assessing the truth-value of its own claims about scientists' institutional interests, controversial views, laboratory practice or discourse' (11). Hence, according to Abraham, relativist studies of science are ultimately self-refuting and analytically impotent because relativism 'denies the possibility of error or misrepresentation' (12). However, the claim is not that there is no basis for the truth or falsity of our statements, but that there is no *external* basis for such judgments. That one word changes everything. The basis for judgment is *internal* to the relevant community of users, who form a consensus on the proper use of terms and concepts.¹⁵ Ironically, after declaring that a relativistic Sociology of Scientific Knowledge is philosophically bankrupt for the sake of his study of bias in pharmaceutical regulation (25), Abraham went on to use a thoroughgoing (and appropriately) relativistic standard of assessing bias — and did it with the help of Barry Barnes.¹⁶ While Abraham's approach turned out to be consistent in general with a relativist view of knowledge, the advantage of the specific framework offered by Barry Barnes and the Sociology of Scientific Knowledge is that it can elucidate social processes of knowledge creation *regardless of whether bias exists*. Hence, the range of

¹⁵ See Bloor's (2007) excellent paper on the nature of relativist knowledge, in which he reminds us that the opposite of relativism is absolutism. If one wishes to refute relativism, then one must propose a basis for knowledge that stands outside of human time, human culture, and human perceptions; it must be immutable, regardless of the changes taking place in human activity, regardless of history. One would have to make recourse to something like God as a source of sure knowledge to make claims to an absolute standard of truth. Human truths, by contrast, are relative to the time and culture in which they are produced, and are always vulnerable to the changing standards of human consensus. It is those standards we humans use on a daily basis to judge truth, falsity, bias, and so on.

¹⁶ Abraham cites Barnes in what he terms an 'uncharacteristically explicit reference to bias' where Barnes argued that scientific work can be criticized 'by reference to its own normal practice' (Abraham 1995, 28). Abraham adopted this standard as the core of his method, which detects bias by using conventional scientific practice as a benchmark and then 'locating and investigating *technical inconsistencies* present in scientists' claims and accounts of their actions' (29, original emphasis).

analytical conclusions available for the Sociology of Scientific Knowledge, properly construed, extend well beyond social measurements of bias.

I should also clarify that Barnes' view of concept application in no way implies an ontological position of anti-realism. This is another charge made by Abraham (1996)¹⁷ and it is also mistaken. Barnes (1977) has written that 'in rejecting a contemplative conception of knowledge and adopting a view which emphasises its social dimension, *it is important not to lose sight of the connection which does exist between knowledge and the real world*' (10, emphasis added). Likewise, on the first page of Barnes, Bloor and Henry (1996), the authors write that one of their aims is 'to show how the sociological analysis of knowledge can and must proceed on the assumption that at the basis of knowledge there lies a causal interaction between the knower and reality' (1). So, for example, when we establish terms and classifications of planets, our basis of classification is open-ended and revisable (Is it a planet, a dwarf planet, or an asteroid? Did it form through accretion or was it a comet-like space chunk caught by gravity into an irregular orbit?), but it is nonetheless a response to the observational stimulus of a big, frigid, rocky object physically orbiting our sun at a great distance away. Whatever we say about it does not alter the physical reality of Pluto's existence, only our own view of it.

In verbally formulated knowledge Barnes (1983) posits two basic stereotypes of speech acts. Natural kind terms, or N-terms, refer to objects in nature (like ducks or planets). For N-terms, concept application proceeds through a process of habituated pattern recognition. Social kind terms (S-terms), by contrast, are 'purely performative' speech acts (526), lacking any referent to the natural world.¹⁸ These are terms like 'money' or 'marriage' which exist because people believe they do and behave accordingly. In this way, S-terms are self-referential. Since there is no external referent, there is the issue of how a social practice such as marriage or money gets started in the first place, and how it grows into widespread acceptance. In Barnes' terms, the system needs 'priming': 'As with a ram-jet there is the problem of how things initially get going' (Barnes 1983, 529). S-terms form the basis for our institutional practices — indeed, they are coextensive with Bloor's (1997) definition of an institution as a 'collective pattern of self-referring activity'

¹⁷ Abraham writes that 'the social constructionists share with the supporters of the Strong Programme and the Empirical Programme of Relativism a commitment to anti-realism, as affirmed by Latour' (9), whom he quotes as if the latter were a representative of all those other schools.

¹⁸ A 'performative' enunciation is one which, simply through an act of declaration, makes the content of the declaration true. See Austin (1962).

(33). An institution can be defined by, and co-equal with, a single S-term, as in the case of money; or an institution can be composed of an interrelated network of S-terms (as in, for instance, the banking system, which relies on money, credit, interest, and a range of other socially created terms).

In general, once consensus is established and stabilized for an S-term, it appears natural and self-evident. I deposit cheques in my local bank without the slightest thought that the cheque is a surrogate for hard currency, which itself is symbolic of purchasing power in a socially constructed economic market. Only in rare moments does conscious awareness of the socially constructed nature of the enterprise intervene like a behind-the-curtain view in a Brechtian drama — as in, for example, the occasional instance when an English shopkeeper refuses to accept Scottish bank notes. How is such consensus achieved?¹⁹ Frequently, particularly in a modern democracy, an awareness of the need for that S-term precedes any performative enunciation establishing that term, although the consensus may only be local and sporadic at first. Indeed, the basic priming device of American government, the U.S. Constitution, was itself primed through a consensus process.²⁰ Thus, another task of this thesis will be to look for the *processes of consensus* underlying the perceived legitimacy of performative enunciation by the FDA. Indeed, we will see that processes of tacit consensus can effectively create ‘rules’ before they are written.

This observation leads us to our final theoretical point in this section. According to Bloor (1997), when we apply a concept, in effect we are using a rule. Likewise when we create a new category; in effect, the category constitutes a set of rules for separating objects or actions according to their character, function, or some other quality. Because rules are defined, used, and judged normatively through social intercourse, we can say that to use a rule is to participate in a custom or institution (Bloor 1997, 5). The idea of an ‘institution’, properly understood as a Barnesian S-term, encompasses the generation and application of rules and concepts.

Bloor contrasted this *collectivist* view of rules with a traditional *individualistic* account of rules. This distinction between individualistic accounts and collectivist or

¹⁹ See Kusch’s (1999) speculative musings on the gradual, step-wise series of transactions that might have eventually led to the widespread use of currency.

²⁰ By the time the Constitution was adopted in 1787, a series of incremental steps had already been taken serving to consolidate the consensus of the colonists for independence, which at first was far from unanimous. See the historical timeline of events available at <http://www.ushistory.org/march/timeline.htm> (accessed 20 December 2007).

‘communitarian’ (Kusch’s [2002] term) descriptions of behaviour is an important one for this thesis. Individualism has dominated much of traditional thought in philosophy and economics, with cascading effects into political science and legal-regulatory theory. However, as we will see below, these accounts of human behaviour tend to rely on assumptions which cannot be easily reconciled with social reality (or, as I will argue below, surreptitiously permit a mediated form of social intercourse to force the model to approximate social reality).²¹ Even sociology sometimes bars individualism at the main entrance but then permits it to be ‘smuggled in by the back door’ in treating behavioural ‘norms’ as individually internalized compulsions rather as socially sustained conventions created and altered in social practice (Bloor 1997, 6). Bloor has challenged individualistic accounts of rule-following²² while Kusch (2002) has challenged the traditional individualistic conception of knowledge in the philosophical field of epistemology, arguing that ‘empirical beliefs’ do not become ‘knowledge’ as such until they are socially shared and consensus is achieved on the status of those beliefs. For Kusch, ‘knowledge’ is an ‘institution’ in Bloor’s sense. Barnes (1995) has likewise challenged the ‘problem of collective action’ — an economically modelled social theory of group action based on

²¹ To be clear, what is invalid is the analysis of hypothetical thoughts and desires of a theoretically discrete, independent individual as the basis for describing social and group behaviour, and phenomena such as consensus or normativity, which can only be understood in terms of group behaviour. This objection in no way applies to individualism in political philosophy, where individuals are very properly accorded various rights, privileges and freedoms.

²² However not without controversy. Bloor (1983; 1997) has argued that the philosopher Ludwig Wittgenstein took a position of rule-scepticism, i.e., the notion that rules are not ultimately determinative of the behaviour they putatively guide, as opposed to the traditional individualist view that the rule itself fixes its own meaning and compels correct usage. Although Bloor is not the only one to have made such an argument (see Kripke 1982), this interpretation of Wittgenstein and view of rules has been rigorously debated within the sociology of science and elsewhere (see especially the debate which began with Lynch 1992 and Bloor 1992, and has been picked up again recently by Kusch 2004). Even aside from what the ‘correct’ interpretation of Wittgenstein might be, debate has raged over the nature and compulsion of rules, in part because Bloor has made his argument for the most difficult case, mathematical rules. An early and persistent critic of Bloor’s, Larry Laudan (1977), has written that anyone who could view a principle like ‘ $2+2=4$ ’ as socially determined or conditioned would ‘betray a remarkable ignorance in the way such beliefs were generated and established’ (200). See Bloor’s response in Barnes, Bloor and Henry 1996. (I would also direct the reader to a recent study by the Columbia University Teacher’s College (2004) which examined the cognitive abilities of an Amazonian tribe whose linguistic conventions only allowed numerical designations for ‘one’ and ‘two’, with all higher quantities designated as ‘many’. The study found that members of this tribe were unable to distinguish between two rows of identical objects, one containing four and the other containing five objects. For this tribe, $2+2$ would not equal 4, which does not exist, but would equal ‘many’ — as would ‘ $2+1$ ’. If such basic concepts were simply a matter of interacting directly with objects in nature and reading them off without any social mediation, the tribe members would have been able to perceive a difference between the two groups of objects.) In any event, my study of regulatory rule-following will demonstrate clearly the social nature of rule-making and rule-following in a legal context, and will therefore justify the validity of Bloor’s ideas as far as their application to this empirical realm, but cannot ultimately settle the argument in mathematics.

analysis of the actions of theoretically independent, rational, calculative individuals. In the next section we will look more closely at the ‘logic’ and Barnes’ response to it, because much of the contemporary theory of regulation builds on assumptions stemming from this model and variants of it.

1.3.2 *The Logic of Collective (In)Action*

Croley (1998) has noted that all the major theories of regulation in the American context are outgrowths of, and reactions to, the pluralist theory of political action which dominated political thinking in the 1950s and early 1960s. Under the pluralist model, private parties (individuals and organizations) form organized interest groups to promote an agenda specific to the collective interests of the members. In this view of political action, such interest groups are driven by narrow, private concerns. Such groups do not promote interests for the general public welfare — or, at least, if such a group promoted a policy which did tend to favour the public welfare, it would be by coincidence, not by design. Since these groups are interested to promote policies serving their own interests, they compete for influence in the political realm, where public decision-makers broker compromises between the competing groups. In this way, the legislative process and administrative decision-making were seen to reflect a balance of competing interests (32). The pluralist view was intended to be both descriptive and prescriptive: pluralists believed that the balance of interests afforded by this system was the proper way for a representative democracy to operate.

The problem with this pluralistic view, according to Mancur Olson in his seminal *The Logic of Collective Action* (1971 [1965]), is that it is a mistake to think of groups as behaving analogously to individuals — something he said many schools of social theory and political science had taken for granted earlier in the century. Individuals, he said, would consistently and ‘rationally’ act to further their own interests, whereas groups will not necessarily do so. They will do so even if it undermines the stated purpose of the group. This phenomenon is fundamentally contradictory to the theories of voluntary action which Olson says dominated social and political theory of previous decades. To prove his point, Olson analyzed group cohesion and action in economic terms, assessing the individual resources required to achieve a collective goal as compared to the individual portion of collective benefit gained. In so doing, Olson concluded that the greatest

hindrance to the formation and function of large groups through voluntary action is the ‘free rider’ problem.

In large groups, members are aware that their individual contributions to the group interests are vanishingly small compared to the aggregate; their individual contributions would not be missed if withheld. For this reason, the individual has no rational incentive to expend effort or resources for the sake of the organization, since the benefits of the organization are a ‘public good’ — meaning in this context that the goods accrue to all members, regardless of individual level of contribution.²³ By contrast, the necessary voluntary individual contribution is a private bad. Under these circumstances, according to Olson, the only ‘rational’ choice to make is to withhold an individual contribution while continuing to receive the group good — free riding.

Indeed, according to Olson, even in smaller groups where voluntary action is feasible, free riding will be a temptation for some members, leading to suboptimal production of the collective good sought. The only way large organizations can cohere is through coercion of its members (e.g. closed-shop unions) or through provision to its members of selective incentives (exclusive goods and services only attainable through membership), or a combination of the two strategies. Indeed, according to Olson, the only successful large special-interest groups are those which provide goods and services as their main activity, with political lobbying taking place as a secondary function or ‘by-product’ of the exercise of the primary function.²⁴ Moreover, even in small or intermediate groups which are able to organize through voluntary action or bargaining, the group good to be pursued will tend to be underproduced because of the free rider problem. The larger the group, the more inefficiency one would see. Accordingly, the interests of large groups of people — i.e., consumers or white collar workers — do not get represented in this way. Such large groups within the citizenry or ‘latent groups’, as he

²³ While often the phrase ‘public good’ can refer to something considered beneficial for the general public or citizenry, in the context of Olson’s theory the ‘public’ at issue is usually the membership of a specific group producing a specific good which benefits only that group.

²⁴ Olson makes an especially powerful argument using the history of labour unions as an example: unions have traditionally compelled membership through a variety of means, including coercion (barring employment for non-union members) and selective incentives such as providing health insurance benefits to members, acting as mediators for employees in labour-management disputes, and providing representation for collective bargaining. Olson argues that these latter function have, in fact, become the primary services offered by unions, with political interest lobbying representing a byproduct (however lobbying was always a secondary function – the main purpose has always been to present a united front to negotiate with industry). He illustrates the free rider problem vividly in noting that when attendance at union meetings is not compulsory (through the levying of fines), a significant number of members choose not to attend.

calls them, cannot effectively form or act on a voluntary basis, whereas small groups ('privileged' or 'oligarchic' groups) such as industry trade groups, can.²⁵ Based on this theory of groups, Olson supposed that the most effective (and therefore most powerful) lobbying groups do not represent large, general interests but are small groups representing narrow political interests, usually business-oriented ones. If so, the pluralist view of interest groups is impoverished and a different view of regulation is needed.

For Barnes (1995), the challenge to theories based on the rational, calculative, informed agents of 'economic rationality' (which Barnes calls 'ER individuals') lies in interactionist social thought.²⁶ ER individuals in so-called rational choice theory operate separately and calculate separately, according to what they consider to be their own 'rational' self-interest; for this reason they are 'incapable of identifying, creating and enforcing norms of action that redound to the good of the whole collective' (77). Individuals as seen by interactionism, on the other hand, live within interconnected networks of human relations, and are mutually susceptible to each other. In such a setting, individuals can affect each other's perceptions of what constitutes a public good, they can constitute such goods as norms, and can 'press each other toward enacting such norms through symbolic communication that threatens or enhances "face" or "dignity"' (77), i.e., through mutual symbolic sanctioning. If the good is *thought* in the collective, says Barnes, 'if it is consistently evident in speech and communication, then it is *sanctioned* in the collective' (84, original emphasis). Agreement is not action. But where agreement exists, the possibility of action follows (85). Barnes asserts that the 'manifest empirical incidence of collective action' supports the interactionist view over rational choice. While there may be a cost to action, there is relatively little cost to agreement and sanctioning. Therefore to 'put a price on the operation of the system is as inappropriate as putting a price on the operation of a single brain, for example as it thinks its way through a cost-benefit analysis' (85).

²⁵ In a privileged group, one member finds it to his benefit to provide most of the resources required to obtain the desired interest. In this way, the great are 'exploited' by the small. The example here is something like a manufacturing trade group representing firms of different sizes, but largely dominated by a few very large firms. An 'oligarchic' group is a medium sized group in which the required contributions would be roughly equal. Note that Olson is discussing only economic and political lobbying groups and their ability to influence policy. He explicitly excludes religious and other supposedly non-political groups from his assessment. I will discuss his exclusions more later in this chapter.

²⁶ Technically, interactionists do not perceive themselves as having a social 'theory', however Barnes skilfully extracts basic social-theoretic principles from the writings of key interactionists to construct what could plausibly be called a theory. See Chapter 3 of Barnes 1995.

In Barnes' hands, generally speaking, individualistic theories of organizational behaviour based on the ER individual have been dealt a severe blow. However, he has not quite dispatched Mancur Olson. In the first place, Olson's empirical accuracy is better than Barnes admits. His core assertion is not that interest groups cannot form and undertake collective action, but rather that small ones form preferentially over large ones, and tend to be more effective in producing group goods. According to the *Logic*, large latent groups (those, like 'blue collar' workers for example, which are only 'groups' in the sense that they share a common mode of existence, putting them in the position of potentially benefiting from the same economic policies) are only able to organize through special circumstances such as compulsory memberships or the offering of selective benefits. Olson's empirical examples are often compelling, especially when talking about 'closed shop' labour union practices and industry trade groups, where the model seems very much to match what has happened historically. One can also see his case in large groups such as the American Association for Retired Persons (AARP), which organizes a latent group through a modest compulsory annual fee and selective incentives such as a well-produced monthly magazine, discounts on car rental and other purchases, group life insurance benefits, and other products and services of interest to pensioners in the U.S. This combination of modest compulsory participation in return for a portfolio of products and services with lobbying as an added benefit is precisely the sort of arrangement Olson predicts for organizing large, latent groups, and many examples of this sort can be found. Examples of the free rider problem and the potential mismatch between individual interests and the public good also abound. People do often make apparently selfish individual choices antithetical to the general good of all.²⁷

²⁷ One example. The state of Georgia has been suffering a drought of historic proportions. The two primary water reservoirs north of Atlanta, on which hundreds of thousands of people depend for clean drinking water, were draining rapidly enough that the local authorities banned outdoor water use on pain of stiff fines (and in some counties, even imprisonment was contemplated). At the time of this writing (2007), no watering is permitted of grass or gardens, no car-washing or use of water on exterior surfaces of homes, etc. Authorities also asked residents to curb indoor water use to the maximum extent possible. Despite the obvious public benefit associated with water conservation, sanctions were necessary to enforce the rule. And even with the sanctions, examples of extravagant outdoor water use could be found — including one household using 260,000 gallons of water a month. For non-compliant residents uncaring of fines, Cobb County resorted to public shaming, publishing the names of extravagant water users in the *Atlanta Journal Constitution*. See 'Cobb Lists the Biggest Residential H2O Users' in published online 20 December 2007, available at <http://www.ajc.com/search/content/metro/northfulton/stories/2007/12/20/cobbwater1220nwns.html> (accessed 20 December 2007).

Even so, over time Olson's account seems increasingly to have deviated from empirical reality. Many contemporary observers note that today's political landscape exhibits more general public interest groups than ever existed when Olson wrote *Logic*. Croley (1998) performed an analysis of the composition of contemporary lobbying groups participating in various lawmaking and regulatory activities. He found that general interest groups are more prevalent than in Olson's time, and generally more prevalent than Olson would have expected. Although narrow business-oriented interests do still tend to dominate Washington lobbying, their mere presence does not unavoidably translate into results, it should be stressed. Overall, when stacked up against empirical measurements of interest group activity, Olson's general programme is still accurate, but less so than it was when he created it; exceptions to his model seem to be accumulating. The question to be answered is: why did his model seem to fit in the 1960s, but fits less so now?

Let us first look at another question. If Olson's rational individual cannot perceive the group good, as Barnes asserts, why does empirical reality match his model within a reasonable approximation? The reason is that Olson allows for a limited form of mutual monitoring to take place between his ER individuals. In a 'reasonably small' organization, if a member 'stops paying for the collective good he enjoys, the costs will rise noticeably for each of the others in the group; accordingly, they may then refuse to continue making their contributions, and the collective good may no longer be provided' (Olson 1971, 43). However, the ER individual may be aware that this will be the outcome of his actions, resulting in a loss of the collective good for all, and she might therefore choose to contribute fully — or not. According to Olson this situation is indeterminate because the 'rational member of such a group faces a strategic problem' (43). While the mutual social awareness Olson describes is not direct, but mediated by the economics of small groups (group member monitoring of cost fluctuations), it nonetheless has at its core a mutual susceptibility which is most pronounced in small groups and becomes increasingly dilute as group size increases. This mutual susceptibility is a fundamental driver of the *Logic*, and is notable for its operational similarity to interactionism. Recall that in Barnes' analysis, we have a network of mutually susceptible individuals who must be in communication with one another to establish and sanction a common conception of what constitutes a mutual good. The logistics of networking and communication are much more manageable in small groups than they are in large ones; therefore, Barnes' interactionist model implicitly tends to privilege smaller groups over large, dispersed groups. The chief reason Olson's

logic seems to work is that in this manner it coincidentally mimics the social logic of interactionism.

In Olson's model, the fundamental distinction between small group and large group dynamics is how significant of an effect free riding has on the cost of a group good to all contributing members. If this view is correct, then one would not expect any externalities to change this situation. Free riders would always be able to proliferate in large groups due to dilute economic effects, and awareness of the potential for widespread free riding would still act as a brake to large group formation and action without coercion or incentives, regardless of any changing external factors. The increasing number of counterexamples evident today therefore undermines Olson's basic premise because it points to factors other than the internal economics of groups. I would argue that one significant external change which has taken place and been influential on the process is the rapid growth in communications technologies, and especially the Internet, over the last two decades. Enhanced networking capabilities and communication over large areas has allowed the voluntary, un-coerced, non-specifically incentivized formation of large groups such as MoveOn.org. Such developments suggest that the central hindrance to voluntary large group formation is not a matter of self-regarding individuals rationally seeking to maximize their 'utility functions' and minimizing costs, but rather the ability of large groups sharing common interests to network, communicate, and form a consensus on a group good. In other words, interactionism not only explains the success of Olson's theory better than the model of the ER individual does, but it also explains why the theory is becoming less empirically valid over time.

Interactionism can also approach the question of free riders and why even when e.g. water conservation is a clear good from which all would benefit, some individuals choose to flout water restrictions in drought. The answer rests in the fact that people are mutually susceptible through interlocking and overlapping social networks, many of which are self-selected. The degree to which this mutual susceptibility influences behaviour is routinely undervalued. Even something like obesity, which we think of as a highly personalized and individualized condition, has recently been shown to spread within social networks (Christakis and Fowler 2007). Our social networks establish common patterns of behaviour, common attitudes, and opportunities for mutual confirmation and sanctioning on every aspect of life. Thus, if the social network with which one primarily identifies encourages recycling or water conservation or civic-minded voluntarism,

assuming sufficient communication to agree that these are public goods and to allow for effective sanctioning, then one will be more likely to do those things. If one is disconnected from networks with such attitudes, free riding is much more likely. Free riding does exist, then, but individuals only behave like Olson's 'rational' decision-makers when they are divorced from social networks in which they are held accountable. While coercion is one way to handle this dysfunctionality, perhaps a more effective solution to free riding is reinstallation of individuals visibly into a responsible and mutually susceptible network, as Cobb County, Georgia successfully did by publicly identifying citizens egregiously flouting water conservation rules (see note 27 above).²⁸

A great deal more can be said about the *Logic* and rational choice theory, especially regarding the definition of 'rational' decision-making. Such utterly self-regarding calculations would only be 'rational' were people truly divorced from one another socially. In actual practice, people's economic calculations can and do include social ones. Indeed, if anything I believe that because the economic system is designed on the assumption that social calculations do *not* take place, such calculations tend to be suppressed artificially, confined by a regime of programmed rationality.^{29,30} I will return to the subject of rationality in the conclusions. For now, I have achieved the goal of bolstering Barnes' argument that interactionism provides a better basis for understanding group behaviour than rational choice theory does. This fundamental observation, in combination with the

²⁸ Water use of the worst violator dropped 70 percent after the publicity.
<http://www.wsbtv.com/news/14587528/detail.html> (accessed 5 February 2008).

²⁹ For example, many years ago I worked in a laboratory in which a young, relatively new employee needed to take several weeks away from work for a problem pregnancy but lacked the accrued leave time to do so. She would be forced to take leave without pay when she could least afford to do it, with medical bills and a new baby on the way. Some co-workers and I approached the management of the company with an offer to donate some of our vacation days to her. We reckoned that a few days from each of us would cover her absence. Our offer was examined and re-examined, but ultimately the accounting systems would not allow it. The proposal might even have run afoul of labour laws. A seemingly simple, 'irrational' idea like donating vacation days to a needy co-worker was something never contemplated by the designers of accounting systems and laws, and was therefore impossible to do under the regime of programmed rationality.

³⁰ See MacKenzie's (2006) study of how financial models performatively shape financial markets. MacKenzie's focus is models which are intended to describe financial activity but end up shaping it. While my comments here do not address models in such a formal sense, one could argue that the assumptions imbedded in accounting systems for payroll and benefits effectively constitute a model. Still, MacKenzie's models were intended as descriptive while the payroll and benefits models are *designed* to constrain behaviour. But I believe the system designers had a mental image of social reality in the workplace which precluded altruism; thus, the designers believed their assumptions reflected existing social reality, but they in fact created and enforced an alternative social reality to the one that would have otherwise existed.

social theory described above and the empirical study to follow, will tend to undermine the fundamental assertions of much (but not all) of contemporary regulatory theory.

1.3.3 *The Logic of Regulatory Theory*

Our final theoretical task is to introduce, briefly and in summary form, the major contemporary theories of regulation as categorized by Croley (1998). For now, I will make only limited comments on these theories, since their evaluation must take place in the light of the empirical material to follow. It is important to say that what is presented below as the ‘four major theories of regulation’ are actually distillations of constellations of theories having similar points of view or attributes. This level of detail, while relatively schematic, should suffice to draw useful contrasts with the empirical work which follows. Moreover, I must underscore that the research questions of regulatory theorists are not my current research questions. Often contemporary regulatory theory seeks to *predict* regulatory outcomes — not only what regulatory agencies will do, but what will be the future social ramifications of the action taken. Who wins and who loses in the game to influence policy? Certain outcomes may be able to be anticipated since, as we will see, decision-making often follows definable trajectories through time. Nevertheless, overall this thesis tends to undermine the idea that regulatory outcomes are routinely and meaningfully predictable, since trends identifiable in retrospect cannot reliably be anticipated on a prospective basis. Additionally, regulatory theories often have a normative or prescriptive purpose; they frequently seek to make policy prescriptions for regulatory reform. By contrast, this thesis is descriptive in its orientation. The central purpose is to show how a finitist conception of knowledge illuminates the social and conceptual dynamics of this period in FDA history, and correlatively to argue for the relevance and validity of finitism in application both to human action in history and to regulatory theory. While it may be able to suggest useful policy prescriptions for regulatory practice, this thesis is simply not designed to propose changes to the regulatory system. What this thesis will show is that most contemporary theories of regulation fail to describe key aspects of the regulatory action we see in this period of FDA history.

While there have been some attempts to modify or improve Olson’s theory, by and large the logic of collective action has been embraced by regulatory theorists (especially economists, but others also) as the fundamental framework to address the problem of collective action in politics and government (Croley 1998). In this view,

public services such as the building and maintenance of roads and sewer systems or the provision of national defence and police cannot be provided through the voluntary action of individual citizens self-organizing collectively. Coercion is necessary to raise the funds for such projects (taxes) and delegation to smaller bodies is required. Thus, citizens delegate to legislators. And legislators ‘face their own set of collective action problems’ (22) by delegating to regulators to accomplish many of the tasks needed to sustain a large technologically advanced democratic system. But delegation has its costs. How can the citizenry know that the legislators are carrying out their duties according to the ‘will of the people?’ How can the legislators know that the regulators are carrying out the will of the legislature, rather than pursuing their own interests? Such ‘principal-agent slack’ can only be reduced through monitoring the activities of the agents by the principals (23). But monitoring is costly and, following Olson’s logic, principal-delegators will only monitor for agent ‘shirking’ or deviation up to the point that the marginal cost of monitoring equals the marginal benefit of doing so (24). As Croley noted, ‘Thus can the regulatory regime be well understood as an institution that simultaneously responds to, and is plagued by, the costs associated with collective action and with the inevitable principal-agent slack created by delegation of regulatory authority’ (24). For Croley, the ideal regime would minimize costs of collective inaction and the costs of delegation such that ‘[w]here collective inaction would be more costly than delegation, regulatory power should be delegated’ (24); but where delegation would ‘generate more ills than it would avoid’, it should not take place. I.e., under these circumstances the ‘market failure commonly said to justify administrative regulation should be tolerated’ (24).³¹ In sum, then, ‘any serious theory must somehow consider that collective goods, including monitoring, will in the absence of some catalyst tend to be underproduced’ (24-5). These are the concepts which tend to underlie contemporary theories of regulation. Let us now briefly turn to each major constellation of perspectives in regulatory theory.

³¹ This view of regulation as serving to protect consumers from market failures is the traditional view of a key purpose of regulation (see Breyer 1982). Indeed, the beginning of market controls were initiated in the late 19th and early 20th century in the U.S., which was characterized by technological advancements, an expanding industrial base and (relatedly) the rise of the ‘Robber Barrons’, an invective for the wealthy, monopolistic industrialists of the era popularized in a 1934 book (Josephson 1995). This period of history is a case study in what happens when there is little economic regulation: monopolies and collusion to increase prices of goods. The government response was ‘trust busting’: the *Sherman Antitrust Act* of 1890 and the *Clayton Antitrust Act* of 1914. In this period, what we now euphemistically call ‘market failures’ would have likely been termed ‘corporate greed and misconduct’.

The so-called *public choice* theory of regulation was first developed in the 1970s, in part as a repudiation of pluralism.³² The seminal work for this approach is Stigler (1971; 1974), however those familiar with critiques of FDA regulation in this period will especially recognize Peltzman's (1973) development and application of the theory (see also Peltzman 1976; 1998).³³ The essence of public choice theory rests in the assumption that the regulatory system can be modeled as a market, with economic principles of supply and demand guiding the provision of regulatory 'goods'. Such regulatory 'goods' are demanded by those who stand to gain from them. Legislators seek to provide those goods to the demanding parties because in so doing they will reap political benefits. The resources necessary to meet the legislator's political needs constitute the 'price' of regulation. Since the logic of collective action dictates that only well organized narrow special interests are able to assert their demands in this way, it is narrow interests which prevail, not those of consumers and other 'general' interests. According to Stigler, individual voter-citizens simply would not have the incentive, for example, to make modest voluntary campaign contributions to politicians since the potential individual returns are too small to justify the cost. Nor would citizens have the incentive to form an organization to represent their interests and perform the monitoring required to see that their agenda is carried out legislatively. Hence, even if legislators were taking into account the interests of the public, there would be little or no monitoring of their activities by the electorate. In the citizen-legislative relationship there is therefore a great deal of slack, and organized interests are able to capitalize on it. This situation leads to 'regulatory capture' — the eventual alignment of Agency goals and interests to those of industry groups. Thus, the primary principal-agent relationship is the one between legislators and regulators. In this relationship, Public Choice treats the regulatory agency as a black box and assumes that legislative mandates (as influenced by small, well organized interest groups) transmit directly into regulatory outcomes. Stigler's ultimate conclusion is that

³² Posner's (1974) older review of theories of economic regulation discusses two main theories: public interest theory, which he defined in terms very similar to Croley's; and 'capture' theory, the economic version of which is equivalent to Croley's public choice theory. In one prominent version of the theory, 'capture' refers to a process whereby regulators increasingly align their interests with those of the industries they are tasked to regulate. The latter version is consistent with the conclusions of the Chicago school of public choice. Hence Levine and Forrence (1990) refer to public choice as 'capture' theory.

³³ See notes 76 and 77 in Croley (1998), which offers a detailed summary of the various branches of public choice theory. In particular, a Virginia school of public choice has conducted significant economic analysis of interest group behaviour in 'rent seeking', meaning the pursuit of government-granted monopoly rights. See the excellent summary in Mitchell and Munger (1991).

since regulation places the power to demand legislative (and consequently regulatory) outcomes in the hands of special interests, it would be better for consumers to limit regulatory power and accept market outcomes, imperfect though they may be.³⁴

Neopluralist theory arose as a critical response to public choice theory (Croley 1998, 58). Like the pluralism of old, this new pluralism holds that interest group competition for regulatory resources will take place resulting in a relative balance of regulatory outcomes, as opposed to the winner-takes-all posture of public choice. However, analysts within this overall point-of-view disagree over the extent to which balance is achieved. Becker's (1983) version is most similar to traditional pluralism, in which interest groups determine the level of effort they can afford to spend in lobbying for a certain good, and their efforts are successful up to the point that resistance is offered by other, opposing groups. Becker's approach is particularly notable since he came out Stigler's Chicago school and used Stiglerian economic analysis of interest group activity as his starting point (see the excellent summary in Mitchell and Munger 1991). In this view, regulators are effectively constrained by the balance achieved between competing interest groups. The most 'efficient' policies (the ones demanded most) will be most successful and winner-take-all can only occur if no other groups resist. Other theorists allow for more uneven distribution of regulatory goods to interest groups (e.g. Stewart 1975).³⁵ As much as legislators would like to cater to them all, it would be impossible to do so. Therefore, legislators act as 'entrepreneurs', brokering compromises and deals in which some groups often do somewhat better than others. According to neopluralists, appropriate regulatory reforms would aim to curb excessive interest-group influence and to require fuller consideration of underrepresented interests.³⁶ (Nevertheless, with respect to the FDA, Becker (2002) ends up taking a position notably similar to his Chicago colleagues in arguing that the best way to reduce drug prices is to 'get the FDA out of the way' by rolling back requirements for effectiveness.)

³⁴ See also Cohen and Stigler (1971), in which can be found a straightforward, non-technical statement of this position.

³⁵ Croley (1998) also places Reich (1985) in this category, but then later cites the same paper in his discussion of Civic Republican theory. Obviously there are shades of grey in these groupings, with overlap between certain categories possible.

³⁶ Notably, Croley (1998) wrote that, 'To overgeneralize a bit, many economists subscribe to one version or another of the Stiglerian model of regulation, while many legal scholars embrace one form or another of the pluralistic neopluralist theory' (58). While a generalization, it seems reasonable to believe that one's posture towards regulation is conditioned by discipline-specific training and professional commitments and corresponding social networks.

Becker is also notable in the context of this thesis because he attempts to incorporate elements of 'social currency' in his economic analysis of, e.g., altruism in family groups (1981) or altruism as an evolutionary adaptation for genetic fitness (1979). While these ideas superficially resemble some of the arguments made in this thesis, there is a fundamental difference. At its root, Becker's approach remains an individualistic one: he posits an isolated individual making calculations to maximize her own utility. That for Becker these calculations may include social considerations makes it no less of an individualistic perspective. To analyze self-interested decision-making as a simple market transaction in social currency is to impoverish the profound inter-susceptibility of social interaction.

Croley's third general grouping of regulatory theories is *public interest theory*, which also seeks to challenge public choice. Proponents of this theory (see especially Levine and Forrence 1990) observe that some public-interest policies do prevail, and ultimately conclude that regulatory outcomes often, if not always, serve to mitigate market failures and serve the interests of citizenry (see Croley 1998, 65-6). Unlike the theories summarized so far, public interest theory considers regulatory outcomes not only in terms of special interest group motivations, but also in terms of the motivations of the regulators and citizenry at large. Because of the focus on relationships between agencies and constituents in this literature, Levine and Forrence (1990) refers to this constellation of perspectives as 'agency theories' (170).

According to public interest theory, citizens desire certain policies, but they also want to pursue many other goals besides those policies. The pursuit of regulatory outcomes is costly in terms of time and resources, and therefore conflicts with the desire to pursue other goals. For this reason, citizens have a reduced stake in the pursuit of specific regulatory outcomes. Interest groups are created specifically to pursue certain policy outcomes, so have a higher stake in the process than citizens. Regulators have many goals, the first of which is job retention, but they can also be seen have having other-regarding aims. In general, they seek to pursue the general interest as they understand it. However, while sometimes regulators may agree on which course of action best represents the public interest, at other times they may disagree. Under these circumstances, the regulator may choose to behave in a 'Burkean' manner, doing what she deems is best for a misguided citizenry.

Public interest theory also attempts to account for environmental factors in this process. The first such factor is political competition. According to the theory, for regulators to remain employed they must garner political support, either from politicians, or special interests, or the citizenry (or a combination). Secondly, regulators must be aware of principal-agent slack. Regulators have a choice of interests with which to align themselves for political purposes, or to act in favour of the citizenry, or to act in a Burkean fashion. Since there is more likely to be slack in the citizen-regulator relationship, acting in a Burkean fashion is unlikely to have political consequences for the regulator. However, a high degree of monitoring by special interests assures that political consequences follow from these groups. Public interest theory predicts that when slack is low (conditions of high visibility), regulators have every interest to pursue the public interest as they conceive it, whether from other-regarding concern or from a desire to reap public support. Thus, the theory predicts that 'hot' public issues are likely to be resolved in a way which favours general interests. Where a great deal of slack exists, so too does ambiguity as to the outcome. Therefore, the theory recommends creating slack-reducing measures to keep regulatory action in the public eye.

The final group of ideas in this scheme is called *civic republican theory*. This set of ideas works on consensus model rather than a competition one, and really constitutes not so much a descriptive-predictive model as a normative one. A core assertion of this approach (which is remarkable for its consonance with interactionism) is that the preferences of actors participating in a negotiated decision-making process take shape in the course of decision-making. Individual preferences for a given solution to a problem are not necessarily well defined and fixed in advance as other major theories of regulation tend to assume, but rather develop and become clarified in the process of defining a problem and negotiating collaboratively to arrive at a mutually agreeable solution (Reich 1985). In the process, regulators act as mediators among the various parties, and indeed their own judgment of the issue can be changed in the process. The result is a sort of equilibrium among participants. However this equilibrium is not a result of conflict, competition, and struggle, as for the other theories, but represents the collective consensus on the best regulatory outcome to pursue. Where and how this negotiation takes place is a matter of some discussion. Sunstein (1990) identified the Office of Management and Budget as one important centre of negotiation. Kelman (1987) similarly pointed to the use of the *Negotiated Rulemaking Act* of 1990 and the use of federal advisory

bodies as important mechanisms by which negotiation can take place. Reich (1985) described a case where regulators actually held public meetings in Washington state in an attempt to educate citizens and ultimately ask them to make the regulatory decision which would affect future activity in their community. Clearly for proponents of civic republicanism, the goal is not to abandon regulation (as in public choice) or to create a low-slack environment in which regulators do not feel free to seek their own self-interest (as public interest advocates), but rather to encourage widespread public participation in regulatory decision-making to give it greater footing with respect to organized interest groups.

This brief description of prevailing theories of regulation gives the reader a sense of the major concerns of these theorists and the approaches they tend to take to the subject. As we wade into the empirical material of the following chapters, I will seek to make observations pertinent to these theories at appropriate points in the thesis. At the same time, I will be testing and developing the finitist framework described above. By the time we reach the conclusions, we should be able to mount a critique of these theories, not in terms of their own research questions, but in terms of mine. In so doing, I can suggest the outlines of a social theory of regulation.

1.4 Plan of thesis

In the next chapter, I will give a detailed explanation and justification for my methods, sources, and theoretical perspective in this thesis. Then, Chapters 3 through 8 present the empirical historical account of recent FDA history from 1962 almost up to the present with emphasis on the development and consolidation of the procedures and standards for the approval of drugs, especially drugs intended to treat life-threatening diseases.

Chapter 3 covers the period from the 1962 Kefauver-Harris amendments to the Food, Drug and Cosmetic Act through the late 1970s. The 1962 amendments required for the first time that drugs be proven both safe and *effective* before they can be marketed. In the aftermath of these amendments, the FDA gradually made the transition from a toothless and docile agency, generally compliant with industry³⁷ to one which could hold

³⁷ According to Hilts (2003), the FDA was especially compliant and submissive under the leadership of George P. Larrick, who was FDA commissioner at the time of the 1962 Amendments. Even with the rather impotent position occupied by the FDA prior to 1962, the FDA reviewer Francis E. Kelsey

its ground against a barrage of challenges to its authority. In this period, a struggle between industry and the FDA over what the 1962 amendments should really mean in practice was repeatedly played out in the courts, with the FDA's position and authority consistently upheld. One result of this struggle was a new set of rules defining what 'substantial evidence' meant — more than one randomized controlled trial demonstrating the safety and effectiveness of the investigational drug. Not surprisingly, meeting the new standard would dramatically increase the time and cost required to bring new drugs to market. In the 1970s, economists and others joined the industry's chorus of critics. The 1970s became a decade of hearings, studies, and reports on FDA performance and regulatory outcomes, culminating in a series of legislative and regulatory proposals for reform — proposals in which a growing tension between data-gathering and early decision-making became increasingly evident. In this chapter, we will see the beginning of the argument that rule-making is retroactively patterned on practice (i.e., on tacit rules around which a significant consensus has formed), and that any socially plausible theory of regulation must account for the formation of this type of consensus and its relation to rule-making.

The reform efforts of this period would likely have progressed at a moderate pace (and did for the next few years) were it not for the collision with the next decade. Chapter 4 moves us into what is often thought of as the 'AIDS era'. In the 1980s, the Chicago-influenced economists and their intellectual kin found a powerful political ally in the newly elected president, Ronald Reagan, who ushered in an era of aggressive deregulation and sought to control the previous discretionary judgment of agencies to make regulations. Then this pro-industry, anti-regulation bloc was ironically joined in their criticism of the FDA later in the decade by militantly vociferous gay men (among others) clamouring for more rapid development and approval of drugs to fight AIDS (Epstein 1996; Edgar and Rothman 1990). The reform efforts of the previous decade, only partially implemented by the mid-1980s, kicked into high gear, resulting in lightning speed approvals, massively expanded programmes for pre-market access to investigational drugs, and a series of regulatory initiatives based on those experiences. Chapter 4 examines in detail the first instance of AIDS drug development and approval in 1987 (AZT) and its aftermath, showing how the first regulation written to expedite development and approval of drugs

nevertheless managed to delay the marketing of thalidomide long enough for revelations about the drug's teratogenic effects to come to light.

to treat life threatening diseases (the Subpart E rule, published in 1988) was retroactively modelled on the experience with AZT. Examination of these events will reveal how processes of practical consensus can create de facto rules which can become the basis for later rule-writing. We also see how key concepts ('treatment IND', 'risk', 'Phase II') changed meaning or became conceptually unstable as a result of the events of this period.

Chapter 5 opens a window to drug development and approval decision-making for cancer drugs during the same period. While the advisory committee making decisions for AIDS drugs found itself in a pressure cooker of public scrutiny, it is instructive to see what types of decisions were being made in the much calmer oncology advisory committee. Three instances of oncology drug approval taking place between 1987 and 1990 will be examined. In so doing, we find similarities to AIDS drug decision-making made in the same period. Indeed, decisions made for these cancer drugs will bend the rules in a manner which anticipates legislative and regulatory reforms not codified until several years *later*, one of which is explicitly modelled on *AIDS drug* decision-making. Clearly, in practice the basis of consensus for these new rules extended beyond AIDS decision-making. Moreover, I demonstrate that some aspects of these decisions contradicted the standards promoted by FDA decision-makers just a few years beforehand. This chapter provides further support for the contention that a consensus or de facto rule for the pertinent laws and regulations was established in practice prior to any formal rule-writing, and that the standards and concepts for drug approval necessarily vary on a case-by-case basis. They did so not because of withering public scrutiny and protest over each meeting, as for AIDS drug decision-making, but out of a growing sense (articulated a decade earlier) that greater risks could be tolerated for the sake of a perceived benefit for seriously ill patients lacking effective therapeutic options.

Chapter 6 returns to AIDS drug decision-making, discussing in detail the next two AIDS drugs to be approved (ddI and ddC, approved in 1991 and 1992 respectively) and their relationship to the next new rule for accelerated approval of drugs on the basis of surrogate endpoints (the Subpart H rule, published in 1992). We find that ddI was approved under circumstances which significantly bent the existing rules and that the Subpart H rule, promulgated a year later, used ddI as a model. Notably, ddC was then the first drug approved under Subpart H accelerated, but the approval went forward when the Subpart H rule was still a proposed rule, with the FDA yet to receive comments and respond to them, suggesting once again the formation of a tacit consensus achieved in

practice. Additionally in this chapter we begin to see how and why Subpart E began to shift in meaning. Originally a specific set of procedures for expediting drug development, I argue that the advance planning required by the rules was impractical when considered in the light of how intermediate phase trials were often conducted for serious diseases (as evidenced in the cases discussed in Chapters 5 and 6). Therefore, it began to be adapted by the FDA to serve more as an abstract set of principles or a philosophy for approaching certain cases of drug approval.

Chapter 7 reviews in detail two more cases of cancer drug approval, this time chosen specifically because they were the first two applications of the Subpart H accelerated approval rule to cancer. These cases demonstrate that as soon as these new procedures were brought to bear on new situations, the application of the rule required significant adaptation and flexibility. By this time in the account, it is clear that the creation of each new rule is founded on at least one existing exemplar (but often more than one), and that subsequent encounters with new prospective applications of the rule invariably require case-by-case adjustment, as would be expected under finitism. Chapters 5 and 7 also demonstrate that regulatory outcomes are in no way a straightforward output of the legislative will. In the consensus-building forum of the advisory committee, regulators have an important role in shepherding the discussion and influencing the opinions of their advisors. Moreover, the ‘interests’ of the participants in these meetings are not reducible merely to ‘bias’, nor to ‘Burkean’ vs. self-interested behaviour.

Chapter 8 tells the story of the circumstances leading to the creation of the *Food and Drug Administration Modernization Act* (FDAMA) of 1997. The events leading to the creation of the FDAMA undermine the notion that special interests always prevail over generalized interests (the FDAMA was demonstrably a product of compromise and moderation) and also that there is anything resembling a univocal ‘will’ of the Congress in its direction and oversight of the FDA. My focus in telling this story is on the provisions of the Act related to so-called Fast Track drug approval and expanded access of investigational drugs. Within these areas of interest, the Act is notable because of the extent to which it effectively directs the FDA to do what it had already been doing. The Act adopted elements of Subpart E and especially Subpart H into its ‘Fast Track’ drug approval provisions and specifically authorized the FDA to expand access to seriously ill patients using all the mechanisms the FDA had created over the previous decades to do so. The Act also specifically authorized the informal FDA practice of permitting drug

approval on the basis of one clinical study, a practice the FDA had long sought to keep informal. In providing 'legislative support' for existing FDA practices, the Congress was essentially sanctioning previously established FDA rules and practices as 'precedent'. Once again, rule-writing was following practice. This chapter further undermines regulatory theories which view the regulatory application of laws as relatively unproblematic. I will show that the FDAMA created a new category of drug approval which overlapped with existing ones such that it was difficult to establish clear boundaries between the new categories and the old. The FDA's solution to this problem was to modify the definitions of related categories. Additionally this chapter discusses one case of failed drug approval to demonstrate that the application of the concept 'Fast Track drug' was anything but routine, even when the users of that concept were FDA drug reviewers.

In Chapter 9 I will endeavour to bring the entire story together, drawing conclusions designed not only to demonstrate the utility of meaning finitism and other social theoretic tools, but also to expand its applicability and develop new tools in thinking about the relationship of regulation, tacit practice and consensus, and knowledge-making. In this chapter, I will also draw some tentative conclusions for theories of regulation in general. A chief conclusion will be that most contemporary theories of regulation suffer a lack of empirical validity because they are mired in the individualistic social theoretic perspectives represented by works like Olson's *Logic*. I will argue that the collectivist-finitist perspective offered in this thesis provides a superior basis for understanding regulatory decision-making and outcomes, and will provide the foundations for a collectivist-finitist theory of regulation.

2. RESEARCH DESIGN AND METHODOLOGY

McDowell (2002) notes that while historians and social scientists both use generalizations, the nature of those generalizations is quite different. Historians tend to focus on specific contingent, unique events, using a generalizing logic only in the sense of placing those events in a larger social and historical context, while sociologists tend to work towards general trends and theories of social structure and behaviour. Sociologists often use surveys, questionnaires, interviewing of a certain sort, participant observation and other methods, creating models which for many historians efface the specificity and contingency of actual human actions, and describing these models in ‘obscure language’ (17). However, argues McDowell, social scientists ‘cannot always focus exclusively on contemporary social structures without considering the historical background of the subjects they investigate’ (19); the ‘strength of explanations in the social sciences is bolstered through a recognition of the interplay between unique events’ (19). In this interplay, certain patterns can sometimes be identified, although many historians have been reluctant to ‘acknowledge the usefulness of the social sciences in helping us to explain the nature of historical development (20).

This research programme embraces the kinds of contrasts discussed by McDowell: I seek to explain regulatory development through examination of particular events. My concern is theoretical in that I hope ultimately to describe the regulatory process in finitist generalities. Nevertheless, as the reader will see in the following chapters, unlike the traditional social scientists characterized by McDowell, I share with historians an intense concern with chronology, as well as a significant concern in maintaining the particularity of the events I describe. Finitism is unlike many social theories that it emphasises contingency. In this aspect of theoretical exposition, finitism is more compatible with the historical perspective than many other sociological frames for generalization. This aspect of finitism perhaps explains the tendency of those of us sharing this theoretical perspective to conduct historical studies. Hence, while my theoretical aims are

sociological in nature, my research tools are primarily historical — appropriately and necessarily so.

2.1 Documentary Sources and Source Selection Methods

The thesis describes events from 1962 to the recent past. It therefore calls for historical research methods to reconstruct bygone happenings. Documents have overwhelmingly served as the basis for research in this thesis. Accordingly, and in keeping with accepted practice in historical research, source documents must be evaluated in terms of the circumstances and purposes of their production, their authenticity and accuracy, authorial authority, and the competence and trustworthiness of the observer whose testimony is recorded within (Howell and Prevenier 2001). The selection of sources, interpretation and use of the material found within documents will be driven by all these concerns. In the subsections which follow, I will briefly discuss the sources of information available to me in each category, considering each in terms of the historical source evaluation criteria noted above. Additionally, for each source used in the thesis I will describe the kind of information sought, i.e. the rationale for selection and use of the document.

2.1.1 *A Note on Electronic Sources of Information*

Electronic copies of laws, regulations, Congressional hearing transcripts, journal articles and many other documents are in abundance on the Internet. Moreover, the contemporary ability to electronically replicate text and post to the Internet is virtually boundless. This relatively new electronic freedom-to-replicate can be both a blessing and a curse. On the one hand, obtaining reliable electronic copies of important documents can be relatively quick and easy; on the other hand, the average researcher does not possess the expertise to establish the provenance of electronic documents found on the Internet. Copies can be made by scanning paper documents, in which case the paper document used could be the original, a photocopy of the original (some of which can be very high quality), or a some version of a fake or tampered document presented as authentic. Scans can be made into image files, into Adobe portable document format (PDF) files, or into text documents readable by word processing software. Errors are particularly likely in the early versions of scanned text documents in which text

recognition software ‘read’ the text and converted it to a word processing document (this technology began to be available in approximately the mid-1990s); files produced in this manner are often littered with misidentified letters and words. Word processing documents can also be directly converted into PDF files (no scanning necessary), allowing endless opportunities for revision, tampering, and typographical errors prior to conversion. With the proper software, PDF files can also be modified or merged with other PDF documents after file conversion.

Faced with this dizzying array of possibilities for document (mis)handling, the researcher’s best defence is to seek databases and websites known for reliability. My particular research programme tends to straddle the decades in which electronic documents began to be available. As described in more detail below, most sources of information from the 1960s and 1970s are not available electronically; sources from the 1980s are mixed, with many key sources still not available electronically; for references from the mid-1990s onwards, a great many of sources of information (but, it is important to stress, not all) are available electronically.

2.1.2 Legal, Regulatory, and Congressional Sources

The general legal framework under which Food and Drug Administration (FDA) and drug sponsors operate was established under the Food, Drug, and Cosmetic Act of 1938 and its legislative amendments. Hence, the most fundamental sources of information regarding drug approval are the texts of these Congressional acts which are written into the U.S. Code as law. As much as these laws form the basis for drug development and approval in the U.S., however, it would be naïve to assume that the letter of the law as written has an unambiguous meaning or is followed without controversy. As we will see in subsequent chapters, the FDA’s regulatory interpretation of the 1962 amendments has been disputed ever since it was written

The authorship of each of these acts is obviously the U.S. Congress, which has authority under the U.S. Constitution to pass laws. Each of the drug laws ultimately serves as a means of regulating U.S. interstate commerce — primarily creating an institutional, rule-guided framework within which to control the purity, quality, and truthfulness of claims associated with therapeutic drugs marketed in the U.S. These laws are ‘factual’ in the sense of having a determinate legislative history and circumstance of production (e.g., the often noted relationship between the 1962 amendments and

thalidomide), and in the sense of providing information about the ‘intended’ framework for FDA action. While present-day versions of these laws (including amendments) can be obtained directly from the U.S. Government Printing Office (www.gpoaccess.gov) or from the Cornell University Law School (<http://www.law.cornell.edu/uscode/>), for my purposes it is more useful to understand the circumstances of their original production.

The Congressional process for creating major laws is rather lengthy and complicated, but produces a series of documents useful for understanding the political issues implicated in lawmaking, among other things. Proposed legislation will typically be associated with Congressional hearings, and the transcripts of those hearings (along with documents and other materials submitted to the Congress as part of the hearings) are published by the Government Printing Office (GPO). Proposed legislation is negotiated and revised behind closed doors by one or more Congressional committees or subcommittees. If the committees can reach consensus on a final version of the proposed legislation, the bill will be returned to the full House or Senate for debate and voting. In so doing, the responsible House and Senate committees will often produce a report to their Congressional colleagues detailing the provisions of the legislation and the intentions or rationale behind the bill. Competing versions of the same bill can be introduced in both the House and Senate, and it is the job of the respective House and Senate committees to reconcile the differences between competing bills within each Congressional arena and, ultimately, to reconcile between Senate and House versions of the bill. A proposed bill can go through multiple iterations before a final agreement is reached.

The outcome of this process for the historian is several types of documents worth consideration: texts of proposed legislation, Congressional hearing transcripts, and House and Senate committee reports. Many committee reports can be found online through the Library of Congress from about 1995 onwards. Reports prior to these dates must be sought in Federal Depository libraries. The Congressional Record contains transcripts of the debates on the floor of the House and Senate, though it must be kept in mind that speeches made from the floor of the House or Senate are consciously made for public consumption, and can easily misrepresent a Congressperson’s behind-the-scenes agenda on a given piece of legislation. The print version of the Congressional Record is widely available at Federal Depository libraries and electronic versions from 1989 can be obtained online through the Library of Congress website. Summaries of House and

Senate reports and legislative histories are also available through the subscription database Lexis-Nexus Congressional from 1969; availability of full-text versions is variable.

Congressional oversight hearings of the FDA can be very rich sources of information, and often allow a rare opportunity to learn about decision-making processes typically hidden behind closed doors. However, interpreting transcripts of Congressional hearings is no straightforward task. The arena is by definition a political one, as well as a highly visible one. One must bear in mind which political party is in power, what circumstances have prompted the hearing, who is asking the questions, and who is answering them. While the rhetoric of bipartisanship is often heard, political agendas frequently condition the nature of questioning. Moreover, one must consider the wider historical and political climate in which the hearings take place. The would-be interpreter of Congressional testimony must also take into account personal agendas of the witnesses. Congressional representatives frequently posture for the media, hoping to get a quotation into print. Additionally, although witnesses testify under oath and are usually called because of their expertise on the subject at hand or because of their position as a direct witness of the events of interest, not all of their statements can be taken at face value. Hence, the political context and actors involved must be well understood for an insightful approach to Congressional hearings. Finally, it should be noted that both the questioners and witnesses in Congressional hearings can cite 'factual' information (statistics, accounts of events, assertions of scientific findings, etc.) which later turns out to be incorrect. Hence, any intriguing 'facts' arising from Congressional hearings should be cross checked for accuracy, if possible.

Congressional hearing transcripts can be found in Federal Depository Libraries, however the collections of transcripts vary widely from library to library. Online versions of recent hearings can also be obtained through Congressional subcommittee websites, and also through the Lexis-Nexus government database. The quality of transcripts can vary widely. In particular, transcripts obtained from Lexis-Nexus (which are often text files, not PDFs or scans) can be riddled with typographical errors and often provide only partial testimony, not full transcriptions. The most reliable versions of these documents are the print versions published by the Government Printing Office and supplied to Federal Depository libraries. These also have the advantage of containing photocopies of appended printed materials submitted to the Committee as supplements to hearing testimony, which the Lexis-Nexus versions do not. Reliable PDFs of more recent hearings

(within the last 5 to 10 years) can sometimes be obtained directly from the website of the Congressional committee which held the hearings.

Acts of Congress create the general framework for action while the specific technical details as to how these laws should be carried out are left to the Food and Drug Administration, which has been empowered by Congress to write regulations. For this reason, the FDA produces detailed instructions intended to comply with the directives given in the U.S. Code. Current versions of these regulations are published and updated regularly in the *Codes of Federal Regulation* (CFR). Title 21 of the CFR is reserved for regulations pertaining to Food and Drugs. The CFR is available online at the Government Printing Office website, www.gpo.gov, as well as at Federal Depository libraries in the U.S.

The FDA carries authority of authorship for these regulations, but it would again be incorrect to believe that it carries out this authority without challenge; or, indeed, that the application of these rules takes place without deviations, without controversy, and without disagreement as to which decision is consistent with the rules as written. Indeed, these deviations and disagreements form part of the basis for this study. Hence, the regulations as written have provided an important source of information for this research because: 1) regulations function as a baseline to which many of the actors in the historical account refer to justify actions; 2) the writing of new regulations can be analyzed in terms of the motivations, goals, and issues most important to the actors at the time of new rule-writing; and 3) the actual application of these rules, and apparent deviations from them, reflect changing priorities for the actors in the historical account.

Another crucial source of regulatory documents for this research has been the *Federal Register*, which is a daily publication of the Federal government in which agencies publish proposals for new rules, responses to comments on rules, announcements of final rules to be published in the CFR, announcements of public meetings, announcements of the publication of new guidance documents or requests for comments on draft guidance, and corrections to previous publications. The *Federal Register* also publishes Executive Orders issued by the U.S. president and other such announcements. While the CFR is a straightforward publication of finalized rules with no commentary or supplemental materials, FDA's publications in the *Federal Register* often have the character of an open forum for debate in which the FDA explains its rationale and attempts to *convince* the regulated community that its actions are reasonable and a fair interpretation of the law.

The FDA likewise publishes responses to public comments it receives on proposed rules, giving answer to criticisms, clarifying ambiguities, defending judgments, and sometimes making adjustments to regulatory wording in a proposed rule as a result of dialog with the regulated community. Hence, the *Federal Register* can provide valuable insight into Agency priorities and points of contention between the FDA and regulated public. Issues of the *Federal Register* from 1994 to the present can be obtained online at www.gpo.gov. Hein Online, a subscription database, very usefully contains scans of *Federal Register* documents going back decades. Otherwise, pre-1994 issues must be obtained from Federal depository libraries in the United States.

The other excellent source of the FDA's current thinking on a given topic is through guidance documents for industry, which provide specific interpretations of how to comply with the regulations to create adequate and well controlled studies, among other topics. Many guidance documents, including outdated archival versions, are available at the FDA website. However, not all historical documents are available online. Documents published before approximately the mid-1990s must be sought in un-indexed binders in the FDA's Dockets Division (discussed below) or through searches of *Federal Register* announcements.

2.1.3 FDA Dockets, Guidance Documents, Advisory Committee Meeting Transcripts, and Online Drug Approval Information

Whenever public notice must be given for a pending FDA action or decision, or when the FDA chooses to solicit public opinion on a particular issue, the FDA publishes an announcement in the Federal Register and opens a file for comments and issue-related documents in the Office of Dockets Management. Similarly, if a public citizen petitions the FDA to take a certain action or make a certain decision, a docket file will be created if one does not already exist on the issue. Dockets might include documents associated with a pending new drug application or supplemental application, requests for patent extensions, solicitation for comments on a draft guidance document for industry, requests for products to be assigned 'generally recognized as safe' (GRAS) status, transcripts and comments from open public meetings, comments on proposed rules, petitions to reject or recall a particular drug product, petitions to change certain policies or procedures, etc. More than that, the Dockets Division is the repository of the transcripts of all advisory committee meetings. Generally speaking, dockets dating from the year 1998 onwards can

be obtained online at www.fda.gov/ohrms/dockets/. All other materials (dating as far back as 1974) must be sought from the docket office in Rockville, Maryland. These materials are not indexed, hence the researcher must be prepared sift through printed lists of docket numbers with associated titles. Access to specific files is made by request, and copies are only allowed to be made by Dockets Division staff for a fee. Advisory committee transcripts are available in microfilm, however the files are somewhat disorderly and specific meetings may be difficult to locate. Some meeting transcripts may also be available in outdated electronic media searchable only by Dockets staff.

Not included in publicly available dockets are copyrighted materials, selected comments from individuals (where privacy concerns are an issue), internal memoranda, proprietary drug information included in drug applications, and other materials exempted by the Federal Freedom of Information Act (FOIA). While many ‘sensitive’ materials can be requested under FOIA, the Act does not require the FDA to turn over internal documents or to divulge information on internal processes used for decision-making.¹ According to the staff at the Dockets Division, the FDA is groaning under a multi-year backlog of FOIA requests. Hence, for the purposes of a doctoral thesis, I deemed it impractical to file requests for restricted data. In any event, many of the materials I would be most interested in seeing are specifically exempted under FOIA (documents which reveal internal debates, such as internal memoranda, internal meeting minutes, etc.) and therefore would not have been released to me.

For comment letters sent directly to Dockets Management, the original signed letter is on file with its envelope bearing the postmark. Hence, issues of reproduction and authenticity are minimal. Where comment letters have been sent to the Commissioner of the FDA or other officials, a photocopy of the original letter is often placed in the docket file. Proposed rules and *Federal Register* notices placed in the Docket are often the original, pre-publication version of the document with a hand-written signature of the FDA Commissioner. Overall, the quality and authenticity of the documents in this archive is high.

When reading comments submitted on the dockets, of course, one must consider the source. Comments from private citizens can include those from concerned citizens or patients having a particular disease, community physicians or pharmacists, academic

¹ For more information on FOIA, see <http://www.fda.gov/foi/foia2.htm>. For an index of the kinds of documents publicly available under FOIA, see <http://www.fda.gov/foi/electr.htm>.

physicians or specialists at prestigious medical centres, or indeed former or current FDA staff writing as private citizens. Clearly, a ‘private citizen’ letter from a well informed cancer patient in Iowa must be read differently from a letter submitted by a now-retired former FDA Commissioner – and, indeed, the latter likely carries more weight than the former in FDA’s weighing of the issues. Congresspersons write to the FDA to advocate for a particular position, or on behalf of constituents who have written to them; just as in the case of Congressional hearings, one must be aware of an individual Congressperson’s political affiliation and historical positioning with respect to medical policy and the FDA.² Other frequent correspondents include disease-based advocacy organizations, trade organizations, physician groups, consumer advocacy groups, individual pharmaceutical companies and consulting firms, etc. The political orientation and (often economic) interests of each of these authors must be taken into account when viewing their comments and opinions. It is not uncommon to find that trade organizations or patient advocacy groups have produced a standard letter template for their members to use in letter writing campaigns to the FDA on a particular issue. (Sometimes the template letter will not be reproduced in its entirety, but salient portions will be reproduced verbatim in an otherwise original text — which, for the researcher sitting in the Dockets office reading these letters in sequence, feels much like reviewing a series of undergraduate papers in which the same bit of text from the Internet has been plagiarized but slightly modified in an effort to pass as original.) In such situations the formal group alliances and parent organization can be easily traced.

For the purposes of this research programme, I viewed the letters in the dockets as a way to gauge the reaction of various parties (primarily) to proposed rules. Caution was necessary because it cannot be assumed that the comments registered in the dockets constituted a uniform sampling of the opinions of interested parties. Nevertheless, the letters were valuable to get a sense of the debates surrounding proposed rules and actions. These debates were especially useful to me when they focused on questions of how clinical trials should be designed and what standards of efficacy should apply. I took these discussions to be reflective of conceptual terrain in negotiation. Additionally, when placed

² For example, Senator Edward Kennedy and Congressman Henry Waxman are both relatively frequent letter-writers to the FDA. Both Democrats, Kennedy has been involved in a range of health issues, including efforts in the 1970s to pass legislation to reduce the burden of evidence required for approval for drugs intended to treat life-threatening diseases. Waxman often plays the role of a medical consumer advocate and is a long-time chairman of a Congressional subcommittee having oversight over the FDA.

in the larger historical context of actions already taken (have the proposed rules already been informally used?) and actions subsequently taken (to what extent did the FDA modify the proposed rule based on these comments?), I was able to attain a better understanding of how this consensus process fits in with other, ongoing forms of consensus-building.

By far, the most useful documents for learning about debates in drug approval are advisory committee transcripts. These documents revealed to me the standards and criteria used for specific instances of drug approval; the factors considered to be mitigating when the data were less than pristine; the criteria committee members use when judging a drug ‘therapeutically important’; and the concerns raised by experts regarding approval of the data. It was here that I was best able to see an expert deliberation process for approving drugs (one not identical to that used by the FDA, but one certainly focused on the questions which motivated the FDA to bring the case to the committee in the first place). The meetings are attended by medical, pharmacological and statistical experts, pharmaceutical company representatives, FDA staff, outside consultants, and patient advocates. While often the committee members who have voting rights are the ones debating the issues, the FDA, pharmaceutical sponsor, and patient advocates typically give presentations on the relevant topic, providing various perspectives for the advisory committee to consider as part of their deliberations. The meeting participants are typically candid and the discussions can be heated. Committees are consciously aware of the social implications of their recommendations, and they consider social and political questions such as the cost of drugs to consumers; the impact of accelerated approval on the ability to conduct additional clinical studies; the potential implications of off-label prescribing for patients not indicated in the labelling; the feasibility or desirability of placing restrictions on which physicians are authorized to administer certain especially toxic drugs or ones requiring special handling procedures, and other concerns. Members sometimes specifically cite such concerns when lodging a ‘protest vote’ against a drug likely to be approved. One also learns from the committees a great deal about the difficulties associated with treating specific disease conditions. This type of knowledge provides important context for understanding why one drug might be given a greater benefit of the doubt than another in approval decisions.

Unfortunately, committees are not used for every drug approval decision. The FDA typically only goes to the committee for recommendations in controversial or

especially technically difficult cases. Hence, ‘routine’ decision-making is not necessarily part of what can be observed in the advisory committee meetings. Moreover, committees sometimes hold closed sessions to protect discussion of proprietary drug information. For obvious reasons, no transcript is made of these sessions. Nevertheless, committee meetings typically provide a window on the most pressing questions facing the FDA and the most ‘cutting edge’ cases of drug approval — the ones which push previous boundaries of technical validity or socially accepted risk, and are therefore invaluable for this study.

However, advisory committee meeting transcriptions are only as good as the transcriptionist. Some transcripts are riddled with mis-transcribed drug names, erroneous medical terminology, and general typographical errors, some of which can completely change the sense of the sentence.³ In some documents the transcriptionist sometimes inserts the word ‘unknown’ in brackets where not even a guess could be made as to the word used. Unfortunately, such instances are often medical keywords crucial to the understanding of the sentence. Therefore, it is important for the researcher to develop a good knowledge of the medical lexicon used by the participants, including statistical terminology used in the assessment of clinical trials.

Summary information on drug approvals, often including the types of studies used for support of the drug application, are available for oncology drugs and some AIDS drugs at the FDA website.⁴ This data can be used to help fill-in some gaps on ‘routine’ drug decisions that cannot be seen through the advisory committees. However, the summary information is often of inferior quality. Materials at the FDA website vary substantially in the level of detail available, sometimes providing virtually blank forms for summary information on drug approval. Where information is available, comparison with other sources of information (advisory committee transcripts and published reports of drug approval) reveals that the summary information is often riddled with errors, sometimes misstating fundamental information such as the numbers and phases of clinical

³ Even relatively common non-technical expressions are often mis-transcribed. For example, in one transcript (FDA 1992b), references to NDAs (new drug applications) are consistently transcribed as ‘MDAs’ (at least eight different instances, pp. 236, 238, 246, 248, 274, 296, 312, 324); references to expanded access programmes are transcribed once as ‘expanded axis’ (252) and another time as ‘standard axis’ (248); a reference to the AIDS Clinical Trial Group was transcribed as the ‘Ace’ Clinical Trials Group (274), etc.

⁴ Information summaries of approved cancer drugs are posted at <http://www.fda.gov/cder/cancer/druglistframe.htm>. More reliable, but for my purposes less useful, information on AIDS drugs approval is posted at <http://www.fda.gov/oashi/aids/virals.html>.

trials forming the basis for approval, or the clinical trial endpoints used to assess efficacy. Hence, caution is necessary in using these sources of information.

2.1.4 *AIDS, Oncology and Clinical Trials*

To understand communally agreed upon criteria for what constituted a ‘valid’ or ‘well designed’ clinical trial prior to the advent of accelerated approval, I consulted 1980s-vintage medical textbooks such as Buyse et. al (1984) and Silverman (1985). Textbooks can likewise be used to trace changing conceptions of clinical trial validity and design through time. Industry standard works such as DeVita et. al (1993) are particularly useful for such a purpose, because the volume has been revised and re-issued repeatedly over two decades.

While textbooks may provide an idealized description of how clinical trials *ought* to be performed in theory, journals provide information on how clinical trials *have been* performed in actual practice. Two different sorts of information are available from the journals: first are published studies of research findings which go into some detail regarding the study design and results obtained; second are announcements of drug approvals giving basic information on the successful trial design and results. The latter are particularly relevant to this research program since they provide information on drug trials leading to drug *approval*; each serves as a case of a successful ‘registration’ study (one used to support a drug application) and tells us something about the criteria used for approval. The former, while not necessarily indicative of what should be considered ‘adequate and well controlled’ studies, are nevertheless examples of clinical trials as routinely practiced and may provide clues for understanding general standards of acceptability of clinical trials within the professional community. Some industry newsletters also publish announcements of drug approvals, though not necessarily including details of clinical trial design.

Trade newsletters and magazines can also be enlightening sources of information on ‘unofficial’ opinions on current topics, or what Collins and Pinch (1979) would call the ‘contingent forum.’ Indeed, these are among the only sources of public information about drugs that have *failed* approval by the FDA. So, for example, in the widely publicized case of a failed drug application for a promising drug sponsored by a company called ImClone (discussed in Chapter 8), a trade newsletter called *The Cancer Letter* obtained a copy of the (proprietary) clinical trial protocol and published critiques by three

oncologists (see the 15 February 2002 issue). Such information on failed applications, while less commonly available than information on approved drugs, can nevertheless be a valuable tool for getting a fuller view of the definition of ‘validity’ in clinical trials (at least, in this case, as deployed polemically).

2.2 Interviewing

Burgess (1996) describes interviews as ‘conversations with a purpose’ (102). Originally, I thought that I would approach the interview analogously to Silverman’s (1993) description of how to view government documents. These, he says, should not be seen as dispensaries of objective information but as expressions of socially and culturally embedded category-making. In this light, one analyzes texts to see how the individuals behind them construct their world. The difficulty with this formulation for my research is that before one can talk about how people create meaning for their particular situations, one must understand the actual events related to the meaning-making. My purpose in interviewing, therefore, was to identify a specific individual who, through personal experience, was involved with a particular FDA decision or who would have personal knowledge of FDA practices in the period of interest, and to try to gain the perspectives and recollections of that person regarding those events and decisions with which they should be familiar.

For me, interviewing served a decidedly secondary role to documentary analysis for a number of reasons. First, memories fade and people often reconstruct past events in unreliable and sometimes fanciful ways. For specific information on what happened and why, documents (properly viewed as described above) are typically more reliable than people. However, the people can provide corroboration of documents and frequently also insightful perspective, which were the main applications of interviewing in this thesis. Second, many of the key decision-makers of the time are high-status individuals, some of whom have already been interviewed numerous times and whose recollections are publicly available through books, articles, online oral history archives, etc. I used such materials to the extent feasible for this research. Third, in some cases requested interviews were not granted.

I used both formal and informal approaches to interviews. Often times it is difficult to get a busy, high-status person to sit for a lengthy interview, but that same

person might (sometimes) respond to quick telephone calls or short emails. Moreover, in some cases I did not have multiple questions I wished to ask, but had a specific situation for which one question would do. In such cases, rather than writing and requesting a formal interview, I used the telephone or email to try to pose my single question. The disadvantage of this approach is that I typically caught the interviewee off-guard with a question about deliberations or events which took place many years ago. The quality of the discussion under such circumstances was variable. More than that, such informal conversations are more in the style of ‘off-the-record’ journalistic interviews and could not ethically be directly quoted without obtaining the consent of the discussants.

In the case of formal interviews, I used a semi-structured approach — one which did not establish a rigid set of questions to follow, but sought to create a ‘conversational guide’ (Rubin 1995), listing key topics with possible follow-up questions, as applicable. Since each person interviewed had a different role to play in the events at issue, the questions were different for each interviewee. In many ways the interview preparation and conduct was more like an oral history interview than a sociological one — or at least, it involved a modification of the latter. In an unstructured sociological interview, the interviewer will need to have in mind ‘a set of themes and topics to form questions in the course of the conversation’ (Burgess 1984, 102) and to assemble these themes sensibly, she will need crucial background information about the interview respondent and the situation at hand. However, for such sociological interviews that background information is the interviewee’s social situation, which is collected by direct observation.⁵ In my case however, the background information collected was knowledge of historical events in which the interviewee participated; hence the ‘observation’, if we can call it that, was observation of the interviewee’s actions or decisions in decades-old documents. Likewise, the questions were often related to the interviewee’s specific actions or perceptions associated with a given event, more like an oral history interview (except the oral history interview is more open-ended, ultimately intended to form a permanent archival record, rather than being associated with the more narrow purposes of a PhD thesis).⁶

⁵ See Burgess’s (1984) discussion of Ferdinand Zweig’s work, p. 103, in which detailed background knowledge is described as crucial to a meaningful interview.

⁶ On the preparation, purpose, and doing of oral histories, see Ritchie 2003; Howarth 1998. Strangely, except for those texts dealing specifically with oral histories, general methodological texts in history often do not broach the subject of interviewing (see for example McDowell 2002; Howell and Prevenier 2001).

2.3 Data Reduction and Analysis

The literature in sociological research often recommends a ‘triangulating’ approach to research questions: using multiple sources of information to approach the same question or fact (see discussion in Blaikie 2000, p. 262 forward). In this way, the researcher builds a ‘convergence of evidence’ (Yin 2003, see Figure 4.2) on key points supporting the argument. Burgess (1984, 145) describes several modes of triangulation in which data can be collected and compared across different times, places, or subjects (‘data triangulation’); or a situation is examined by more than one investigator (‘investigator triangulation’) or using more than one theory (theory triangulation); or using a ‘methodological triangulation’ in which either the same method is used in different circumstances (‘within method’) or in which different techniques are applied to the same object of study (‘between method’).

Texts on historical methods also refer to a similar process forthrightly termed ‘comparison of sources’ (Howell and Prevenier 2001, 69). While similar, there are differences in goals between the sociological and historical procedures: the sociological version tends toward gathering evidence of a certain social trend or phenomena across disparate data sources; the historian, by contrast, is concerned to piece together an accurate picture of events within an appropriately construed historical context. In the historical version, therefore, the details of events are corroborated across multiple data sources. Where no such corroboration can be achieved, the source of a given ‘fact’ itself must be evaluated for credibility and validity, and this ‘fact’ is accepted provisionally if it is consistent with information already known about the historical situation of interest. Similarly, if sources contradict each other about a given circumstance, it may be possible to discern the most likely version of events through comparative contextualization: testing for consistency with other information accepted to be valid about the actors, culture, region and local practices, etc. (Ibid, chapter 3).

For this research program, I have used two primary forms of triangulation: what Burgess calls data triangulation and methodological triangulation. The documentary data from the sources described above were assembled in a historically comparative manner to piece together an understanding of the events in question (the term ‘events’ used loosely here to refer not only to actual happenings, but the attitudes of participants, and information contributing to an understanding of drug approval in practice). I integrated

the interview data with the documentary data in a comparative way, using it as a corroborative source to the documents in a 'between method' strategy. These sources were used to triangulate around the questions discussed in the next section.

Since many of the key events and decisions I would hope to study took place behind closed doors, it is often the case that sources of information on a given event are limited. In such cases, 'triangulation' as described above may not be possible. When a single view or account of events may be the only source available, that account need not be dismissed automatically, but consistent with established historical methodology was judged for credibility, as well as consistency with what is generally known. In these cases, as for all information presented in this thesis, I have sought to document clearly all sources of information so that a knowledgeable, critical reader can readily assess the supporting data for herself.

2.3.1 *Reduction and Analysis of Documentary Sources*

'Data reduction' for this research programme primarily involved studious note-taking, as well as filing and organization of original source documents in such a way that information is grouped logically and readily retrievable. For the purposes of this study, I considered the information contained in documents as falling into one of two related categories: What happened? And what were the important circumstances or context surrounding what happened? Then, the analysis of the data involved asking the related questions: What do peoples' actions and attitudes towards a given rule or procedure say about their view of the meaning of the rule or procedure? How do the uses of key concepts as reflected in people's actions, statements, and decisions change over time? Are patterns evident in the way concepts are applied over time? How do people agree on what constitutes a valid usage of a concept or rule? What standards of validity are reflected in the actions taken or rules written, do these two sets of standards diverge, and how do these standards change over time? What conclusions can be drawn?

The first question is relatively straightforward. Who wrote the law; what were the written provisions? What drug was submitted for approval; what were the committee's recommendations? These are the sort of events which have been recorded in documents generally agreed to be credible and valid, and therefore the recounting of these events as such would be uncontroversial. The central matter of judgment for the analyst is which events to include in the account. In a study of FDA drug policy in the 1980s and 1990s,

does the analyst need to give an account of *every* drug considered for approval by the FDA? Most reasonable observers would answer negatively. Even in a study of the development of the policy for accelerated approval, it should be apparent that there are key decision points and key events which fed into the development of the policy and these are the ones which the analyst should see to recount. Chapter 1 of this thesis includes some explanation of why certain cases of drug approval were chosen for close study.

The second question about important context or circumstances is really an extension or subset of the first question – it extends the question of which increasingly minor details are important to the account. For my purposes, it is important to know not only that the advisory committee voted a certain way, but why they voted in that manner. What were the data they considered as a basis for approval? What were their stated impressions about the data? Did they consider it ‘valid’ and by what criteria did they judge? There is a world of secondary information available to the analyst and only a small fraction of it is necessary or desirable to recount in detail. It is the job of the analyst to provide as much secondary detail and context as required to address the research questions and give the reader a clear and hopefully vivid picture of each situation while not littering the account with so many superfluous details that main argument is obscured or confused. To some degree this is obviously a matter of individual judgement and even individual style. However, as a matter of research ethics, it is also incumbent on the researcher to include *all* information she encounters which impinges on the questions at hand, not merely the data which supports her theory. It is here where the researcher has the opportunity to reject, revise, or improve empirically based theories. The biggest fear of the researcher in addressing the first two questions is that some crucial piece of information will be missed. There are no guarantees that this kind of error can be avoided completely, however allowing review of one’s work by knowledgeable persons is an effective form of quality assurance. For this reason, I have taken opportunities to present portions of this work at conferences and workshops, and have sought review of my written materials by advisors and others in a position to judge my work.

Finally, in the analysis step, the salient features of events are compared and themes begin to emerge from the pages of historical exposition. This process actually begins even as data collection takes place, with the formation of partial answers to these questions forming the basis for seeking additional material with the potential either to substantiate or invalidate a perceived emerging pattern. Blaikie (2000) terms this process ‘retroductive’

research. This is where individual investigator observation is most essential and most subject to between-investigator variations. I am not referring necessarily to ‘bias’ or the operation of investigator ‘interests’ to prove a certain point, though of course these factors may impinge on the process. Even more than those possibilities is simply the fact that due to variations in personal training or experience or thought process, different investigators may ‘see’ the data differently, with some putting more emphasis on some concepts than others, perhaps spending more time on technical details of clinical trial design criteria than others would. Such differences of emphasis or perspective are not necessarily ‘wrong’ when based on data elements that are carefully researched and considered in context according to the principles of scholarship discussed here. Indeed, the idea of such scholarship is that others should examine my view of the material and add their own perspectives to mine in the literature, leading to a more fruitful and multi-dimensional understanding of the underlying empirical material. I leave it to the reader to do just that.

2.3.2 Reduction and Analysis of Interviews

For formal interviews, I made a recording of the interview which I then transcribed. I used a style of transcription in which significant pauses, interruptions, and false starts were indicated in the transcript, however since the main purpose of the interview was historical rather than ethnographic, I made no attempt to time the length of pauses, transliterate words stretched for emphasis, or use other such conventions which seek to capture as closely as possible fine details of expression and interaction. Since interviewees would be reviewing the finished transcripts, the overall goal was to edit as little as possible while maintaining readability and intelligibility. The degree to which interviewees felt that the transcripts were readable appeared to vary: my first interviewee, Dr. Temple, returned his transcript to me fully edited to reflect a smoothed-out, journalistic transcript. Thoroughly abashed, I quickly responded to Dr. Temple thanking him for his comments and attempting to explain that the transcript style I had chosen was only lightly edited by design. If only I had first read Ritchie’s (2003) warning to fledgling transcribers of oral histories: ‘Educated interviewees who say “yeah” will insist on altering the transcript to “yes.” They are sorely displeased when transcripts show them saying “gonna” or “talkin” and would prefer to see their spoken words reproduced as they would write them’ (57). Given that Dr. Temple is a high-level FDA administrator with a more

than thirty-year career with the Agency, I suppose his reaction is understandable. For the subsequent formal interviews conducted (two, both with the same person, Dr. Ellen Cooper, former head of the antiviral division of the FDA), I used slightly more aggressive editing, though still left many pauses and false starts.

Analysis of the interviews involved comparing interview material to information in documents, using the interviews primarily for corroboration and clarification of documentary sources with respect to the questions discussed above. However, as already noted, the interviewee's memories were understandably not always clear on the details and chronology of the events in question. Therefore, where interviews conflicted with documentary sources on such historical details, if additional corroborating evidence could not be obtained, the documents were taken to be more likely to be valid. With this caveat about memory, as well recalling Löwy's (1996) warning about the interviewee's possible desire to reconstruct events in a certain light, I also found that the most interesting and useful information provided by my informants had to do with attitudes to the events in question (e.g., AZT was a success and something worthy of holding up as an example) and background information and insight that would be difficult to obtain without an 'insider' perspective (e.g., what the FDA was like culturally in the 1970s).

2.4 Research Ethics

The University of Edinburgh's School of Social and Political Studies (SSPS) has developed a set of procedures to review proposed research for ethical accountability. The School has developed a three-level approach beginning with a checklist for self-audit (Level 1) to be used for proposed research for which no ethical risks have yet been identified. Higher levels of assessment may be required if ethical risks are revealed in the self-audit. In the proposal for this research I completed a self-assessment checklist and obtained the review and signature of my supervisors. In the subsequent carrying out of the programme, I have made no changes to my purposes or approach that would in any way modify the original self-assessment.

In this research programme, the documentary sources used are publicly available and not ethically sensitive. Given that I am not seeking sensitive, personal, or proprietary information, I have not requested, nor have I encountered, ethically sensitive data. Likewise in interviews, I am not collecting personally sensitive information about my

respondents. In each formal interview, whether in person or over the telephone, I have asked permission to record, offered the option of anonymization in publications and presentations, offered the ability to review resulting interview transcripts for accuracy, given the option of reading and responding to papers prior to publication, and of refusing recorded interviews. Interviewing for this research has been relatively limited; however, both major interviewees waived the option of anonymization, requested to see the completed transcripts, and requested to see my use of their comments in written material. I complied by sending completed transcripts and an early draft of material using their comments. One interviewee responded with comments; the other did not. None of the comments received contained an objection to my use of quoted material, but were remarks on other aspects of the work. Out of respect for the interviewee's time, I have not sent later drafts of the same work if the use and context for the already-reviewed, quoted material was not altered.

In addition to these formal interviews, I did have a number of informal telephone and personal conversations with a former FDA NDA reviewer, and two former Senate staffers, one of whom is a well known regulatory attorney in Washington. Since these conversations were 'off-the-record', I have not quoted nor directly used any of the comments made to me by these respondents, but have rather used those conversations to inform my sense of context.

Given the relatively uncontroversial nature of my research, there was no justifiable reason to conceal my purpose from prospective interviewees, although I believe it was important to position myself as an uncritical researcher with no political 'axe to grind'. Given that my 'axe grinding' is of a theoretical nature (i.e., picking bones with other social scientists and making a case for a certain theoretical position) rather than prescriptive, I believe that the characterization as a largely uncritical researcher is an accurate one. Although, as discussed below, a position of absolute neutrality is of course impossible.

2.5 Research Topic, Researcher Motivations, and Method Choice

2.5.1 Motivations

In some ways, explication of researcher motivation at the *end* of a research programme is more difficult than it appears. The motives prevailing at the end of the programme, driving the effort to a conclusion, are likely to be somewhat different from

the ones which originally launched the project — or, at least, related in obscure ways. Originally, of course (as is often the case), my topic was a subject of general interest in need of focus and framing. The initial motivation to study the procedures for accelerated approval arose out of a personal experience: waiting for Fast Track approval of what was thought to be a highly promising oncology drug (one of the new molecularly targeted therapies), making unsuccessful attempts to gain pre-market access to the drug, and ultimately seeing the drug application bungled by the sponsor and rejected by the FDA, meanwhile watching my father progressively succumb to the grip of cancer. That experience created in me a strong desire to understand what went wrong with this Fast Track drug application, which then led to questions about what happens when new regulatory categories are created. It is important to say that at no point was my motivation to identify a villain and assign blame for the failure of this drug to make it through approval on the first attempt. After studying the case, it seemed to me that, at least in part, the FDA and drug sponsor became tangled-up in the very phenomenon I have sought to study: instability in the meaning of a new drug approval category. Nevertheless, in the interest of full disclosure I should note that the drug Erbitux, discussed in Chapter 8, is the drug we sought to obtain for my father's cancer.

Additionally, I should note that this personal experience shades my view of patient advocacy in FDA advisory committee meetings, particularly patient demands for pre-market access to experimental drugs. As a scientist myself, I am sometimes caught between, on the one hand, the FDA's desire to have some scientific understanding of the potential risks and benefits of a drug prior to therapeutic use and, on the other hand, terminally ill patients fighting for a last-ditch hope to live. The same sorts of issues are at stake in accelerated approval: in the case of terminal disease where therapeutic options are inadequate, how much evidence is enough for approval? Under such circumstances, the concept of 'scientific evidence', which sounds so concrete in the mouths of Congresspersons, attorneys or media presenters, becomes a fragile and malleable thing. As we will see throughout this thesis, ultimately, the question of 'adequacy' of evidence in clinical trials is a political one: adequate for what social goal? What level of safety, or of certainty, is adequate? Of course my task in this research is not to take sides with either the FDA or patient advocates (who, by the way, are in no way speaking with one voice on the issue of pre-market access), but to trace the conceptual shifts attendant to new category formation and application.

2.5.2 *The Interrelation of Research Questions and Method Choice*

In Blaikie's terms, the strategy or logic of the research design for this study is primarily a deductive one. I have examined a particular set of events using a given theoretical model and have derived research questions in the language and terms of that model. If I investigate my questions only to find that the empirical data cannot be accommodated to the framework of the theory (something which is possible, despite the development of the research questions with the theory, because of the flexibility and revisability of the qualitative approach), then it will be clear that this theory is not adequate to describe the situation at hand. In other words, there is an element of theory-testing to the strategy for this research. Crucially, however, this theory-testing is not a straight up-or-down set of judgments that will either validate the overall theory or invalidate it. Rather, as Blaikie (2000) discusses in his 'cycle of theory construction and testing' (158), after the empirical material is assembled, an essentially inductive process begins in which refinements are made to the original theory. Of this inductive phase, Blaikie says that '[w]hat should occur is a complex trial and error process, more akin to that used in the abductive research strategy and by grounded theorists' (p. 159). My adjustments or additions to the theoretical frame will be offered throughout this thesis, and especially in the conclusions.

In the previous chapter I explained and defended the specific theoretical framework to be employed in this investigation. Here it may be worthwhile to offer a more general rationale for the choice of this particular perspective. Yin (2003) proposes a scheme for determining which kinds of research strategy are most appropriate for which kinds of research questions (5). While one can criticise aspects of his breakdown of questions and strategies, a very useful idea arises from Yin's scheme: that the kinds of research questions one can ask and fruitfully answer are necessarily bound-up in the methods and (correspondingly, according to Crotty) theoretical perspectives used.

When it comes to research questions and strategies for answering those questions, it seems we have a 'chicken-or-the-egg' situation. Which came first, the research questions or the methodology? In practice, I suspect the theoretical commitments and training of the researcher condition the kinds of questions that might be asked, even as the researcher should be seeking the methodological approach most apt to elucidate the topic at hand. (Or, as in a famous statement attributed to psychologist Abraham Maslow, if one is adept at using a hammer, everything tends to look like a nail.) In my case as a

student, the research topic and the epistemological point of view were developed in tandem — each playing against the other during the preliminary research into the topic. Once the research programme matured, I then tended to treat finitism as a focusing and framing machine into which I feed a set of historical events and out of which drop research questions and answers. However, the development of the questions was not such a clean, stepwise process. Indeed, the best reason to use finitism and Edinburgh SSK as an explanatory frame for my empirical material is that the material and the frame were developed together and are therefore intertwined in the research questions. Could other framing devices or methodological approaches be used? Certainly, but the resulting research questions and answers would not be the same.

2.5.3 *Qualitative vs. Quantitative Approaches*

One might ask why I have not employed a quantitative strategy to approach the subject of Fast Track drug approval. Even a cursory review of the social science literature on drug approval will reveal a vast array of metrics and models designed to describe the characteristics of FDA decision-making and drug approval actions. Why not gather data on the characteristics of instances of drug approval associated with the new rules and see how these characteristics change and stabilize (or not) over time? To the minds of many, it would be in my best interest to do so. Indeed, in the pilot study to propose this research, I did gather such data to try to track the changing profile of evidence proffered in successful instances of drug approval over time. However, I found that such data was only available in detail for oncology drugs, and that the available data was often erroneous or incomplete. More than that, I found that the aggregation of data often led to the conflation or comparison of drug approval cases which were incommensurable in their specifics. Hence, the procedure was hazardous and could only provide broad generalities based on rather tenuous comparisons. The better way to get more highly refined and detailed results was to conduct a qualitative study.

Nevertheless, in the social sciences in general, and policy-oriented studies in particular (perhaps especially in the U.S. context), quantification is perceived as producing ‘hard’ data, as opposed to the ‘soft’ data of qualitative studies.⁷ As a result, when authors

⁷ See Geiryn’s (1999) discussion of Congressional debates over whether social science research should receive government funding through the National Science Foundation, a separate governmental organization – or whether it should receive government funding at all. Many of the arguments marshalled

of quantitative studies cite work done in a qualitative genre of social science, it usually comes with a dose of scepticism. For example, in a 2002 study of FDA approval practices (Carpenter 2002), the author notes that a series of observers including Epstein (1996) have ‘argued’ (as opposed to ‘documented’ or ‘demonstrated’) that organized patient advocacy groups and media attention have influenced FDA drug review. ‘This claim has not, however, been subjected to a quantitative test’ (491). Ironically Carpenter’s quantitative approach, while impressive in many ways for its innovation and sophistication, begins with a fundamentally flawed assumption: that the procedures for reviewing and approving all new chemical entities (NCEs) are uniform. In fact, some NCE are given expedited handling under the rules examined later in this thesis. As a result, his model aggregates cases which are not necessarily comparable in terms of specific relevant characteristics. This pared-down view of drug review procedures carries over into his interpretation of findings as well. For instance, to explain why drugs which target large populations tend to be approved more slowly, Carpenter wrote that ‘conditions such as the common cold, influenza and Poison Ivy are extremely common but command less political emphasis than do rarer but deadlier conditions’ (502). While true, this explanation overlooks a more immediate cause: the risk-benefit approach used by the FDA to evaluate NCEs works in that manner by design. In drugs for life-threatening diseases where existing therapeutic options are lacking, the potential for benefit is seen as outweighing the risk of expedited decision-making, even when data collection on the drug is incomplete; by contrast, the risks of exposing large populations to a new drug speedily do not outweigh the relatively small benefit of, say, alleviating common cold symptoms. Knowledge the FDA’s risk-benefit policy would have gone some way towards answering Carpenter’s research question, ‘why does an agency that approves some drugs very quickly take so much time with others?’ (Carpenter 2002, 490).

More could be said, however my point is to underscore the nature and limitations of quantitative modelling: in terms of the reliability and accuracy of conclusions, even a well constructed quantitative model of FDA decision-making is not necessarily better than a well researched qualitative case study. For all his mathematical elegance, and despite many useful and valid observations about the FDA’s decision-making process, should we be gratified that Carpenter ultimately concludes that ‘citizen groups and the media have a

in support of government funding for the social sciences tended to wed them to the natural sciences, among other ways by emphasizing ‘objective’ quantitative research methods.

powerful impact on regulatory outcomes' (503)?⁸ His quantitative testing has verified a phenomenon already historically documented, even while it allows for misinterpretation of some data trends due to exclusion of detailed situational knowledge of the process to be modelled.

Quantitative research ultimately requires a solid founding in qualitative social and historical research methods if it is to model reality meaningfully. While we could say that qualitative research also creates a model of reality (not unlike that created in quantitative modelling), the qualitative approach affords more opportunities for flexibility and revision of the model. The gathering of information typically leads to refinement of the ideas in play and the recognition that more information is required to complete the picture realistically and meaningfully. This process obviously leads to a 'stopping problem' of the kind posed by Carpenter in this model of FDA decision-making (how much information is enough?), but the result is ultimately richer in comprehension and explanatory power — which, Yin might agree, is what is needed for an open-ended, process oriented, essentially historical research question like the one posed by Carpenter.

I was once asked by a quantitative policy analyst how he could know that the quotations I had selected from historical documents for my empirical exposition would have been the same quotations he would have chosen. The question cuts to the quick of the positivistic perception of qualitative work as 'subjective' as opposed to the 'objectivity' of quantitative study. In light of my discussion of Carpenter's paper, I could turn the question around and ask: how can I know that the assumptions, variables, and relationships you choose for your model would be the same ones I would choose? If this question does not seem equivalent to the positivist's question about my quotations, consider this: what is my basis for disagreement with Carpenter's assumptions and analysis? It is that a detailed knowledge of approval history reveals flaws in his model assumptions and conclusions. To make an effective critique of a model of a socio-historical process, it is not enough to understand mathematics; one must be intimately familiar with the social and historical setting to be modelled. Likewise, to object to my choice of quotations, one must be familiar with the source documents and the social and historical context of their creation. What is called for in either situation is a detailed knowledge of the case at hand. Whether the issue is quotation choice or model

⁸ Indeed, since he was not able to analyze which factors contribute to the differing approval criteria used on a case-by-case basis for NCEs, his estimation of the affect of media and citizen groups could be exaggerated.

assumptions, ultimately the selection is a matter of individual judgement — both types of choices are epistemologically equivalent.

Does this assertion imply that the choice of quotations or model assumptions is random or baseless? Absolutely not! In both cases the selection process is guided by institutionally derived principles of sound scholarship; conventionally agreed-upon practices which one accepts as part of membership in a given academic group and which, when followed scrupulously, serve to justify one's selections as valid. Discussion of the academic conventions by which to judge my own work are the subject of much of this chapter. Here, however, it seemed useful to bring my full methodological rationale into the light of day, including my conviction that for the kinds of questions I wish to pursue in this research, a qualitative, finitist approach is useful and appropriate.

2.5.4 *Other Intellectual Orientations*

To complete this discussion of research strategies and theoretical commitments, I will make explicit (though it is likely clear enough to the reader already) that finitism views the creation of categories as a social process of consensus-building wherein stimuli from the social and physical environment (concepts and objects) are assigned place and meaning relative to one another. As such, we could say that finitism is epistemologically constructivist (as opposed to objectivist or subjectivist); it posits a model in which, in Crotty's words, '[m]eaning is not discovered, but constructed' and where 'subject and object emerge as partners in the generation of meaning' (1998, p. 9). As David Bloor (2007) makes clear, such a position carries no necessary implications with regard to realism or anti-realism. I touched on this point in the last chapter and do not wish to belabour it. However, even the most senior and insightful observers in our field continue to perpetuate the false idea that the opposite of relativism is realism, and therefore tend to conflate realism with absolutism.⁹ Suffice it to say that this thesis will use an approach which treats the natural world in realist terms (chemicals have an external existence; we can manipulate them, but only in finite ways, to create therapeutic drugs; tumours have an

⁹ Consider this statement from Jasanoff's (2005) most recent book. In discussing what she sees as 'two profoundly destabilizing changes in the way we view the world: one cognitive, the other political', she observes: 'On the cognitive front, the shift is from a realist to constructivist view of knowledge. Years of work on the social construction of science and technology, and the contingency of similarity and difference judgements, have taught us to be sceptical of absolutist claims concerning objectivity and progress' (13). By positioning constructivism and realism as opposing and mutually exclusive terms, she creates false opposites and perpetuates the improper conflation of absolutism with realism and (by extension) constructivism with anti-realism.

independent existence which appears to be affected by the use of certain drugs, etc.), even while recognizing that the whole notion of a 'drug' is a social construction. (One can envision an alternative society in which such chemicals exist, but they are not swallowed or injected by people having illnesses. In such a society, although the chemicals exist, 'drugs' as such do not.)

3. BACKGROUND: THE ROOTS OF REFORM

For a full understanding of the development of rules for accelerated approval and expanded access, one must look back at least a decade-and-a-half in an effort to recreate the historical context for the decisions and actions taken in the late 1980s and 1990s. Often observers mistakenly begin the story in the middle, with the advent of AIDS. Certainly one cannot give a full account of the story without reference to the rise of the AIDS epidemic and the efforts of AIDS activists, whose persistent calls for access to experimental drugs and changes in clinical protocol design for drug evaluation were highly influential in regulatory reforms implemented in the late 1980s and 1990s. On this aspect of the events of the period, the work of Epstein (1995; 1996; 1997) is authoritative. However, as influential and important as this work is, it would be a mistake to assume that activism was the only force to shape events in this period, or that prior to AIDS there was no pre-market access to investigational drugs or significant efforts to expedite development and approval of important drugs. Patient lobbying for the right to choose their own therapy pre-dated AIDS by at least a decade (Markle and Petersen 1980)¹ — although certainly not with the same stridency or radicalism as the AIDS activists — and the U.S. Food and Drug Administration (FDA) had been informally granting seriously ill patients access to experimental drugs on a pre-market basis for decades (see note 19, this chapter). Nevertheless, many authors suggest that reforms for expanded (pre-market) access to investigational drugs ‘began in the late 1980s’ (Orlando 1999, 543) or bluntly characterize reforms in this period as ‘The AIDS Reforms’ (Siegel and Roberts 1991, 1).

¹ On the controversy in the 1970s over whether Laetrile was an effective anti-cancer therapy, Markle and Petersen note that a significant aspect of the controversy was ‘value disputes’ in which health advocates, Laetrile proponents, and political philosophers joined forces to argue that cancer patients ‘have a right to choose their own form of cancer therapy without interference from the medical community or the government’ (7). According to Petersen and Markle the issue, termed ‘freedom of choice’, was ‘the single most effective argument that Laetrile proponents have used in the courts, state legislatures and media’ (7). It can also be argued that this movement towards increased patient involvement in what had been traditionally the physician’s domain of decision-making stems from a fundamental shift in the doctor-patient relationship beginning in the mid-1960s (Rothman 1991).

While such shorthand might be useful for the specific purposes of law or policy, it is difficult to see how we can effectively learn ‘lessons from the AIDS experience’ (Siegel and Roberts 1991) without a clear understanding of what that experience was. In this chapter I will provide the necessary background material to show that while AIDS provided a critical spur to action, the questions of how to accelerate approval and expand access pre-existed the 1980s, with concrete efforts at reform already underway. These informal practices and conceptual tools already in circulation at the FDA were the ones initially brought to bear on the new challenge of AIDS.

3.1 The 1962 Drug Amendments

In 1959 Estes Kefauver, a U.S. Senator from Tennessee, began conducting a series of hearings on the pharmaceutical industry.² The Senator’s staff uncovered evidence that pharmaceutical companies were colluding to fix drug prices, and were selling some drugs to consumers with profits of hundreds or even thousands of percent. Thus, in December 1959 the Subcommittee on Antitrust and Monopoly began investigative hearings beginning with the issue of price gouging and collusion. In time, however, the Subcommittee began to examine the marketing claims made by drug-makers, and the methods for drug promotion. In so doing, the Subcommittee eventually began to question the quality of new drugs, and the ability of physicians to adjudicate between various drug products and the claims made by pharmaceutical ‘detail men’ of safety and efficacy of the drugs. The senator introduced legislation in 1961 seeking to foster competition between pharmaceutical companies while increasing FDA surveillance over new drug production and introduction to the marketplace (Temin 1980, 122). In addition to provisions to lower the prices of brand-name drugs and to require companies to license their patents to other firms after three years, the Senate bill also changed the requirement for marketed drugs from being ‘safe’ to being ‘safe and efficacious in use’ (Ibid, 122).

² See Harris (1964) for an excellent journalistic account of Kefauver’s investigation, the hearings, and the subsequent legislative process. The book is based on a series of articles published in the *New Yorker* by Harris during the hearings. However the work is augmented by subsequent interviews with Kefauver and others. Another journalist, Morton Mintz, also produced books related to the hearings. For more academic accounts, see Bud (2005); Edgar and Rothman (1990); Merrill (1994); Temin (1980). Lasagna (1989) has also written on the hearings. Considered to be the father of clinical pharmacology, Lasagna gave testimony at the hearings and went on to become a recognized authority on FDA regulation, having founded in 1976 the Center for the Study of Drug Development at the University of Rochester which subsequently moved with Lasagna to Tufts University.

Initially, drug company executives dismissed the Senator as ‘nothing more than a meddling hillybilly seeking some publicity’ (Harris 1964, 55). By the end of the hearings in February 1962, however, Kefauver was receiving bags of supportive mail from the public, consumer purchases of many drugs had slackened considerably, and Kefauver was aggressively pilloried by pharmaceutical industry partisans, who among other things called him a socialist (Harris 1964). Nevertheless, in spite of this public support and the often sensational testimony on the apparent bad faith of many major pharmaceutical companies, Kefauver’s bill lacked support of the Senate and the White House. The turning point came in 1962 when the United States narrowly avoided a medical calamity then afflicting Europe and Canada: thalidomide babies. An FDA reviewer named Francis Kelsey was unconvinced of the safety of the drug, and had delayed approval of thalidomide in the U.S., repeatedly returning the application to the drug sponsor for more information, while the drug was used widely in other countries.³ In the interim, reports began to surface from Europe and Canada of babies being born with severely stunted arms and legs when their mothers had been given thalidomide during pregnancy. Morton Mintz, a journalist with a flair for the dramatic, declared Kelsey the ‘heroine’ of the FDA whose ‘vital’ delay kept a ‘bad drug’ off the market (Mintz 1962). Initially it was assumed that since the drug had not been approved for use in the U.S., few Americans would have been exposed to it. However, under the 1938 Act the FDA did not have authority to oversee testing of investigational drugs. It soon became apparent that over 2.5 million tablets had been distributed to more than 1,200 physicians in the U.S. for investigational purposes (Temin 1980, 198, 123-4), and that recordkeeping and reporting associated with the drug had been haphazard such that it was impossible to account for all of the drug distributed and to locate all of the patients to whom it had been given (Hilts 2003, 157).

Kefauver’s Senate bill was (once again) renegotiated in a Senate committee and eventually combined with the Kennedy administration’s alternative version introduced by Congressman Harris in the House. The final version, ultimately passed as the Kefauver-Harris Act, lacked provisions for price controls and mandatory licensing of patents, but contained provisions granting FDA authority over clinical study of investigational drugs. In addition, the 1962 Act required sponsors of new drugs to present evidence not only of safety, as required under the original act, but also of effectiveness. Whereas previously a

³ See Sjöström and Nilsson (1972) for a comparative account of Swedish, Canadian, German, and American policy and approach to thalidomide during this time, and their respective responses to the crisis.

new drug would automatically become legally marketable within 60 days following submission of safety data unless the FDA specifically intervened, now there would be no acceptance by default; the drug sponsor would have to wait for review of the safety and efficacy data provided and receive an approval from the FDA prior to marketing the drug. The final version of the Act used the word 'effective' rather than 'efficacious', and required that effectiveness be proven with 'substantial evidence', which meant:

evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could be fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof (P.L. 87-781, §102).

Some observers such as Abraham (1995) consider the Act to have been weakened by replacing 'preponderant' evidence by 'substantial' evidence. However, Democratic negotiators at the time felt that they had won a victory in the inclusion of language specifying the nature of the evidence required (Harris 1964). Controlled trials meant that the responses of a group of patients receiving the investigational drug needed to be compared to another group of patients who were not receiving the drug. 'Well-controlled' trials would need to be designed carefully, such that the patient groups used in the study were comparable in terms of disease status and other variables. Well-controlled trials would need to use various techniques (such as randomization and blinding) to eliminate bias in the study. This phrasing opened the door to requirements for statistically valid clinical studies in drug approval. Temin (1980) also points out that this is a more rigorous standard than the original version of the bill calling for marketed drugs to be 'safe and efficacious in use', since the latter provision did not specify the type of evidence required for such a judgement (see Temin's sixth chapter). In retrospect we can also note that ultimately a Federal advisory committee system developed in which drug approval recommendations are often made by panels of experts. Hence, in at least one way the system evolved to resemble the concept of 'preponderant evidence' since, in practice, expert consensus is often sought by the FDA for evaluation of clinical evidence offered by drug sponsors.

Observers sometimes point out that the requirement for effectiveness contained in the Harris-Kefauver Act would not have prevented the thalidomide tragedy — and that, indeed, the drug was being kept off the U.S. market by existing regulations (e.g. Lasagna 1989). While true, to say so without qualification leads to the impression that this provision was written solely as a response to thalidomide when, as we have seen, an efficacy provision was included in the original bill *prior* to the revelations about thalidomide. Kefauver included this provision because in Congressional testimony it was clear that drug companies often misled physicians on the safety and efficacy of marketed drugs, exaggerating both, and that doctors found it difficult to assess the information provided by drug companies and to adjudicate between competing products. Hence, one effect of the thalidomide situation was to put Democratic negotiators in a good position to strengthen already existing language for certain provisions. More than that, as some authors have noted (Temin 1980; Hilts 2003), in considering the influence of thalidomide on the shape of the final legislation, one must look to the new law's provisions for FDA oversight of the investigational drug process. Such oversight was lacking but clearly needed in the case of thalidomide, since the drug was distributed widely and incautiously in the U.S. while unapproved.

When discussing the thalidomide tragedy and its effect on this legislation, observers tend to focus on positive effects — the provisions included the bill — while passing over negative effects, meaning the provisions left out. It seems clear from the account in Harris (1964), that one effect of the thalidomide situation was to promote a shift of emphasis in the kinds of consumer protection sought: moving from protection from exorbitant drug prices to protection from products having unknown or misrepresented effects. The patent provisions, already flagging before thalidomide, were easily jettisoned in its wake. The power of the FDA to control how drugs are tested and which drugs appear on the market was greatly enhanced, but how much pharmaceutical companies would charge for those products and why American consumers pay so much more for drugs than patients in, say, Canada or Britain, remains a matter of heated debate to the present day.

3.2 FDA Growing Pains

By 1962, the basic elements of the legal regime overseeing the production and marketing of new drugs in the U.S. was in place. However, as Merrill (1994) observes, key details were left to subsequent legislation or to agency initiative to be worked out (51). The FDA moved quickly in 1962 to clarify what kinds of clinical studies they expected to see in drug development. Proposed in 1962 (FDA 1962) and published as a final rule in 1963 (FDA 1963), the investigational new drug (IND) rules were among the first to be written with the FDA's new authority over the investigational drug process granted by the 1962 Amendments. Under the final rule, the IND submission would include an 'outline of any planned phase or phases of the planned investigations' (FDA 1963, 179), and the FDA provided a description of what kinds of phases they expected to be typical of such investigations (more on the clinical phases follows below). Yet there was still a world of specifics left undefined. While the statistical techniques associated with randomization and controlled trials had been in development for some time, and while the RCT was generally recognized by the end of the 1950s as a powerful technique for evaluating the efficacy of drugs, the procedures for designing adequate and well-controlled trials were not widely understood or practiced (Meldrum 2000). Moreover, even as the procedures and techniques for clinical trials were becoming more clearly defined, there were no standards for judging the 'adequacy' of a given clinical trial in proving efficacy. Indeed, Temin's (1980) discussion of physician prescribing practices demonstrates that the scientific literature on drug evaluation from this period is replete with inconclusive studies (see Temin, chapter 5) and there was still a strong reliance on physicians' clinical judgment from individual case reports.⁴ Therefore, in an awareness that the new requirements would necessitate the organization of a new regime of expertise and practice, implementation of the new standard for effectiveness in the 1962 Act was delayed two years 'to allow the drug firms to adjust to the new conditions' (Ibid 125) .

⁴ An excellent discussion of the traditional importance of the 'case' in medical training and clinical judgement can be found in the first chapter of Rothman (1991). Also, Marks (1988, 1997) argues that the barrier to accepting statistical techniques for therapeutic evaluation was social and organizational at least as much as it was cognitive. While Rothman's book is concerned with the rise of medical ethics and the increasing oversight of medical research in the 20th century, his work intersects with Marks' in interesting ways. Rothman traces a history of medical decision-making in which the individualistic bedside judgement of the physician of the first half of the century is by the 1970s replaced with 'debate and review by colleagues and laypeople' (2) — second opinions, institutional review boards, and ultimately the movement towards evidence-based medicine (Berg 1997; Timmermans and Angell 2001; Timmermans and Berg 2003). Marks traces a parallel transition, during which time the individual assessment of therapeutic efficacy gave way to 'rational therapeutics'; i.e., the individual accrual of knowledge to manage chance and randomness in clinical therapeutic testing gave way to an ensemble of methods and techniques.

The ensuing years would be characterized by an effort to define in practice the principles for clinical drug evaluation established in statute. Moreover, the development of this definition would become intertwined with the effort to judge the adequacy of already-marketed pre-1962 drug products. Under the new law, the pre-market approval requirement applied to ‘new drugs’; however, there were a great many existing drugs on the market for which no evidence of efficacy had ever been proffered, or in some cases even existed. The new law defined a ‘new drug’ as one ‘not generally recognized by experts as safe and effective for its labeled uses’ (Merrill 1994, 51) — a standard leaving open to debate *which* pre-1962 drugs were generally recognized as safe and effective.

In 1966 the FDA contracted the National Academy of Sciences/National Research Council (NAS/NRC) to examine the evidence for efficacy of over three thousand drug products in what was known as the Drug Efficacy Study Implementation (DESI). According to Hilts’ (2003) journalistic account of the study, the NAS/NRC panel was often aghast at the lack of evidence supporting the use of many common drugs and found a great many marketed drugs for which there was no evidence to support any claims of efficacy whatsoever. Among the first of the panel’s conclusions transmitted to the FDA in 1968 was that certain antibiotic products fell into the category ‘ineffective as a fixed combination for the indications specified in the labeling’ because the components of such products had not been proven to contribute to overall effectiveness individually.⁵ In December of the same year the FDA announced its intention to remove a combination product, Upjohn’s Panalba, from the market unless interested parties could produce data supporting the effectiveness of the combination. Then in May 1969 it revoked the regulations written in the 1950s for certifying batches of fixed-combination antibiotics (FDA 1969a). The manufacturer of the product, Upjohn, submitted to the FDA a vigorous formal objection to the order, claiming that the NAS/NRC report was in error, that existing scientific literature established the effectiveness of the product, and requesting a formal hearing on the company’s objections. FDA responded in August

⁵ Temin (1980) refers to this category of drugs as ‘an ideological opposition to fixed-combination drugs that was a clear expression of a particular academic view of medicine’ (129) while Abraham (1995) treats this category as sensible and criticizes the FDA for not moving more quickly against combination antibiotics. . Certainly for the case of antibiotics the category makes sense. Even at the time it was understood that overuse of antibiotics could lead to resistant bacteria strains (see, e.g., Mintz 1965). Hence, the unwarranted use of antibiotic combination products was ill advised. Moreover, the risk of adverse effects and drug interaction increases with each added component in a combination product; therefore, from the perspective of risk management it is sensible to restrict drug products to as few components as possible generally, and to require justification for each component.

1969 by publishing a detailed critique of the literature submitted by Upjohn in support of its objections and granting a hearing (FDA 1969b).

The FDA's critique of Upjohn's evidence for Panalba illustrates the difficulty of enforcing a new concept of 'evidence' in drug evaluation and approval. One of the published studies submitted by Upjohn in support of Panalba was not even a study of Panalba; according to the FDA critique, the antibiotic product tested in the study had 'a different composition entirely' (FDA 1969b, 12966). Of the same study, the FDA added that 'Upjohn has previously been informed that Dr. Carter is not a qualified investigator, is ineligible to receive investigational new drugs, and that his studies cannot be used to support claims of safety and efficacy' (12966). Another published article submitted by Upjohn was a case study of a single patient in which Panalba was administered following penicillin. The patient improved, but the FDA commented that this account of a single patient's experience offered no evidence that *both* of the antibiotic components of Panalba contributed to the improvement (12965). Additionally, the FDA frequently noted a lack of appropriate laboratory cultures or controls. Indeed, the FDA found that of the 52 studies from the literature submitted by Upjohn, only two had any sort of controls, but were not necessarily blinded or randomized. Additionally, in the review of Upjohn's originally submitted data to support the approval of combination antibiotics, the FDA 'concluded that the clinical studies included in this material consisted of individual case reports, some in the form of testimonial letters, frequently without laboratory cultures, without protocol, or uniform dosage', were not controlled and did not compare the activity of individual components to the combination product (12966). Hence, seven years after passage of the 1962 amendments and five years after the two-year grace period allowed for implementation of the new requirement, manufacturers like Upjohn still hoped to pass with the forms of evidence characteristic of the pre-1962 era — case reports, testimonial letters, uncontrolled and often poorly executed studies. When the FDA rejected their arguments and moved to withdraw the drug from the market, Upjohn sued the FDA (Upjohn vs. Finch).⁶ The FDA's position was ultimately upheld by the courts.

In the meanwhile, the FDA acted to specify the rules for hearings in such cases, moving simultaneously to clarify and more explicitly define the evidence required to prove

⁶ This was only the beginning of a series of challenges to FDA's authority and basis for removing drugs from the market. See the FDA 1974 for list of eight unsuccessful suits decided between 1970 and 1973.

efficacy claims. In September 1969, the FDA published a series of ‘principles’ which it said had been ‘developed over a period of years and are recognized by the scientific community as the essentials of adequate and well controlled clinical investigations’ (FDA 1969c, 14596). These principles included, among other things: a test subject selection procedure which would allow ‘[a]ssignment of the patients to test groups without bias’; a description of ‘the steps taken to document comparability of variables’ such as age, sex., etc.; a description of the methods of recording and analyzing patient response variables and ‘the means of excluding or minimizing bias from the observations’; and a ‘precise statement of the nature of the control group against which the effects of the new treatment modality can be compared’ (14596-7). The three types of possible controls were listed as placebo, active drug control (comparison with another drug known to be effective in the condition under study), and historical control. Notably, the definition of ‘historical control’ in this document is written relatively broadly. The regulation specified that for ‘diseases with high and predictable mortality’ or for diseases having ‘signs and symptoms of predictable duration or severity’, the results of testing of a new drug ‘may be compared quantitatively with prior experience historically derived from the adequately documented natural history of the disease in comparable patients with no treatment or with treatment with an established effective therapeutic regimen’ (14597). Hence, while ‘historical control’ could mean comparison of a contemporary treatment group with a specifically identified previously treated (or untreated) group of similar patients, this definition appears to open the door to ‘prior experience’ so long as that experience is ‘adequately’ documented and can form the basis for quantitative comparison. The regulation was clear that ‘[u]ncontrolled studies or partially controlled studies are not acceptable evidence to support claims of effectiveness’ (14597). An ‘inadequately controlled’ study was defined as one in which ‘patient selection criteria are not adequately defined, investigator bias is not minimized, or an inadequately sensitive method of observation and evaluation of results is employed’ (14597).

These ‘principles’ were challenged in court because they had been issued without a notice of proposed rule-making and comment period. A proposal for the same set of rules was therefore published again in February 1970 (FDA 1970a) and the Final Rule with responses to received comments was published in May 1970 (FDA 1970b). The stated criteria for adequate and well controlled studies in this May 1970 version of the rules was virtually identical to the 1969 version with two significant exceptions. First, the

final rule contained provisions for drug sponsors to apply for a waiver stating reasons why the specified criteria ‘are not reasonably applicable’ to the clinical investigation in question, and describing alternative procedures used and results obtained (FDA 1970b, 7252). Second, uncontrolled or partially controlled studies were now unacceptable as the *sole* basis for approval of claims of effectiveness. Such studies, ‘carefully conducted and documented, may provide corroborative support of well-controlled studies’ and may yield ‘valuable data regarding safety of the test drug’ (7252). Nevertheless, ‘[i]solated case reports, random experience, and reports lacking the details which permit scientific evaluation will not be considered’ (7252). In these modifications, the FDA made reasonable allowances for existing practice to provide support for the approval of a drug, while also upholding the primacy of controlled studies for decision-making.

Together, the 1962 amendment and the 1970 *Federal Register* publication created the basic regulatory framework guiding drug development and approval in the United States: the need to demonstrate *both* safety and efficacy in humans in a controlled and statistically calculable way has been a defining characteristic of the clinical trial system ever since — although, as Meldrum (1994; 2000) notes, departures from the ideal design for the sake of goal oriented problem-solving have been common. As we will see, the refinement of these rules and standards will continue. It is important to underscore here, however, that the procedures and standards for conducting clinical trials to generate evidence of effectiveness did not come fully formed from the 1962 Act; rather, these rules were forged and clarified through a social process of challenge and conflict. The Harris-Kefauver amendment, so often treated as a *conclusion* point in history, in fact marked an *inflection* point in a confluence of oppositions.

One such opposition rests the rise in the 20th century of reformers working to establish an ‘independent science of drug evaluation’ — a ‘rational therapeutics’ based on scientific comparisons (Marks 1997). Controlled trials have existed at least since the 1740s when James Lind famously conducted an experiment to determine the most effective treatment for scurvy (Ibid, 5; Meldrum 2000). However, Marks shows how the changing material circumstances of experimentation, and consequent changing ideas about experimentation (including changing ideas about what it meant to do a ‘well-controlled’ study), induced shifts in perceptions of which groups were best suited for the work: from community-based physicians of the 19th century to university-based ‘cooperative studies’ of the 1930s and 1940s to the realm of the statistician beginning in

the 1950s and beyond (6).⁷ The growing dominance of statistics in controlled studies would require a displacement (or appropriation and modification) of previously endorsed modes of evaluation. The value of individual medical judgement would necessarily undergo somewhat of a devaluation and compromise to accommodate new forms of scientific and ethical decision-making in therapeutic evaluation

A second historical opposition feeding into this inflectional moment involves effectively a power struggle between industry and those who would seek to regulate it. The 1962 introduction of a requirement for drug effectiveness was by no means the first attempt to establish this criterion as a condition for marketing. As Temin (1980) noted, the first attempt to assure that all marketed drugs are effective was overruled by the U.S. Supreme Court in 1910 (*U.S. v. Johnson*) on the grounds that effectiveness of drugs ‘was a matter of opinion’ (Temin 1980, 125). Moreover, according to Temin, an early draft of the 1951 Humphrey-Durham Amendment to the *Federal Food, Drug and Cosmetic Act* (the statute creating the category of ‘prescription-only’ drugs) contained an effectiveness provision which was removed in the legislative process prior to passage of the bill (125-6). Nevertheless, the pre-1962 FDA considered efficacy part of its judgment of safety, reasoning that any assessment of the acceptable level of risk posed by any drug (safety) had to be made in light of that drug’s potential benefits (efficacy), although the FDA’s ability to block market entry of new drugs on this basis was limited. Meanwhile, in the post-World War II era, the pharmaceutical companies made the transition from small firms to large ones (Temin 1980), growing rapidly in no small measure on the strength of their potent antibiotic products. This greater size and affluence translated into greater political power and, as the Kefauver hearings exposed for all to see, these large firms acted with virtual impunity when it came to pricing, testing, and marketing their products. That Kefauver’s efforts to rein in the industry were nevertheless on the verge of failure, lacking support from many in Congress and the White House prior to the revelations about thalidomide, demonstrates the sort of high threshold of urgency and public outrage needed at this time to challenge the industry. This observation would tend to support the idea (from the summary of regulatory theory discussed in Chapter 1) that when narrow interests are set on securing a particular policy, the public mood must be highly charged (i.e., low slack in the citizen-legislator relationship) to oppose it. Even so, we will see

⁷ Cooperative group studies are still done, but such groups now also include biostatisticians. On the development of oncology cooperative groups, see Keating and Cambrosio 2002, 2007.

another example in Chapter 8 where this principle appeared to be at least partly contradicted.

All of the above themes and oppositions intersect in the Kefauver hearings and the post-1962 conflicts over the standards which should be applied in practice. Throughout the 1970s, the FDA would continue to augment and clarify its positions through the rule-making process, turning its eye to issues such as human research ethics and obligations of clinical investigators. It would also seek to improve the organization of its own activities and rulemaking (including a massive restructuring of the *Codes of Federal Regulation* to accommodate existing regulations and allow room for new categories of rule writing). Some of this refinement-in-practice was simply a continuation of the process of clarifying standing rules and (inevitably) making room for new ones. However, many of FDA's actions during this period can be seen as defensive, seeking to mitigate the growing chorus of complaints that would come not only from industry during this period, but from critics in academia and government as well.

3.3 Rising Criticism and Efforts for Reform

While large controlled trials were generally required as proof of efficacy, getting to the point that one can perform such studies is no simple matter. The system has evolved over time to balance the often contrary demands of ethics and statistical validity. To demonstrate the efficacy of a drug in controlled studies, a group of subjects receiving the investigational drug must typically be compared to a group of subjects receiving either a placebo or an alternative therapy. Generally speaking, the smaller the response difference to be discerned between the 'active' and control arms of the study (e.g., in a comparison of two cancer treatments having similar response effects), and the greater degree of certainty desired, the larger the study population needs to be. However, before one can ethically use an investigational drug on a large study population, clinical investigators must understand the toxic effects of the drug and must have determined the highest dose level at which the drug is tolerated by patients (the so-called maximum tolerated dose, or MTD). Even once that information is in hand, investigators are still not ethically justified in using a potentially dangerous drug on a relatively large patient population unless there is some evidence to suggest that the drug is effective at a safe dose level. Accordingly, a three-phase drug development process has evolved for drug evaluation, with each phase

of evaluation building on, and justified by, the information gleaned from previous phases. As we will see below, the definitions of these phases and their use in practice has changed over time, with consensus being somewhat elusive. For now, the issue to underscore is that the drug sponsors were not merely expected to perform controlled studies in the simple sense represented by James Lind for scurvy. Rather, the controlled studies needed for drug approval were the culmination of stages or phases of investigation which could take some years to complete.

Given that proof of efficacy for new drugs was *not* required before 1962, any new such requirements would be expected to result in a longer and more involved drug development process. As Merrill (1994) has observed, ‘That the 1962 amendments would increase the cost and time required to introduce new drugs was surely predictable’ (53). Nevertheless, some observers seemed taken aback by the entire process. Some pharmaceutical representatives complained (somewhat disingenuously, it seems to me) that they were ‘taken by surprise’ by the FDA’s interpretation of the new statute, expecting that older forms of evidence would still be accepted (FDA 1974, 9752). Others expressed alarm at the degree to which drug development was lengthened. The time required to get a drug through the development pipeline escalated steeply in the 1960s and 70s, going from roughly two years prior to the 1962 drug amendments to eight years or more by 1980. This substantial slow-down of the drugs ‘pipeline’, combined with what many saw as an adversarial and obstructionist FDA, led to harsh criticism for, among other things, hampering innovation in drug development, inducing a decline in the introduction of new drugs, unnecessarily delaying the introduction of important new drugs into the U.S., and putting U.S. drug firms at a competitive disadvantage to companies overseas, which were gaining the upper hand in introducing drugs to the marketplace more quickly than their American competitors.⁸ I will sketch the chief features of these complaints and the FDA response in the following sections.

3.3.1 *FDA Under a Microscope*

Peltzman (1973, 1974), a devotee of the public choice school of the theory of regulation (Chapter 1), provided an early, influential criticism of the 1962 amendments based on a comparison of the quantity of new chemical entities (NCEs) introduced before and after the amendments. On this basis, he argued that the requirement for effectiveness

⁸ See also Grabowski 1982; Gelijns 1990.

and need to await FDA approval was more costly to patients in terms of delay of new drugs than it was beneficial in preventing disasters like thalidomide. Wardell (1973; 1978) introduced the idea of ‘drug lag’ — delay in introduction of new drugs compared to other industrialized nations.⁹ Wardell’s analysis was widely aired. He was invited to testify in joint hearings before two Senate subcommittees investigating new drug research and FDA decision-making (U.S. Senate 1974), among other venues. Calls for repeal of the 1962 amendment began to be heard even from highly distinguished figures such as the economist Milton Friedman (HEW 1977a, 1). Meanwhile, government regulation was often represented as bumbling and inefficient.¹⁰ Indeed, President Jimmy Carter, in his 1978 State of the Union Address, referred to his Administration’s efforts to turn ‘the gobbledygook of Federal regulations into plain English’.¹¹ In a post-Watergate era characterized by cynicism towards government and poor morale at the FDA (Hilts 2003), an anti-regulatory backlash began to reverberate; these criticisms and others like them were repeated and elaborated throughout the decade.

This was a time when investigations and inquiries multiplied dizzyingly, much of which would ultimately feed into legislative efforts.¹² Senator Gaylord Nelson (D-WI)¹³

⁹ Temin (1980) addressed some of these criticisms. Among other things, he argued that the fall-off in the introduction of new chemical entities preceded 1962 and was concentrated in central nervous system drugs, especially tranquilizers. He speculated that the difficulties with thalidomide might have discouraged the introduction of other drugs with similar effects. Alternatively, he noted that Hoffman LaRoche’s phenomenal success with Librium and Valium were ‘reportedly’ due in part to their having patented the chemicals that might have been able to compete with these drugs, effectively foreclosing on the ability of competitors to develop tranquilizers (143). He also argued that foreign drug firms were already in ascendance prior to 1962. For the full argument, see Temin (1980) chapter 7.

¹⁰ For example, in a remarkable advertising campaign paid for by the petroleum giant Mobil, a cartoon depicts two children (a girl and a boy) playing ‘telephone’ with two cans attached by a long string. A man wearing a suit and carrying a briefcase takes the can from the girl. With the boy still listening on the other end, the man says into the can, ‘I’m with the Federal Communications Commission. Do you have a license?’ (Display Ad 667, *New York Times*, 8 August 1976, p. 154.) This cartoon is imbedded in a series of paragraphs, each of which describes one vignette of regulatory blundering or excess, including one concerning the FDA. (Each anecdote is taken from a book called *The Trenton Pickle Ordinance and Other Boneheaded Legislation*)(Hyman 1976). The final paragraph of the advertisement says, ‘What’s your favorite horror story? We know we’re not alone in the desire to stem the tide of over-regulation.’ The advertisement then invites readers to post anecdotes to the company, with a reward of a \$25 U.S. Savings Bond if the entry is selected for publication.

¹¹ For the text of the speech, see the Jimmy Carter Presidential Library website, <http://www.jimmycarterlibrary.gov/documents/speeches/su78jec.phtml> (accessed 12 May 2007). The speech is also available in *Compilation of Presidential Documents* Vol. 14 no. 3 (1978).

¹² A search in Lexis Nexis Congressional for hearings on the ‘Food and Drug Administration’ between 1970 and 1980 resulted in 327 hearings, 65 of which (depending on how one categorizes them) were directly related to oversight of FDA policies, procedures, and drug-related decision-making; pharmaceutical company practices in preclinical and clinical study; proposed legislation for FDA reform; and the safety and effectiveness of marketed drugs.

began congressional hearings in 1967 which would extend to 37 hearings over 150 hearing days (Nightingale 1981). While the Nelson hearings were focused mainly on the pharmaceutical industry, Senator Edward Kennedy (D-MA) spearheaded another set of hearings into the system for drug approval which spanned 35 hearing days between 1973 and 1978 (Nightingale 1981). The Kennedy hearings partly overlapped with an investigation conducted by an expert review panel assembled by the Department of Health, Education and Welfare (HEW), the Review Panel on New Drug Regulation, also called the Dorsen panel for its Chairman, Norman Dorsen, who was a professor of law at New York University.¹⁴ In May 1977 the Review Panel concluded a two-year study and published four volumes of reports refuting the analyses of critics like Peltzman and Wardell, and making recommendations to strengthen and improve the FDA's drug review process (HEW 1977a; 1977b; 1977c; 1977d).

The influence of the Dorsen report is notable throughout this period in efforts to write legislation. Moreover, while none of the proposed legislation discussed in this chapter (below) became law, many of the report's recommendations nevertheless carried through into regulatory reform and actions often associated with the ensuing decade. The Dorsen panel was clearly concerned not only to examine the claims of critics like Peltzman and Wardell and respond to them, but in having decided that the benefits of the current regime of drug regulation were worth the cost, the panel sought to improve and strengthen the FDA. To this end, the panel recommended programmes for enhancing the scientific reputation, training and involvement of FDA staff. It recommended greater openness and public accountability of agency decisions to foster public confidence, suggesting that the FDA should include a voting 'public interest' representative on each advisory committee to 'have a voice in deciding whether the social benefits of new drugs outweigh their risks' (HEW 1977d, 5). It also sought to adjudicate a controversy over whether FDA reviewers had been pressured by senior staff to make decisions in favour of industry.¹⁵ Most notable for this thesis is its assertion that while completion of all phases

¹³ Conventionally, Congresspersons are identified parenthetically by their party, 'R' for Republican or 'D' for Democrat, and the state they represent (in this case, 'WI' for Wisconsin. Standard postal codes are typically used to identify the state.). I will use this convention when first identifying Congressional actors in this account.

¹⁴ The HEW is now called the Department of Health and Human Services and is the parent organization of the FDA, National Institutes of Health, and Centers for Disease Control, among others.

¹⁵ According to Hilts (2003), the Nixon administration felt the FDA to be composed almost entirely of Democrats and sought to place Republicans in high-level Agency positions. With industry-cooperative Republicans in key positions, rank-and-file medical reviewers felt harassed by senior staff to approve drug

of clinical testing should be required for ‘most’ drugs, ‘*accelerated approval of new drugs*’ should be permitted in certain exceptional cases, such as those involving a drug that represents a major therapeutic breakthrough’ (HEW 1977d, 6, emphasis added). Here the Dorsen panel clearly contemplated the feasibility of approving certain ‘breakthrough’ drugs before all of the clinical drug testing was completed. More than that, the report urged Congress to give the FDA greater *postmarket authority* to limit the distribution of certain highly toxic drugs and ‘to require sponsors to conduct additional research either as a condition for approval or after a drug has been marketed’ (9). The report envisioned such postmarketing authority to be necessary only in ‘exceptional cases’ (9). Hence, clearly the review panel anticipated certain approval situations, however unusual, in which important questions about the drug would still need to be answered at the time of approval. In this thesis we will trace the development (and also the linking) of these concepts of accelerated approval and postmarket study from this time in the mid-1970s through various forms over the next decade-and-a-half. The Dorsen report also encouraged a more collaborative approach between the FDA and drug sponsors, recommending that the FDA hold periodic conferences with drug sponsors ‘during clinical testing of a drug, before the filing of a new drug application, and at key times during review of the application’ (7). We will see that the FDA did implement this recommendation and that this kind of collaborative approach will be used to great effect for development of the first AIDS drug, AZT. Finally, I should note that the Dorsen report asserted that the requirement that drugs be proven ‘safe’ and ‘effective’ was ‘imprecise’ since ‘no drug is absolutely safe or always effective’ (6). Therefore, it suggested that the statutory standard be amended to reflect explicitly the kind of risk-benefit assessment required for new drugs — one which takes

applications — and in time some of these reviewers were ‘neutralized’ by transferring them to other posts within the FDA. These disgruntled staff made their story public in one of the Kennedy hearings in August 1974. The new FDA Commissioner, Alexander Schmidt, was called to testify in this hearing and was blindsided by these revelations (see his side of the story at the FDA oral history website, <http://www.fda.gov/oc/history/oralhistories/schmidt/default.htm>). Schmidt launched his own investigation and found no wrongdoing, although the Dorsen report found Schmidt’s enquiry inadequate. Nevertheless, the Dorsen panel (who hired a Special Counsel for investigative purposes) found no wrongdoing or undue industry influence. Rather, it concluded that since 1970 the agency had been making a conscious effort to be more cooperative with drug-makers and that FDA management did in fact seek to neutralize medical officers perceived as uncooperative, doing so in a fashion which, while not illegal, was ‘unprofessional and discourteous’ (HEW 1977d, 17). Abraham’s (1995) account of this episode makes much of the neutralization policy of the time, quoting from the Dorsen report frequently to establish the inadequacy of Schmidt’s investigation. However, Abraham dismissed the conclusion that no specific bias was found in drug approval decision-making on the grounds that the Special Counsel did not assess the scientific basis for decisions made.

into consideration not only clinical data, but also ‘the severity of the disease to be treated, the availability of alternative remedies, and the drug’s public health implications’ (6).

In the aftermath of the Dorsen report, Senator Kennedy held another hearing on proposed legislation (U.S. Senate 1977). A total of 17 Senate bills were discussed in two days of hearings held on 26 and 27 July 1977. The most substantive of these proposals, Senate bill S. 1831, was the primary focus of the hearings. Called the *1977 Amendments to the Federal Food, Drug, and Cosmetic Act*,¹⁶ the bill introduced some far reaching proposals to strengthen the FDA, including a direct statutory directive authorizing the existence of the FDA and the creation of explicit FDA powers for postmarketing surveillance and withdrawal of approved drugs. More important for this study, the bill sought to create a form of expedited drug approval and to require postmarket studies — two among many provisions of the bill taking up recommendations from the Dorsen report. I will describe some key provisions of this bill in the sections which follow.

Also in May 1977, the Chairman of the Congressional Subcommittee on Domestic and International Scientific Planning, Analysis, and Cooperation, a subunit of the House Committee on Science and Technology, requested that the General Accounting Office (GAO) undertake an investigation of the FDA approval process (GAO 1980). The research questions put before the GAO by the subcommittee chairman clearly reflected the arguments of critics like Peltzman and Wardell:

- (1) whether there are inordinate delays in processing and approving new drugs for marketing in the United States;
- (2) whether delays in approving new drugs adversely impact on the introduction into the United States of therapeutically important drugs that are available in other countries;
- (3) how FDA’s drug approval process compares with approval processes of

¹⁶ In the U.S., references to legislative proposals are typically identified with the title they would have once made into law along with a numeric designation for the bill (See, for example, listings of current bills provided at the Library of Congress website for legislative information, www.thomas.gov.) The way to distinguish laws from bills is through the numerical designations assigned to each. Legislative bills introduced into the U.S. House of Representatives are assigned a number starting with the abbreviation ‘H.R.’, which stands for House Resolution. Numbering for Senate bills begins with the abbreviation ‘S.’, standing for Senate. Proposals which have actually been made law are numbered beginning with ‘P.L.’, which stands for Public Law. The numbering for public laws includes also a numerical descriptor for which Congress passed the law. Hence, P.L. 105-78 refers to the 78th law passed by the 105th congress, which served during the 1997-1998 legislative session. I will provide the appropriate numbering designation parenthetically for each bill or law discussed. There are some other types of congressional documents (concurrent resolutions, joint resolutions, etc.) which will not appear in this thesis. For more information, see the congressional glossary and other detailed information on the U.S. lawmaking system at www.thomas.gov.

other technologically developed countries; and (4) whether innovative use in computer technology could eliminate inordinate delays in the drug approval process (Ahart 1979, 1).

Two years later, in June 1979, the investigation was nearing completion and a GAO representative, Gregory Ahart, gave testimony before the Congressional subcommittee (now merged into the Subcommittee on Research, Science, and Technology) to summarize the findings. Ahart testified that new drug application (NDA) approval generally took many times longer than the statutorily required 180 days; that a drug lag existed such that therapeutic drugs, many of them considered significant, were sometimes getting approved in other countries years before they were approved for use in the U.S.; that FDA guidelines were vague and subject to various interpretations; and that no mechanism existed for resolution of scientific and professional disagreements between industry and the FDA.¹⁷

Meanwhile, President Carter made a pledge in his January 1978 ‘State of the Union Message’ (a written message sent to Congress, delivered in tandem with his State of the Union Address) to propose legislation ‘to reform regulation of the drug industry, which will protect the consumer and make regulations fairer and less burdensome’ (President 1978a, 109). Later that year, on 16 March, proposed legislation called the *Drug Regulation Reform Act of 1978* (S.2755) was introduced in the Senate by Edward Kennedy while a similar bill (H.R. 11611) was introduced in the House by Rep. Paul Rogers (D-FL). The HEW (later to become the Department of Health and Human Services, or HHS) characterized the bill as the ‘fulfilment of President Carter’s plan for human drugs’ (HEW 1978, preface). The bill contained substantial changes to existing practices for drug development and approval, including a ‘monograph’ system under which a distinction was made between individual drug ‘entities’ and drug ‘products’.¹⁸ For this thesis, the most important reforms proposed in this bill have to do with its definitions of safety and efficacy for approval, access of ill patients to investigational drugs, and proposals for

¹⁷ The GAO’s final report resulting from this investigation was published the following year (GAO 1980).

¹⁸ Sponsors of drug products containing new chemical entities would need to submit to the FDA a monograph containing chemical-specific data on the drug entities and also a new drug application (NDA) containing safety and efficacy data for the proposed drug product. While the NDA and monograph can contain similar information, the monographs would be public documents while the New Drug Application (NDA) for a particular product would contain manufacturing and other proprietary data and be closed to public scrutiny. Hence various products containing the same entity could be approved under the same monograph, but each new product would require a separately approved NDA.

‘provisional approval’ for drugs intended to treat life threatening diseases. In these provisions, we see tensions emerging between the goals of research and therapy, and also between the need for information weighed against the desire for decision-making to take place earlier in the drug development process for certain drugs perceived to be therapeutically important. How much evidence is sufficient to approve a drug on an expedited basis for very ill patients? How much information on efficacy should be acquired before allowing patients to gain access to investigational drugs? How can a drug even be perceived as ‘therapeutically important’ without a certain, relatively advanced body of clinical study? These types of questions, which were pressed to the breaking point during the AIDS crisis of the subsequent decade, find clear expression in the 1970s in response to the issues of the day.

3.3.2 Proposed Legislation and Conditions for Approval

President Carter’s plan for human drugs, as written into the 1978 Senate bill S.2755, essentially retained the definitions of ‘effectiveness’ and ‘substantial evidence’ already written into the law. An ‘effective’ drug was still considered to be one which, when used as directed, had the effect purported in the labelling; and this effectiveness would have to be proven with ‘adequate and well-controlled investigations’. However, according to the HEW the bill included an ‘important improvement’ over the current law because the ‘*significance* of the effectiveness of the drug is expressly considered for the first time in making decisions regarding the approvability of a drug’ (HEW 1978, 30; emphasis added). If this bill were passed, then for the first time the law would treat drug ‘safety’ as requiring an assessment of the net benefit of the drug entity weighed against its net known risks. The HEW wrote that the weighing of risk and benefit ‘reflects long-standard practice in evaluating drugs for approval’ (33). While true, it is doubtful that the FDA was free to consider as many factors as expressed in this bill. George Larrick, Commissioner of Food and Drugs, stated the traditional view in testimony before a 1964 Congressional subcommittee: ‘There is no such thing as absolute safety in drugs. There are some drugs that are less liable to cause harmful reaction than others, but people die every year from drugs generally regarded as innocuous’ (U.S. House 1964, 147). Accordingly, said Larrick, quoting a decade-old statement from Dr. Torald Sollman, ‘the administration of potent drugs involves a “calculated risk” where the presumptive benefit is balanced against the possibility of toxic effects’ (147). Hence, even before proof of efficacy was

required by statute, a ‘presumptive benefit’ had to be taken into account to assess the relative safety of a drug. Ultimately, however, the range of considerations available to the FDA for this judgement was limited. By contrast, in determining whether a drug is ‘safe’, S.2755 directs the FDA to consider, among other things:

- the health benefits of the drug and the *significance* of those benefits;
- known risks of the drug;
- benefits and risks associated with *other, existing forms of therapy* for the target disease; and
- the adequacy of label information to assure the correct use of the drug.

According to the HEW, the significance of the effectiveness of drugs is ‘often debatable, particularly in cases where the long-terms [*sic.*] effects in reducing morbidity of patients are not clear from physiological effects which the drug is shown to produce in the body’ (34-5). Hence, this assessment of the significance of benefit seems to imply a preference for drug entities for which clinical endpoints have been established, meaning direct measures of patient benefit (i.e., tracking the improvement of clinical symptoms of a disease or underlying condition, as opposed to surrogate endpoints, which are a measure of some physiological or laboratory parameter thought to be related to patient condition, e.g., blood pressure as a measure of cardiovascular health). The significance of benefit may also include consideration as to whether the new drug product offers advantages over other similar products in terms of dosage, route or frequency of administration, or other factors which might increase patient compliance.

In the proposed legislation, evaluation of the known risks of the drug entailed more than simply accounting for the adverse effects likely to be experienced by the patient. In some cases there were also possible risks to persons handling the drug. Moreover, there were potential societal risks. Was the drug entity likely to be abused intentionally? Was there a potential public health risk if the drug were not used as set out in the labelling? According to the HEW, ‘a drug that may be acceptable on a risk-benefit basis in a few individuals may be unacceptable when viewed from the standpoint of society as a whole’ (36). All of these types of ‘risks’ were eligible for consideration as part of the risk-benefit assessment.

While the HEW summary document stated that the provision ‘does not require that a drug offer a better risk-benefit ratio than currently available drugs in order to be approved’ (36), it clearly indicated that the risks and benefits of other drugs intended to

treat the same disease should be a matter of consideration in assessing the safety of a given drug entity and the significance of its benefits. The HEW cited a recent example of the FDA removing from the market a series of older products when an equally effective new product with a significantly better safety profile was introduced. Presumably the proposal would also allow the FDA to deny approval of a product having a risk-benefit ratio substantially inferior to already marketed products.

3.3.3 Proposed Legislation and Investigational Use of Drugs

The existing *Food, Drug and Cosmetic Act* with its amendments dealt with the subject of investigational drugs ‘in only four sentences’ (HEW 1978, 83). It authorized the promulgation of regulations for the investigational use of drugs and the imposition of conditions of use; it included requirements for the control and use of investigational drugs, for protection of human subjects through informed consent; and it contained reporting requirements. It had been incumbent on the FDA to fill in the details in regulation. Both the 1977 and 1978 bills (S.1831 and S.2755 respectively) attempted to address some of these details. Before examining these proposals, we must describe the existing IND regulations.

As previously noted, the original version of the IND regulations was proposed in 1962 (FDA 1962), as one of the first new regulations to be written with the FDA’s new authority over the investigational drug process granted by the 1962 Amendments, and was finalized in 1963 (FDA 1963). Under the final rule, the IND submission would include an ‘outline of any planned phase or phases of the planned investigations’ (FDA 1963, 179). The rule defined the first two phases of study as ‘clinical pharmacology’ while efficacy was only evaluated in the third phase. Phase I was the initial introduction into humans (only *in vitro* and animal data would exist at this time) in which investigators were to study human toxicity, metabolism, absorption, elimination, the preferred route of administration, and safe dosage range (180). Phase II covered ‘the initial trials on a limited number of patients for specific disease control or prophylaxis purposes’ (180). In the rule, only Phase III was designated a ‘clinical trial’, and was designed to provide ‘the assessment of the drug’s safety and effectiveness and optimum dosage schedules in the diagnosis, treatment, or prophylaxis of groups of subjects involving a given disease or condition’ (180). Hence, as summarized later by the FDA, these regulations ‘suggest[ed] that the real

demonstration of effectiveness does not occur until phase III, when controlled clinical trials are carried out' (FDA 1979a, 38).

The 1977 proposed legislation, S.1831, reiterated verbatim the descriptions of the clinical phases provided in the 1963 regulation, retaining the earlier scheme of clinical drug development, but adding a 'Phase IV' for postmarket study, discussed in the next section. However, the 1978 proposed legislation, S.2755, revised the formulation of the clinical phases, creating 'three distinct types or categories of clinical investigations involving unapproved drugs' (HEW 1978, 83). The first category was to be called the 'drug innovation investigation', during which a drug is first introduced into a small number of human subjects to study the clinical pharmacology of the drug entity. This investigational category essentially defined what was called a 'Phase I' clinical study in the 1963 IND regulations. The second investigational category proposed by S. 2755 was the 'drug development investigation', during which the effectiveness of a drug would be evaluated along with 'short-term and relatively common adverse effects of the drug' (84). Significantly, unlike the 1963 regulations, S.2755 did not identify a third phase of study in which effectiveness was determined; rather, effectiveness was to be studied in the drug development investigation, which was defined in terms clearly associated with what was previously called Phase II — a point to which we will return below.

S.2755 did propose a third category of drug investigation but it was not analogous to the 'Phase III' of the IND regulations. This proposed IND category was called 'drug treatment investigations', in which the 'small numbers of human participants' are given an investigational drug 'to treat a serious disease, injury, or condition that is not satisfactorily treated by other forms of therapy' (HEW 1978, 104, emphasis added). In such cases, the 'use of the unapproved drug is primarily intended to provide treatment to the participants rather than to assess its risks or effectiveness, although information relevant to such an assessment may be derived during the investigation' (84-5). The proposal would formalize in law a category of investigational drug use under which patients with *a demonstrable lack of therapeutic options* could gain access to potentially therapeutically important drugs before they were approved for marketing. This use of the investigational drug was to be seen *not as research*, but primarily as treatment, although certainly useful information could be derived in the process. Here we see developing a nascent tension between conducting research, on the one hand, and using investigational drugs merely as therapy, on the other.

The procedure described in S.2755 was a practice sometimes called ‘compassionate use’ or ‘compassionate IND’ (FDA 1979a) which had been informally used by the FDA since the 1938 act was first passed.¹⁹ This is our first example of (attempted) rule-writing following in the wake of informal practice, and the circumstances of it are notable. While a demand for investigational drugs long existed, it is not unreasonable to suppose that with the longer drug approval times resulting from the 1962 drug amendments, demand was growing for investigational drugs by patients lacking therapeutic options. If so, then the attempt to write legislation at this time was a recognition of, and response to, that growing demand. There was, in other words, an effective consensus over the informal practice in question; a *de facto rule existed*.

Under the proposed requirements for drug treatment investigations in S.2755, the drug product could only be distributed to ‘individual physicians with a limited number of patients with the disease, injury, or other condition to be treated with the [investigational] drug product’ (HEW 1978, 129). Hence, it was not necessary for the physician to be ‘an expert qualified by training and experience to investigate the drug’, as was required for registration as an investigator; rather, participation would be allowed for ‘any practitioner who is willing to accept the responsibility for use of the drug and to comply with all of the requirements of a drug treatment investigation’ (129). For a patient to be accepted for drug treatment investigation, the potential benefits needed to be judged to outweigh the risks of participation.

According to the HEW (1978), two situations would be most commonly expected for drug treatment investigations. The first situation was when patients who lacked therapeutic alternatives were believed to have ‘a reasonable possibility of being successfully treated with the drug product’ (85). In such cases, although adequate and well-controlled studies had not yet been completed, there should be ‘sufficient’ evidence of effectiveness to justify therapeutic use. If deemed appropriate, the therapy would be administered ‘subject to the normal requirements for a clinical investigation’ including requirements for informed consent of the patient (85). Although not passed as law at this time, this is essentially the version of pre-market access to investigational drugs we will see the FDA proposing as a regulation in 1979 into the early 1980s. Notably, in this proposal we see an awareness that granting access to still-investigational drugs means using the

¹⁹ According to the Senate report on the FDA Modernization Act of 1997 (U.S. Senate 1996), discussed in Chapter 8 of this thesis, the FDA had begun to practice some form of compassionate use not long after the passage of the 1938 Food, Drug and Cosmetic Act.

drug for therapy *before* the drug has been fully investigated. There is therefore a concern to make decisions on the basis of *less information than desired*, but with knowledge still ‘sufficient’ to justify therapeutic use of the drug. The question of ‘how much information is sufficient to administer the drug?’ was of course at the heart of the 1962 drug amendments. It will continue to be at issue as the FDA seeks to grant access to investigational drugs for certain patients and to expedite approval of certain drugs. Again, we have a nascent tension developing between sufficient information-gathering and earlier decision-making.

The second situation envisioned by HEW (1978) for drug treatment investigation was one in which a drug entity was denied approval for commercial marketing or withdrawn on the judgment that its risks outweighed its benefits, but there nevertheless remained a ‘very small group of patients to whom these risks may be acceptable’ (85-6). In such cases, the drug could either be administered as a drug treatment investigation or under special requirements for limited distribution and dispensing of the drug (provided for in another section of the bill). This version of treatment investigation is notable because it highlights a conflict the FDA has long struggled to resolve. Sometimes there is only a very small group of patients thought to benefit from a drug which is highly toxic or dangerous for other reasons. Some drugs require highly specialized training or equipment for proper administration, and pose an added risk to the patient or staff if not dispensed properly. Moreover, the ability of physicians to prescribe medications for conditions other than those indicated in the product labelling (a practice called ‘off-label prescribing’) has sometimes given the FDA pause when approving such drugs. Accordingly, the FDA has sought at various times to limit the distribution of certain drugs. However, they have never been able to achieve a consensus for these types of limitations. Such restrictions have typically been opposed as an attempt to regulate the practice of medicine. This provision was never made into law – and lacking a law such as this, opponents consistently argue that the FDA has no legislative authority for such restrictions.

The issue of consensus is key here. The FDA lacked specific ‘legislative authority’ for compassionate use, but was able to practice it informally and would eventually codify it because there had been no significant opposition to the practice. If anything, the evidence suggests there was a growing demand for it. By contrast the FDA has never been successful in developing a rule for limited distribution of dangerous therapeutic

drugs in large part because physicians, pharmacists and pharmaceutical companies all oppose it.

3.3.4 Proposed Legislation and 'Provisional' Approval

As noted above, the 1977 proposed legislation (S.1831) created a fourth phase of study. This Phase IV was by definition postmarket study which could be used for a range of purposes, including further controlled studies of effectiveness, tracking adverse reactions, tracking off-label prescribing practices and the patterns of use of the drug within various medical specialities (U.S. Senate 1977, 19-20). The FDA Commissioner would have the authority to institute Phase IV either at the completion of three phases of study or 'in combination with an abbreviated Phase III' (21). Among other conditions, Phase IV would only be initiated if the Commissioner determined an 'urgent need' for the new drug 'where there is no other comparable alternative drug on the market' (20), or if the drug were for a chronic illness where no suitable alternative therapies existed and 'short-term studies have been completed' (20). In other words, consistent with the recommendations of the Dorsen report, S.1831 introduced the idea of approving a drug on the basis of an 'abbreviated Phase III' study, or on the basis of 'short-term studies', when the drug was intended for a *serious disease* for which *therapeutic options were lacking*, on condition that the drug sponsor conduct postmarket investigations to fill information gaps apparent in the abbreviated data used for approval. While this bill was not passed into law, this provision substantially anticipates the approach that would be used to develop and approve the first AIDS drug, AZT, as well as regulations for expedited approval of drugs to be written more than a decade in the future.

The 1978 bill, S.2755, also contained a mechanism for expediting approval of certain drugs. According to the HEW, the bill recognized that sometimes 'it may become apparent, even before adequate and well-controlled clinical studies are completed, that a new drug constitutes a major therapeutic advance' (HEW 1978, 42). The bill envisioned that in such cases it would be desirable to allow patients 'to obtain the life saving or other major health benefits of a breakthrough drug earlier than is possible under the current law' while still assuring that 'if at all possible, the adequate and well-controlled trials regarding effectiveness will be completed' (42). To address such situations, the bill proposed a 'provisional monograph' which could be issued for 'certain drugs without completion of

well-controlled studies in very limited and narrowly drawn circumstances' (43). The provisional monograph would only be allowed:

- for drug entities intended for use in a 'life-threatening or severely debilitating or disabling disease or injury' (40);
- if the drug entity 'offers a major therapeutic advantage to patients with that disease or injury' (40);
- if the patient population at issue cannot be adequately served through drug treatment investigation ('compassionate IND');
- if delaying approval of the drug entity would pose a greater risk to patients than approving the drug;
- if the drug entity has been evaluated for risk;
- if there is significant evidence for the effectiveness of the drug entity, and if well-controlled studies are underway (if ethically and methodologically possible)
- and if the benefits of provisional approval clearly outweigh the risks from such approval.

A provisionally approved monograph would automatically expire after three years unless new evidence of effectiveness were presented. The provisionally approved product could also be more easily withdrawn from the market than a drug having 'full' approval (42-3).

The bill acknowledged that at the time of provisional approval, the FDA would have 'less evidence of effectiveness' than normally required, but still evidence 'sufficient to justify a finding that the drug entity offers major therapeutic advantages for patients' (43). The bill therefore attempted to define what this lesser standard of evidence would be, and created the category of 'significant evidence' of effectiveness (44), which it defined as:

Evidence consisting of valid and meaningful scientific investigations, including investigations in animals and well-documented clinical experience and clinical investigations (unless such investigations are not feasible due to an absence of methodology or due to ethical proscriptions on such investigations), conducted by experts qualified by scientific training and experience to evaluate the effectiveness of the drug entity involved, on the basis of which it could fairly and reasonably be concluded by experts, qualified by scientific training and experience to evaluate the effectiveness of the drug entity involved, that the drug entity will have the

effect represented in information labelling for drug products eligible to be licensed under the monograph (44).

The HEW explanation of the bill added for emphasis that under this test, ‘drug entities for which there is little or no scientific evidence of effectiveness could not be approved’ (44).

Like S. 1831 from the previous year, this bill sought to expedite the approval of certain drugs on the basis of abbreviated clinical data. However S.2755 was much more aware of the implications of doing so. This passage of the proposal explicitly recognized that making decisions for approval earlier in the drug approval process necessarily involves an *information deficit*. Although this bill was not made into law, is the first formal attempt to define a category of FDA drug approval decision-making based on consciously limited information, and to write a standard for the level of information deficit to be tolerated. Given the concerns such an information deficit raises, the bill sought to assure that study of the drug would be completed during the provisional period, and created a process to revoke the approval failing completion of the studies or in the event that the study results were unfavourable. We will see these basic concepts integrated into future regulatory reforms. We will also see in the next chapter (in the discussion of the first AIDS drug approval and its aftermath) that the challenges associated with this type of drug approval will lead to a new conception of *risk*. Whereas in S.2755 the type of ‘risk’ addressed has to do mainly with drug toxicity or ‘social’ risks, consistent with the view of risk reflected in the Dorsen report, in the future a new conception of risk will develop which specifically includes the *risk of information deficit*.

3.3.5 Aftermath

As already noted, neither S.1831 nor S. 2755 were made into law. The following year in the aftermath of Congressional midterm elections, new versions of the *Drug Regulation Reform Act* were introduced into the House and Senate with the hope of accomplishing in the ninety-sixth Congress what there had not been time to achieve in the previous one. Ultimately, however, the bill sponsors were never able to muster the support they needed for the proposed legislation; the 1979 versions of the bills stalled in their respective committees as well. At the same time as these bills were under negotiation in the Congress, the FDA was preparing its own plan for reform — a plan in

many ways consonant with the ill-fated S.2755. Through the FDA plan, a version of some of the concepts embodied in previous failed legislation would be written into regulation. Others would remain latent until the advent of AIDS.

3.4 FDA Response to Criticism: *Concept Document*

In the late 1970s, as major legislation for FDA reform rose and fell in the Congress, the FDA assembled a series of task forces to review the current state of drug review and regulation. Under the oversight of a steering committee, thirteen FDA task forces produced 14 concept papers making recommendations for changes to the IND and NDA regulations. These papers were bundled together into a comprehensive proposal for reform called *Investigational and New Drug Regulation Revisions: Concept Document* (FDA 1979a) (or *Concept Document* for short), published by the FDA in October 1979. Publication of the *Concept Document* began a protracted revision process that included public hearings, publication of a spate of new guidelines encompassing a range of topics and procedures, and a series of proposed revised regulations associated with the submission of IND applications (an application required to begin clinical evaluation of drugs on human test subjects) and New Drug Application (NDA) submissions (applications to submit clinical evidence of efficacy and safety of a new drug, and to request marketing approval for a specified indication).

The introduction to the *Concept Document* discussed the need for updating the rules and augmenting efficiency. The FDA wrote that the proposed revisions were necessary because the new drug approval process had ‘continued to evolve and become more complex’ while the associated regulations ‘remained essentially constant’ (4). Accordingly, the proposed revisions were designed to ‘expedite the review process, reduce paper work, and redefine the IND and NDA requirements in line with FDA experiences in current practices’ (5-6).²⁰ Similarly, the FDA’s Dr. Robert Temple stated in a November 1979 public meeting to discuss the *Concept Document*: ‘It is apparent that the phases as defined in

²⁰ According to the *Concept Document*, influential events feeding into these proposals were the ‘Auerbach effort’ in 1972 and 1973, which recommended changes in the paperwork and recordkeeping systems associated with IND and NDA submissions; conferences and workshops held in 1974 and 1975 which brought new thoughts to the Agency regarding new drug development and the controlled clinical trial; and the Dorsen Review Panel (1975-1977), which offered recommendations to make the drug review process more open and accountable, to increase FDA’s scientific capabilities, to improve the standards and procedures for pre-market approval of new drugs, and for increasing FDA’s authority in the post-marketing period. (FDA 1979, 4-5).

current regulations . . . no longer reflect the way we think of them' (FDA 1979b, 10). The same appears to have been true for the proposals for treatment IND, as stated by the FDA's Dr. Crout at the same public meeting, 'the intent is to change the regulations basically to legitimize what we do now' (36). Hence, with regard to these issues the *Concept Document* can be considered an attempt to reflect what the FDA felt was contemporary practice.

FDA actions for reform during this period were also undertaken with an awareness of the criticism directed against it and a desire to quell the controversy. In the same public meeting, the FDA's new Commissioner, Dr. Jere Goyan, gave a short but often eloquent introductory address, characterizing the development of human drugs as a balance between the hope of curing 'the ailments that beset us' and the fear that 'new, powerful, biologically active substances can do our bodies harm' (FDA 1979b, 2). Both the hope and the fear are real and rational, said Goyan, and the entire history of American drug law 'has been nothing more than an effort to assure that we are blinded neither by fear nor hope, and that approval has not been accorded to a drug because we hoped too much, nor denied because we feared too strongly' (2). He acknowledged the FDA's critics, saying that many people 'feel that our present drug approval process is governed by excessive deference to fear' (2). Listing a number of industry's key complaints, and noting the 'controversial and often emotional tone of such criticism', Goyan called for the Agency's 'very best response' to be carried out calmly and judiciously, with a willingness to accept necessary changes even while maintaining 'necessary protections' (3).

The document contained a wide-ranging set of proposals, addressing issues such as the format and content of IND and NDA submissions; the conditions and procedures associated with disqualification and reinstatement of clinical investigators; provisions for the termination of inactive INDs and NDAs; clarification and refinement of the regulatory definition of adequate and well-controlled studies; inclusion and use of foreign clinical data in NDAs; requirements for Abbreviated New Drug Applications (applications for 'me-too' drugs); postmarket surveillance of approved drugs, including procedures and requirements for adverse drug experience reporting; procedures for 'NDA supplements' to amend pending NDAs; plus requirements for recordkeeping by drug sponsors, procedures for hearings, and other topics. For the purposes of this thesis, the most relevant proposals in the *Concept Document* are those which redefined the clinical phases and kinds of investigation applicable under an IND; those which established a

formal basis for pre-market access to investigational drugs; and those which pertained to the definition of ‘adequate and well-controlled’ and other criteria to be considered in the approval of new drugs.

3.4.1 *The Concept Document and the Definitions of Clinical Trials*

As we have seen, the IND²¹ regulations included a description of the phases of clinical study that should be reflected in the investigational process. According to the *Concept Document*, the ‘working definitions of investigations have evolved to the point that they are significantly different from the definitions in the regulations’ (30). As noted above, the 1963 version of the IND rules defined the first two phases of study as ‘clinical pharmacology’ while efficacy was only evaluated in the third phase. The authors of the *Concept Document* wrote that these regulations ‘suggest[ed] that the real demonstration of effectiveness does not occur until phase III, when controlled clinical trials are carried out’ (FDA 1979a, 38). However, according to the IND Task Force responsible for this section of the document, actual practice had gradually diverged from this characterization. In contemporary (as of 1979) practice, Phase I was still for studying clinical pharmacology, while Phase II was ‘the time of the earliest, very rigidly controlled, clinical trials, often conducted in an in-patient setting, that provide evidence, meeting statutory requirements, of effectiveness’ (38). Awkward prose notwithstanding, it is important to note that the IND Task Force’s view of Phase II in actual practice was as a clinical study capable of producing efficacy data *defensible in light of statutory requirements* (i.e., in light of the definition of substantial evidence given in the 1938 *Food, Drug and Cosmetic Act*, as amended by the 1962 Kefauver-Harris amendment). Yet this was still considered the ‘earliest’ evidence of efficacy, often conducted in an ‘in-patient setting’, presumably meaning that the enrolled patients were relatively ill and the studies were relatively small and short in duration. In this revised version of the clinical trial system, Phase III was then the time when ‘expanded clinical trials in multiple settings and long-term studies are carried out and such questions as drug interactions are explored’ (38).

For its revised definition of the clinical phases, the *Concept Document* authors referred the reader to the then-recently prepared FDA publication, *General Considerations for*

²¹ Prior to 1979, IND stood for ‘Notice of Claimed Investigational Exemption for a New Drug’. The Task Force authoring the IND section of the *Concept Document* declared this name to be ‘a clumsy mouthful that appears in full only in regulations and in FDA-industry correspondence’ (46). The Task Force sensibly recommended the name of the application be changed to ‘Investigational New Drug Application, to be abbreviated as (fittingly) ‘IND’ (46).

the Clinical Evaluation of Drugs (FDA 1977)(here called *General Considerations* for short), which ‘redefined the phases to make them more compatible with current use’ (38). The *General Considerations* guidelines likewise characterized Phase II clinical investigations as controlled studies to demonstrate efficacy and safety ‘performed on closely monitored patients of limited number’, while Phase III were ‘expanded’ trials which could be either ‘controlled or uncontrolled’, performed ‘after effectiveness has been basically established, at least to a certain degree, and are intended to gather additional evidence of effectiveness for certain indications, and more precise definition of drug-related adverse effects’ (FDA 1977, 6).

In both the *Concept Document* and the *General Considerations* guidance document, we see consonance with the version of clinical development proposed in the 1978 Senate bill S.2755, where the analogue to the earlier phase III was omitted. Compared to the 1963 IND regulations (and also to the 1977 Senate bill S.1831) , all of these later documents redefined Phase II: it shifted from a secondary pharmacological study to a trial capable of providing ‘statutorily defensible’ proof of efficacy. Meanwhile, Phase III was effectively demoted, moving from the first real evidence of efficacy in the 1963 version of the clinical phases to something more of an operation in data refinement and supplementation in the 1979 version. This is not to imply that the study of long-term exposure and drug interactions are unimportant, nor is it to say that the FDA or Congress considered Phase III unnecessary. (Although S.2755 did not create a specific analogue to Phase III, the types of studies included in Phase III would obviously be subsumed under a ‘drug development investigation’. However for the proposed legislation, the investigation of efficacy was concurrent with study of ‘short-term and relatively common adverse effects’ [HEW 1978, 84], which we can safely equate with Phase II.) The point is that at this time Phase II had grown in importance in practice as the time when the effectiveness of a drug began to be established, and could at least in theory meet the statutory requirement for substantial evidence. In this way, the authors of the *Concept Document*, the *General Considerations* guidance document, and S.2755 were responding to changes in practice taking place in the research community as the techniques for clinical study were refined. We can observe once again that the FDA and the Congress were acting on what was, in effect, a consensus in practice — new rules in all but name.

In recommending the *General Considerations* guidance document, the authors of the *Concept Document* wrote that no further changes would likely be required to the definition

of clinical phases contained in *General Considerations* (38-9). In other words, they believed clinical practice had largely stabilized. However, in an FDA guidance document on clinical trials for anti-tumour agents published in 1981, the FDA advised the oncology investigator to ‘get early observations on therapeutic activity’ of the drug in *Phase I* (FDA 1981, 2, emphasis added), thus borrowing from Phase II in a modest redefinition of Phase I. Meanwhile, Phase II was characterized generally as a study in which the anti-tumour activity of the investigational drug (measured as tumour shrinkage, response duration, survival time, etc.) should be assessed against either varying dose levels or varying schedules of administration — i.e., the goal for Phase II was to identify a tumour type, dose level, or route of administration for which the drug is most efficacious. In Phase III, then, the effectiveness of the investigational drug in the most promising patient groups at optimal dosing level or routes of administration (as determined in Phase II) was compared to that of an existing cancer treatment or a placebo in a randomized, controlled study (5). Such disciplinary differences will continue to problematise generalizations about the clinical phases. Nevertheless, the movement towards gathering information on a drug’s therapeutic value earlier in the clinical development process is evident here as elsewhere.

The observations made in this section should come with one caveat. One wonders to what extent clinical drug development in practice ever conformed to the 1963 version of the clinical phases. The FDA clearly believed their 1979 revision to correspond with a change in practice. Indeed, it is reasonable to assume that changes in practice did occur over the period from 1963 to 1979, when use of the randomized controlled trial would have become more familiar to many researchers and became more refined through use. However, it is also clear that in 1963 the FDA’s goal was not to mimic previous widespread practice in drug development, which was then considered inadequate, but to create guidelines and standards for a *new* way of approaching clinical drug development. I therefore propose that we should think of the 1963 rules as an early ideal for how clinical studies should be designed (although no doubt there was some empirical basis for this ideal) and the 1979 revision as reflecting the practices which actually developed on a widespread basis over time.

To what extent can we see drug approval being granted on the basis of Phase II studies at this time? Publicly available drug approval data from this period tends to be fragmented and incomplete, but certainly examples of Phase II-based approvals can be found. According to FDA records, the first approval of the drug tamoxifen for breast

cancer in 1977 was based on four small (approximately 30 to 80 patients) Phase II studies. From the limited information available it is not clear what, if any, controls were used.²² A series of small studies characteristic of Phase II also formed the basis for the approval of cisplatin to treat both testicular and ovarian cancers in 1978. The same set of studies is listed by the FDA as supporting approval for both indications.²³ Even more notably, in 1983 etoposide appears to have been approved for use in twice-relapsed (third-line) testicular cancer patients on the basis of *one* Phase II study containing 53 patients.²⁴ We should note by way of context that testicular cancer is a relatively rare disease for which accrual to clinical studies is problematic. We will see another detailed example of the same researcher who conducted the etoposide study coming to the FDA with a similar application for treating third-line testicular cancer in Chapter 5 of this thesis. Here we can observe that at least in the realm of drugs for serious and life-threatening diseases, the changes made to the descriptions of clinical phases in 1979 did indeed reflect actual practice. Significantly, none of the approval decisions noted above appear to have taken place on the basis of *controlled* trials. As we will see in much more detail in Chapters 5 through 7 of this thesis, such exceptions to the rules come about as a pragmatic compromise necessitated by the nature of Phase II testing in oncology (and no doubt applicable to other medical specialties as well).

This apparent shift in practice and the effort to incorporate it into the formal rules represents, in retrospect, an important historical transition point. The outcome of this period, as reflected in new IND regulations (below) will be a definition of the clinical phases which still included Phase III, and the FDA would still generally require the kind of large randomized controlled study represented by Phase III to support NDAs.²⁵ However, we will see repeatedly that what is true as a *general* orientation is often contradicted in specific cases where extenuating circumstances are deemed to apply. The growing *formal* recognition at this time that Phase II could in theory provide sufficient

²² <http://www.accessdata.fda.gov/scripts/cder/onctools/studies.cfm?ID=61>.

²³ Compare the 'study results' listing at <http://www.accessdata.fda.gov/scripts/cder/onctools/summary.cfm?ID=71> and <http://www.accessdata.fda.gov/scripts/cder/onctools/summary.cfm?ID=73>.

²⁴ <http://www.accessdata.fda.gov/scripts/cder/onctools/studies.cfm?ID=82>.

²⁵ Almost a decade after the *Concept Document*, the FDA's description of Phase II was that it contained 50 to 200 patients and was considered to be an intermediate study 'in which the safety and efficacy of the drug are first evaluated in controlled trials' and Phase III contained from 200 to 1,000 or more patients (FDA 1988, 41518).

information about the effectiveness of a drug to serve as the basis for decision-making will open the door to a series of rule changes over the following decade-and-a-half.

3.4.2 *The Concept Document and Treatment IND*

The FDA's proposed reallocation of the initial demonstration of efficacy from Phase III to Phase II plainly resonated with concerns at that time about providing earlier access to promising new drugs, particularly for serious conditions where no other treatment was available. We have already seen this concern expressed in S.2755, with its proposal for a class of 'drug treatment investigations' for such cases. This concern was likewise reflected in the *Concept Document*, which recommended codifying existing informal practices for making investigational drugs available to very ill patients on a premarket basis when therapeutic alternatives were lacking.

As far as I have been able to determine, specific figures on the number and types of requests made for informal treatment access to investigational drugs are not publicly available. However, the measures taken by the FDA and the National Cancer Institute (NCI) at this time can be adduced to support the contention that demand for treatment access was rising through the 1970s. In the mid-1970s the FDA allowed large 'open-protocol' (unblinded) studies of cardio-selective beta blockers with the idea of expanding access to these experimental drugs for patients having heart disease (Temple 2005; U.S. House 1987). Moreover, in the mid-1970s the National Cancer Institute (NCI), in cooperation with the FDA, created a classification of experimental cancer drugs called 'Group C', under which patients were allowed to enrol in ongoing clinical study protocols. Notably, both of these mechanisms were designed to accommodate *groups* of patients (with no accommodation for individuals who did not meet the group profile, even if the group profile was somewhat relaxed to allow greater enrolment)²⁶ and ultimately had the primary purpose of experimental research and data collection, not therapy, despite their more open character.²⁷ For individual patients not able to participate in these clinical studies due to eligibility or geographic limitation, the only available mechanism for individual access to experimental drugs was the informal access

²⁶ See the discussion of individual vs. group expanded access in a meeting of the FDA Oncology Drugs Advisory Committee meeting (FDA 2000).

²⁷ This supposition was confirmed for me by Ellen Cooper in a telephone interview, 10 August 2006 (Cooper 2006).

discussed in Section 3.3.3 , sometimes referred to as ‘compassionate IND’ or ‘emergency IND’, and proposed in the *Concept Document* as ‘treatment IND’ (FDA 1979a).²⁸

From the FDA proposals to codify the ‘treatment IND’ regulations (FDA 1979a; 1983) it is clear that the FDA practiced its informal version of treatment IND as an ‘exemption’ under the rules prohibiting the sale or commercialization of a drug prior to marketing approval. Under the law, unapproved drugs could not be shipped over state lines unless they had received an exemption for investigational use; this is the exemption granted through the IND application process. If drugs granted investigational status were to be allowed for treatment use, the FDA reasoned that drug sponsors participating in the treatment IND program should supply the drug to patients without charge. In this way, ‘commercialization’ of the drug was barred and the FDA effectively mandated corporate altruism as a feature of ‘compassionate’ IND. Moreover, drugs used under an IND were, by definition, *investigational*; in stretching the regulation to fit a therapeutic purpose, the FDA required that the drug be distributed under an investigational new drug protocol: either under a drug sponsor’s IND and (ideally) existing treatment protocol; or through a physician-sponsored IND, under which a patient’s personal physician was treated effectively as a physician-investigator and required to create a treatment protocol to administer the drug. Under this mechanism, the FDA allowed experimental treatment of ‘hundreds’ of patients per year (FDA 1983) — a quantity sufficient for the 1979 *Concept Document* to recommend that the FDA ‘specifically recognize the concept of “treatment IND” in the regulations’ (FDA 1979a, 6).

The document specified that under a treatment IND, ‘the primary purpose of drug use is not investigation of its effectiveness, but the use of the drug (ordinarily one in Phase III) to treat patients with a serious illness not treated satisfactorily with alternative therapy’ (18). The FDA subsequently published specific proposals for the formalization of treatment IND in the 1983 ‘IND rewrite’ proposal (FDA 1983) which, consistent with the *Concept Document*, stated that the purpose of treatment IND should be primarily therapeutic; that the programme was designed to address the needs of relatively small

²⁸ The introduction to the *Concept Document* strings these terms together as if they refer to essentially the same practices. In other later materials, however, ‘emergency IND’ was defined as an individual situation so dire as to warrant distributing the drug first and sorting out the administrative paperwork later (see ‘Emergency IND’ provisions in FDA 1983, p. 26743). ‘Compassionate IND’ has often been used synonymously to ‘treatment IND’, but on at least one occasion was treated like ‘Emergency IND’, requiring less evidence of efficacy for action than a treatment IND (see the testimony of Dr. Maureen W. Myers of the National Institute for Allergy and Infectious Diseases before a Congressional subcommittee in U.S. House 1987, 34).

populations of individual patients who lacked adequate therapeutic options for serious diseases and had no access to investigational drugs through clinical trials; that it was preferable for patients to receive the investigational drugs under a sponsor's existing protocol, but that alternatively physicians representing patients could design individual protocols; and that to avoid commercialization of investigational drugs, participating drug sponsors who elected to provide investigational drugs to patients under this programme would be expected to do so free of charge. These provisions are also consistent with the proposals for 'drug treatment investigation' in S.2755.²⁹

Note that in the version of treatment IND envisioned by the *Concept Document*, the time for distributing an investigational drug to a very ill patient would 'ordinarily' be Phase III. While, as we have seen, the door had been opened to earlier decision-making through the view that Phase II could be used to generate the first adequate evidence of effectiveness, the formal expectation was that in most cases, distribution would not take place until relatively mature data could be obtained. In the late 1970s and early 1980s, this was the shape of the compromise achieved in the regulations between knowledge of effectiveness and decision-making for early access to investigational drugs. We will see in the next chapter that as this set of proposals collided with the contingencies of the 1980s, the scale of the programme, and ultimately aspects of its purpose were significantly modified in practice.

3.4.3 *The Concept Document and Substantial Evidence*

The *Concept Document* included a series of proposals to specify more fully the types of evidence the FDA wished to see in support of drug applications. While the regulations already required that patient populations in clinical studies be selected to assure comparability of variables such as age and sex (FDA 1970b), the authors of the *Concept Document* now recommended that sponsors be required to conduct an analysis of the study data for comparability, including an assessment of the impact of any non-comparability. Moreover, they recommended that all regulations explicitly identify 'randomization as the method of subject assignment most likely to assure comparability of groups and freedom from biased assignment' (FDA 1979a, 20). Similarly, while the current regulations

²⁹ The bill provided for access to investigational drugs only through personal physicians, and not through the institutional treatment protocols of existing investigational programs. Other than this difference, however, the bill and the *Concept Document* present very similar views of how the programme should work, including the defining expectations that treatment IND would apply only to very limited groups of patients having serious medical needs unmet by existing approved therapies.

required an explanation of procedures employed to minimize bias in the clinical study, the authors suggested that sponsors should include an explanation ‘of how the steps taken *were adequate* to minimize that bias’ of the subject, observer, and analyst (21, original emphasis).

The FDA task force members who authored this section of the *Concept Document* also sought to establish an order of preference for study designs. The 1970 regulations made clear that controls were required, but now the task force recommended that the FDA establish in the rules a hierarchy of controlled studies in which placebo control was first preference. The second preference was the use of a control group receiving either no treatment at all (‘no treatment’ control) or treatment using an existing therapy for the condition (‘active treatment’ control). The last preference was to use a historical control. ‘Historical control’ is not specifically defined, however the Task Force indicated that a sponsor proposing a historically controlled study ‘should be required to identify an explicit control group so that the results can be compared quantitatively with the study group’ (21). Under this set of recommendations, a drug sponsor would have to supply a justifying rationale to use a study design other than a placebo control. Further, specific justifications would be sought by the FDA for whichever control method was to be used as an alternative, with assurances that the proposed method of study would be able to generate the kind of evidence required to establish efficacy (21). This order of preference for the design of clinical studies did not ultimately get encoded into the regulatory proposals. However, the impulse to specify a hierarchy of controls and to force drug sponsors to justify their approach to bias minimization can be seen as a reflection of the FDA’s continued struggle to define and refine a more rigorous approach to conducting randomized controlled trials. For the convenience of the reader, I have summarized key features of the proposals and regulations discussed thus far in Tables 3.1 and 3.2.

3.5 IND and NDA rewrite proposals

While Senator Kennedy’s drug reform bill failed in the Congress, the reforms suggested in the *Concept Document* were carried through into regulatory proposals published

Table 3.1
The Clinical Phases in Selected Regulations and FDA Reform Proposals

	Phase I	Phase II	Phase III
1963 IND Regulations	Clinical pharmacology (toxicity, metabolism, absorption, elimination, etc.)	Clinical pharmacology (initial limited trials 'for specific disease control or prophylaxis purposes')	Controlled clinical trial safety and effectiveness; optimum dosage schedule
1977 Bill (S.1831)	Same as above	Same as above	Same as above
1978 Bill (S.2755)	Drug Innovation Investigation (clinical pharmacology)		Drug Development Investigation (evaluation of effectiveness and 'short-term and relatively common adverse effects of the drug')
1979 Concept Document/ General Considerations	Clinical pharmacology	Earliest controlled statutorily defensible evidence of effectiveness	Expanded trials in multiple and long-term settings, controlled and uncontrolled

Table 3.2
Expedited Approval, Phase IV Commitments and Pre-Market Access to Drugs in FDA Reform Proposals

	Expedited Approval	Phase IV (Postmarket)	Notes on Evidence and Risk/Benefit	Pre-Market Access to Drugs
1977 Bill (S.1831)	'Abbreviated' Phase III or 'short-term' w/ postmarket study (cf. Dorsen report).	Postmarket study for multiple situations, including expedited approval.		
1978 Bill (S.2755)	Provisional monograph; three year expiration; easily withdrawn from market	Additional study required for permanent approval	Consider <i>significance</i> of effectiveness; risk/benefit includes 'social' risk. 'Significant' evidence for provisional monograph	Drug Treatment Investigation (compassionate use or when drug denied approval benefits small group of patients)
1979 Concept Document/ General Considerations			Require randomization for mitigation of bias; preference for placebo and concurrent controls	Individual treatment IND; 'ordinarily' in Phase III; emphasis on treatment, not investigation

in 1982 and 1983. Referred to as the 'IND rewrite' (FDA 1982) and 'NDA rewrite' (FDA 1983), elements of these proposals would later be reflected in the approach taken to develop the first successful AIDS drug, AZT. In the next chapter we will see that at the time these proposals were published, AIDS was beginning to be recognized as a potential public health crisis but therapies were not yet developed. As much as the conceptual trajectory established in the 1970s would still be evident in the 1980s, at this time also the Reagan administration assumed power and would begin to make its influence felt in the regulatory process.

The move towards regulatory reassessment had already begun during the Carter administration. In a 1978 executive order (E.O. 12044)(President 1978b), President Carter sought to assure that all new regulations would be developed in such a way that the need for the regulation would be clearly defined, that meaningful alternatives would be considered prior to issuing new regulations, and that compliance costs would be minimized. The President ordered the agencies to publish semi-annual summaries of 'significant' regulations under development or review, including a justification and statement of the legal basis of the proposed action. The heads of agencies were to define which rules were 'significant', which of these significant rules might have major economic impacts when implemented, and to review and approve all proposed rules, taking into consideration whether the rules met the requirements of E.O. 12044 and whether public response was adequately considered. The Order also mandated periodic reviews of existing regulations. The Office of Management and Budget (OMB) was to have general oversight of the process, reporting to the President on the effective implementation of the order. The OMB's other major task was to review reports prepared by the agencies within 60 days of the issuance of the order. These reports outlined the current processes for developing regulations and proposed changes to be made to procedures and criteria for rule-making to comply with the Order.

In February 1981, President Reagan revoked the previous order and issued E.O. 12291 (President 1981). This new Order retained many of the features of the Carter policy on regulatory review and procedural reform, however there were notable shifts of emphasis and discretionary power. Carter sought to make the regulatory process efficient, effective, and user-friendly, including requirements that the regulation be written in 'plain English' for ready comprehension, and emphasising the need for public comments to be taken into account. In Carter's view of regulatory reform, the Agency heads were to be

held responsible for the effective conduct of their organizations. Reagan, by contrast, defined more restrictive criteria for what might be considered a 'needed' regulation, instructing that '[r]egulatory action shall not be taken unless the potential benefits to society from the regulation outweigh the potential costs' (125). The sole criterion for choice among alternative regulations would be which one involved 'the least net cost to society' (125). The Order specified that new regulations must take into account 'the condition of the particular industries affected by regulations [and] the condition of the national economy' (125). This formulation required calculation of 'net cost to society' on the basis of what I would argue was an incommensurable combination of heterogeneous elements, some quantifiable and others not, leaving a great deal of interpretive latitude. Further, the Reagan order shifted substantial discretionary power from the agency heads to the OMB. The latter would now review all regulatory impact analyses as well as all proposed and final rules. The OMB was specifically authorized, among other things, to redefine whether a rule was a 'major rule' (subject to regulatory impact analysis), to waive requirements of the Order at will, to require agencies to 'evaluate, in connection with a regulation, any additional relevant data from any appropriate source' (127), and to 'prepare for consideration by the President recommendations for changes in the agencies' statutes' (127). Hence, whatever interpretive latitude was permitted under this Order was available to the OMB, not agency leaders. The OMB operated under the direction of the newly formed 'Task Force on Regulatory Relief', headed by Vice President George H.W. Bush. This Order was part of a larger campaign of the Reagan Administration, through the OMB, to control the rule-writing process for all agencies and was designed to discourage new rule-writing in the first place (Morrison 1986).

Nevertheless, the FDA continued the reform efforts from the last decade. It is unclear whether the task forces for regulatory review organized by the FDA in the late-1970s were prompted by Carter's executive order, or if they were in fact already operative when the order was issued. Whatever the case, it is clear that by the time the FDA was publishing proposed rules in the early 1980s, it was expedient for the Agency to characterize the action more as a revision of existing rules (which, in part, it was) rather than as an enactment of any new rules (although new policies and procedures were in fact being introduced). Hence, in its 1982 publication of the proposed new drug and antibiotic regulations (the 'NDA rewrite', FDA 1982), the introductory paragraphs describe this proposal as part of the regulatory review and revision required by Reagan's EO 12291.

The emphasis of the NDA rewrite is on better service through streamlined procedures and reduced paperwork burdens. Hence, the NDA rewrite contained provisions to streamline the NDA format and to reduce recordkeeping and reporting requirements on approved applications (even while strengthening procedures for postmarketing surveillance of approved drugs and establishing a new requirement for drug sponsors to provide periodic updated safety data on pending applications).

In this document, the FDA was clearly thinking in terms of ways to expedite the NDA process, however modestly, suggesting for example that approval decisions could proceed when the only outstanding issue was review of the final printed labelling for the drug. The FDA was likewise thinking in terms of reducing the amount of data required for an NDA or supplemental application (an application requesting changes for the labelling or manufacture of a marketed drug), giving specific consideration to what kinds of information was superfluous or unnecessary. For example, FDA proposed that case report forms for all patients would no longer be required as part of an NDA; rather, the sponsor could submit tabulations of ‘essential individual patient information’ (FDA 1982, 46623) and only submit case reports for patients who dropped out of the study or who died while on study. Similarly, the FDA would now be more selective about what kinds of data were needed for supplemental applications, noting that pre-market application and approval would only still be required for changes to marketed drugs ‘which would be expected to affect adversely the agency’s previous conclusions about safety and effectiveness’ (46623).

The NDA rewrite also contained a definition of ‘adequate and well-controlled investigations’ similar to that offered in the *Concept Document*. It affirmed that the study design must allow comparison with a control; that the methods of assigning patients to groups and of assessing response must be designed to minimize bias (i.e., randomization and blinding respectively); and that the analysis of the results should include explanations of the methods used, criteria used to assess response, and should demonstrate the comparability of test and control groups. This passage also included a description of the various types of controls recognized for study (placebo, no-treatment, active treatment, and historical), adding a fifth category, ‘dose-comparison’, in which different test groups receive different doses of the same drug. In a departure from the *Concept Document*, however, this passage proposed no ‘hierarchy’ of control methods nor required any

justification for the method chosen. Likewise there was no requirement to show how steps taken to minimize bias were *adequate* to do so.

Significantly for the discussion in the next chapter, in the IND rewrite published the following year (FDA 1983) the emphasis was on increasing FDA-sponsor collaboration and also allowing greater flexibility in the early stages of clinical research. The FDA wrote that although 'FDA has for several years offered "end of phase 2" conferences for drugs likely to provide significant and modest therapeutic advances, FDA now proposes to give the sponsor of any IND an opportunity to hold such a conference with the agency' (FDA 1983, 26721). According to the FDA, the purpose of this meeting was to discuss the design and conduct of the Phase III trials in the hope that such a 'meeting of the minds' would 'significantly reduce the possibility of disputes' upon submission of the NDA (26721). The FDA would also offer sponsors a 'pre-NDA' meeting to discuss preparation of the marketing application.

Although this is the first time that end of Phase II conferences appear formally in the regulations, the FDA had developed procedures for such meetings in 1976 and began holding meetings in 1978 (GAO 1980), with the result that, according to one official from the General Accounting Office, the drug companies 'strongly support end of Phase II conferences, and indeed, those who have participated in the conferences characterized them as excellent and helpful' (Grant 1981, 4). Hence, by 1983 FDA had already been actively engaging in such meetings for sponsors of therapeutically important drugs, and was now offering the possibility of meeting with *any* sponsor to help produce Phase III studies suitable for supporting a marketing application.

The IND rewrite also contained proposals for codifying treatment IND. The version of the rule appearing in this 1983 proposal is consistent with the view of treatment IND proffered four years beforehand in the *Concept Document*, and also in important ways with the Senate bill S.2755. In the preamble to the proposed IND rewrite, the FDA specified that the treatment IND 'is intended for a serious disease condition in patients for whom no satisfactory approved drug or other therapy is available' (26729). Moreover, the potential benefits of using the experimental drug must outweigh the risks. The Agency reasoned that the only way one can weigh the benefits of an experimental drug against the risks is to have already collected sufficient evidence to adjudge the safety and efficacy of the unapproved drug. Hence, 'FDA believes that the model for treatment protocol/IND use should be a drug in Phase 3 when the major clinical trials are

completed or underway’ and where the evidence collected to date pointed towards subsequent marketing approval (26729). Requests for treatment IND earlier in the development process would only be warranted in a ‘compelling circumstance’ (26729). When discussing the proposed treatment IND rules, the Agency made a point of noting the consistency of these proposed regulations with the recently passed 1983 Orphan Drug Act,³⁰ and created ‘emergency IND’ provisions for situations dire enough that the drug must be shipped first, and formal IND paperwork requirements handled later. The Agency also urged drug companies to create treatment protocols under which physician requests for treatment INDs could be processed, rather than forcing physicians to submit individual treatment IND requests (with individual treatment protocols) to the FDA, since for some investigational drugs, requests can ‘extend into the hundreds’ (26729).

It is also significant to note that the question of the *sale* of the investigational drugs administered under treatment IND is, in this proposal, treated perfunctorily; almost as an afterthought. Appearing in the ‘Miscellaneous Provisions’ section towards the end of the preamble, the Agency wrote that the ‘proposal would retain, essentially unchanged, the current provisions prohibiting promotion and commercialization of investigational drugs’ (26734). Drug sponsors responding to requests for experimental drugs under treatment IND must either supply the drugs without charge or must provide ‘a full and satisfactory explanation’ as to ‘why the sale should not be regarded as commercializing the drug’ (26734). We will see this concept of treatment IND change substantially in the next chapter.

3.6 Two Decades in Review

The Kefauver hearings and legislation represent a significant moment in which countervailing forces in play throughout the first half of the century came together. In this historical moment, with the unlikely-but-crucial midwife of thalidomide, ‘scientific’ therapeutic evaluation won out over physician opinion and case reports, while the machinery was put in place for an unprecedented level of government oversight of the drug development process. According to Harris (1964), the vote by voice to approve the

³⁰ Among other provisions, the Orphan Drug Act (1983) specified that sponsors of orphan drugs should design clinical studies in such a way as to allow entry of patients who desire therapeutic access to the experimental drug. Orphan drugs are defined as drugs having a U.S. target population of no more than 200,000 patients.

bill in the House was unanimous, ‘whereupon the members burst into spontaneous applause — an unusual demonstration that led a veteran member of the House press corps to remark, “I guess they must be applauding themselves. It’s not often they’re able to do something for the people”’ (233). Still the specific nature of the evidence to be required for approval, and the procedures necessary to design ‘adequate and well-controlled investigations’, would have to be worked out through conflict and challenge. It would also be worked out in daily practice.

In this period, we saw in various proposals for reform a movement towards making decisions earlier in the clinical development process. This movement was expressed most obviously in proposals to approve drugs intended for life-threatening diseases on a conditional basis, using a standard of evidence consciously reduced from that normally required so that decisions could be made earlier in the process. In this proposed rule, and also in the question of when to permit treatment IND, there was an emerging tension between information-gathering and early decision-making. This tension also fed into the long time conflict between research and therapy. The specific form it took in this period centred around how much data was sufficient for expedited decision-making. At the end of two decades during which the overall movement was towards requiring *more* information for drug approval, legislators and regulators (among others) were now grappling with a standard for making decisions with *less* information. In treatment IND, we also see research-therapy tension in that while pre-market access was explicitly to be granted for the primary purpose of therapy, the desire to find ways to collect information from that therapy are never far away. For example, S.2755’s description of the drug treatment investigation included the statement that ‘information relevant to such an assessment may be derived during the investigation’ (HEW 1978, 84-5). Indeed, the very phrase ‘drug treatment investigation’ contains the tension of which I speak. It is both ‘treatment’ and ‘investigation’ in some sense. Here in the late 1970s and early 1980s, the emphasis is on ‘treatment’. But in the future in application to AIDS we will see this balance shift dramatically.

The movement towards earlier decision-making is also reflected in the recognition by regulators that regardless of what they had written in 1963, in practice early evidence of efficacy was being produced in Phase II studies, not exclusively Phase III. This recognition would set the stage for the creation of formal rules for earlier decision-making in the next decade. Indeed, for drugs intended to treat life-threatening diseases, some

approval decisions were already being made on the basis of Phase II studies. FDA regulatory guidance and rule-making was therefore coming in the wake of practice, as a reflection of it. More than that, future rules written to accommodate this practice will constitute a formalisation of actions already taken (Chapter 4). We likewise saw formal rule-making following practical action for treatment IND, where I have argued that a growing demand for pre-market access to investigational drugs likely spurred the impulse to formalise the practice at this time. If so, then a clear consensus had already been formed around the need for such a rule; in effect, a de facto rule was in existence which the *Concept Document* was proposing to codify. The same is true for the definition of clinical phases.

These observations point to a tacit process of consensus feeding into the creation of new rules outside of the formal processes of consensus required by law (under the *Administrative Procedures Act* of 1946) — a consensus which, if not essential for new rule-writing, certainly inspires and facilitates the creation of new rules. This observation highlights a key difference between any socially sensitive theory of regulation and the types of theories discussed in Chapter 1. Those theories of regulation consider only formal processes of consensus in their evaluation of regulatory outcomes, if they consider it at all. Even the civic republican theory, which is more ‘social’ in perspective, contemplates consensus processes which are ‘formal’ in that they are consciously orchestrated through meetings and other planned mechanisms. However, this account demonstrates a tacit consensus process taking place on the level of practice. Ultimately, in these examples, it was the informal consensus and growing frequency of practice which created the perception of the need for a formal rule. We could also think of it as a set of rules, {I}, becoming primed. How can a regulatory agency allow the practices of a fully primed {I} flourish without specifying parameters for future I₂? It is their job to coordinate action. Formal procedures must be created for practices around which such a consensus has condensed. Any ‘social’ theory of regulation must see regulatory agencies not merely as providers of regulatory ‘goods’, but as organisers and directors of coordinated social action. Any social theory of regulation will need to recognize the powerful role of tacit, practice-based consensus in the creation of rules, and as well in outcomes for rules not-yet-written.

4. AZT: REFLECTIONS OF THE PAST AND A MODEL FOR THE FUTURE

The story of the rise of AIDS is well known. A report of the first handful of cases of gay men suffering from immunosuppression and opportunistic diseases appeared in the Centers for Disease Control (CDC) Morbidity and Mortality Weekly Report (MMWR) in June 1981.¹ Just two months later, in August 1981, the CDC reported 108 more cases of immunocompromised patients suffering from opportunistic infections. From this beginning in 1981, the death toll mounted rapidly, doubling from 1982 to 1983, and then tripling from 1983 to 1984.² A diagnosis of AIDS was effectively a death sentence, and the steep rise in reported cases and deaths was alarming. Under these circumstances, researchers and regulators had a clear incentive to innovate and stretch conventional boundaries to address what was poised to become a global public health issue.³

After three years of intensive research, a consensus was finally achieved on the identity of the causal agent for AIDS in 1984, allowing researchers to begin the search for a chemical that might interfere with the activity of the retrovirus. Dr. Samuel Broder, chief of the clinical oncology program at the National Institutes of Health (NIH), teamed with the drug company Burroughs Wellcome in an arrangement offering preclinical testing services in National Cancer Institute (NCI) laboratories in exchange for access to some of the company's stock of experimental compounds known to have some anti-viral

¹ Michael S. Gottlieb and Wayne Sandera published their findings in the *MMWR* of June 6 1981 (vol. 30 iss. 21) under the title, simply, 'Pneumocystis Pneumonia -- Los Angeles.' In a section of the website offering a retrospective look at AIDS, the CDC has posted an html version of the original publication. See http://www.cdc.gov/mmwr/preview/mmwrhtml/june_5.htm.

² See the figures for numbers of reported AIDS cases and deaths helpfully provided in the 'AIDS timeline' at the National Institutes of Health AIDS history website, <http://aidshistory.nih.gov/timeline/index.html>.

³ Some researchers seemed to have a greater awareness than others of the ultimate potential for the disease to become a global problem. Fauci recalls thinking that, given the transmission routes used by the disease, there was no reason it shouldn't spread unless checked. Dr. Samuel Broder similarly recalls a sense of urgency. Dr. Henry Masur, by contrast, an expert in *Pneumocystis carinii* pneumonia, recalls being drawn into the situation out of scientific interest rather than any awareness of the potential seriousness of the situation. See the fascinating collection of interviews at the National Institutes of Health AIDS history website: <http://aidshistory.nih.gov/transcripts/index.html> (accessed 3 October 2006).

activity. Among the anonymized compounds supplied by Burroughs to NCI, the one labelled compound ‘S’ increasingly showed promise against the AIDS virus. This compound, AZT, coincidentally turned out to have been developed more than 20 years earlier on an NCI grant.⁴

With the drug’s promising pre-clinical performance, researchers scrambled to get the drug into Phase I testing — the first introduction of the drug into human subjects — which began in June 1985. Nineteen patients were enrolled in the study, with eleven treated at NIH and nine at Duke University Medical Center. All of the patients in the Phase I study were very ill when testing began, and all experienced small but statistically significant increases in CD4 cell counts, or helper T cells (Yarchoan 1998). Since these important immune system cells were destroyed by the virus, an increase in the number of cells per cubic centimetre of blood was taken to be a sign that the drug was interfering with the action of the virus — i.e., CD4 counts were taken to be a ‘surrogate marker’, a laboratory measure of patient response to the test drug thought to be meaningful for ultimate clinical benefit.

Researchers proceeded rapidly to a placebo-controlled Phase II study with 282 patients enrolled at 12 medical centres across the United States. After four months, nine patients on placebo had died while the AZT cohort saw only one death. The data review took place on 19 September 1986. The study was discontinued the next day and the drug was made available to AIDS patients almost immediately — not only to the patients on the study, but to seriously ill patients nationwide. As we have seen, such pre-market access of very ill patients to investigational drugs (‘treatment IND’) had been practiced informally by the FDA in various modes (Chapter 3). Although, as I will discuss in more detail below, the application to AZT took on an expanded scope and significance from previous routine practice. At the time of this pre-market distribution of AZT, treatment IND was still considered informal.

⁴ In 1964, a young biochemist named Jerome P. Horowitz custom-designed the drug as a prospective cancer treatment. Unfortunately, AZT turned out not to be effective against cancer and the drug was shelved; as one article put it, the drug went ‘unused, a curiosity in the literature of chemistry, for 15 years’ (Hilts 1986, A11).

The Anti-Infective Drugs Advisory Committee (ADAC)⁵ met four months later, in January 1987, to consider a new drug application (NDA) based on this single study. The Committee was faced with clear deficits of information on drug toxicity and long-term effects. Indeed there was no information at all regarding whether less ill or non-symptomatic patients would respond to the drug, nor what the effects of longer-term administration might be. At the same time, the Committee understood that upon marketing approval, personal physicians would begin to prescribe this therapy to *all* their patients having AIDS, including those less ill patients for whom information was lacking. Hence, among other risk-related knowledge deficits, it was clear that approving the drug for seriously ill patients created an indirect risk to less seriously ill patients who were likely to receive the drug without sufficient information to evaluate risk or benefit. Nevertheless, the Committee consciously weighed those uncertainties against the potential benefits of approving the drug and voted for approval of the drug with the understanding that the sponsor would conduct additional studies on less seriously ill patients to fill the information gaps as quickly as possible (see FDA 1987a).

As a result, AZT was developed and approved in a mere two years — lightning speed by the standards of the day, or even today — and was approved on the basis of a single, Phase II study. While the alacrity with which this drug was developed and approved was exceptional, in Chapter 3 we saw that Phase II trials were considered able to produce the first valid evidence of effectiveness and that, notwithstanding formal guidelines for three phases of investigation, in practice Phase II studies had been used as the basis for approval of some cancer drugs.⁶ Moreover, we saw various proposals for expediting drug development through decision-making earlier in the development process. It should therefore not be surprising that a Phase II study might have been deemed acceptable as the basis for approval for AZT. In this sense, the approval of AZT

⁵ The first new drug applications for AIDS were reviewed by the Anti-Infective Drugs Advisory Committee. An Antiviral Drugs Advisory Committee was subsequently established and began meeting to discuss AIDS-related drugs and issues in January 1990. For the convenience of the reader, I will use the abbreviation ‘ADAC’ in the text to refer to both committees. If needed, the actual committee referenced can be determined either by the date of meeting or by reference to the cited committee transcript in the bibliography.

⁶ Although I should note that the 1986 version of the IND regulations published in the *Codes of Federal Regulation* (CFR) did not define Phase II as producing the first statutorily defensible evidence of efficacy, as the 1979 *Concept Document* did (Chapter 3). It retained the 1963 definition of Phases I and II as ‘clinical pharmacology’ while only Phase III was the ‘clinical trial’. See the 1986 version of the CFR Title 21 Section 310.1 (2)(10)(a) and (b) (page 63). Hence, what the *Concept Document* and *General Considerations* guidance document acknowledged as de facto practice, the CFR still failed to recognize.

was fully consistent with past practice.⁷ This clinical study was somewhat unusual in that it was relatively large for a Phase II trial, and by the standards of the time the multicentre design was more characteristic of larger Phase III study designs. (We will see later in this chapter that because of its design and placebo control, at least one prominent member of the medical establishment considered this a Phase III study masquerading as Phase II.) Nevertheless, characteristic of Phase II studies, there was still a great deal not known about the drug.

What was more unusual about this case was the approval on the basis of a *single* Phase II study. Single-study approval was a rare practice which contravened the Food and Drug Administration's (FDA) traditional view of the statutory 'substantial evidence' requirement (Chapter 3), which called for 'adequate and well-controlled investigations' — plural. The FDA held this phrasing to mean that *at least two* such investigations would be required as a form of scientific replication. The FDA did make exceptions for extenuating circumstances, however, and AZT was one of them.⁸ The primary novelty of AZT rested in the speed with which it was able to be developed combined with the fact that it was the first drug approved for AIDS at a time when this uniformly fatal disease appeared unstoppable. This cavalry-charging-over-the-hill moment thrust AZT — and the approach used to develop it — into the spotlight. Development and approval of AZT was deemed a resounding success. At the January 1987 ADAC meeting to consider AZT approval, the FDA's Dr. Ellen Cooper characterized the experience with AZT as 'a remarkable story in drug development' (FDA 1987, 8). Years later, this event is still characterized by people like the NCI's Broder and FDA's Dr. Robert Temple as a success story (Broder 1997; Temple 2005⁹). Perhaps unsurprisingly, then, when Vice President

⁷ Rothman and Edgar (1992) have noted that for cancer drug development, NCI's procedures often departed from FDA requirements which led to clashes between the FDA and NCI in the late 1970s and early 1980s. The NCI was engaged in 'more patient-centered risk-taking' while the FDA was more 'paternalistic' and concerned with 'sound science' (197). Nevertheless (despite NCI's involvement in AZT development), according to Rothman and Edgar, the placebo-controlled AZT trials were based 'not on the cancer model, but on a more generalized medicine model, really an infectious disease model' (204). While true, we will see that as AZT resistance became apparent and the need for other therapeutic options became increasingly evident, the development of subsequent AIDS drugs began to follow something akin to (or, indeed, exceeding) the cancer model as described by Rothman and Edgar.

⁸ See the FDA's discussion of the interpretation of this clause in the Subpart E rule publication, FDA 1988, p. 41521.

⁹ In an interview with the author Dr. Temple, who at the time was acting Director of the Center for Drug Evaluation and Research and is now the Director of the Office of Medical Policy at the FDA, commented that AZT is 'an example we like to give because everybody's instinct early on is to just do a single-arm study. But the first study of AZT was a randomized trial with 16 deaths in the placebo group and one in

George H.W. Bush, in his capacity as the chairman of the Task Force on Regulatory Relief, appealed to the FDA to develop new procedures for expediting new therapies for life-threatening diseases, the FDA turned to the experience with AZT for a blueprint.

4.1 AZT as a Model for New Rule-Making: Subpart E

In response to the Vice President's appeal, the FDA rapidly published an 'interim rule' — an exceptional rule-making procedure used by regulatory agencies when the proposed rules are deemed too time-critical to go through the usual public comment period. An interim rule is treated as a final rule unless subsequent amendments are published (which, to my knowledge, never took place for these rules). In their 1988 *Federal Register* publication of the new 'Subpart E' interim rules (FDA 1988a) (so named as a shorthand for the applicable subpart in the *Codes of Federal Regulations*), the Agency wrote, 'These procedures are modelled after the highly successful development, evaluation, and approval of zidovudine, the first drug approved to treat the AIDS virus' (41517). Subpart E was to apply to new chemical or biological products 'that are being studied for their safety and effectiveness in treating life-threatening or severely debilitating illnesses' (41517).¹⁰ The main provisions of Subpart E were that, first, the drug sponsor should confer with the FDA at the end of Phase I to consult on a design for the Phase II study suitable for producing the kinds of evidence normally derived from a Phase III study (thereby mitigating the need to perform a Phase III study). Second, when Subpart E drugs showed 'early evidence' in Phase II testing of being 'promising' (41529) for life-threatening or severely debilitating illnesses, the FDA would work with the sponsor to develop an appropriate therapeutic regimen under the 'treatment IND' programme. Third, once the Phase II study was complete, the FDA would use a risk-benefit analysis to 'consider whether the benefits of the drug outweigh the known and potential risks of the drug' (41529). This risk-benefit analysis would determine whether the drug should be approved on the basis of Phase II data.¹¹ Finally, following market approval, the 'FDA

the treated group and it was so obvious that this drug needed to be approved right away. If they'd just given it, we wouldn't have known what to make of it. So that's a message we like' (Temple 2005).

¹⁰ The language 'life-threatening or severely debilitating' was used here in this interim rule, but in subsequent rules was modified to 'serious and life-threatening', with corresponding changes and refinements of the definition of eligible disease types.

¹¹ This was not a fixed formula, but a matter of clinical judgment on the part of decision-makers. To this day, while various statistical and quantitative tools are used to assess the adequacy of clinical data, the level of adverse effects posed by investigational drugs, and the response of patients to treatment, the weighing

may seek agreement from the sponsor to conduct certain postmarketing studies (phase 4)' (41517) since approval will have been often been granted 'on the basis of limited, but sufficient, clinical trials' (41521).

Significantly, although the FDA approved AZT on the basis of a single Phase II study, the Agency specifically excluded this feature from the Subpart E rules. The FDA noted in the publication of Subpart E that the statutory requirement for drug applications to be supported by adequate and well-controlled clinical investigations 'has long been interpreted to mean that the effectiveness of a drug should be supported by *more than one* well-controlled clinical trial and carried out by independent investigators' (FDA 1988a, 41521, emphasis added). This requirement was considered to be 'consistent with the general scientific demand for replicability to ensure reliability of study results' (41521). Nevertheless, there had been 'a few unusual circumstances in which a particularly persuasive multi-center study has been accepted in support of a claim of increased survival' (41521) such as AZT. However, the Agency cautioned drug sponsors to plan for study replication because such 'persuasively dramatic results are rare' (41521); the results of 'two entirely independent studies will generally be required' for approval (41521).

While explicitly based on the procedures used for AZT, clearly many of these provisions can be seen as extrapolations of reforms implemented by the FDA a decade or more beforehand. Consistent with the recommendations of Dorsen panel (Chapter 3) the FDA had already been offering Phase II meetings to discuss the design of Phase III trials; now for the sake of Subpart E approval the meeting would take place at the end of Phase I to design a plan for the Phase II trial, thereby moving a key conference for design to an earlier point in the process with the intention of allowing earlier approval decision-making. Additionally, since the FDA would now allow treatment IND to proceed as a provision of Subpart E, and since one phase of study was effectively being eliminated for approval, the decision for treatment IND would also come at an earlier point than recommended in previous versions of the practice – now in Phase II rather than (ideally) in Phase III. These rules also clearly reflect the Dorsen panel recommendation and the related 1977 Senate bill (S.1831)(Chapter 3) proposing to create an expedited form of approval based on short-term studies or abbreviated Phase III data with postmarket study to fill the gaps. (Later in this chapter we will see more than one observer interpret the

of foreseeable (and possible unknown) risks against the perceived benefits of therapy is a matter of informed medical judgment. We will see this decision-making process in some detail in the cases of drug approval discussed in subsequent chapters.

AZT trial as an expanded Phase II study or as something more akin to Phase III — a blurring of Phase II and III already evident in S.1831, and also in the ‘drug development investigation’ phase in the proposed *Drug Regulation Reform Act of 1978* discussed in Chapter 3.)

While the FDA had always used a form of risk-benefit assessment for decision-making on drugs, the version of it prescribed in Subpart E represents a shift in the application of the concept. Historically, the FDA recognized that any drug might cause an adverse effect and, as we saw in Chapter 3, even before proof of efficacy was required by statute, a ‘presumptive benefit’ had to be taken into account to assess the relative safety of a drug. It was in this sense that the concept of risk-benefit assessment had been used for many years. Although the 1978 drug regulation reform bill, S.2755, sought to expand the realm of possible risks to include social risks (e.g., the possibility of intentional substance abuse) or toxicity risks to those who would administer the drug, it did not explicitly include consideration of the risk from knowledge deficit. The Subpart E conception of risk is more akin to the concept of ‘provisional’ approval articulated in the 1978 bill — a concept where a decision is consciously made with less information than normally desired. Still, S.2755 never explicitly envisioned including the information deficit as part of the risk assessment, even though it did create a standard of evidence to reflect an acceptable level of information deficit. In the 1988 Subpart E rule, however, the risk-benefit approach is explicitly modelled on the procedure used by the ADAC to evaluate AZT: when a decision must be made on the basis of less information than would normally be desired for approval, the risks of both drug toxicity and information deficit should be weighed against the anticipated benefits of approval. As stated in the Subpart E publication, ‘FDA will consider whether the benefits of the drug outweigh the known and potential risks of the drug *and the need to answer remaining questions about risks and benefits of the drug*, taking into consideration the severity of the disease and the absence of satisfactory alternative therapy’ (FDA 1988a, 41520, emphasis added). Here, for the first time, we see the FDA articulating in the regulations a principle of decision-making on the basis of limited information with the possibility that sponsors would be required to collect additional data *after the approval had taken place*.

Clearly full and perfect information had never been available for any instance of drug approval. A perusal of pre-1988 cancer drug approval data will show that it was not unusual for the FDA to seek postmarket study commitments from sponsors. However,

the posture reflected in this rule is different. When asked in a 1987 interview if the post-market studies required for AZT were ‘usual’, the FDA’s Dr. Ellen Cooper responded that, no, ‘their degree and intensity are certainly unusual. Important questions about zidovudine remain unanswered. If we were dealing with a less severe illness that did not require such urgent action, these questions might have been resolved before the drug was approved’ (Stone 1987). Subpart E reflects a willingness to push the comfort level of decision-making into a zone where there is a conscious lack of important information. It is a zone in which decision-makers are aware that under other circumstances, a decision would be delayed pending completion of additional studies. This was now a codified principle.

In summary, although other examples of Phase II drug approval existed, Subpart E was clearly patterned specifically on the approach to AZT. Since the approach to AZT was rooted in prior experience, we can see in Subpart E echoes and extrapolation of already existing concepts and practices. The move to meet with drug sponsors prior to Phase II was a reflection and extension of the Dorsen report and ‘IND rewrite’ proposals discussed in the last chapter. The reliance on a Phase II study for approval decision-making was not unprecedented when seen in light of earlier practice, and the decision was consistent with the 1979 *Concept Document* in which Phase II was described as producing the first legally defensible evidence of effectiveness. The simultaneous decision to allow treatment IND at the end of Phase II (discussed more in the next section) likewise reflected a concept already proposed, while also extending the application of the concept beyond the original conception. Moreover, the notion of consciously approving drugs perceived to be clinically important for serious diseases on the basis of less information than normally desired had been contemplated in the late 1970s, although it was not made into law at that time. While the articulation of a risk-benefit concept for drug approval decisions under subpart E clearly reflected actual practice and a long-standing attitude toward drug approval, it also expressed what was for regulation (but not necessarily for drug approval in practice) a prescription for a new form of risk: the risk of lack of information for decision-making.

These historical threads of continuity intertwine to account for the initial approach taken with AZT. They also serve as a reminder that in a growing crisis, the foundation for decision-making must necessarily begin with already existing experience and ideas. However, even while Subpart E agrees with the spirit of earlier reforms, many of its

provisions extrapolate beyond their initial boundaries. This extremity, of course, was catalyzed by the extremity of the disease itself. It was also catalyzed by a historically contingent confluence of forces: the most strident FDA critics from the previous decade found support for their views in the deregulationist Reagan administration, which was ironically bolstered by the AIDS activism movement (Epstein 1996). As Edgar and Rothman (1990) have noted, ‘Sick gay men, abandoned by a president who refused publicly to acknowledge their disease on all but one occasion, provided the shock troops to move forward his administration’s deregulatory drug control program’ (124); meanwhile, ‘large parts of the AIDS advocates’ critique of the FDA could have been scripted by the Pharmaceutical Manufacturer’s Association’ (125). We will see in the next section how this successful experience also added to the knowledge base for future decision-making and showed a way (or multiple ways) forward based on a past success.

The Subpart E provisions not only reproduced the approach taken with AZT, but pointed towards future directions in four notable ways: 1) in its movement of key meetings and decisions to earlier points in the drug approval process; 2) in its codification of a possible requirement for supplemental post-market data collection; 3) in its movement towards reducing the burden of evidence required for approval;¹² and 4) in its use of a single study as the basis for a drug approval. I will return to these themes later in the thesis.

4.2 Treatment IND and the Collision with AIDS

As we saw in the last chapter, allowing very ill patients to have access to still-investigational drugs on a limited basis was a procedure informally practiced by the FDA at least as early as the 1970s. The 1983 proposals to formalize the procedure clearly envisioned relatively small-scale programmes. The Agency made a point of noting the consistency of the proposed regulations with the recently passed 1983 Orphan Drug Act

¹² In the preface to the Subpart E rules (FDA 1988), the FDA stressed that the issue was not how many phases of study have been completed, but whether convincing data had been generated from the studies done. On that basis, the FDA would deny that any real reduction is evident in the burden of proof required for Subpart E approval. Nevertheless, a reading of the advisory committee meeting transcript to approve AZT (FDA 1987a) makes it clear that the Committee was faced with important information gaps, many of which would have been filled by a Phase III study (and Committee members were aware of it). Furthermore as this account of events advances, the trend toward less evidence will become increasingly apparent, especially in the approval of the AIDS drug ddI, for which there was a great quantity of data but, as we will see, undeniably poor quality.

(which encouraged research on drugs for rare diseases which, by definition, have relatively small target patient populations of less than 200,000). The FDA additionally urged drug companies to create treatment protocols under which physician requests for treatment INDs could be processed, and expected drug sponsors to supply the drugs without charge in most circumstances (FDA 1983, 26734).

In 1985, when the FDA moved to finalize the 1983 treatment IND proposals (Chapter 3), the Agency submitted the final rules to the Office of Management and Budget (OMB) for review as required under Executive Order 12291. The OMB rejected key aspects of the FDA version of the rule, seeking to make the granting of a treatment IND request automatic unless the FDA could show that the investigational drug was unfit for use. An OMB background paper obtained by a Congressional subcommittee made the Office's intentions clear: it wanted to alter the wording of the rule to say that the FDA Commissioner '*shall grant*' a treatment IND request, rather than '*may grant*' the request; it sought to eliminate requirements that the investigational drug's potential benefits be shown to outweigh potential risks, and that there be 'sufficient evidence' of the drug's safety and effectiveness; and it sought to reverse FDA's 'presumption against selling investigational drugs' (U.S. House 1987, 95). Accordingly, the OMB version of the rule would now authorize sponsors 'to charge for investigational drugs made available to patients under a treatment protocol/IND, so long as the sponsor complies with certain designated safeguards against commercialization and notifies FDA 10 days prior to the commencement of such sale' (FDA 1987b, 8850); 'Prior FDA approval of the sale in this context *would not be required*' (8850, emphasis added).

The FDA Commissioner, Dr. Frank Young, negotiated with the OMB in an effort to retain more stringent requirements for demonstration of safety and efficacy, to emphasize the need to have progressed at least into Phase II to make such judgments possible (a compromise on the language in the original proposed rule calling for the drug normally to have progressed into Phase III), and to uphold the ability of the FDA to deny applications on a case-by-case basis (U.S. House 1987). Young's persistent negotiation over regulatory language led one OMB functionary to complain that 'FDA refuses to budge' on its proposal which gives 'total discretion to the agency' to approve or disapprove treatment use (U.S. House 1987, 282). After at least a year of largely ineffectual negotiation, Young was overruled. According to one press account, in March 1987 the OMB remanded the unyielding Commissioner to the Vice President's Task

Force, which summoned Dr. Young to an audience much like ‘commanding a recalcitrant priest to defend his views before the College of Cardinals’ (Havemann 1987, A22). The meeting with the Task Force took place on 9 March and a ‘reproposed’ rule containing the OMB’s modifications was published ten days later (FDA 1987b).¹³

The timing of this meeting and publication is notable and suggestive. The AZT Phase II clinical trial had been discontinued the previous September. Between that time and March, when the drug was approved, enrolment in the treatment IND programme for AZT had swelled to four thousand patients. We can get a sense of the challenge posed by the size of this programme from a discussion of the ADAC, which met just two months earlier, in January 1987, to discuss the approval of AZT (FDA 1987a). In that meeting, the Committee considered *not* approving the drug because AZT had only been proven to have efficacy in a narrow patient population over a relatively short study period. However, as noted earlier, once the drug was approved, physicians would prescribe AZT for *all* their AIDS patients, not just the most ill ones, even though the risk/benefit profile for less ill patients was unknown. As one committee member observed, ‘On the one hand, to deny a drug which decreases mortality in a population such as this would be inappropriate. On the other hand, to use this drug widely, for areas where efficacy has not been demonstrated with a potentially toxic agent, might be disastrous’ (145). Still, as one participant noted, Burroughs Wellcome was not being compensated for the drug administered under treatment IND, even though in the expansive growth of that programme following the discontinuation of the Phase II AZT study, almost *25 percent of the living AIDS population* was receiving therapy through treatment IND (145-6). If the drug were not approved pending further study, Burroughs Wellcome would be faced with the prospect of treating a substantial portion of the AIDS patient population for free, leading one Committee member to assert that this ‘is the whole reason for this particular meeting, to see whether the company becomes compensated, within a defined framework’

¹³ According to Havemann (1987) Young held out against the OMB for two years. If so, it would imply that there had been negotiations between the FDA and OMB a year before the OMB formally rejected the FDA’s version of the rule in 1986. Havemann also indicated that the publication of the rule as a reproposal was a technical victory for Young, since it allowed a public comment period. Indeed, many in the clinical research community were highly critical of the rule (see testimony and appended documents in U.S. House 1987), believing that wide availability of investigational drugs would ultimately hinder clinical research; that sale of investigational drugs created a class bias, since it favoured access for the rich and forced the poor into clinical trials; and that weak standards of safety and efficacy for therapeutic use of investigational drugs would allow disreputable drug sponsors to profit from selling ‘quack’ medicines on a pre-approval basis. The publication of the reproposal did generate a great deal of opposition and discussion, including the April 1987 Congressional hearings from which some of this account is drawn (U.S. House 1987).

(147). However, no well-defined framework existed for compensation under treatment IND. When asked if Burroughs Wellcome could receive remuneration for drug administered under treatment IND, FDA's Dr. Tabor responded that allowing cost recovery under an investigational IND 'is not something that is normally done' (160-1).

Shortly after this reassertion of the policy disallowing cost recovery, the FDA included cost remuneration in the reproposal for the treatment IND, but apparently not willingly. What is significant here is not simply that the FDA relented on this issue, but that the treatment IND programme for AZT appears to provide a significant element of context within which to consider this reversal; the original significance of 'treatment IND' seems to have shifted in its application to AZT. In the past, and certainly up to the time that the 1983 IND rewrite proposal was written, treatment IND had typically been used as a convenient stop gap, a bridge between existing therapies and soon-to-be available drugs for relatively small groups of patients. In this context, it seemed reasonable to ask drug companies to bear the cost of supplying the experimental drug. However, as the 1980s progressed, treatment IND began to change. Indeed, the size of the AZT treatment IND programme was not entirely unprecedented. In the mid-1980s, the FDA had allowed an even larger programme for antiarrhythmic drugs (U.S. House 1987). Still, the overall proportion of existing patients receiving AZT under treatment IND was notable. More than that, the specific application of treatment IND to AZT differed from previous uses of the procedure because AZT was the only drug effective at that time against an infectious, fatal, and rapidly spreading disease. In other words, in contrast to previous applications of treatment IND, the programme for AZT could be seen as *holding an epidemic at bay* pending approval of therapeutic drugs.

The importance of the experience with AZT as a potent element in the debate over the repropounded rule can perhaps best be seen in the manner in which observers upheld the approach to AZT as a model to follow. The difficulty was in establishing which aspects of the experience with AZT were the ones to be upheld as exemplary. In the April 1987 Congressional hearings to investigate OMB involvement in FDA rule-writing (U.S. House 1987), participants opposed to the repropounded rule frequently pointed to AZT as the model of how treatment IND should be done. A key witness was Dr. Martin S. Hirsch of Harvard Medical School, who gave testimony in opposition to the repropounded rule, arguing that '[w]hat we do *not* need is premature release of unproven agents in what have become known as "compassionate plea" or "treatment IND"

programs. Release of any drug prior to demonstration of clear clinical benefit would make it impossible to conduct properly controlled trials of drug safety and efficacy' (U.S. House 1987, 58, original emphasis). In this regard, he added, 'the recently completed AZT trials were remarkably successful. Since no AZT was available off this carefully designed protocol, enrollment of patients was completed within a few weeks, results were obtained rapidly, and the drug was quickly licensed' (58). Hirsch's clearly stated view is that the AZT development experience was successful because no drug was allowed to patients off-trial until the clinical study was complete. By contrast Young (who took a position of soft advocacy following the publication of the reproposal, downplaying the difference between the proposed and re-proposed sets of rules) testified in the same Congressional hearing using the example of AZT to emphasize the abbreviation of the clinical trial process and the rapidity of the approval, reinterpreting Hirsch's testimony in the process:

AZT was just the example that we would like to see used for this process. After the phase II clinical trials were over, we did not go on to phase III. Dr. Hirsh spoke to that this morning and applauded it. And then we used this period of time prior to the approval to treat thousands of patients. Actually, in this case, instead of a 7-year period, it was 22 months in which the entire clinical trial and approval process took place (U.S. House 1987, 70).

Young contrasted this case with the re-proposed treatment IND procedure for serious disease (which differed from the procedure for 'life-threatening' disease), noting that 'with serious diseases, we would expect that the treatment IND would not start until after the clinical trials had been completed' (70). Hence, in Young's view, since Phase III was never done, the AZT clinical trials were *not complete* when the decision was made for pre-market distribution. Hirsch, by contrast, was viewing clinical testing as having been completed for AZT (the Phase II study provided a demonstration of clear clinical benefit) before non-experimental subjects were allowed access to the drug; hence, he was *not* applauding the lack of Phase III in the sense of a truncation of data-gathering.

Similarly, in an address to the Institute of Medicine on 20 March, the day after the publication of the re-proposed rule, Dr. Young discussed the new rule in light of the example of AZT, stressing the efficiency and rapidity of the experience with AZT, saying that '[o]ur best success story to date is the rapid handling of the treatment IND for AZT'

(U.S. House 1987, 227). He also noted that the treatment IND for AZT was expedited by waiving the usual requirements for local Institutional Review Board (IRB) review¹⁴ — a feature of the treatment IND reproposal favoured by the OMB and opposed by many participants of the hearings. Hence, while critics of the reproposal could point to AZT's almost-approved status prior to implementation of the treatment IND programme as an example of how treatment IND should be conducted, supporters of the new rules could use the example of AZT to support decision-making earlier in the drug development process or to argue for IRB waivers to expedite early patient access to investigational drugs. In this way, elements from the experience with AZT were available for appropriation by people on opposite sides of the debate. Significantly, throughout this debate, no one contested the notion that AZT should be used as a model for action.

Again, as for the case of Subpart E rule-writing, in treatment IND we see rule-making following in the wake of practice. When routine (informal) practice involved relatively few requests for treatment access to investigational drugs, rule-making reflected those procedures. When the nature of practice began to change, rule-making changed (however reluctantly) as a reflection of it. In this example, we can also see a combination of historically contingent and interrelated forces colliding: a dire and rapidly spreading disease; an interventionist Task Force on Regulatory Relief; and FDA's informal practices and existing proposals for reform being pressed into service to address the needs of the moment, with the practices themselves being modified in the process. Additionally, it is no coincidence that it was at this time, in 1987 and 1988, that AIDS activists initiated and escalated their cries to get 'drugs into bodies' (Epstein 1996, 222). Indeed, without a drug like AZT, without a tangible, clearly visible hope of an effective therapy, what impetus would there have been for AIDS activists to protest in the streets and demand access to experimental drugs? The highly successful early attempt at development and approval of an anti-retroviral drug¹⁵ seems to have fuelled and helped to refine the activist agenda (as in, e.g., opposition to placebo controlled studies). Clearly whether AZT had been developed or not, the OMB would have insinuated itself into the FDA's rule-writing process. Nevertheless, it seems reasonable to suggest that the rule-writing process and

¹⁴ In the IND process, local Institutional Review Boards assure patient safety, among other ways, by reviewing procedures for informed consent of patients prior to drug administration.

¹⁵ AZT came early, but it was not the first drug to be tried in clinical trials against AIDS. According to Dr. Robert Yarchoan (1998), suramin had been unsuccessfully tested in a clinical trial, and interleukin-2 had also been tried with some patients.

ultimate shape of treatment IND might have been quite different without the intervening development of a widely hailed drug and the direct experience of thousands of AIDS patients receiving the drug to deter an epidemic in the interim prior to drug approval.

Finally, in the provisions of treatment IND we see key decision points slipping earlier in the drug development process (with correlatively less information for decision-making): from mid-Phase III in the 1983 IND rewrite, to ‘after Phase 2 investigations have been completed’ in Dr. Young’s interpretation of the 1987 reproposal (FDA 1987b, 8856), to the 1988 Subpart E rules, which included a provision to initiate treatment IND ‘when *early* evidence from phase 2 indicates that a drug for a life-threatening or severely debilitating illness is promising’ (FDA 1988a, 41520, emphasis added). The fingerprints of the OMB are clear enough on this trend. Even so, despite FDA’s objections to this conceptual slippage to earlier decision points in the process, in the next chapter we will see the FDA initiate even earlier treatment IND decision-making on another antiretroviral drug, ddI. First, however, we will turn to the reaction of the regulated community to the Subpart E rules.

4.3 Comments on Subpart E¹⁶

Overall, the Subpart E rules appear to have received broad support from physicians, patient groups representing AIDS, cancer and other diseases, and drug-makers. However, even on a procedure for which there was broad support, one can readily see how the FDA was pinched between ideals and practical application, and caught between contradictory interests, incompatible views on the role of government agencies, and conflicting interpretations of law and scientific evidence. When viewed from the perspective of the FDA, sometimes a palpable irony rises from the pages. A federal agency long criticized for being antagonistic towards drug sponsors collaborated with the sponsor to design a successful Phase II protocol for the first AIDS treatment and approved the drug for marketing in 107 days. For the FDA, this was a standard-setting achievement worth using as a model for future drug development. Yet when the FDA attempted to codify these successful practices, and even as observers largely praised the effort, no aspect of the rules went untouched by negative comments. Reservations were

¹⁶ Unless otherwise noted, all comments quoted in this section can be found in Docket 1988N-0359. Each document in a docket bears an item number. Hence, parenthetical citations in this section give the item number in the docket followed by the page number of the quotation within the item.

expressed over ‘FDA objectivity in evaluating the results of studies it has helped design’ (C00032, 2); over whether the FDA had enough resources to implement the programme proposed;¹⁷ and over whether the FDA should have its own programme of focussed research as proposed in the interim rule.¹⁸ Observers likewise disagreed generally on the role of government and regulation: many commentators, especially AIDS advocacy organizations, lamented the general tone of voluntarism throughout the regulations,¹⁹ while the Pharmaceutical Manufacturer’s Association (PMA) and other aligned interests objected to the idea of mandatory procedures.

More significant for the present study are the many observers who questioned the meaning of the categories underlying the new rules — especially what Subpart E did to the definition of the clinical phases. Expressions of confusion and requests for clarification in the comments point to an underlying instability of meaning induced in related categories when a when the new category of drug approval, Subpart E, was created. The features of this instability will be discussed in the next section.

4.3.1 *Subpart E and New Definitions*

In the opening paragraph of a stinging letter in opposition to the interim rule, Dr. Charles Moertel from the prestigious Mayo Clinic wrote that the ‘FDA has undoubtedly received a flood of responses from the scientific community protesting the ambiguity of this proposal’ (C00012, 1). In a more measured tone, the American Association of Pharmaceutical Scientists asked what exactly constitutes a ‘life threatening’ or ‘debilitating’ disease? Who makes the determination? What is an ‘acceptable alternative therapy’ and who makes the determination? How does the new therapy relate to existing therapy and what constitutes an improvement? (C00016). Similarly, the National Organization for

¹⁷ Such comments were made by Hoffman-LaRoche (C00020), the AIDS Action Council (C00032), Gay Men’s Health Crisis (C00030), Pharmaceutical Manufacturer’s Association (PMA)(which was concerned that resources might be diverted from other drug review functions)(C00017), and the National Coalition for Cancer Research (C00013).

¹⁸ The Pharmaceutical Manufacturer’s Association was against it (C00017), while the AIDS Action Council thought it might offset some of the bias problem associated with FDA judgment of studies it helped to design (C00032).

¹⁹ The Gay Men’s Health Crisis wrote that early cooperation with FDA should be mandatory; that postmarketing surveillance should be mandatory; and that participation in treatment IND should be mandatory for any drug sponsors wishing to obtain Subpart E accelerated approval (C00030). The Lambda Legal Defense and Education Fund argued that outside expert consultants should be required to be used throughout the research process (C00034). Moreover, some groups complained that as written the rules ‘lack the teeth’ to be effective; ‘where the supporting language flourishes the enacting language fails’ (National Gay Rights Advocates, C00022, 1).

Rare Disorders argued that the definition of the term ‘life-threatening’ should not include the phrase ‘especially in a short period of time’ (C00011, 2) and that the term ‘severely debilitating’ encompasses ‘a vast array of chronic and debilitating illnesses that are and are not necessarily degenerative’ (2). The Gay Men’s Health Crisis commented that that it was unclear how assessment of risk/benefit would proceed in practice (C00030). The Candlelighters Childhood Cancer Foundation asked what it meant for the regulation to require a ‘meaningful’ duration of response from a clinical study. Many other disease specific organizations wrote to the FDA to confirm that the conditions represented by their organizations (psychiatric conditions, epilepsy, narcolepsy, premature babies, etc.) were definable as ‘life-threatening’ or ‘severely debilitating’ diseases, and individual letters from private citizens were forwarded by their Congressional representatives to the FDA, often expressing a desire for their own disease or that of a loved one to be included in the categories accepted for Subpart E approval.

Hence, key concepts in the regulation, including the applicable disease classifications, the criteria for judging the importance of the experimental drug in a therapeutic context, the risk/benefit analysis, and judgment of a successful clinical study, were all questioned and criticized as ill-defined by observers. These types of complaints had been voiced about the FDA for many years. A 1980 report of the Government Accounting Office cited unclear and inadequate FDA guidelines as a chief complaint among the drug industry and even among FDA reviewers (GAO 1980). We can see in these comments, and in the history of the FDA, the iterative process necessary to clarify and refine guidelines and regulations. While we may be able to point to certain provisions or guidance documents which are poorly worded or subject to multiple interpretations, there is a more fundamental issue at stake: the meaning of any rule or standard is ultimately a matter to be worked out in application.

In some ways, the uncertainty expressed by these observers is surprising. Take for instance the comment that it was unclear how the risk-benefit assessment would proceed in practice. This complaint was made despite the obvious example of risk-benefit rationale available in the case of AZT. As we have seen, other examples also existed in recent FDA decision-making. Hence, for those willing to do the research, there already existed a set of instances which could serve as exemplars for the new category. Nevertheless, although exemplars were available as a resource, the characteristics of the next ‘I,’ to be assigned to the relatively small set {I} (see Chapter 1) were open to

question. How would risk-benefit assessment proceed for a drug less effective or less safe than AZT? How would it proceed for a drug having less convincing evidence than AZT? How would it proceed for a disease other than AIDS? Ultimately, as we would expect in finitism, the next application of the concept was open-ended. These questions and comments reflect the uncertainty inherent in creating new categories of rules, and constitute an engagement in what we could call *articulation-practice* — a verbal rehearsal of proposed exemplars in preparation for contingencies of practice that can be envisioned based on the experience and viewpoint of the observer doing the articulation. Articulation-practice is an essential strategy for reducing uncertainty of concept application in a regulatory environment.

In a regulatory environment in which boundaries were already being pushed, many observers wanted to extend the concepts behind this regulation even further. Hoffmann-La Roche's Philip Del Vecchio argued that the Subpart E provisions should apply to *all* new drugs: 'The concepts of FDA participation in the planning of drug development, close coordination with the sponsor, evaluation of data in terms of risk-benefit analysis, and approvals based on commitments for extensive postmarketing studies seem to be appropriate for all drugs' (C00020, 2). Similarly, the PMA wrote that they 'believe the Agency has taken an unnecessarily narrow view of the applicability of the rule by suggesting that it is especially relevant to situations where no satisfactory alternative therapies exist. PMA suggests that there are likely to be many examples of drugs under development to treat life-threatening and severely debilitating diseases where the advantage to the patient population will be a substantial improvement in the quality of life compared to currently available therapy' (C00017, 1). One private citizen wrote to say that the provisions should not apply only to drugs to treat life-threatening diseases, but should also include products to diagnose, cure, mitigate, or prevent life-threatening diseases (C00015). A Congresswoman from South Carolina forwarded a letter from a constituent arguing that the approach used for Subpart E should be made available for non-drug therapies (C00035). The American Society of Clinical Oncology was concerned that the discussion of valid clinical endpoints seemed to be limited to survival: 'In addition to the currently required criteria for survival, tumor regression, palliation of symptoms and avoidance of toxicity are valid endpoints for evaluating new agents' (C00036, 1). The National Coalition for Cancer Research made a similar comment (C00013).

At the same time, other observers were alarmed or baffled at the suggestion that Phase II studies should stand as sufficient for marketing approval, often taking this measure to represent a weakening of the traditional interpretation of statutory and scientific standards of efficacy. Significantly, many of the comments struggle with what these proposals do to traditional categories of drug approval and point towards at least a partial redefinition of Phase II in terms of Phase III. The Mayo Clinic's Dr. Moertel, a gastrointestinal oncologist and no stranger to FDA regulatory practices,²⁰ wrote that the 'term "well controlled Phase 2 study" is an internal contradiction' (C00012, 1). In his comments, Moertel reviewed what he considered to be the 'traditional and well established' phases of drug development. Phase I is the 'initial experience with a new agent in humans' and as such is 'designed to explore early toxic reactions' and to establish 'the most appropriate schedules and routes of administration' of the drug (1). What results from Phase I is an 'early estimate' of clinically tolerable dosage. In Phase II, therefore, it is necessary to 'expand toxicity knowledge and to redefine appropriate dosage' (1). Also in Phase II, said Moertel, an important major objective is to 'obtain early evidence of possible therapeutic activity' (1). 'A Phase 2 study is rarely controlled at all and is never "well" controlled' (1). Phase III is the 'only phase of drug development that is traditionally well controlled' (1). Moertel went on to note that the FDA cited the clinical studies of zidovudine and timolol (an anti-arrhythmic drug) as examples of well-controlled studies, but he argued that clinical trial scientists 'would characterize these as Phase III trials; and, indeed, this is the way they were characterized by the investigators' (2). He insisted that if by 'well-controlled Phase 2 study' the FDA meant 'the traditional Phase 3', then he had 'no quarrel whatsoever' (2). But if so, the wording of the rule should be changed. For Moertel, the only way that Subpart E accelerated approval could be considered scientifically valid is to run a Phase III study and call it Phase II — which is precisely what he contended was done with AZT.²¹ But, in fact, the transcript of the January 1987 advisory committee meeting to consider AZT (FDA 1987a) approval makes it clear that the trial produced far from adequate data in many respects. The committee

²⁰ In an interview, Dr. Robert Temple mentioned that as the head of the oncologic drugs advisory committee, Moertel had been an ardent advocate of better clinical trial design. According to Temple, until roughly the mid-1980s, cancer trials were thought to be successful if there was evidence of tumor shrinkage. In opposition to this practice, Moertel argued that clinical benefit must be demonstrated for a trial to be valid. We will see Dr. Moertel in action on the committee in the next chapter of this thesis.

²¹ Contrary to a comment he made about AZT in a December 1987 oncology advisory committee meeting, where he characterized the AZT approval decision as more political than scientific. See the discussion of the meeting for the cancer drug mitoxantrone in Section 5.1 of this thesis.

was faced with having to make a decision to approve the drug without even knowing the optimal dose for the narrow patient group tested in the study. Moreover, there was no data for the drug on long-term exposure, on states of disease progression other than the most severe, on interactions with other common drugs patients were taking to treat symptoms of opportunistic diseases, and since a good method for culturing HIV was still lacking, there was no virologic data on in vivo efficacy.²² These are all the types of data deficits Moertel would presumably expect to see when one uses a Phase II study to support a marketing application. Hence, while the AZT trial was large, multi-centred, and placebo-controlled (features which Moertel obviously considered uncharacteristic for a Phase II study), it still suffered many of the information deficits typical of early clinical studies.

Moertel's description of the clinical phases clearly differs from the version proposed in 1979 and ultimately incorporated into the IND rules, which characterized Phase II as capable of producing early but legally defensible evidence of efficacy and indicated that a Phase III trial could be controlled or uncontrolled. In part this discrepancy is likely rooted in disciplinary differences. We can see such differences in, for example, the National Coalition for Cancer Research's comment that '[t]raditional Phase II studies for cancer agents are used to define the spectrum of an agents' [*sic*] anti-tumour activity. As a result, redefining Phase III as Phase II is unlikely to speed up the approval process for cancer agents' (C00013, 1). Indeed, an FDA guidance document for antineoplastic agents (FDA 1981) indicated that Phase II studies for anti-cancer agents are used to test the activity of the drug against a number of different disease types and tumour types. Larger, controlled studies on a particular disease type cannot be accomplished until appropriate disease targets are established. Hence, even if one designs a Phase II study to be large and controlled, there will still be the issue in oncology trials of which disease entity to target. Of course, in the case of the AZT trials, there was no need to test the experimental drug on a range of diseases; the target agent was singular and the clinical trial was designed around what, at the time, was considered a meaningful distinction between AIDS and ARC patients.²³

²² See the beginning of the afternoon session where the FDA's Dr. Cooper summarizes the strengths and weaknesses of the trial (pp. 124-128)

²³ ARC stood for AIDS Related Complex and was considered a preliminary condition to AIDS, characterized by swollen lymph glands, night sweats, fevers, and unexplained weight loss. The advisory committee transcript indicates that the AZT pivotal trials tested 'relatively advanced stages of disease, late

These comments highlight the practical difficulty of designing a general rule based on a specific experience: the model experience tends to establish a conceptual frame that generates default boundaries. The design and use of Phases II and III in practice are variable and disease-specific, while the version written into the regulations represented a generalized, standardized conception not applicable to every case.

The comments provided by attorney William Schultz²⁴ for Public Citizen likewise underscored the ambiguity between Phase II and Phase III trials created by the interim rule (C00018). Noting that the new regulations ‘provide for the elimination of Phase III of drug testing’, Schultz remarked, ‘we agree that the issue is not which phases have been completed, but whether a new drug application satisfies the safety and efficacy requirements of the Food, Drug and Cosmetic Act’ (3). Following a description of the phases of clinical testing in recent practice,²⁵ Schultz agreed that ‘the safety and efficacy of certain drugs may be supported by evidence gathered during an expanded Phase II, in lieu of the evidence ordinarily collected during Phase III’ (4). He went on to say that if his interpretation of the proposed agency action is correct, then it is ‘not objectionable’, however ‘it is misleading since the agency claims to be eliminating the most resource intensive phase of drug testing (Phase III) when in fact the regulation simply codifies the FDA’s past practice of substituting an expanded Phase II for Phase III’ (4). Schultz clearly saw the FDA’s proposal in practical terms of how much evidence had been generated. Based on his view of the requirements of the Food, Drug, and Cosmetic Act, he dismissed out-of-hand the idea that a marketing application could be based on a Phase II trial as defined by recent FDA publications. Instead, following FDA’s lead in taking

ARC and AIDS post-PCP’ (14), where the latter category referred to AIDS patients who had been afflicted by *pneumocystis carinii* pneumonia, a common opportunistic disease for AIDS patients.

²⁴ Schultz subsequently became the FDA’s Deputy Commissioner for Policy and was a key negotiator for the FDA in crafting the 1997 Food and Drug Administration Modernization Act.

²⁵ Schultz wrote that ‘[i]n the past, Phases I, II, and III had no precise definition and no particular regulatory significance’ (3). In recent years, however, he writes that the Agency defines Phase I as testing for safety and dosage, but not efficacy, on 10 to 50 test subjects; Phase II as collecting continued evidence on safety and preliminary data on efficacy using 20 to 200 test subjects; and Phase III as gathering more data on safety and efficacy in a controlled study of roughly 1,000 or more subjects (3-4). Interestingly, both this and Moertel’s definitions of the phases differ somewhat from the one provided by the FDA in their 1983 publication of the proposed IND rewrite (see FDA 1983, 26723). Two major differences are that in the 1983 publication, Phase II trials are ‘the early controlled clinical studies conducted to evaluate the effectiveness of the drug for a particular indication’, and they are ‘typically well controlled, closely monitored, and conducted in a relatively small number of patients’; Phase III studies are ‘the expanded controlled and uncontrolled trials’ which are ‘performed after preliminary evidence of the drug has been established’ and are conducted on a test group of ‘from several hundred to several thousand patients’ (26723).

AZT to be the paradigm case, he assumed that what FDA would actually approve under Subpart E was a type of ‘expanded’ Phase II.

By contrast, James T. Doluisio, Ph.D., President of the American Association of Pharmaceutical Scientists, clearly perceived the rule change as constituting drug approval on the basis of less evidence (C00016). (It might be worth pointing out that as a representative of scientists who worked for drug-makers, Doluisio had more of an interest in promoting this view of the rule change than Schultz, who represented a well known public interest group which with a history of opposing rule changes perceived as compromising drug safety.) One of Doluisio’s comments expressed concern that expedited review procedures would expose FDA reviewers to greater risk of litigation since the ‘reviewer is asked to make regulatory decisions on less scientific information’ (1). Doluisio did not contest the idea of approval taking place based on Phase II studies, but he registered concern that serious side effects may not be seen in the smaller patient populations typical of Phase II testing, thus taking for granted that smaller patient groups characteristic of Phase II studies would be a feature of Subpart E accelerated drug approval.²⁶ Doluisio further expressed concern that benefit could not be adequately weighed in any risk-benefit assessment under Subpart E, since in this situation benefit was not fully known. Accordingly, he suggested that the ‘design of the Phase II studies may need to be altered to provide a more balanced assessment’ (2). If one needs better efficacy data, of course, one needs to run a controlled study with a reasonable number of patients in each arm. So although Doluisio is more direct (and seemingly also more sanguine) than other observers about Subpart E approval proceeding with less data, he nevertheless anticipated adapting Phase II in a manner not unlike Schultz’s ‘expanded’ Phase II.

Significantly, this and most of the other comments addressed the problem only in terms of Phase II design as it traditionally follows from Phase I. It did not occur to anyone at this point that re-designed Phase I trials could fill some of the data gaps that Phase II trials formerly addressed (such as response of a range of tumour types). Eventually, as we will see, FDA guidance on clinical trial design did reach back progressively to include Phase I and even pre-clinical studies. But at this juncture, only

²⁶ Two other comments made the same point (see item nos. C00006 and C00015), with one (C00015) arguing that the previous practice of including phase 3 testing ‘provided a “standard” of expectation by the prescriber and patient’ (1). Hence, he suggested, ‘Products marketed without Phase 3 testing should be required to be labeled to inform prescribers and patients that no phase 3 testing was performed’ (1).

one comment suggested that FDA consultation should extend beyond Phase II to Phase I and pre-clinical study design (C00015) (and it is not clear if this observer anticipated an overall redesign of the clinical trial system to produce adequate evidence of safety and efficacy by the end of Phase II). Hence, the rules for Subpart E and the comments received on them reflect the same kinds of conceptual boundary-drawing and negotiation evident in the definition and practice of the clinical phases apparent in the 1970s, when over time the meaning of Phase I remained more stable while the boundaries between Phases II and III appeared more flexible. It is clear from the comments that the introduction of the new regulatory category, Subpart E, created new instability in the meanings of these clinical phases.

The boundary-drawing I discuss here should not be confused with the types of boundaries at issue in Gieryn's (1983) boundary-work, in which effort is made to separate 'science' from 'non-science'; or with the related concept of 'ethical boundary-work' used by Wainwright et. al. (2006), in which scientists seek to separate ethical issues from scientific work having profound social valence, such as in embryonic stem cell research. By contrast, I am speaking of definitional boundaries within categories of action in drug evaluation — boundaries all of which are considered firmly within the realm of 'science' by the actors negotiating the meaning of the categories. While this definitional instability resonates with the concept of 'interpretive flexibility' used in the 'Empirical Programme of Relativism' (EPOR) (Pinch and Bijker, 1987) in which differing interpretations of nature are shown to be available to scientists, here the issue is not really interpretive flexibility so much as interpretive *confusion*. The point of the discussion is to demonstrate the definitional instability introduced into existing related categories (and subcategories) of action when a new category is created, as would be predicted by finitism (Chapter 1).

4.3.2 *Subpart E and the Regulatory Importation of Treatment IND*

One other set of comments should be noted here. The AIDS Action Council made an intriguing argument in which they compared the standards for accepting a treatment IND application as written into the 1987 final rule and the standards for 'inviting' a drug sponsor to submit a treatment IND protocol under Subpart E (the invitation comes from the FDA once it seems clear that a marketing application will be successful under Subpart E approval) (C00032). This comment from Executive Director Jean F. McGuire argued that the 1987 standard was less stringent than the 1988 version

written for the Subpart E rules. The problem, according to McGuire, was that the 1987 final rule on treatment IND ‘clearly contemplates’ (3) the use of treatment protocols for experimental drugs in cases where additional studies are required; i.e., the rules applied to a drug ‘under investigation in a controlled clinical trial under an IND’ *or* could also apply to a drug for which ‘all clinical trials have been completed’ (3). The Subpart E rules, by contrast, speak of the treatment IND merely as a ‘bridge between phase 2 trials and the point of marketing approval’ (3).²⁷ Under these circumstances, McGuire argued, ‘more people will benefit from access to drugs during phase 3 than from the early licensing anticipated by the interim regulation’ because ‘only rare cases will produce phase 2 data sufficient for market approval — even with the best design of studies’ (3).

Is treatment IND different in the two contexts? While the 1987 final rule on treatment IND specifies that the sponsor of the experimental drug must be ‘actively pursuing marketing approval of the investigational drug with all due diligence’ (FDA 1987c, 19476), the rules do make allowance for drugs that may still require additional testing. In comparing this rule to the discussion of treatment IND appearing in the Subpart E rules, it must be said the Subpart E rules *do not* specifically contradict this allowance. Moreover, no discussion of treatment IND in the Subpart E rules supersedes any portion of the 1987 final rule. Indeed, the section of the Subpart E rules at issue specifically references the sections of the *Codes of Federal Regulation* in which the 1987 treatment IND provisions are codified.²⁸

Nevertheless, having made these points clear, one can still say that there seems to be a subtle shift of emphasis from the 1987 standalone version of the treatment IND rules and the 1988 discussion of treatment IND as imported into Subpart E approval procedures. As the above-noted comments adamantly charge, the ‘bridge’ language appearing in Subpart E (FDA 1988a) (that treatment IND is a bridge between ‘completion and initial analysis of promising phase 2 studies and the point of marketing approval’ [41520]) is suggestive. This passage makes it seem as if a treatment IND will only be considered where the clinical trial data is in processing for a marketing application, ignoring situations where Phase II data is promising enough to write a treatment protocol

²⁷ The treatment IND quotation comes from FDA 1987c, 19476. The Subpart E quotation is from FDA 1988a, 41520. This quotation is slightly different from the original passage, which reads, ‘Within the drug development process, treatment IND’s can provide a bridge between the completion and initial analysis of promising phase 2 studies and the point of marketing approval’ (41520).

²⁸ *Codes of Federal Regulation* Part 21 Sections 312.34 and 312.35.

but additional clinical study is desirable before marketing is granted — a scenario the 1987 final rule appears to accept.

Hence, the apparent function of treatment IND within the context of the Subpart E rules *does* appear to have shifted, with the apparently unintended consequence of equating the criteria for a Phase II trial adequate for Subpart E approval with those for allowing treatment distribution of an investigational drug. In my opinion, it would be mistaken to suggest that the FDA intended to narrow the definition of treatment IND in their discussion of Subpart E accelerated approval. What is reasonable to suggest, however, is that the experience with AZT shaded the thinking that went into writing this section of the rule. Indeed, as with many sections of the Subpart E rules, AZT is never far away. Consider the full ‘bridge’ quotation in context:

Within the drug development process, treatment IND’s [*sic.*] can provide a bridge between the completion and initial analysis of promising phase 2 studies and the point of marketing approval. Thus, when early evidence from phase 2 indicates that a drug for a life-threatening or severely debilitating illness is promising, FDA will actively work with the sponsor to evaluate the appropriateness of a treatment protocol. This approach was used during the development of zidovudine, and allowed wide availability of the drug to over 4,000 patients while the marketing application was being assembled by the sponsor and reviewed by the FDA’ (FDA 1988a, 41520).

Here treatment IND takes on yet another meaning in yet another context. The first major transition of meaning, as described above, came in the collision with AIDS and AZT development. The second, more subtle, shift of definition came with its importation into the Subpart E rules in an attempt to create procedures for accelerated drug approval based on the AZT experience. One shift of meaning comes about as a result of practical application of the rule to a specific, historically contingent circumstance. The other shift of meaning comes about through regulatory articulation-practice, this time with unconscious or unintentional shading of the original concept. More transitions for the meaning of this concept are yet to come.

4.4 Preliminary Interpretations

Prior to AZT there had been many examples of the FDA using its discretionary power to hurry along drugs perceived to be important. The FDA used some of these cases as evidence to support the validity of the interim Subpart E rule, noting that '[t]here have been other circumstances, particularly in the oncology area, where early (phase 2) results were such that additional studies were not needed to conclude that the drug was effective and that its benefits outweighed its risks' (FDA 1988a, 41519). We saw some examples of this practice in Chapter 3. Additionally, while the Agency was very clear in Subpart E that at least two clinical studies would still be required to satisfy the statutory mandate for substantial evidence of efficacy, it nevertheless noted that AZT was not the only drug approved on the basis of a single study. The drug timolol had been approved for reduction of post-infarction mortality on the basis of a single, large study (FDA 1988a, 41521). (And as we saw in Chapter 3, etoposide was approved in 1983 for testicular cancer on the basis of a not-so-large single study — an example the FDA chose not to highlight in the Subpart E publication.)

Given the FDA's actions prior to AIDS and AZT, we can suppose that if the only goal were to expedite drugs to AIDS patients, the FDA could have done so by continuing to exercise the same kind of discretion it had used in the past for other situations. Instead, due to various considerations including political pressure, they chose to define the past exceptions to the rules as 'precedent' and build new rules on that basis.²⁹ What we see here, in effect, is a new regulatory category being 'primed' (Barnes 1983) through performative enunciation (Chapter 1). The new category then became defined in relation to existing categories, and in terms of those characteristics which made it deviant in the first place. Notably, however, not *all* of the deviant characteristics became the basis for the new category by default; a selection process took place. Thus, for example, the FDA

²⁹ Here I come close to Temin's (1985) argument that in times of crisis, governmental agencies make a transition from 'customary' behaviour to 'command' behaviour; i.e., whether needed or not, the response to crisis is to make new laws and regulations. Temin asserts that one can see this pattern throughout 20th century drug regulation. Moreover, he notes, the statutory and regulatory reforms often fail to address the provoking crisis; public outrage gets used as leverage to implement other agendas. While I agree in outline form, and indeed have made a similar argument (Messner 2006), Temin assumes that the proliferation of rules always results in 'tightening the FDA's control' (438). To be fair, writing in 1985, Temin did not have the hindsight of the AIDS crisis. Even so, one could easily imagine an argument that this period represented an increased 'scope' or 'reach' of the FDA consistent with Temin's model. However, in these examples the FDA claims no new scope — they already had jurisdiction over drugs to treat life-threatening diseases — and we see no tightening in terms of the effect on drug sponsors and consumers. Rule-making during this period significantly liberalized FDA standards for approving certain drugs. Additionally, whatever the written rules may be, agencies such as the FDA have broad discretion to improvise as needed in a goal-oriented fashion — and one can point to examples of rule-bending with both restrictive and liberalizing effects.

excluded the deviant characteristic ‘single-study drug application’ from its new category ‘Subpart E’. Through its selection criteria for Subpart E, the FDA expressed a willingness to continue reforming its practice along the lines already traced by earlier actions, but had a clear desire to retain the traditional interpretation of ‘substantial evidence’ as requiring multiple clinical studies. As we will see later in this account, the conceptual element ‘single-study drug application’ will nevertheless be appropriated by others and eventually formalized.

Importantly, this account adds more examples of a principle identified in Chapter 3: *formal rules followed practice*; they did not anticipate situations to come, but rather were a reflection of actions already taken and decisions already made. More than that, rule-‘breaking’ was *constitutive* of rulemaking: the way in which the existing rules were adapted to contingencies of the moment became the (selective) basis for designing the new rules. For actions to have already been taken, there clearly existed a perceived need for such action — a tacit consensus constituting a *de facto* rule. The body of action already taken served as the primer for the acceptance of future similar action, and ultimately as justification for the new rule, even as the new rule legitimized the previous decisions to take action. This function of rule-making as legitimating past action (which will be seen again in Chapter 6) is clearly not contemplated in any direct way by contemporary theories of regulation, which tend to treat new rule-writing as merely a response to legislative directives.

Subpart E was issued as an ‘interim’ rule and, as such, was effective immediately. While a response to comments and revision of the rule could, theoretically, have been made, to my knowledge no final rule for Subpart E was ever published. Therefore, the comments reviewed in this chapter were never directly addressed. The implication is that the FDA was confident that a significant consensus had already been achieved on the issue — or at least that consensus was achieved among the political and medical actors who mattered most in the negotiation. The development of AZT was accomplished through collaboration with scientists from the NIH and Burroughs-Wellcome, and the approval had gone forth on the recommendation of an advisory committee composed of individuals considered to be experts in the field. Everyone from physicians to politicians to AIDS patients agreed that the development and approval of AZT was a significant achievement and, indeed, AZT was universally agreed to be the model to follow even for those who disagreed on what that meant. A demand for a new rule had come from the

Vice President and the rule itself had to be reviewed by the OMB prior to publication. The rapidity with which the last step took place is an indication that the White House fully endorsed this action. Hence, for Subpart E, any comments received following the publication of the rules were secondary to the consensus process which had already taken place by the time Subpart E was promulgated. This is not to imply that the FDA ignored or dismissed the comments received on Subpart E (indeed, FDA did try to clarify some of these issues, e.g. the definition of 'serious or life-threatening' and which diseases fell into those categories, in subsequent rule-writing for Subpart H, discussed in Chapter 6), however they certainly were not deemed significant enough to merit any revision of the interim rule.

If the formal 'notice-and-comment' rule-making required by U.S. law does not serve as the primary consensus process, then what is its purpose? To be clear, I am not suggesting that notice-and-comment rule-making does not serve *any* consensus-making role. But this is a process generally assumed to be of primary importance to consensus-building and I am asking what role it really serves in practice. Notice-and-comment rule-making has a specific legal function made clear by the *Administrative Procedures Act*, which is that Agency responses to comments and justifications for new rules must be made in sufficient detail to provide the basis for a defence should a legal challenge of the new rule be mounted. Hence, these procedures serve to assure that Agencies will not gratuitously dismiss opposing positions. However, this latter function is not necessarily the same as consensus; the law does not require that agencies find agreement with their notice-and-comment interlocutors, only that they provide a well reasoned basis for disagreement. Providing such a reasoned argument *could* help to win over opposition. However, the fact that the reasoned argument must be sufficiently detailed to endure legal challenge means that the law does not anticipate consensus to be achieved in every case. Moreover, however unusual the practice may be, the fact that interim rules can be published tends to undermine the notion that notice-and-comment rule-making always serves a consensus-building role. In the case of Subpart E, the consensus clearly already existed and served as the unspoken justification for the issuance of the interim rule. Nevertheless, notice-and-comment rule-making is not always as straightforward as this example, sometimes requiring multiple iterations of comment, response, and reproposal (Breyer, 1982).

The observations made to this point suggest that the function of notice-and-comment rule-making varies in practice according to how well primed the new category of

rule-making had been prior to the attempted performative enunciation — what we should really call a *contingent performative enunciation*, since the FDA’s authority is by definition circumscribed and subject to legal challenge. The FDA exercises contingent authority in a knowledge-making community. So, as a preliminary proposition, it is reasonable to suggest that the number of iterations required in notice-and-comment rule-making is inversely proportional to the level of consensus achieved in practice prior to rule-writing. Correlatively, I would suggest that to the extent the rules have been based on tacit practices, notice-and-comment rule-making amounts to an exercise in articulation-practice, where uncertainties regarding the next addition of I_n to the {I} represented by the rule are mitigated. If this is true then it suggests another reason why tacit, practice-based consensus should be acknowledged by the theory of regulation: the degree to which such consensus exists will tend to be predictive of the level of challenge agencies can expect during the formal rule-making processes.³⁰

In this account AZT began as a meaningless physical substance (or more accurately, a substance having the meaning ‘failed cancer drug’, or at most, ‘possibly bioactive molecule’), and then was transformed: it was re-identified as a research object, characterized with respect to its activity against AIDS, and assigned a value in terms of ‘safety’ and ‘efficacy’ with correlative associations such as ‘promising’, ‘toxic,’ etc. This process is describable by the term ‘interpretive flexibility’ (Pinch and Bijker, 1987). However, although this term underscores that physical objects can be interpreted in various ways thereby opening the way to a sociological understanding of scientific knowledge and technological development, it tends to leave unspecified the role of such reinterpretation in the process of concept application. In this account, we can see that through this reinterpetive process, AZT became intelligible as a conceptually and culturally significant object and, as such, one which could be appropriated as a resource implicated in the development of certain attitudes, decisions, and actions — and one which was appropriated heterogeneously, to support contradictory positions. AZT was not merely reinterpreted, but was *resignified* as a term for heterogeneous social appropriation and application. Such resignification and appropriation of an originally

³⁰ Here I mean ‘formal’ as opposed to tacit processes of goal oriented rule-bending. In the terminology of administrative law, ‘formal’ rulemaking refers to a rarely used procedure involving formally conducted evidentiary hearings, while the rule-and-comment rulemaking under discussion in this chapter is called ‘informal’ rulemaking. See the description of the process in Breyer (1982) and Croley (1998), among many other sources.

(relatively) meaning-neutral object is a supremely human achievement. Clearly this process, which we might term *transformative resource creation*, is fundamentally characteristic of the way scientific knowledge is created.

This concept of transformative resource creation may seem similar to actor-network theory's 'translation' (Callon 1986; Latour 1987; 2005), but there are fundamental differences between the two concepts. Translation involves the establishment of an actor network in which both human and non-human actors could conceivably be doing the work of translation. The great insight of actor-network theory is that non-human actors have an important role to play in accounts such as these, and that meanings are generated in the interaction between actors in a given situation, and should not be presupposed. Indeed, in the account given here, the physical nature of AZT itself (in combination with the retrovirus) led to certain clinical results which created the conditions for AZT to become an important conceptual and cultural resource. Even so, Barnesian N-terms and S-terms (Chapter 1) cannot choose their own significance. Although meanings are indeed generated in the 'network', not prior to it, only humans can create and sustain meaning. In the terms I use here, it is the perception of humans, and humans alone, which does the work of resignification. If AZT was influential, it is because of the decisions of individual people who chose to work with it, and thereby had their view of things changed. In this context, I am reminded of Pickering's 'mangle of practice' (Pickering 1999) in which both matter and the manipulators of matter are altered in the process of interaction. However, as much as Pickering does construe knowledge as a social product, his concept of the 'mangle' was designed to elaborate the nature of the interaction of humans with matter, and in so doing social processes of knowledge-making tend to be eclipsed by consideration of the scientist with her object of study. To be sure, physical objects impose themselves on our being in ways not of our own choosing, such as the emergence and devastation of the AIDS virus itself. However, outside of human social context, even the AIDS virus has no *meaning*. What is considered to be knowledge of the AIDS virus is derived and agreed upon through social processes (Kusch 2002). Hence the idea of 'transformative resource creation' emphasizes the social-consensual derivation of knowledge in a way that the 'mangle' does not and 'translation' cannot. The physical characteristics and behaviour of a particular natural object may specially qualify it to be perceived as socially or culturally significant (which is certainly the case for both AIDS and AZT), but those objects are obdurately insensible to whatever significance we would

attach to them; transformation of such objects into an important cognitive and social resource for humankind can only be accomplished by humans.

5. WAR ON CANCER: THREE CASES

On the topic of drug development for serious disease — and the regulations created to expedite their approval — one typically associates the 1980s with AIDS. However, in the last two chapters I have been building the argument that many of the approaches to AIDS drug development and approval seen in the 1980s and 1990s had their roots in previous decades. Which is to say that the original concepts of accelerated approval and expanded access proposed and informally practiced prior to the 1980s were developed to address life-threatening diseases *other* than AIDS, especially heart disease and cancer. The latter disease was perhaps especially motivational in the movement towards earlier regulatory decision-making since it is, as former FDA commissioner Donald Kennedy reportedly said, ‘the disease Americans fear more than they fear war’ (Anonymous 1979, 46D). Declarations of wars against cancer had been made and sustained in the U.S. throughout the century.¹ However the most expansive war on cancer would be initiated by President Nixon in 1971, with the announcement of his 1972 fiscal year budget which included a \$100 million increase in funding ‘to accelerate greatly the search for a means of preventing and curing cancer’ (President 1971, 136). In tandem with this increased funding came the *National Cancer Act* of 1971, which among other things allowed the National Cancer Institute (NCI) director to bring funding requests directly to the office of the President. The Nixon initiative and the *National Cancer Act* represented an unprecedented infusion of public funds specifically earmarked for cancer

¹ In 1913, distinguished physicians and well known women of wealth met to establish a ‘National Anti-Cancer Association’. Reported in the *New York Times* under the headline ‘Rich Women Begin a War on Cancer’ (Anonymous 1913), the organization was intended to promote public education on the symptoms, early detection, and prevention of cancer on the model of successful public awareness campaigns on tuberculosis. Hopes for early detection and treatment of cancer were bolstered by the discovery and use of radium and x-rays, which in 1921 the ‘famous’ Dr. Charles Mayo (one of the founders of the Mayo Clinic in Rochester, Minnesota) declared ‘valuable aids’ in cancer treatment and diagnosis (Anonymous 1921). Congress established the National Cancer Institute in 1937 with the charge to conduct research into the causes, diagnosis and treatment of cancer, and to foster such research in both private and public agencies. Hopes of a cure were fostered from the late 1940s due to the development of chemical agents having some effect against certain forms of cancer (Anonymous 1949), which became a focus of aggressive effort thereafter (Schmeck 1959).

research and granted the NCI broad discretion in deciding how those funds would be allocated. Hence, during the period of the 1970s when the first attempts were made to formalize accelerated approval and treatment IND, the emphasis on speedily developing and approving drugs for cancer patients is clear.

We have already noted in Chapter 3 some examples of cancer drugs approved in the late 1970s or early 1980s on the basis of one or more Phase II studies. These approval decisions help to demonstrate that the experience with AZT reflected *both* continuity and change: other drugs perceived to be therapeutically important had been approved at an earlier point in the drug development process than ideally desirable; what made AZT different was the rapidity of the overall development process and its status as the first therapy for a new, frighteningly lethal disease. In this chapter I will continue to build the case that the type of decision-making used for AIDS drugs was not unique. I will examine three cases of approval decision-making for cancer drugs selected for their temporal proximity to the approval of AZT and to that of two other AIDSs drugs, ddI and ddC, discussed in the next chapter. In so doing, we will see that although much of the media attention for accelerating approval of important drugs became focused on AIDS in this period, the Oncologic Drugs Advisory Committee (ODAC) quietly made decisions which, like those made for AIDS drugs, anticipated key features of yet-to-be written laws and regulations. This is not to imply that the ODAC made these decisions in a vacuum, unaware of the events surrounding AIDS drug approval and regulation. However the ODAC often had the opportunity to contemplate these decisions away from the limelight and without throngs of activists clamouring for access to the podium and observers filling stuffy meeting rooms to capacity, as they increasingly did for Antiviral Drugs Advisory Committee meetings. Hence, although future rule-writing (particularly the Subpart H rule for accelerated approval discussed in the next chapter) will be explicitly based on the experience with AIDS drugs, key practices embodied in those rules were also evident outside the realm of AIDS. We will also have an opportunity to study the role of FDA regulators in advisory committee decision-making and compare it to the role of the FDA in the AIDS drug decision-making discussed in the next chapter. We will find that in important ways, the story of AIDS decision-making is not so different from the deliberation taking place in the quieter oncology meeting rooms.

5.1 Mitoxantrone: Shades of Future Legislation

5.1.1 Study Design: *The Ambiguity of Equivalence Trials*

In December 1987, nine months after the AZT approval decision, the ODAC met to discuss the new drug application (NDA) for mitoxantrone (trade name Novantrone) for patients having acute non-lymphocytic leukaemia (ANLL) (now more commonly called acute myeloid leukaemia), a relatively rare progressive malignant disease in which immature blood-forming cells in the blood and bone marrow proliferate abnormally (FDA 1987d).² The drug sponsor, Lederle Labs, requested in the NDA that the drug be indicated in the labelling for first-line adult patients having ANLL — ‘first line’ meaning that the drug would be used as part of the first treatment regimen attempted on chemotherapy-naïve patients. If the first-line treatment failed, then a second-line therapy would be attempted, and so on. Generally speaking, the more chemotherapy a patient has received, the more compromised is a patient’s physical condition and the less likely are subsequent courses of therapy to succeed., so the first-line treatment should be the most effective option available.

In the case of ANLL, the standard (first-line) therapy at the time was the administration of cytosine arabinoside (or ‘Ara-C’) for seven days followed by infusion of daunorubicin (also called cerubidine)³ for another three days. This therapeutic regimen is sometimes abbreviated ‘7 and 3’ (and at times the drug sponsor’s presenters called the same regimen ‘3 and 7’). The sponsor had conducted two randomized, multicentre studies designed to substitute the investigational agent, mitoxantrone, for daunorubicin in this regimen (Figure 5.1). Patients who failed the first course of therapy were given a second course, this time as ‘five and two’: five days of Ara-C followed by two days of either daunorubicin or mitoxantrone (patients who initially received daunorubicin would continue to receive that drug and likewise for mitoxantrone). Patients who failed the second course were dropped from the study. Patients who achieved complete remission on either the first or second course of therapy would receive two more courses of therapy using the administration schedule on which remission occurred (either 7-and-3 or 5-and-

² According to one advisory committee member, in 1987 there were approximately 13,500 patients having the disease in the U.S. (FDA 1987d, 83).

³ Daunorubicin is also known as cerubidine. In the advisory committee transcript, the drug is alternatively called daunorubicin or cerubidine. Likewise, Ara-C is alternatively called by its nickname, or by arabinoside or by cytosine. To avoid confusion, I will refer to these drugs as ‘Ara-C’ and as ‘daunorubicin’ through the text.

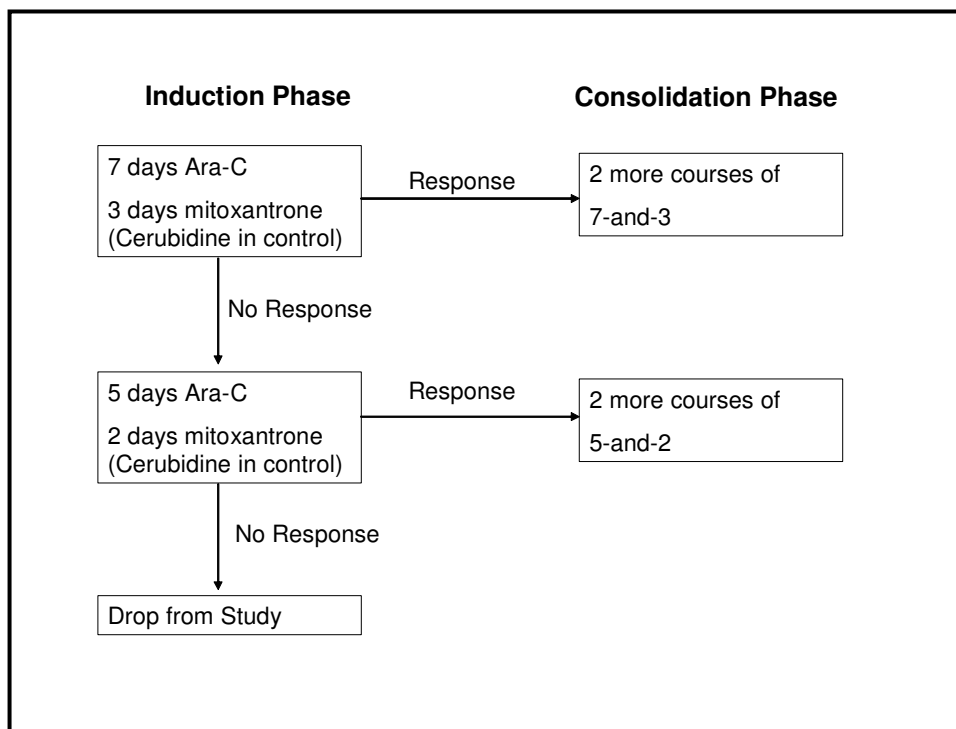


Figure 5.1
Therapeutic Scheme for Mitoxantrone Clinical Trials

2). The period of drug administration during which a remission was sought was called ‘induction’ therapy while the post-remission follow-up was called ‘consolidation’ therapy.

The trials of mitoxantrone were designed to be ‘equivalence’ studies in which the goal is not to demonstrate the superiority of the investigational regimen over the standard one, but merely to show that it is neither better nor worse than the standard therapy within a reasonable tolerance.⁴ Although equivalence trial designs require larger patient populations than the more efficient ‘superiority trial’ design, from the drug sponsor’s perspective the strategic advantage of the equivalence trial (or of a ‘non-inferiority trial’) is clear enough: the investigational drug need not outperform the standard therapy to be demonstrated to have effectiveness. The disadvantage from the perspective of the FDA and advisory committees is that equivalence trials can be devilishly difficult to interpret. In the first place, an investigational drug can be less effective than the control drug and still pass the test for ‘equivalence’. The smaller the patient population used, the less the precision with which one can characterize the gap in efficacy between the investigational and control arms. Hence, except for situations in which there are pronounced

⁴ For a useful discussion of the basic statistical differences between placebo-controlled, superiority, equivalence, and non-inferiority trials, see Streiner 2007.

discrepancies between effectiveness or toxicity, the degree to which one drug is less efficacious than another can be difficult to discern in medium-to-small trials.⁵ More than that, even if the investigational drug performs equivalently to the control drug, there is still the possibility that *neither* drug is very effective in a given study. For this reason, the effects of the standard treatment must be well documented, ideally against placebo. Accordingly, in this case the drug sponsor presented the results of two historical studies of the standard therapy which were judged to use patient populations similar to the studies in the NDA, and thought to give ‘a true picture of the current approach to acute nonlymphocytic leukemia, and what one can expect’ (FDA 1987d, 15). Those reference studies resulted in a complete remission rate of between 53 and 58 percent and a median duration of remission (i.e., time to relapse) of one year using the standard therapy.

5.1.2 Study Results: A Lack of Replication

Lederle’s Dr. Z. Arlin presented the results of the first of the two studies included in the NDA. This trial, referred to as study 374, was conducted entirely in the US with 216 patients enrolled through 25 medical centres across the country. According to Arlin, for the 98 patients ultimately judged eligible for efficacy analysis, there were 63% complete remissions on the treatment arm containing mitoxantrone and 58% complete remissions on the daunorubicin treatment arm with a ‘hazard ratio’ of 1.14. The hazard ratio is a comparison of the ‘hazard rates’ of the treatment and control groups. The hazard rate is a probabilistic time-to-event calculation. Traditionally, the hazard rate is calculated as the probability that a patient who has survived up to a certain time will fail within the next time interval — hence the term ‘hazard’ rate. However, the same calculation could be made using an event such as time-to-remission.⁶ The ratio is usually expressed with the hazard rate of the investigational agent in the numerator and that of

⁵ According to Streiner (2007) equivalence studies should have at least 229 subjects *per treatment arm* – twice the size of the studies presented in this case. To be fair, Streiner’s paper was written 20 years after this advisory committee meeting and, although statisticians were present at the time, nowhere in this advisory transcript do we see significant discussion of the sample size. Still, looking at these studies in the retrospective light provided by Streiner, we can better understand some of the frustrations the committee had in interpreting the data in these studies.

⁶ Spruance et. al. 2004 express the ratio as a time-to-cure while Box-Steffensmeier 2004 expresses it in the original sense of a time-to-failure. Spruance et. al. (2004) is highly useful for understanding the probabilistic reasoning behind the ratio, however in reading this paper it is also important to remember that the authors are discussing curable diseases. In cancer clinical trials, time-to-failure is often a more useful measure. Also, the authors of the paper are discussing placebo controlled trials, not positive controls. In their examples, the control cannot be more efficacious or more toxic than the investigational agent, as can occur in positive control studies.

the control in the denominator (see, e.g., Box-Steffensmeier 2004; Spruance et. al. 2004). However in the Lederle presentation, the ratio was consistently expressed as ‘C/N’ (Cerubidine-to-Novantrone, or daunorubicin-to-mitoxantrone), putting the control in the numerator (e.g. FDA 1987d, 18). So in this case a hazard ratio greater than one would mean that the standard therapy patients were at a greater risk of failure than the mitoxantrone patients, although the result was not statistically significant. The patients on mitoxantrone also achieved remission somewhat more quickly than those on daunorubicin (35 vs. 43 days), although not to a statistically significant degree. The duration of remission and survival (with a median follow-up time of approximately 14 months) were comparable for both study arms. The toxicities experienced by the patients on the two arms of the study appeared to be comparable as well. The 95% confidence limits for the hazard ratio were reported to be 0.80 and 1.62. This result means that 95% of the time, the hazard ratio would be expected to fall within these limits.⁷ Since this study was designed to show ‘equivalence’, the lower confidence limit of 0.8 seemed reasonable to the committee given the apparently reliable data and overall lack of statistically significant difference between the two arms of the study. In this case, the advisory committee was satisfied that within reasonable limits, the clinical trial appeared to demonstrate equivalence.

Lederle’s Dr. Saletan presented the results of the second study to the committee, study 3-603, an international trial composed of patients from Europe, Latin America, and the Pacific. In this clinical trial, which employed the same treatment regimen as the U.S. study, 116 patients were randomized to the investigational (mitoxantrone) arm of the study while 123 patients were randomized to the control (daunorubicin) arm. Fifty percent of the mitoxantrone patients experienced complete remissions, as did 51% of the daunorubicin patients. While these response rates were lower than those found in the American study, they were still considered by the committee to be reasonably comparable. What was more notable was the difference in hazard ratio: in this international study, the hazard ratio was 0.84 with a 95% confidence interval of between 0.6 and 1.16. In other words, in this study the patients appeared to have rather less chance of success on the investigational agent than on the standard therapy.

⁷ Or to put it in more precise statistical terms: if one repeatedly draws sample groups (e.g. clinical study groups) from an overall sample population (e.g., adults with ANLL), then the hazard ratio for 95% of those groups would be expected to fall within those confidence limits.

Why the difference between the two studies? Notably, in the international study a number of patients on the investigational therapy died not of leukaemia, but of adverse effects of the drugs. Both treatment arms in this study were highly myelosuppressive, leaving the patient at risk for infection and haemorrhage. However in the international study, 14% (eight out of 56) of the patients who had complete responses died during consolidation therapy (55) — meaning that they died *after* having gone into remission. Two of these deaths were due to haemorrhage, while the remainder were due to sepsis. In the U.S. study, by contrast, there was one sepsis-related death on daunorubicin and two on mitoxantrone. The drug sponsor attributed this contrast at least in part to a difference in supportive care. Since the treatment can temporarily devastate a patient's ability for natural immune response, aggressive antibiotic treatment is considered an important part of supportive care at a certain stage of therapy. The adequacy of such care in practice varies. While in the U.S. and Europe, patients on study received what the sponsor considered 'inadequate antibiotic coverage' approximately 25% of the time, in Latin America (where five of the eight deaths took place) the proportion of patients receiving inadequate antibiotic coverage approached 100% (54). Thus, in consolidation therapy, just after remission had been induced when patients were at their most vulnerable to infection, the lack of proper antibiotic care became apparent.

There were other problems with the international study besides. According to the FDA's Dr. Burke, there were irregularities in patient randomization in the Hong Kong and Taiwan study centres, potentially biasing 13% of the sample. Moreover, Burke reminded committee members of the positive control design of this study. The sponsor provided historical data on the activity of the combination of Ara-C with daunorubicin. However, that data said nothing about the individual contribution daunorubicin made to the overall therapy. The assumption of the study design was that daunorubicin did indeed contribute to the combination and that if a similar response rate were obtained by substituting mitoxantrone for daunorubicin, then the former must be making a similar contribution as the latter agent. However, Burke reminded the committee, the somewhat inferior hazard ratio of 0.84 with a weak lower confidence limit of 0.6 made quantitative comparisons more difficult.

Hence, strictly speaking, the two studies presented in this NDA did not provide scientific replication of the results. The committee's choices appeared to be to reject the application due to this lack of replication, ignore one of the studies, or find a way to

adjudicate the differences between the two. As we will see in the next section, their solution effectively combined elements of the latter two options.

5.1.3 *A Judgment of 'Guilty with Explanation'*

In addition to the apparent lack of replication, one committee member, Dr. Charles Moertel of the Mayo Clinic, was vexed by what he considered the immaturity of the data. What happens to patients in terms of long-term survival? The median follow-up time was approximately 14 months. Would the two- or three-year survival rate for the mitoxantrone group drop off relative to that of the doxorubicin group? There was no way to know from these studies. When asked what 'survival' really meant for these patients, Arlin told Moertel that no more than 25% of patients live beyond two or three years (FDA 1987d, 63)⁸ and suggested that both the literature and his experience with leukaemia would indicate that both survival curves would 'continue to drift downwards' (61); i.e. that patients on both the investigational and treatment arms would be unlikely to live beyond two to three years. Moertel was unsatisfied and pressed the point, asking, 'what assurance do we have that Novantrone [mitoxantrone] won't do worse?' (64). For Moertel, more time was needed to make the data meaningful. He noted repeatedly that one study was initiated in 1985 with only a year-and-a-half between patient enrolment and data reporting (see pp. 60, 66, 67-8) — a duration insufficient to show long-term survival effects, even with the somewhat feeble definition of long-term survival as two or three years. 'I am certain that there have been very few drugs ever brought to NDA, that have gone through this quickly, with a study just started in '85, and, you know, just a very short period of follow, where suddenly it is on the market. I am anxious about this short-term follow-up' (67-8). In response to this assertion, one of his colleagues on the committee, Dr. Canellos, said, 'What about AZT, Charlie? That made it very quickly — the AIDS drug. It made it very quickly' (68). 'Yes', Moertel replied, 'but I think, other than scientific issues might have been involved in that one' (68).⁹

This lack of follow-up was compounded by the weakness of the international study combined with its design as an equivalence trial. As Fleming argued, 'the way I review the results now, based on what information I have on the active control, is that the

⁸ Later in the discussion, this figure dropped to 10-15% (78).

⁹ Note that this is the same Dr. Moertel who submitted comments on Subpart E to the FDA arguing that the AZT trial was tantamount to a Phase III trial (Chapter 4). Unfortunately, Dr. Moertel passed away in 1994 and therefore cannot be asked for clarification of his attitude towards the AZT trial.

first [U.S.] study really is a positive study. The second [international] study is not; and if that is the consensus, then the question to the FDA, or to my colleagues here, is, is ANLL a rare enough disease that one positive study is enough for approval?' (90). Fleming expressed the additional concern that the toxicity differences between the investigational and control agents evidenced in the international study could continue to become more pronounced with continued follow-up. Indeed, the international study results 'are already sufficiently negative, in the sense that the lower bound of the confidence interval is already .6 . . . if we did obtain the additional data that Dr. Moertel is talking about, we could go from a position of saying that it is a concern, to it being definitively negative' (103). Moertel agreed with Fleming's negative view of the international trial (95), adding that the equivalence study design is a 'compromise with practicality' which 'creates a lot of problems for committees like this' (96).

Given the uncertainty of the data, the reliance on response data rather than survival, and the existence of accepted therapies, Moertel questioned the wisdom of approving the mitoxantrone regimen. 'I find it very difficult to see what need this [drug] serves to the American public. Daunorubicin is already out there. So is 6-thioguanine. So are a number of other drugs. What does this particular drug serve to the American public, so that it should be pushed forward, on the basis of equivalency, and on data that really seems, at this point, immature?' (96). Indeed, he declared, 'without evidence of equivalency in long-term survival, and without this drug meeting some great need, I am very reluctant to, you know, go down in favor of this' (97).

Dr. Robert Bast, then director of Duke University's Comprehensive Cancer Center, countered by suggesting, in effect, that the committee should ignore the international study. Noting that that in the U.S. trial, seventy percent of the patients studied had already passed away, he said, 'we are looking at the last 20th percentile in that study, which is mature, and which I think, by consensus, is the one study on which we are probably basing equivalency at the present time' (99). However, even the U.S.-based trial gave some observers pause because of the equivalency design. In a later discussion, Dr. Albert Bernath, Director of Hematology at Geisinger Medical Center (Pennsylvania) expressed concern over how the practicing oncologist would choose between mitoxantrone and other agents for use with Ara-C, saying that it is a 'fundamental question about what equivalency means. And as Dr. Moertel says, having another agent out there, without superiority, with roughly equal toxicity, at least in the American trial,

and with no data on survival, does give one concern' about choosing a drug 'based on a guess' of equivalency (119). He urged caution in assuming 'where survival curves will go, just because they have gone some way in the past' (119).

Seeing that the committee was weighing issues concerning approval on the basis of the one 'good' study, the FDA's Dr. Temple intervened in the discussion to clarify the applicable FDA policies and regulations. Temple pointed out that while the FDA had previously accepted a 'single persuasive study' as the basis for approving a new drug, the rarity of the underlying disease to be treated was not usually the basis of the decision. Rather, he said, 'the circumstances that have forced that conclusion was, where the results were dramatic, usually showing some effect on survival, and it was felt that it was not reasonable to try to — or ethical to try to randomize patients back to a confirmatory trial' (103-4). Even so, he felt that it was 'not a good idea' to accept a single trial in cancer studies 'because of the uncertainty attached to those trials'; 'one wants the usual scientific standard of confirmation' (104). Even so, he said, 'I don't think that necessarily implies that the confirmatory study has to be as persuasive as the other one. There is room for some flexibility' (104). So the question amounted to how persuasive as confirmatory the committee found the international study to be (104) — an assessment Temple colourfully called 'guilty with an explanation' (105), meaning that if the committee found the results of the second study to be reasonably supportive of the first study, a vote for approval was appropriate.

Temple also clarified the FDA's view of equivalence studies, noting that the FDA does not — and cannot — require a demonstration of superiority for approval. The question is not whether the investigational therapy shows an advantage, but 'whether there might be a hidden disadvantage' (120). When a well-accepted standard therapy already exists, 'one wants a reasonably high level of assurance that you are not going to do worse' with the investigational therapy. According to Temple, it is therefore fair to ask drug companies to present 'two reasonable studies' to build confidence that there are no hidden problems. 'I'm not advocating a demonstration of superiority. I don't think the law allows us to advocate that, and I don't think that is appropriate. But one wants a reasonable level of assurance. And I think that's what we're seeking' (120-121).

With Temple's clarifying comments, nine out of nine committee members voted to find the second trial confirmatory of the first (121), and seven voted (with two abstentions) to approve the drug for use with Ara-C in first-line treatment of ANLL (122).

Formal FDA approval came a little more than two weeks later (on 23 December), with the approval ‘contingent’ on a commitment by the drug sponsor, Lederle Labs, to continue to follow-up on patients in both clinical trials and to submit yearly updates of survival data as well as data regarding possible chronic toxicities.¹⁰

5.1.4 Discussion

The approval of mitoxantrone illustrates some key issues for drug approval, then and now. The ambiguity of the equivalence study design left the advisory committee with a difficult interpretive task. This task was made more difficult by the historical and ongoing inclination towards accelerating drug development through the abbreviation of clinical studies. As we have seen in the case of AZT, a year-and-a-half for clinical testing of a drug was considered light speed. Moertel’s scepticism notwithstanding, clearly at least one other advisory committee member was aware of what had just taken place with AZT and was willing to offer it as a counterexample to Moertel’s insistence that drugs never get tested and approved so quickly. Of course, in the case of AZT survival was the clinical trial endpoint. In ANLL, however, the shorter follow-up time meant that survival could not be the basis of decision-making for evaluating the effectiveness of the drug; rather response rates would be used. As we have seen, a complete response in this disease did not mean that patients were cured. With proper supportive care, 60 percent of the patients would achieve a complete response and then most of those would relapse, so that overall only 15 to 25 percent of patients could be expected to live between two and three years — and even fewer beyond that. More than that, many on the committee questioned whether toxicity differences in the two drugs would become more pronounced in follow-up, or whether the survival curves for patients in the two treatment arms would diverge over time. As Moertel repeatedly noted, there was no way to know if the long-term survival benefit of one therapeutic regimen exceeded that of the other.

This is a fundamental quandary of using surrogate rather than clinical endpoints as a basis of measuring drug effectiveness in clinical trials. The most desired indicator of patient benefit in life-threatening diseases, survival, usually takes a number of years to measure. The AZT trial was somewhat unusual in that even after a relatively short follow-up, the drug’s effect on survival was pronounced. More often with serious disease, if one

¹⁰ Source of this information is the FDA’s online compilation of data on approved oncology drugs, <http://www.fda.gov/cder/cancer/druglistframe.htm>.

wishes to accelerate drug development, one must find other measurements thought to be meaningful indicators of patient condition and long-term benefit. However, the degree to which these alternative measurements for various diseases are meaningful varies greatly; indeed, the degree to which alternative measurements are even available varies. As we will see in the next chapter, after the initial approval of AZT, a chorus of voices joined together calling for the evaluation of drugs on the basis of measures other than survival at a time when meaningful surrogate measures for AIDS patient response to therapies were not yet well developed. With this approval of mitoxantrone, we can see that the question of whether to wait for survival data or to move forward with more rapid measures of patient response was an issue in cancer clinical trials as well as AIDS.

We can also see in this case one of the FDA's strategies for managing the uncertainties stemming from approval on the basis of data other than survival: require additional data-gathering *following* the approval of the drug. We saw in Chapter 3 a number of proposals which would have legislated postmarket study as a condition of earlier approval. Now here, as for AZT, the Agency had no legally binding authority to hold the drug sponsor to their portion of the agreement for postmarket study. Once a drug was on the market, the FDA could not demand a withdrawal unless there was evidence to suggest that the drug posed a heretofore unrecognized hazard to patients. Nevertheless, when making a decision such as this one, the FDA would seek agreements for additional data-gathering, relying on the professionalism of the drug sponsor (and, perhaps more realistically, on the desire of the sponsor to maintain sanguine relations with the FDA for the sake of future drug development programmes) to carry out its part of the agreement.

Clearly the practice of seeking postmarket study commitments *preceded* the Subpart E rulemaking under which FDA authority to require such studies began to be articulated in the regulations (Chapter 4). Likewise, the procedure used for this 1987 drug approval — approving on the basis of a surrogate endpoint and then continuing study to confirm the clinical validity of that result — is a practice which will be codified *five years later* in the Subpart H rules for accelerated approval (Chapter 6). Furthermore, this drug approval decision anticipated a law which would not be passed for another decade, the *Food and Drug Administration Modernization Act of 1997* (FDAMA). The FDAMA modified the substantial evidence clause to authorize the FDA to admit as evidence for approval a

single study plus confirmatory data¹¹ — and, as well will see in Chapter 8, this standard of evidence applies not only to especially promising or clinically meaningful drugs, but in principle to any new drug.

The acceptance of this therapy on the basis of ‘guilty with an explanation’ is particularly noteworthy since, as Moertel and others pointed out, there is no unmet medical need here as there was for AZT. As we have seen, the ODAC’s repeated mention of considerations related to the rarity of the underlying disease or the availability of other therapies motivated Dr. Temple to insist that acceptance of single-study data was related to the *quality* of the data, rather than to these other circumstances (a position we will see him modify for another NDA later in this chapter). Nevertheless, these types of considerations were clearly also in play for the AZT decision on which the Subpart E rules were modelled. It seems that the rarity of the underlying disease and the scarcity (or not) of alternative therapies can count in favour of a drug (on the benefit side of the equation), but not necessarily against it. In this case, although other comparatively suitable therapies already existed, this drug was approved on the basis of a single adequate study with a second trial having weakly *unfavourable* results but nevertheless counted as confirmatory.

5.2 Ifosfamide: Constitution of a Single-Study Drug Approval

5.2.1 *A Pivotal Trial by Serendipity*

Four months after the mitoxantrone approval, on 19 April 1988, the Oncologic Drugs Advisory Committee met to consider an NDA for use of the drug ifosfamide in combination with other, already marketed oncology drugs in the third-line treatment of refractory testicular cancer (FDA 1988b). Significantly, while the sponsor of the new drug application was Bristol-Myers, the presentation of the ifosfamide NDA data was not made by a company representative but by the academic researcher who developed and tested the investigational regimen, Dr. Lawrence H. Einhorn of the University of Indiana.

¹¹ In this case the degree to which the international study was actually confirmatory is debatable. Some of the committee members felt the international study to be so inferior that it could not even be considered as providing support. When asked to weigh in on the question of ‘guilty with explanation’, Dr. Elaine M. Smith commented that while she accepted that there were ‘probably’ equal response rates between the two therapeutic regimens, ‘[d]efinitely, no superiority is shown in the new drug, and I strongly do not understand why the supportive care issue seemed to select out mitoxantrone.... Specifically, I accept one study and not the other’ (111).

Einhorn was well known to the committee as the researcher who in 1974 developed the first chemotherapeutic regimen effective against testicular cancer. Prior to 1974, only five to ten percent of patients with this disease could expect to be cured with the standard therapy. Einhorn's innovation of using a platinum-based therapy, cisplatin plus vinblastine plus bleomycin (PVB), increased that cure rate to 57 percent (Einhorn 1990). Subsequent refinements improved the cure rate still further. In 1978, Einhorn achieved another first, using cisplatin plus etoposide (VP-16) as a salvage therapy for patients not cured by first-line therapy. Twenty-five percent of these patients were thus cured — an achievement which represented the first time an adult solid tumour had been cured with a second-line regimen. Einhorn continued to refine the therapeutic regimens for testicular cancer while also turning his attention to patients who had failed both the first-line and second-line therapies. In 1983, he began studying the use of ifosfamide with other agents in the third-line setting. The results of the latter study were the basis of this new NDA. Hence, Bristol-Myers had as an advocate for the product the oncologist who the previous year had been named Distinguished Professor of Medicine at the University of Indiana, in no small part because of his reputation as the man who had virtually *cured* testicular cancer.¹²

Einhorn began his presentation by describing the disease, which was rare but amenable to therapy by surgery and by platinum-based chemotherapy. Of the 5,500 cases of this disease in the U.S. a year, 1,700 would be treated by chemotherapy. Of those 1,700 cases, 70% would be cured (this was Einhorn's word) in first-line therapy while an additional 10% would be cured by second-line therapy, leaving only about 340 patients a year who would require third-line treatment. Hence, third-line treatment was an especially rare clinical circumstance. This third-line patient group was said to possess the most recalcitrant forms of the disease. Also, this highly pre-treated group of patients (some of them were in fourth- or fifth-line treatment) was likely to have developed some resistance to platinum-based therapies and also likely to be more susceptible to the adverse effects of any further treatment. According to Einhorn, re-treatment of these patients with already-administered therapies consistently failed. For these reasons, the University of Indiana

¹² In the 1990s, Einhorn would become famous as the oncologist who cured the professional cyclist Lance Armstrong. Einhorn subsequently received a \$1.5 million endowment to be the Lance Armstrong Foundation Chair in Oncology at Indiana University. See http://www.walther.org/wcf_blogs/WCFNewsReleases/Walther_Research_Veteran_Dr._Lawrence_Einhorn_Aw.html (accessed 18 July 2007).

had traditionally moved third-line patients to Phase II investigational agents: there were no therapeutic options available after second-line therapy.

Initially, therefore, the activity of ifosfamide in testicular cancer was observed when it was a Phase II drug being evaluated for single-agent activity. As Einhorn reported, ‘We first did single agent ifosfamide studies, like we do any Phase II agent, and, quite frankly, most of our Phase II trials are zero for 14 and we move on to the next Phase II drug’ (37). However ifosfamide exhibited clinical activity so Einhorn’s group expanded the Phase II trial. Einhorn’s team initiated a six-arm study in which ifosfamide was tested in a variety of chemical combinations on groups of between three and five patients (58). Finding most of these combinations to be ineffective in the third-line setting, Einhorn dropped the other arms of the study and continued accruing patients to the single remaining treatment group. This group was what would become the uncontrolled, unrandomized study of 59 patients submitted to the FDA as the single pivotal study serving as the basis of this NDA.

The patients in this study were treated with a combination of platinum (cisplatinum), ifosfamide, and either vinblastine (Velban) or VP-16 in a regimen referred to as ‘VIP’. Which ‘V’ was used (vinblastine or VP-16) depended on which regimens the patient had previously received. In his presentation, Einhorn reported that 15 patients (25%) achieved a complete response (remission) while another eight patients achieved disease-free status with VIP followed by surgery. Overall, 23 patients (39%) achieved disease-free status. An additional nine patients (15%) had partial remissions which lasted a median of six months with median survival time of one year. Twenty-seven patients failed to achieve an objective response. For these non-responding patients, the median survival was eight months. After a follow-up of two years (follow-up ranged from two years to 244 weeks), 13 of 59 patients (22 percent) were disease free. According to Einhorn, in this form of cancer ‘a two-year disease free survival is tantamount to a five-year disease free survival in any other solid tumor as relapses beyond two years are purely anecdotal’ (14) and even relapses after only one year of being disease-free are ‘exceedingly rare’ (15). If a relapse was to happen, said Einhorn, they would see it quickly. For this reason, Einhorn referred to these patients who were disease-free for more than two years as ‘cured’.

This was not a controlled study. Therefore, at the request of the FDA, Bristol Myers and Einhorn’s group had searched for an existing study to act as an historical

control for this data. Einhorn's University of Indiana colleague Dr. Rozenzweig admitted that the historical control had 'clear limitations', but that for what it was worth, they had 'performed subset analyses which largely confirm the superiority of the ifosfamide regimens' (28). Given the weakness of the historical control, Einhorn and Rozenzweig suggested that the patients effectively served as their own controls (29). Einhorn argued that 50 of the 59 patients had previously received as second-line therapy the 'V' and 'P' components of the VIP therapy — i.e., VIP without ifosfamide. Since these patients had already failed the 'VP' combination when they responded to VIP, they served as their own controls to demonstrate the effectiveness of adding ifosfamide to the combination.

5.2.2 *Committee Discussion: Students Assessing the Teacher*

In FDA advisory committee meetings to consider an NDA, a question-and-answer period follows the sponsor's presentation, which is then followed by the FDA's analysis of the data. While advisory committees are typically respectful of the sponsor's representatives during the question period, even when challenging the data validity or the company's interpretation of it, in this meeting the tenor of the questioning seemed unusually respectful. As a recognized authority on testicular cancer, Einhorn frequently offered his personal experience as an authoritative source of context for the data, and the Committee tended to accept him at his word. In some cases, they directly asked him for such opinions, often flattering Einhorn in the process. For example, in an attempt to understand better the interpretive context of the data, Dr. Robert Capezzi, acting chairman of the committee, asked if 'in your experience, Larry, at the University of Indiana, which is certainly one of the more extensive ones in the nation, as well as your literature experience, the durability of response in this clinical setting is something that is unique' (FDA 1988b, 78). (The answer was yes.) Even Dr. Moertel, a highly opinionated and trenchant advocate for methodological rigor, expressed his qualms in a way which demonstrated deference to Einhorn. Moertel was concerned about releasing the drug in Phase II with so little data supporting the decision. This was especially a concern since the drug was 'extraordinarily toxic, not gentle toxic' (39), requiring specialized support for administration, although testicular cancer was something most oncologists would have had little experience treating. To pose his question, Moertel said, 'I know you are here as an advocate, in a sense, but I don't feel that strong an advocate', and asked Einhorn to 'philosophize about development of a drug in this particular setting' and to consider 'what

is the opportune time for a drug like this to be released' (39) (Einhorn thought it should be approved in Phase II).

A notable shift in tenor took place when the FDA's Dr. Sokol gave the FDA presentation (86-99). Sokol's discussion highlighted some of weaknesses one can often expect when a Phase II investigation becomes a pivotal trial for an NDA, not least of which was the retrospective search for an historical control. Sokol enumerated a series of problems with the historical control beyond those admitted by the sponsor. More than that, he did an analysis of the time to next therapy for first-line and second-line testicular cancer patients suggesting, contrary to Einhorn's claim, that relapses *do* take place after two years and casting doubt on the suggestion that Einhorn's disease-free patients are 'cured'. Through a detailed analysis of the individual cases in the study, Sokol suggested that some of the patients with complete responses may have had adjuvant chemotherapy, i.e., chemotherapy as a follow-up to surgery rather than as a standalone treatment. He also identified patients whose tumour masses did not appear to be inoperable as Einhorn had claimed, commenting, 'Now, I don't know about Indiana University but I know that ten years of practice in our particular facility has resulted in a far more aggressive surgical service when metastatic disease is addressed.' (95). Consequently in some cases it was difficult to tell if remissions were due to ifosfamide or due to surgery. Moreover, not all patients who had surgery received the same follow-up, or 'adjuvant', chemotherapy (98). Sokol also demonstrated from the case reports that nine patients were given salvage therapy *after* VIP using what was then the first-line therapy, PVB, despite Einhorn's claim that PVB is not effective and therefore not used in third-line treatment. Sokol commented that 'we heard today, at least anecdotally, that none of those patients responded but, again, somebody believed that there was some retained activity of PVB and it would be very interesting to actually look at those patients and see how they responded' (97). In addition, there were data omissions that made interpretation more difficult. For instance, none of the data from the other discontinued arms of the original six-arm study was submitted with the NDA. Sokol also noted that, although Einhorn had referred to these patients as the 'bad players' in terms of their disease (71), there seemed to be a trend that patients who are 'born to be complete responders will continue to complete respond' (102). In other words, close inspection of the data revealed five patients listed as first-time complete responders (CRs) but who probably did have previous responses which were 'either short remissions or remissions that were called into

question because of other considerations' (102). This finding would further problematize whether the complete responses identified by Einhorn really were 'cures'.

These and other discrepancies discussed by Dr. Sokol placed the NDA in a more critical light. While committee members were aware of the status of this NDA as a single Phase II study and had asked questions concerning, e.g., whether it would be possible to obtain supporting data from other, ongoing trials (45), Sokol's presentation cast doubt on the validity of individual data points and on specific aspects of Einhorn's presentation, going well beyond the level of critique offered by the Committee to this point. In what followed, we may perhaps suppose that as a newcomer to the FDA, Dr. Sokol was overly enthusiastic in his recitation of the study's flaws. What is clear is that Dr. Einhorn was not accustomed to being contradicted. He responded with an exceptional assault on the credentials and experience of the FDA presenter:

DR. EINHORN: You wax elegantly and eloquently and I don't have as good a vocabulary as you have but I would like to think I have a little more knowledge about testicular cancer from my personal experience You said you came from an institution where surgery was frequently used, may I ask you what that institution was?

DR. SOKOL: Harvard Medical School.

DR. EINHORN: Okay. At Harvard Medical School, in testes cancer I would guess you would probably see -- what? Three patients a year?

DR. SOKOL: That would be an underestimation I believe.

DR. EINHORN: I bet it isn't. (103).

Einhorn and Sokol clashed over the role of surgery and chemotherapy in the treatment of testes cancer, leading to a heated exchange over the number of patients who could actually be classified as 'cured' in the study: Einhorn asked, 'Doesn't it perturb you minimally that we cannot document in the literature any patient in a comparable setting . . . who has ever been a two-year disease-free survivor in that setting?' Sokol retorted that 'if, indeed, the data you are presenting had long-term survivors that numbered more than three or four with the accomplishment of that "long-term cure" without aggressive surgery under very aggressive circumstances, I would say that that is a valid point' (104-5). Einhorn protested that there were 13 patients who were two-year disease-free survivors,

but Sokol pointed out that not all these patients were fully evaluable. Some of those patients got a different 'V' (V-16 or vinblastine) than the one previously administered, so it was not clear whether a response was due to the 'V' or the ifosfamide component of the therapy (104-5).

Sokol eventually sought to summarize the issue in a manner useful to the committee, and in a way which would mollify the ruffled Dr. Einhorn: 'Dr. Moertel, in many of his recitations, has taught me to be very, very wary of anecdotal circumstances. And whether we are talking about three patients or ten patients in the world's literature, the issue that I raised, not arrogantly, but the issue to be considered with sobriety by this Committee is whether that is adequate or not' (106).

While the committee continued to be deferential towards Einhorn, ODAC members responded to the more critical light provided by Sokol's presentation. For one thing, although Einhorn had argued that the patients could serve as their own controls, the committee did not appear to accept this judgement. Second-line therapy is not the same as third-line therapy — the latter was uncharted territory, in which patients had a different disease burden (extent of disease), a greater level of pretreatment, and perhaps other, not-yet-realized differences compared to second-line therapy. Worse still was the historical control which exhibited so many differences from the situation in the investigational study, most of the committee members dismissed it, in the words of Dr. Krown, as 'largely irrelevant' (130). Since the historical control was not adequate, the only other basis for judgment was comparison of the treated group response to what was known about the natural history of the disease (128) — something unknown for the third-line setting without ifosfamide. On this matter, Einhorn had given his opinion, but as Bernath noted, 'once again, we are being told about it without having read about it in the submission. Is that equivalent to anecdotal data?' (121). As for the mitoxantrone case above, it was also difficult to judge whether the follow-up was long enough to declare the disease-free patients as 'cures'. Dr. Bernath asked the committee, 'Are we all in uniform agreement on what the natural history is? We have heard Dr. Sokol rattle off at least eight or nine long-term time to next treatments for this group of patients, even long-term ones after failure of second-line treatment. Does that rattle our confidence somewhat in calling these two-year disease-free survivors cures? Yes, it rattles my confidence somewhat' (121).

Another issue was the single-study status of this NDA. Many committee members expressed a desire for supplemental information, especially considering that there were only 59 patients in this study and a great deal of data was missing. Bernath commented that for this NDA ‘[w]e have to look at individual patients. We have not been able to do that effectively. We don’t have marker data on all these patients. We don’t have complete data on all of the 13 long-term responders. We don’t have accurate data, as we have just heard. Patient number 20 looks one way on paper; sounds another way’ (119-120). More than that, this was a single study from a *single institution*. The chances of a biased data set were high. Dr. Brenner noted that this was especially a concern ‘when the original data come from such a renowned place as Indiana, with such an incredible experience — really the world’s leader in the treatment of the disease. Can these data be duplicated in a cooperative group in other institutions?’ (123-4). Dr. Moertel agreed: ‘It would be very reassuring to me to have this particular drug tested under these circumstances in a broader and more representative selection of oncologists than the very unique Dr. Einhorn’ (133).

The FDA’s Dr. Temple once again weighed in on the question of single study drug applications, giving a notably different account than that offered in the discussion of mitroxitron. In that meeting, he said the justification for single-study acceptance was typically related to convincing data and compelling results. In this meeting, by contrast, he said that the issue ‘comes up for us whenever a drug seems to have life-extending potential. That is probably the only circumstance in which we would willingly agree to accept an unreplicated study. But it has been done in that situation in the cardiovascular area and, to some extent, in the oncology area’ (136). He noted, however, that the study did not stand alone; that there was supporting data from other sources.

Ultimately, the committee members were inclined to look past many of these flaws because of the nature of the disease and the nature of the expertise behind the data. Dr. Capizzi’s comments were typical: ‘We have a relatively rare disease. There are only some 340 patients eligible for such studies in the nation and when is enough enough? Do we need another randomized trial to carry us to the next three, four or five years before we are able to make a decision? I don’t think so’ (135-6). Capizzi also added, ‘I think the data presented by Dr. Einhorn — and I echo every laudatory statement made about him. I too consult him over the phone and he has been very helpful regarding patient problems etc. The data is very compelling’ (136). Dr. Fleming likewise said that although there

were concerns about the application, '[c]learly, Dr. Einhorn's knowledge is impressive. His presentation is influential in trying to establish whether or not the ifosfamide regimen is, in fact, contributing' (129). Perhaps most telling was Dr. Krown's remark:

'I guess when I walked in here this morning I was ready to say something very different from what I am going to say right now. And Bristol should be for ever grateful for having Dr. Einhorn speaking on their behalf because I think what he does is to show us all how important it is to have a clinician who is familiar with his patients and with the natural history of the disease. That sometimes tells us a lot more than things you can put in tables and statistics and percentages (130).

Even Dr. Moertel, while conceding that this was a very difficult decision for him and that he still had reservations about the NDA, said:

It is, I think, one of the occasions in which I do believe in anecdotes. I think the anecdotes that might be told by a thoroughly knowledgeable investigator of high integrity, with unparalleled experience in a disease, frequently can become much more convincing than two randomized, controlled studies conducted in a sloppy fashion with 20 percent ineligibles and 20 percent loss to follow up, and what-not. So I do not necessarily denigrate the anecdotes in these situations. As a matter of fact, I consider the anecdotes very persuasive (132).

5.2.3 *NDA Dependence on an Unapproved Drug*

The disadvantage to this apparently efficacious treatment regimen was that it was severely toxic. While non-hematologic toxicity (especially nausea and vomiting) was contributed by all the components of the regimen, the ifosfamide was solely responsible for a number of highly serious toxic effects, including encephalitis, suppression of white blood cell counts and platelet counts, and a form of urological toxicity called hematuria, the presence of blood in the urine. Einhorn's group at the University of Indiana controlled this latter condition using 'uroprotector' drugs. At first they used a substance called N-acetyl cysteine (NAC), which helped with the condition but induced yet more nausea and vomiting for these already nauseous patients. Then, a new drug called mesna became available through treatment IND. Mesna appeared to be more effective than

NAC at controlling the hematuria without inducing any additional nausea, so Einhorn's group began to use this still-investigational drug to control blood loss of patients on VIP.

The problem was that at the time of this application, mesna was still not available commercially. It could only be obtained through treatment IND, although an NDA for mesna was purportedly forthcoming within a few months. Nevertheless, according to Dr. Einhorn in his presentation, '[i]n the oncology community mesna is very widely considered as the preferred uroprotector' (FDA 1988b, 32). Given its unapproved status, however, Einhorn's Indiana colleague Dr. Rozenzweig tried to downplay the importance of mesna in supportive care by pointing out that with vigorous hydration, most patients can tolerate the VIP therapy (63). However, others like Moertel (63-4) and Sokol (67-8) expressed concern over patient welfare without uroprotection. More than that, Moertel pointed out that no data had been presented to the committee on hydration — that the application was based on the use of mesna to alleviate urotoxicity (63-4).

Given the need for mesna as part of supportive care, and given the lack of commercial availability, Moertel asked what benefit approving ifosfamide would do patients in far flung corners of the country who would be obliged to go through the 'Mickey Mouse' routine of applying for mesna through treatment IND as part of using a legally marketed therapy (83). Bristol's Dr. Gill dismissed the comment as 'philosophy of health care', insisting that the drug had been proven effective and safe, and therefore should be commercially available to all physicians (83-4). Nevertheless, Moertel reminded his colleagues that, once on the open market, ifosfamide would be fair game for off-label prescribing (134). Community oncologists lacking experience with this toxic agent would 'push buttons on PDQs [drug reference guides] and see ifosfamide mentioned in this experimental regimen or that, or maybe read about ifosfamide giving response rates in colon cancer in some obscure medical journal and decide to try it as third or fourth line. That is what happens in oncology practice' (134). He concluded that he 'would like to approve [ifosfamide] six months from now when mesna is available' (135).

Dr. Temple once again weighed in, saying that if the Committee considered ifosfamide approvable and agreed that mesna was a necessary part of the therapy, the drug label could specify that ifosfamide should be administered with an investigational agent obtained under treatment IND (137-8). He acknowledged that such labelling was unusual, but said it was not entirely unprecedented (138). Moertel reminded Temple that ifosfamide could be used for more than this single indication and demand could easily

outstrip the limited supply of mesna available for use under treatment IND (139). The mesna NDA was expected soon, however, and Temple assured Moertel and the committee that there was always a ‘take-off period’ for drugs (140) — a lag time between approval of a drug and widespread awareness and use of the new drug. More than that, with sufficiently strong wording in the label for ifosfamide, he reasoned, those looking to use ifosfamide on an off-label basis might be deterred until mesna was widely available (140).

With this assurance, the committee voted to approve ifosfamide contingent on verification of certain data to be pursued by the FDA, including further examination of the patient characteristics of the historical control group, and confirmation that ifosfamide appeared to contribute to the complete responses seen in the Phase II study. Approval was also contingent on the label specification of concomitant uroprotection to be used with treatment. Dr. Capizzi, the acting chair of the committee, jokingly referred to mesna, saying, ‘I suspect we have had a historic occasion this morning. We have already approved yet another drug that hasn’t even come before the Committee’ (140). Ifosfamide and mesna were both approved eight months later on 30 December 1988.

5.2.4 Discussion

This single-study drug approval went forward two months after the publication of the Subpart E interim rule in which, as we have seen, the FDA refused to codify single-study drug approval as a way to accelerate drug development and approval. Instead, the FDA characterized single-study approval as an exceptional circumstance, acceptable only when there is compelling evidence. AZT was held up as an example in which a single, large, multi-centred, placebo-controlled trial produced adequate, well-controlled and dramatic evidence of effectiveness. As we have seen, that principle was re-stated by Dr. Temple in the 1987 ODAC meeting to discuss mitoxantrone: compelling evidence should be the spur to acceptance of single studies for drug approval, not the rarity of the disease to be treated, nor lack of other therapeutic options.

In the case of ifosfamide, however, we have the antithesis of the kind of single study exemplified by AZT. This was a very small, uncontrolled and non-randomized Phase II trial conducted at a single institution well known for having special expertise in

testicular cancer and an especially aggressive approach to it.¹³ In no small way, the rarity of the disease and the associated difficulty of mustering a second, confirmatory study was a central consideration in the vote to approve this drug. The committee members did ultimately feel that the evidence presented in the study pointed towards the drug's effectiveness, but for many the evidence was not particularly compelling — especially not after hearing Dr. Sokol's analysis of the data.

Clearly, another substantial contributing factor to the committee's positive view of this drug had to do with the esteem they accorded Dr. Einhorn. I have quoted ODAC members at some length to demonstrate the high regard they held for Dr. Einhorn, and to establish his high-level status in the oncology world. More than that, in these quotations and others, ODAC members made it clear that, even with the faults described in some detail by Dr. Sokol, some of which contradicted Einhorn's presentation, they nevertheless found the data more convincing because of the extensive knowledge and reputation of the man who did the study and presented the data to them.

Importantly, combined with the persuasive power of expertise is the role of anecdote. Physicians are traditionally trained through case reports — the detailed recording and recounting of the examination, diagnosis, and treatment of individual patients (Rothman 1991). Obviously, this case-oriented method of teaching is explicitly a system of ostension — learning by example — and as such depends on the authority of the teacher for its validity (Barnes 1982). While all learning ultimately rests on authority to pass information to the next generation of students, in medicine the connection between authority and the case report is especially notable, since this didactic method constitutes such a prominent and rigorous form of training for young doctors. Hence, in medicine, despite many years of reform towards 'scientific' decision-making (Marks 1997; Timmermans and Berg 2003), the anecdote can clearly still be highly persuasive in certain contexts, and especially when the anecdote comes from a source of acknowledged expertise. In this case, even Charles Moertel — widely known as a staunch advocate for scientific rigor in clinical study — was willing to accept 'anecdotes that might be told by a thoroughly knowledgeable investigator of high integrity, with unparalleled experience in a disease' (132). This is not to denigrate expertise nor to stand in judgement of the practice of accepting the experience of a trusted source as a form of evidence. Indeed, as is clear

¹³ Another quotation from Dr. Bernath: 'This institution, Indiana University, has a reputation for non-wimp therapy' (120).

from Ravetz (1979), with experience comes tacit knowledge gleaned in practice and transmitted primarily by example. The committee's acceptance of Einhorn's opinions and anecdotes represents an acknowledgement of Einhorn's tacit knowledge. It also signals a willingness on the part of the committee, in effect, to assume temporarily the role of students — a politically awkward position for a group established to stand in judgement of the teacher's data. This latter fact may go some way towards explaining the acceptance of this data as sufficient to support a new drug application. Another factor may simply have been the frequent use of the word 'cure' in Einhorn's presentation, an unusual and startling word in a medical world where phrases like 'time to treatment failure' or, at best, 'median survival time' are much more commonly heard. There can be no doubt, however, that the esteem of the researcher was an important lens through which the data were viewed. Even in an arena circumscribed by law, regulations, and standards of evidence for approval (including a regulation specifying the use of randomized controlled trials as the basis for approval), anecdote can assume larger importance and validity depending on who is telling the story. Essentially, the committee accepted Einhorn's description of the natural history of the disease and how the drug intervened to modify it.

Another aspect of this approval worth consideration is the unanticipated consequences of the newly promulgated treatment IND rules, which were published only the year before this ODAC meeting. It is unlikely that anyone involved with rule-writing would have consciously contemplated a scenario in which a treatment IND drug could become the industry standard treatment for a particular condition before it was even approved for commercial distribution. Yet here we have mesna, with a treatment IND granted in December 1987,¹⁴ only five months later referred to as the established standard of supportive care to control urological adverse effects. Clearly mesna must have been available as one of the 'exceptions' prior to the promulgation of the treatment IND rule. Indeed, a review of the medical literature shows that mesna was described as an effective uroprotector for use with ifosfamide in the German literature in 1979 (Brock, Stekar, and Pohl 1979) and in English the following year by one of the same authors (Brock 1980). While it is difficult to determine exactly when mesna began to be distributed informally, it had clearly been used supportively for some years without approval for commercial distribution, becoming the standard for urological protection in cancer chemotherapy in

¹⁴ See the treatment IND list at <http://www.fda.gov/bbs/topics/ANSWERS/ANS00094.html> (accessed 2 Feb 2008).

the process. This fact could have served as a potential stumbling block for the approval of ifosfamide. Instead, the approval of ifosfamide virtually guaranteed the simultaneous approval of mesna. What if there turned out to be problems with the mesna drug application? Obviously the FDA had some confidence, perhaps from discussion with the sponsor, that the application would be acceptable — but small problems with the application would no doubt be overcome simply by virtue of the fact that mesna was already established as the standard for uroprotection in practice. Practical experience (and the consensus arising from it — another type of de facto rule) would provide support for the NDA, regardless of technical details.

According to materials posted at the FDA website,¹⁵ mesna was approved on the basis of two pivotal studies: a multicentre Phase III placebo-controlled study of 101 lung cancer patients receiving ifosfamide, and a Phase II study of 16 patients on ifosfamide, nine of whom received mesna for uroprotection while the remainder received conventional therapy. The second study was small enough that we could reasonably consider this another case of ‘single study plus confirmatory evidence’.¹⁶ A number of uncontrolled studies were also submitted as supportive evidence. Unfortunately, we do not have the benefit of an ODAC discussion to understand their perspective of the case. The FDA apparently felt the approval decision to be uncontroversial and did not convene an advisory committee to discuss the NDA.

The ifosfamide case also illustrates many of the problems associated with using Phase II cancer studies as pivotal studies — a theme that will arise again in this thesis. We have seen that for cancer drugs, Phase II is traditionally a time for exploring the activity of a drug in various tumour types and various chemical combinations. Such exploration often implies less rigor than would normally be required for an NDA, and more flexibility. Indeed, this ‘study’ began with six treatment arms, each composed of only a few patients, none of which were control groups, and five of which were dropped when the investigators saw no positive results. Judging from the advisory committee transcript, it is doubtful that the decision point for discontinuing these study arms was based on any prospectively designed protocol. Rather, as an unblinded study, the investigators simply watched patient responses to the treatment and made a judgment that to continue those

¹⁵ <http://www.fda.gov/cder/cancer/druglistframe.htm> (accessed 2 Feb 2008).

¹⁶ Indeed, the Phase III study is quite small for Phase III. Without more detailed information on the study supporting the application, it would be difficult to know what characteristics of the study qualified it to be considered Phase III.

treatment arms would be fruitless. Indeed, it seems clear that the purpose here was as much therapeutic as it was investigational. Patients who had failed second-line therapy were given Phase II investigational agents, and investigators varied the salvage therapy beyond third-line treatment according to an individualized judgment of what might work best for certain patients. The lack of a rigid protocol no doubt facilitated this investigation-as-therapy approach. However, as laudable as the goals of the institution might be, such an approach tends to allow many ‘exceptions’ to the investigational procedures — exceptions which, when a study undergoes the category shift to a pivotal study, become ‘protocol violations’. Such studies are also often too small in size to produce statistically significant data for a pivotal study. Here, nevertheless, a decision was made for third-line treatment on the basis of, at most, thirteen patients who had complete responses.

5.3 Fludara: Retrospective Data Mining and the Meaning of Clinical Benefit

5.3.1 *A Retrospectively Assembled NDA*

In September of 1990, the ODAC met to discuss the NDA for Fludara (generic name fludarabine) for the palliative treatment of chronic lymphocytic leukaemia (CLL) in patients refractory to other treatments (FDA 1990d). CLL tends to be ‘an indolent disorder of elderly patients’ (9). However, it is biologically heterogeneous such that some patients who get the disease can be expected to live a normal lifespan, while others die within months of being diagnosed. For patients with more threatening forms of the disease, biologically active agents had begun to be developed and by this time there were a number of drugs which could produce responses in at-risk patients.

This NDA was sponsored by Triton Biosciences, however much of the development work appears to have been performed or sponsored by the NCI.¹⁷ In 1980, laboratory researchers at the NCI found that fludarabine was active against certain strains of leukaemia. On the strength of this laboratory finding, the NCI sponsored Phase I and II trials carried out at various academic institutions in the early 1980s. These trials

¹⁷ Indeed, in the ODAC meeting under discussion, the presentation on fludarabine’s background and early clinical development was given by Dr. Bruce Cheson of the NCI. According to Angell (2004), drug companies have been able to make arrangements to patent the products of government research since the passage of the 1980 *Bayh-Dole Act*.

appeared to support the conclusion that fludarabine was active in CLL patients who had become refractory to the standard therapies. At this point, apparently, Triton became involved and made licensing arrangements to complete the drug's development and bring it to market. According to the company's Dr. Jeffrey Latts, Triton approached the FDA about the promising data coming from the Phase I and II studies, and whether they would support an NDA. Notably, the company was not proposing to conduct any new trials, but wished to base an application on NCI's existing studies.

A chief difficulty in interpreting the clinical trials performed up to this time, however, was that the defining parameters of the trials varied so widely, there was no common basis for comparison of results between studies. Eligibility criteria differed from trial to trial, as did the measures of drug toxicity, the procedures for modifying dosage when patients experienced toxicity, the criteria to define refractory disease, and the dose intensity used in the study. Defining the clinical meaning of a 'response' was likewise problematic and inconsistent. Since leukaemia is essentially a cancer of the blood, a range of haematological parameters could be used to gauge a patient's reaction to therapy and there had been no consistency as to what levels of which blood constituents (lymphocytes, neutrophils, haemoglobin, etc.) constituted a response. For this reason, the National Cancer Institute eventually appointed a panel to standardize the clinical parameters to be used in the design of clinical trials for this disease — but not until the literature was replete with inconsistent and non-comparable studies.

In collaboration with FDA, Triton selected two studies to be retrospectively analyzed — one performed by the Southwest Oncology Cooperative Group (SWOG)¹⁸ and the other by the M.D. Anderson Cancer Center in Houston. Whereas most of the studies already completed on fludarabine employed continuous intravenous infusion (not a user-friendly form of administration for a marketed drug), these two studies administered fludarabine as a single agent in a brief infusion once a day for five consecutive days. These two studies would be analyzed retrospectively using an agreed-upon protocol based on the new objective response criteria for CLL developed by the NCI (the studies were originally performed using other response criteria, but the data were available for reinterpretation by the new criteria). According to the sponsor, the FDA further agreed with Triton that these objective response measures, coupled with the

¹⁸ A consortium of cancer centres, hospitals, and community oncology groups which participate in clinical trials under centralized organizational administration.

predictable course of the disease, would allow the use of the patients in the study as their own controls. If the analysis produced positive results, then the FDA agreed it would serve as substantial evidence for approval. Finally, it was agreed that Triton would evaluate safety data from other NCI studies for supplemental support of the application (7-8).

The two trials re-analyzed to serve as pivotal studies were both described as ‘Phase I/II’ studies (22), indicating that the studies incorporated features characteristic of both Phase I and Phase II studies. Both trials were single-arm, uncontrolled (and therefore unrandomized), unblinded (also referred to as ‘open-label’), and both used small test populations. In the SWOG group, 31 out of 32 patients were refractory to previous treatment (meaning that they no longer responded to previous therapies) and of these 31 refractory patients, four achieved a complete remission. Overall, 10 patients achieved an objective response (using the NCI criteria) for an overall response rate of 31 percent. In the MD Anderson group, 48 of 101 patients were demonstrably refractory. Of these 48 refractory patients, 23 responded (including 13 complete remissions) for an overall response rate of 48 percent. The median time to relapse for the responders in the MD Anderson group was two years; all of the non-responding patients passed away within 18 months.

5.3.2 Donning a ‘Clinician’s Hat’

While these results seemed promising, there were two major related problems. First, responses did not reliably translate into survival benefits (FDA 1990d, 13). Second, without a control group, how could one judge the significance of these responses? According to Dr. Latts, no comparable control group could be found in the literature because previous studies often did not distinguish between previously treated patients who were refractory to therapy and those who were not. The patients in these pivotal studies, by contrast, were all refractory (64-5). Dr. Keating of M.D. Anderson (who implemented one of the pivotal studies) confirmed that ‘there is very little to go on’ in the literature for a control group (65), but ‘we all have anecdotes’ about patients who are ‘still going’ (66). He then told of one patient whose story was published in *The National Enquirer* (an American tabloid newspaper famous for reporting encounters with aliens, Elvis sightings, ‘miracles’, etc.):

This is a lady who came down with stage IV disease, massive splenomegaly. She arrived on a Monday, having had emergency radiation therapy for possible impending splenic rupture at the time. She came in a wheelchair and after three courses of treatment she was fully ambulant and went shopping at the local high-price shopping center. Now she has decided that she can walk so well that she is having a hip replacement. She is one of the patients out in the two-plus year category (66).

After recounting this dramatic case, he nevertheless concluded by saying that it is 'is premature for us to conclude that there is any survival advantage of the drug' (66).

If no survival benefit could be claimed, asked Dr. Temple, then what response-related clinical benefit do these patients experience? Are there quality-of-life improvements? (66-7). No quality of life measures were systematically taken. However, Dr. Keating noted, responding patients 'feel much better' and benefit from a 'change in collar size' (67), meaning that palpable lumps and swollen lymph nodes in the neck begin to soften and recede, which has 'an impact on their attitude to life' (67). Committee members asked if the drug had been found active in other types of cancers. Dr. Greaver of SWOG described activity in a number of tumour types, including non-Hodgkin's lymphoma. Indeed, one of the two earliest Phase I patients in whom drug activity was observed been treated for non-Hodgkin's lymphoma and was still alive after five years. Greaver prefaced discussing this clinical case by saying, 'I know anecdotes are not exactly [the] most important substantive evidence but . . . I think it is important for you to know that that person is still alive and functioning as a school teacher' (71). Dr. Temple quipped, 'That is not an anecdote; that is an N of 1 study with the patient as their own control' (72).

In the FDA's review of the NDA, the Agency's Dr. Anthony Murgo concurred that there was nothing in the literature that would serve as adequate historical controls for these pivotal studies (80). As an alternative, he agreed that since each one of these patients was refractory to previous treatment, a response in this group is 'evidence of activity' (81) and the patients acted, in effect, as their own controls. Murgo made a closer evaluation of the background data on the refractory patients who responded to therapy in the pivotal studies. While some patients may not have been truly refractory, all of these patients were nevertheless 'heavily pre-treated' and for these patients a response rate in the range of 32 to 48 percent 'clearly indicates that the drug has activity' (48). What about clinical benefit? According to Murgo, the question of survival prolongation could not be answered definitively from the data presented. Instead, Murgo took off his 'regulatory

hat' and put on his 'clinician's hat' to ask 'what types of benefit we would like to see in patients in CLL' (83). What are the major complications of this disease? The list included infections, febrile episodes, and bone marrow failure. While there were no data to support improvement in infections or fever, Murgo performed an analysis showing that patients did improve in symptoms associated with bone marrow suppression (primarily anaemia and thrombocytopenia), despite the fact that fludarabine is a myelosuppressive agent (85-6), again suggesting positive effects of the drug.

Murgo's analysis was compelling, but ODAC members struggled to fit the data to the ideal of 'adequate and well-controlled investigations'. Dr. Piantidosi of Johns Hopkins University asserted that a controlled study is one which 'renders groups comparable on all bases except for the treatment received' (88). In this case, however, since a concurrent comparison could not be made between survival on the investigational drug and survival on other drugs to which this form of cancer had become resistant, the studies should be thought of as 'partially controlled' (89) or even not controlled at all. The FDA's Dr. Temple responded that, 'for better or for worse, our definition of a controlled trial, which has been in the regulations for a long time, includes the historically controlled trial' which means trials 'in which a historical experience is identified and those in which the patient's baseline is used as his own control' (89). It was, therefore, 'regulatorily' speaking, 'potentially a well-controlled study' (89).

There was also the question of what response rates meant in this disease. Dr. Bernath pointed out, among other things, that in the French cooperative group experience early treatment of low-risk patients seemed to enhance response rates while worsening survival, 'which is a sort of disturbing possibility for those of us who are ready to jump on and worship at the altar of CRs [complete responses] and PRs [partial responses] and accept response rates only as a measure of efficacy in this disease' (96). Nevertheless, he said, the committee should focus on those patients for whom there was not a reasonable expectation of any response based on the literature. In other words, based on what was known about the natural progression of refractory disease, there were patients in these studies who clearly would not have been expected to have any beneficial response to additional therapy. For Bernath, this aspect of the NDA hearkened back to the ODAC's deliberations on ifosfamide two years earlier. In a point-by-point comparison to the ifosfamide case (96) Bernath drew many parallels between the two situations, concluding that '[w]e are left primarily with response data in a setting where, I believe, responses of

this quality were not previously expected, and that is sort of what I am coming down to' (99). Moreover, there was nothing else to offer patients with refractory CLL. So although 'the place of fludarabine in CLL therapy . . . must obviously evolve from randomized Phase III trials' (97), the question was whether to approve fludarabine in the interim. For Bernath, the lack of therapeutic options for refractory CLL patients combined with the relative safety of the drug tipped the balance 'slightly in favor of approval' (98). Committee member Dr. Craig Henderson of Harvard Medical School asked whether, in retrospect, it had been a wise decision to approve ifosfamide when they did. Bernath responded enthusiastically: 'In the interest of brevity, I would say that I am proud that we approved that drug' (100). He must have said it with some drama, because the committee responded with laughter. Indeed, Bernath considered the approval of ifosfamide to be an important turning point generally. A few minutes earlier in the meeting, he remembered:

When I joined this Committee three there was a flurry of harsh-spoken criticism against the FDA, primarily about the following issues: perception that the FDA was stubbornly clinging to a notion that you had to have two randomized, well-controlled Phase III studies showing not only efficacy and safety but improvement in survival duration in order to get a drug approved.

When this group recommended approval of ifosfamide based on 13 individual refractory testicular cancer patients, in an uncontrolled Phase II trial, I believe the old illusion was shattered. But now we are asked to consider fludarabine on the basis of two fairly small, uncontrolled, in my mind, phase I/II trials where efficacy considerations must rest solely on response rate. Survival cannot even be seriously contemplated and quality of life data were not measured. Times have certainly changed (93).

Even so, the significance of the response rates was still in question (100). Dr. Temple commented that while response rate often simply meant tumour regression (or in this case, measurement of certain parameters in the blood or marrow), the NCI definition of complete response included some measurements relating to the clinical condition of the patient. So, he reasoned, the response rates used in this NDA could translate into clinical improvement for the patient (101-2) — i.e., the surrogate endpoint contained elements of clinical endpoints which one might wish to see when wearing a clinician's hat. For that reason, Temple did not see this situation as 'all that precedent-setting' because 'if you feel tumor free and you are happy about it and know that it is not there, even that for a drug that is not excessively toxic might be considered important' (102).

The committee, as always, was consciously aware of the larger implications of their actions. Dr. Henderson commented that ‘we often look at just two endpoints, that is, response rate or the classic Phase III randomized trials’ (102), however, in this case, another view of the data was required. Henderson complimented the drug sponsor on a willingness to take a different approach than usual, because ‘we have to try and find new and innovative ways of evaluating drugs’ (102). As will be discussed in the next chapter, the so-called Lasagna Report had just been released — a report written by a panel of experts having close political connections to the administration of George H.W. Bush and recommending greater flexibility on the part of the FDA and an increased willingness to accept NDAs on the basis of surrogate endpoints. The committee was well aware of the release of this report. Grace Monaco, an attorney sitting on the committee as a patient representative, commented that Dr. Murgó’s analysis was heartening because it gave some basis to believe there was a benefit to patients from this therapy. However, she urged caution in how these results were represented publicly. ‘We have a charge from the Lasagna committee and a cast of millions to try to find new and flexible ways to get helpful things out there to patients, and we want to do that’ (104-5). But she wanted to avoid the perception that the committee had ‘fallen victim’ to the notion that there was a firm correlation between complete response and survival, and also to avoid getting ‘mousetrapped’ by future drug sponsors expecting approval decisions for endpoints of dubious merit (105). Indeed, Dr. Temple later commented, the FDA has been pushing for better measures of clinical benefit. ‘We like easy decisions, you know, so we have been pressing harder for better measurements of these things. We did not need the Lasagna Committee to tell us, but there has never been a conclusion that you have to improve survival. There are other relevant endpoints. It is just that they are not often demonstrated’ (114).

Ultimately, the idea of alternative clinical benefit to the patient became the reasoning underpinning the committee’s vote for approval. When the committee’s Dr. Kemeny (of the Memorial Sloan-Kettering Cancer Center) asked, ‘Let’s say there was a randomized study that demonstrated that there was no effect on survival with this drug and, yet, we know the drug is effective, what would we do with that information?’ (110), Dr. Temple’s response was persuasive. He said that perhaps there was some clinical benefit to a reduction in lymph nodes and, if so, he would defer to the oncologists in the room who deal clinically with the issue to make that judgment. ‘There is nothing in our

law that says that those things cannot be considered valuable. There is not even anything in our law that says that response rate could not be an acceptable measure, as it was for many years. We just came to think that there ought to be a defined clinical benefit' (110). Real clinical benefit is a matter of clinical judgment. 'You can have such a benefit without improved survival' (110). Temple indicated that he had been pressing for better quality-of-life indicators six or seven years earlier, but that drug sponsors rarely, if ever, offer such data to the FDA. This response led Dr. Bernath to ask what, precisely, 'efficacy' really means. Temple responded that the answer is 'fairly complicated' but that legally, 'the statute says that what a drug has to show is that it can do what its labeling says it can do. In theory, that could be nothing more than "I shrink tumors"' (113). However, the courts had upheld the notion that claims in the labelling must make medical sense. 'The courts have sustained the idea that clinical usefulness is relevant' (113). The idea of clinical benefit must be weighed against the toxicity of the drug. However, within these parameters Temple indicated that there is 'a fair amount of judgment' (114).

While some committee members continued to yearn for a randomized study, most appear to have accepted that approval could be granted because many of these responses were surprisingly durable (compared to what would normally have been expected) and that while 'response' did not mean 'survival', the definition of 'response' included elements that could be viewed as clinical improvement for an enhanced quality of life for the patient. The committee voted unanimously to approve the drug on condition that the labelling specify the definition of a 'refractory' patient and that the sponsor perform additional studies to define more clearly the safety and dosing information for patients having impaired renal function.

5.3.3 Discussion

By now the theme of Committee decision-making on the basis of less-than-ideal evidence has become very familiar. As for ifosfamide, the studies offered in this NDA were not prospectively designed to serve as pivotal clinical trials for approval, but were early- and intermediate-phase uncontrolled trials which produced in patients responses thought to be unusual for the natural history of the disease as understood from experience and the literature. In this case, in fact, the studies were 'Phase I/II' trials, as opposed to

the ‘Phase II’ trial used for ifosfamide.¹⁹ A search for historical controls in this case was even less successful than in the ifosfamide case. (One ODAC member in the fludarabine meeting referred to the literature review as a ‘fishing expedition’ [FDA 1990d, 104].) In both cases, the sponsor encouraged the FDA to think of patients as their own controls. For ifosfamide, that argument was not convincing. Here, however, there was more acceptance of that position — and Dr. Temple even suggested that a patient could serve as his own ‘historical’ control.²⁰

Fludarabine was clearly perceived as a safer drug to use than ifosfamide, and this approval demonstrates how much more comfortable the ODAC is with flexible decision-making when a drug appears to be relatively safe and is intended for refractory patients lacking other therapeutic options. Although not required by law or regulation to do so, physicians on the committee could not seem to avoid putting on their ‘clinician’s hats’ to identify with the situation as physicians treating patients. Dr. Bernath’s comments included a discussion of what one does with refractory patients (94-96) in which Bernath expressed a clear desire for having more treatment options. He asked what one should do with an early stage patient who was treated and became refractory. What next? ‘Would you just tell the patient, well, there’s not much to do. I know you’re at stage I now but there’s not much to do; let’s just wait? No, you would treat them’ (95).

This is an often unrecognized aspect of the decision-making for cancer drugs. It is not only patients who push for additional therapeutic options, but the oncologists themselves who in their own way look to expand the options for patients. This is likely part of the reason that the NCI formally organized the ‘Group C’ cancer drug program for expanded access many years before the FDA formally established treatment IND. For NCI oncologists dealing with patients in clinical settings, it was an ever-present concern. Meanwhile, the FDA was more cautious in its approach.²¹ This fundamental cultural difference between the FDA and the NCI became especially apparent in the

¹⁹ While the specific difference in the designations is not completely clear in these cases, we can assume that the pivotal fludarabine trials included elements usually associated with early study of a drug such as dosing, toxic effects, or how the drug is metabolized and excreted, in addition to measurements associated with the activity of the drug in this disease setting.

²⁰ Presumably, Temple made the suggestion here but not for ifosfamide because physicians on the panel were more comfortable that they understood the natural history of the disease in the setting studied in the fludarabine trials than that of the third-line (and fourth-line and fifth-line) setting of the ifosfamide study.

²¹ In drug development, the FDA wished to promote ‘sound science’. For a discussion of some of the differences between the FDA and NCI in their attitudes to drug development, see Rothman and Edgar (1992).

1970s. Key FDA decision-makers believed that NCI physicians were too hasty in trying investigational drugs on patients, neglecting ethical standards of patient safety in the process. From the NCI perspective, the FDA was overcautious. ‘It is not bad to go to clinical trials early when life is ebbing away’, said the NCI’s Vincent DeVita (Blakeslee 1977, A3). When, in 1977, the FDA proposed restrictions on the investigational use of combinations of drugs, the American Cancer Society lobbied Congress to shift oversight of cancer drug testing from the FDA to the NCI (Blakeslee 1977). The effort was unsuccessful, but controversy continued.²²

By 1990 when the approval of fludarabine took place, the relationship between the FDA and NCI was calmer. This very brief recounting of the tension between the two organizations nevertheless provides insight into the attitudes of NCI oncologists and their differences from FDA staff. It is therefore notable in this account of fludarabine that it was in no small part the intervention and advice of the FDA that facilitated approval of the drug. By this time the FDA had clearly accepted a more flexible posture towards development for certain types of drugs. Although Dr. Henderson congratulated the drug sponsor for taking an ‘innovative’ approach, in fact the sponsor’s analytical approach — based on the idea that some patients experienced transition to earlier stages of disease as a result of the therapy — was dismissed by most of the observers on the committee (as well as by Dr. Temple) as tantamount to a reshuffling of response data. The suggestion that quality-of-life parameters could be considered as an alternative endpoint for an NDA came from the FDA. Moreover, more than one committee member cited as influential the FDA’s Dr. Murgo’s presentation, in which elements of response were related to clinical improvement.

Again, we have a case of drug approval pre-dating the 1992 Subpart H rules for accelerated approval (discussed in the next chapter), but anticipating some of its features.

²² Reports of patients who died on investigational cancer therapy began to surface in the media (including some on early-phase investigation of mitoxantrone, the first of the cancer drugs discussed in this chapter). Asked to comment on the situation, FDA staffer Dr. Robert Young noted the frequently-made analogies between the ‘War on Cancer’ and the War in Vietnam. Calling these analogies ‘scary’, he said, ‘You’ve got the generals, the NCI. And you have this attitude among the generals: “We’ve got to burn the village to save it.” The generals keep saying, “We’ve invested so much; we can see the light at the end of the tunnel.” But you’ve got to keep asking yourself – what’s the gain? I mean, what about the patient?’ (Gup and Neumann, 1981). Meanwhile, amid revelations that the NCI delayed notifying the FDA and physicians of toxic effects of their investigational drugs, DeVita was called before a Senate hearing to explain the NCI’s actions. DeVita admitted a lapse in NCI protocol and promised more timely notifications in the future (Kurtz 1981). *Mea culpa* notwithstanding, resentment lingered between the two agencies. As DeVita conceded regarding his staff’s attitude to the FDA, ‘It is not rare for anyone, including NCI staff, not to like to be investigated and regulated’ (Kurtz 1981, A11).

I earlier quoted Dr. Bernath as saying that ‘the place of fludarabine in CLL therapy . . . must obviously evolve from randomized Phase III trials’ (97), and that the question for the ODAC was whether to approve fludarabine in the interim. Bernath was clearly thinking of this case of drug approval as one requiring additional study for a full understanding of its therapeutic utility; he was thinking of this as an interim approval.

Moreover, where data is sparse or inconclusive, we once again see anecdotes of individual cases tending to fill the void, even while disclaimers are made about the lack of statistical validity of such stories. Indeed, it seems useful for drug companies to provide access to the investigators who performed the studies in question, not merely for informational purposes, but because these investigators remember specific details of the individual lives of patients, adding vibrancy and verisimilitude to anecdotal accounts. While these investigators were perhaps not as renowned as the esteemed Dr. Einhorn who presented the data for ifosfamide, their opinions, observations and recollections nevertheless carried the weight of well respected colleagues. In clinical trials composed of small patient groups having serious diseases, it becomes not only possible but often desirable for Committee physicians to be able to review case-specific data for patients. We will see a yet more pronounced example of this type of situation in our discussion in Chapter 7 of the first oncology drugs approved under the new rules for accelerated approval.

5.4 Conclusions: The ODAC as a ‘Microclimate’ of Standards-Making

In these three cases of oncology drug approval, we can see all the trends identified earlier in our discussion of Subpart E: key decisions taking place earlier in the drug approval process; movement towards reducing the burden of evidence required for approval while often shifting some of that burden in the form of post-market data collection requirements (or assumptions that continued study will take place); acceptance in principle of single-study evidence with ‘confirmatory’ data as the basis for drug approval. Notably, although all of these conceptual elements identified in the previous chapter are evident in the drug approval decisions discussed in this chapter, none of these approval decisions fit the model established by Subpart E. The mitoxantrone NDA was

formally categorized as approved on the basis of two Phase III studies,²³ but in reality the weakness of the second study forced a decision based on one study plus confirmatory data. The ifosfamide approval was based on a Phase II study, but not one prospectively designed in coordination with the FDA to mitigate the need for Phase III; rather, the use of the data to support an NDA was a retrospective decision and the data were extracted from what was originally an uncontrolled six-arm investigation (in the ‘shotgun’ approach often characteristic of Phase II study). The data supporting the NDA for fludarabine were retrospectively selected and analyzed.

These approval decisions underscore the adaptive and often retrospectively justificatory nature of oncology drug approval. The type of development and approval process contemplated in Subpart E requires forethought and planning in coordination with the FDA. In real-life oncology, researchers appear to be caught off guard and pleasantly surprised when a Phase II drug is significantly active. As Dr. Einhorn commented, ‘quite frankly, most of our Phase II trials are zero for 14 and we move on to the next Phase II drug’ (FDA 1988b, 37). Obviously when seriously ill patients lack therapeutic alternatives and an investigational drug surprises researchers and *is* found to be effective, the impulse is to legalize the therapy as quickly as possible, even though it will require some retrospective analysis and flexible application of the rules to do so. In contrast to Subpart E, the soon-to-be-written Subpart H rule for accelerated approval (discussed in the next chapter) was such that it could be readily applied to retrospectively selected and analyzed data. As this elaboration of FDA rules and decisions continues through the next chapter, this difference will become increasingly salient as we study the shuffling and redefinition of the rules in practical application.

Also note that none of these decisions were ultimately made on the basis of survival. To Dr. Moertel’s vexation, mitoxantrone was approved on the basis of response rates while the long-term equivalency of the survival curves was unproven. Mitoxantrone was approved on the basis of a surrogate endpoint thought reasonably likely to correlate with long-term patient benefit, but not proven to do so. In the case of fludarabine, the committee essentially ignored the response rates and approved the drug primarily on the

²³ See the summary data for mitoxantrone at <http://www.accessdata.fda.gov/scripts/cder/onctools/summary.cfm?ID=90>. This is a good example of why summary data at the FDA website should be consulted with caution and should not be directly imported as aggregate data into quantitative models of FDA decision-making: the summaries often oversimplify or idealize the true basis of decision-making.

belief that the therapy helped to improve patients' quality of life — which is a measure of clinical benefit to be sure, but not one directly correlated (necessarily) with long-term outcomes such as survival. The study closest to being judged on survival was the one for ifosfamide. However, although Dr. Einhorn frequently used the word 'cure', the FDA analysis of the data cast doubt on whether those responses were actually cures. The drug was approved, not because ifosfamide cured people, but because the drug combination induced impressively durable responses in a small group of patients who had relapsed more than once on previous therapies.

Dr. Temple noted in the discussion of fludarabine that at one time tumour regression had been commonly used as a measure of drug effectiveness against cancer manifested in solid tumours. While this endpoint was not formally 'validated' in the sense of having been specifically tested to establish its relationship to survival, the general consensus of the oncology community was that tumour regression was meaningful and helpful for the patient. Hence, the validity of tumour regression was assumed as a matter of community consensus. However with the general movement towards more scientifically rigorous clinical studies, regulators increasingly recognized that the clinical significance of tumour regression was not clear. Drs. Temple and John R. Johnson wrote in a 1985 publication: 'In the past, new anticancer drugs were approved solely on the basis of objective tumor response, but this is no longer the case. Biological activity (tumor shrinkage) and efficacy (patient benefit) are not necessarily synonymous' (Johnson and Temple 1985, 2). For this reason, in the mid 1980s the ODAC had made a recommendation to require survival and quality of life data for cancer drug NDAs (2).²⁴ From the examples offered in this chapter, we can see that the ideal previously upheld by the Committee was a difficult one to maintain. The move towards formalizing the acceptance of surrogate endpoints for drug approval will be discussed more in the following chapter. Here it is simply useful to note that surrogate and other non-survival endpoints had already been in use at the time that political attention became focused on the issue.

Notably, this publication by Johnson and Temple — a report from a symposium on FDA requirements for the approval of new anti-cancer drugs — discouraged precisely the kinds of studies used as the basis of approval in the cases above. The authors wrote

²⁴ Dr. Charles Moertel of the Mayo Clinic had substantial influence on this movement towards clinical endpoints. See Chapter 4, note 20.

that the approval of ‘any’ new drug required ‘a minimum of two independent well-controlled clinical studies demonstrating the drug’s safety and efficacy for each proposed indication’ consistent with ‘the accepted scientific principle of replicability’ (1). This description fits none of the cases discussed in this chapter. As for single-study drug applications, they wrote: ‘In rare instances new drugs have been approved on the basis of one multicenter study, usually because the results of the study were so striking that repetition would raise serious ethical problems. To be convincing, such a multicenter study should be of prospective randomized design, and results across centers should show appropriate consistency’ (1). Additionally, they said that the results from such a study ‘must be clear and convincing and must show an important therapeutic gain. For cancer drugs, this means cure or probable cure in some of the study patients’ (1). Strictly speaking none of these criteria applied to the single-study NDA for ifosfamide, although the ODAC seemed to accept that a cure was possible for a handful of patients. On the design of positive control studies, the authors wrote (anticipating one problem with the mitoxantrone studies) that such studies may be required for ethical reasons, but warned that ‘these studies pose major difficulties when the test drug is not superior to the control drug, because the effectiveness of the control drug is often not well documented in the population to be studied’ (1). More than that, regarding the comparison of responders to non-responders in single-arm studies (as in the ifosfamide study), Johnson and Temple wrote that the ‘FDA has not accepted a survival advantage in responders as evidence of drug benefit, because a capacity to respond (rather than the response itself) may be the cause of the better survival. In general the only way to assess the drug contribution is by comparison with an independent control group that did not receive the drug’ (3). This was essentially Dr. Sokol’s comment when he said of Einhorn’s study that some patients appeared to be ‘born to respond’. Two years after penning this refusal to accept such comparisons without an independent control group, Dr. Temple presided over the acceptance of ifosfamide.

My purpose is not to accuse the FDA of hypocrisy or inconsistency, but rather to demonstrate how theory and practice diverge in the face of real world situations. To paraphrase myself from the end of the last chapter, consistent with a finitist view of concept application, ultimately the standards for drug approval are determined moment-by-moment in practice. This is not to say that they are arbitrary. Clearly those moment-by-moment calculations are based on a range of considerations, including most

fundamentally an experienced assessment of the level of risk one should hazard to achieve a perceived benefit. The reason that these cases of drug approval anticipate key features of rules yet-to-be written (Subpart H and the FDAMA) is that it is problem-solving in real-life cases which effectively sets the standard, creates the de facto rule, and forms the basis for consolidation or ‘priming’ of that rule (typically as I_n accumulate). In a regulatory environment, if existing rules such as Subpart E do not seem to fit the steps necessary to get promising drugs to seriously ill patients, the perception of a need for other rules will grow.

In the three examples of cancer drug approval discussed in this chapter, Dr. Temple’s role is notable. In the mitoxantrone case, he guided the committee away from the idea that the inadequacy of the second study left them with a single-study drug approval, instead suggesting that the committee could make a judgment of ‘guilty, with an explanation’, i.e., a single-study with confirmatory evidence from a second study. In the ifosfamide meeting, he assuaged the committee’s concerns about approving a drug which required a not-yet-approved uroprotector for safe administration and did not gainsay the single-study status of this NDA, although it did not appear to meet the criteria he described in the mitoxantrone case, nor those he wrote about in his paper with Johnson. In fact, he appeared to revise his criteria for a single-study NDA to suit the situation. For fludarabine, while the drug sponsor floundered through the literature hunting for historical controls and ultimately finding none, Temple was the one who suggested that a form of alternative clinical benefit could be considered, and suggested that clinical improvements such as shrinking lymph nodes would be unexpected for those patients (hence, potentially a kind of historical control). In Chapter 7, we will see Dr. Temple use the concept of historical controls with an even greater adaptability.

In these examples, Temple was a voice of adaptive, flexible moderation, consistently removing obstacles in a manner strikingly inconsistent with the stereotype of the obstructive FDA bureaucrat. Temple’s role in these advisory committee meetings (especially when compared with his earlier writing) highlights the inadequacy of theories of regulation which ignore the role of individual regulators (Chapter 1). Regulations are not simple reflections of legislative will, nor is regulatory decision-making a simple reflection of the rules on the books. In this realm of higher ‘slack’ decision-making, away from the lights and cameras of AIDS, even as Dr. Moertel derided the rapid development and approval of AZT as politically motivated, ODAC decision-making reflected a flexible

adaptivity not so different from that exercised in the lower slack antiviral committee meetings, in part because of Dr. Temple.

For some observers such as Abraham (1995), this deviation from standards might be considered ‘bias’. If one’s theoretical tools are designed only to see bias, then bias is what one will reliably find whenever there is a separation of theory and practice. However, this analytical category is much too limited to describe the process whereby, as a matter of social practice and on-the-spot consensus-making, standards are adaptively applied on a case-by-case basis. In effect, each advisory committee meeting represents a microclimate of standards-making, influenced by a range of factors including the nature of the disease and the needs of the patients. To be sure, the personal interests of participants in these meetings *can* involve pecuniary ones (most obviously associated with industry representatives presenting the NDA, but also potentially attributable to committee members having industry ties). However, a fully engaged sociology of scientific knowledge must recognize the full range of interests clearly operative in decision-making (Barnes 1977), including medical professionalism and maintaining the esteem of one’s colleagues (or wishing to stand apart from one’s colleagues in a bid for recognition), commitments to promoting ‘scientific’ standards of research, professional experience-based preferences for some types of therapeutic approaches over others, and old-fashioned ‘doctoring’, i.e. concern for the patient. All of these interests variously wax and wane in importance during the course of negotiations over specific cases. Even within an individual decision-maker, the set of interests predominating can vary through the course of discussion as part of inter-committee negotiation and interaction, with the perceived significance of the new drug application changing in the process. For these reasons, I would suggest that a deviation from standards is a ‘bias’ (in the narrow sense of favouring pecuniary or industrial interests) only if pecuniary interests can be seen ultimately to dominate the decision-making process or if a significant proportion of medical experts clearly believe the standards or medical judgements at issue to be patently deviant from sound medical reasoning — a situation clearly not evident in any of the cases studies here. In other words, a deviation of standards in favour of approval does not automatically imply undue industry influence. Other forces can be at work to produce the same result. Indeed, I would suggest that if one views bias more broadly — as the predominance of *any* interest — then another picture often emerges. It is notable how often, when the data break down, FDA regulators and ODAC committee members don their ‘clinician hats’. It

is notable how earnestly oncologists desire their patients to have more options for therapy. If these deliberations can be accepted as largely genuine and not taken as an elaborate ruse to cater to industry desires, then the cases of ODAC decision-making examined here and also in Chapter 7 often reveal, if anything, a patient-oriented ‘bias’.

In this chapter, we have seen examples of drug approval which contradict the standard wisdom for what constitutes evidence sufficient for an NDA, even while they anticipate yet-to-be written regulations and laws. This is not to imply that all cases of drug approval fit this anomalous pattern, nor even to imply that most cases of AIDS and cancer drug approvals do. Rather, I wish to demonstrate that at this time such ‘anomalies’ constituted at least a visible minority of drug approval cases; to demonstrate that these types of decisions were being made under conditions of both higher slack and lower slack, for cancer as well as AIDS; to highlight the individual care invested in each decision by regulators and ODAC members, illustrating the case-by-case variability in applicable standards-making which resulted; and to underscore an important theme of this thesis, that rules do not drive practice as much as practice drives standards-making and new rule-writing.

6. ACCELERATED APPROVAL

Following the approval of AZT and concurrent regulatory reforms, drug development and rule-making became even more aggressive in pushing the boundaries of earlier decision-making. In this chapter we will see the Food and Drug Administration's (FDA) role as a facilitator continue to expand in its involvement with the AIDS drugs ddI and ddC, where the FDA served as a more effective advocate even than the drug sponsor for approval of these new drug applications (NDAs). During the period leading up to the promulgation of the 'Subpart H' rules for accelerated approval discussed in this chapter, AIDS-era politics were clearly a powerful factor shaping the regulation of drugs for life-threatening diseases at this time — a topic already well described by Epstein (1995; 1996). My task here will be to continue the study of individual drug approval cases and the creation of new rules, examining the social processes of rule-writing and rule-following.

One of the chief ways that earlier decision-making was accomplished for AIDS drugs in this period was through the use of surrogate endpoints in clinical studies. In the last two chapters, we encountered the issue of using clinical vs. surrogate endpoints in drug evaluation and approval decision-making, but here in the context of AIDS drugs the issue will come to a head. The controversy over surrogate endpoints for AIDS drugs in this period has also been described in detail by Epstein (1997). My purpose here will not be to rehearse the political arguments of the day, but to examine how the issue of surrogate markers insinuated itself into the already existing tension between information-gathering and earlier decision-making. In so doing we will continue to see the themes explored in previous chapters.

6.1 Surrogate Endpoints and ddI

6.1.1 *'Treacherous' Surrogate Endpoints*

We have seen that clinical endpoints are a direct measure of patient benefit, and therefore reliable and readily interpretable, but the course of the disease has to be followed to observe them — sometimes a lengthy process, especially for measurement of survival. Hence, the use of laboratory measures or other ‘surrogate’ markers of patient condition thought to be related to clinical benefit represent a significant time saver. More than that, for many in the AIDS community the idea of waiting for patient deaths in clinical studies was morally repugnant — a practice some AIDS activists referred to as taking ‘body counts’ (Epstein 1996). Unfortunately, we will see that in 1990 AIDS researchers as yet had not been able to establish the validity of various laboratory measures then being considered (and used) as possible endpoints for study. Worse yet, surrogate endpoints could be tragically misleading. For example, the cardiology community had been ‘badly burned’ (Temple 2005) in a well-known case of a surrogate endpoint failure. Two drugs, flecainide and encainide, were thought to reduce the risk of sudden death in post-myocardial infarction (post-MI) patients based on their ability to suppress ventricular arrhythmias. These two drugs became the centrepiece of an aggressive treatment policy for ventricular arrhythmia in post-MI patients, and as such, were two of the drugs the FDA had chosen to make available in pre-1987 versions of informal expanded access (FDA 1990a). Tragically, in the Cardiac Arrhythmia Suppression Trial (CAST), initiated in 1987 and discontinued prematurely by an independent safety monitoring board, these drugs were found actually to *increase* mortality in these patients.¹

This and other conspicuous failures of surrogate endpoints in the study of cardiovascular drugs led researchers in this field to view the use of surrogate markers with a jaundiced eye (Temple 2005).² This lesson was not lost on members of the Antiviral

¹ A good summary of the trial results can be found in Anonymous (1991). See the preliminary investigation report in CAST Investigators (1989). The FDA also made reference to this event when discussing surrogate endpoints in their publication of the Subpart H rules (discussed later in this paper) (FDA 1992b).

² In a personal interview with the author, Dr. Temple commented that the best known example of the cardiovascular community being ‘burned’ was CAST. But ‘there are other examples. Drugs for heart failure which improved cardiac function (actually they were to treat symptoms, too, so it wasn’t just a surrogate) turned out to be lethal. A whole bunch of them. And the cardiovascular community discovered that they could do large trials quickly, and get the real answers instead of the “fake” answers. So they’re unbelievably enthusiastic and they’ve translated that over to devices. And it’s because they can. Well, there’s a number of reasons. For the most part the conditions are asymptomatic. So you can do large, long-term, placebo-controlled trials. Whereas, if a person had symptoms they won’t let you do a large, long-term, placebo-controlled trial. They’ll leave the study. So the cardiovascular community has a lot of people who believe that you’re just stupid if you accept a surrogate.’ Temple’s comments highlight

Drugs Advisory Committee (ADAC), who met in September 1990 to discuss the criteria to be used for design of pivotal studies to support NDAs. Dr. Paul Meier noted with apparent alarm that '[a]rrhythmia is such an obvious endpoint'; the CAST study, he said, must be seen as a cautionary tale. 'Surrogate endpoints are, by their nature, treacherous.' (FDA 1990c, 93). More than that, there did not yet seem to be a reliable surrogate marker to use for AIDS. Dr. Daniel Hoth of the National Institute for Allergies and Infectious Disease commented that 'everybody would desperately like to identify CD4 as a marker sufficient to demonstrate drug efficacy' but that the state of research for such a demonstration was not there (69). Dr. Fred Valentine of the New York University Medical Center agreed, saying, 'At present, I feel that we need clinical endpoints' (71). Valentine also noted with frustration a recent Harvard study of p24 antigen as a surrogate marker in which p24 was found not to predict clinical outcome. 'That is upsetting to me', he said, 'because this is something that we can measure very accurately' (72). This was the awareness of all involved on the ADAC as discussions of AIDS drug approval proceeded.

6.1.2 A 'Very Unusual' NDA: *ddl*

Fresh from the success of AZT, testing rapidly got underway with nucleoside analogue 'cousins' to AZT. Clinical testing of dideoxycytidine (ddC) began in 1987. The drug had a great deal of activity against HIV and, according to Dr. Robert Yarchoan, an AIDS researcher in the NCI's Dr. Broder's Clinical Oncology Program, 'was active at lower doses than what we would have anticipated. But the toxicity prevented us from going much higher on the doses, and it did not induce as many immunologic benefits as AZT except in doses that over a period of weeks wound up being toxic' (Yarchoan 1998, 30). A year later, the study had to be suspended because of severe, painful peripheral neuropathy in the patients on study (FDA 1990c). A new study was not initiated until 1989.

Given this experience with ddC, researchers were more conservative in their approach to dosing when they began studying dideoxyinosine (ddI) clinically in 1988, leading to a reputation for ddI as 'AZT without tears' (FDA 1990a; 1990c). Even so, by the time the Phase I trials for ddI were completed in 1989, it was clear that the drug caused significant toxicities, especially peripheral neuropathy and pancreatitis.

not only the potential pitfalls of using surrogate endpoints, but also the important role that disease- and discipline-specific factors can play in how drugs are evaluated and clinical studies are designed.

Nevertheless, it was equally clear that AZT was no miracle cure.³ The drug was toxic enough that many patients were intolerant to the standard dose schedule. Moreover, there was both laboratory and clinical evidence that the virus became resistant to AZT after prolonged therapy (FDA 1990c). The obvious answer was to get more therapeutic options quickly, and with ddC delayed for toxicity problems, ddI emerged as the focus for desperate AIDS patients in 1989.

For these reasons, ddI was made available to patients through treatment IND in 1989 — on the basis of *Phase I data*. In a different context, any information on efficacy from such a study would have been considered at best preliminary, but here it became part of the justification for an expanded access treatment protocol (see FDA 1990a). Only two years after formalization of treatment IND, the parameters for implementation were being stretched aggressively. Moreover, the treatment IND protocol for ddI took on the aspect of a clinical trial. In an attempt to collect meaningful data from the substantial patient population receiving therapy under the treatment IND, eligibility requirements were established as part of the treatment protocol, which was now being called ‘Study 001’ (FDA 1990a). The treatment protocol would only accept AZT-intolerant AIDS patients having CD4 cell counts less than 200. In this way, the treatment IND targeted patients who were not eligible to participate in the ongoing Phase II clinical studies for patients pretreated with AZT. But now since treatment IND had eligibility requirements beyond mere therapeutic need, some other mechanism needed to be established for expanded access for patients who did not qualify for either clinical trials or treatment IND. This new programme would be called ‘parallel track’ (FDA 1990e). The FDA envisioned the programme as being carried out under the existing IND rules but taking place ‘somewhat earlier in the process than treatment IND, i.e., requiring less data to support effectiveness’ (FDA 1990a, 153). Thus did the treatment IND program for ddI revert to the pattern more characteristic of the large cardiac trials of the 1970s while ‘parallel track’ was poised to serve as a shadow treatment IND. At the same time, the sponsor began another ‘open-label’ (unblinded) study for AIDS patients whose condition was deteriorating on AZT. These arrangements were made even though, as noted by

³ Indeed, when asked if the first Phase II trial for AZT was a greater success than he anticipated, Yarchoan responded, ‘No. I was glad it was as clean as it was, but we were not surprised. The thing that we were struggling with at the time was how long AZT was going to work. We were continuing to follow these patients on AZT. What we were seeing was that the CD4 count was going up and then it was coming down. For me it was something like Flowers for Algernon’ (Yarchoan 1998, 28).

ADAC Chairman Dr. Henry Masur, for the subpopulations in question (patients intolerant to AZT or failing therapy on it), ‘we really do not know whether ddI has a glimmer of efficacy or safety’ (FDA 1990a, 191).

Meanwhile, paediatric studies of AZT had been underway and the results of one Phase I and one Phase II study were presented to the ADAC in March 1990 (FDA 1990b). Remarkably, although the Phase I study had been initiated three years beforehand and the Phase II study had started almost two years beforehand, most of the children on the study were still alive. Rather than waiting for survival data (a rather gruesome prospect in a paediatric study), the sponsor presented CD4 cell counts as endpoints along with survival probability calculations from the available data. The subsequent approval of AZT for paediatric use was the first case of CD4 lymphocytes being accepted by the FDA as an endpoint in trials to support a drug application (Ezzell 1989). This approval took place just a few months before the ADAC meeting in which surrogate endpoints were described as ‘treacherous’. It should be noted here, however, that by this time the FDA had a relatively strong database of information on AZT in adults, including clinical endpoints, as well as recently completed data from trials for less ill (adult) patients. Although there were differences between adult and paediatric AIDS treatment that needed to be taken into account (for example, ‘normal’ levels of CD4 and p24 antigen were known to be somewhat different in children than adults), generally speaking the knowledge already gleaned from adult studies bolstered confidence that AZT could be used safely and effectively in children. Hence for this situation, accepting CD4 cell counts as a surrogate endpoint was not as dramatic of a step as it might initially appear.

Accepting CD4 counts as a surrogate marker the following year to approve ddI, however, was a much bigger leap of faith. In July of 1991, the ADAC considered an NDA for ddI composed of evidence that in a more traditional context would surely have been considered inadequate (FDA 1991a; 1991b). The drug sponsor, Bristol Myers Squibb (BMS) presented a sprawling patchwork of data from a total of 12 different studies: eight treatment group studies, most of which were either Phase I trials or expanded access ‘trials’; plus four trials of patients conducted from 1986 to 1989 to be used as historical controls which, as we saw in the last chapter, is often a weak basis for comparison (see also the ADAC’s discussion of the inferiority of historical controls in FDA 1990c). Where efficacy-related response rates were calculated, they were defined as some combination of changes in CD4 cell counts, p24 antigen, increases in patient weight,

and other measures. Variations in protocol between the studies made comparisons difficult: just within the four studies constituting the core of the evidence, there were twenty-four dose levels, three different schedules of administration, two routes of administration, and four different drug formulations, among other variables. The problem was not that the individual studies presented in the NDA were flawed. As the FDA's medical reviewer, Dr. Rachael Berman, noted, each study met its stated goal but the studies were now being used for purposes other than their original design: 'The data were looked at using criteria which appear arbitrary and were developed after the studies were completed' (FDA 1991a, 245-6). This patchwork of retroactively assembled data resulted in a drug application was unorthodox not only for the lack of mature clinical study data, but for the lack of concurrent controls as well: 'What is very unusual about this NDA, if you look at the treatment group, is that there are no randomized, concurrently-controlled, trials contained in the NDA' (220-1).

The FDA's approach to this data was also unusual. When the FDA was dissatisfied with the sponsor's inconclusive analysis, which found no significant dose-response relationship for ddI, they did their own analysis using a different technique in an effort to 'separate signal from noise' (261-277). When the FDA found the efficacy data inconclusive even after this additional analysis, the Agency requested interim CD4 data from another, ongoing study of ddI conducted by the AIDS Clinical Trial Group (ACTG) called study ACTG 117. This was done even though, as recalled the following year by the FDA's Antiviral Division head Dr. David Feigal, there was a great deal not understood about the significance of CD4 cells counts at the time of the ddI approval (FDA 1992c, 11-12). Although from epidemiologic studies CD4 cell counts were known to be an indicator of health risk and were a predictor of duration of survival, it was unknown whether drug-induced CD4 cell count changes were predictive of clinical outcomes for ddI. Indeed, according to Feigal, the survival rate in the AZT trial could not be explained solely on the basis of CD4 cell counts. More than that, CD4 measurements could be highly variable and scientists had yet to work out how best to measure CD4 levels and how to use that information in clinical trials. Even so, the FDA's institutional investment in this particular drug was literally embodied through the highly unusual presence of FDA Commissioner Dr. David Kessler, who was in attendance for the entire two days of meetings.

Kessler, who took the reins of the FDA on 8 November 1990, had been called on almost immediately to participate in a Working Group on the Drug Approval Process under the purview of the new ‘Council on Competitiveness’ which was, in effect, the new incarnation of the Task Force on Regulatory Relief now under the direction of Vice President Dan Quayle.⁴ The Working Group was specifically charged with the task of reviewing the 1990 *Final Report of the Committee to Review Current Procedures for Approval of New Drugs for Cancer and AIDS*, usually called the ‘Lasagna Report’ after the committee chair, Dr. Louis Lasagna. The latter report contained a series of recommendations, including that the FDA adopt an accelerated approval procedure using surrogate endpoints as evidence of efficacy and that the FDA exercise more flexibility generally in application of the standards for substantial evidence.⁵ The recommendations ultimately issued by the Working Group to the Council on Competitiveness embraced those recommendations from the Lasagna Report (U.S. House 1992, 264-265). While the committee’s report was completed after the approval of dDI — and actually cites dDI as a favourable example of approval on the basis of surrogate endpoints — Kessler’s unusual presence at the advisory committee discussions for dDI clearly represented not only his interest as the FDA Commissioner, but as an active member of the Working Group, whose recommendations were published in a Council on Competitiveness ‘Fact Sheet’ released on 13 November 1991 (U.S. House 1992, 200-209).

6.1.3 The Actual Basis for dDI Approval

⁴ President George H.W. Bush asked Quayle’s Council on Competitiveness to resume the responsibilities of the former Task Force under the original executive order issued by Ronald Reagan. See documents appended to U.S. House (1992), including copies of the executive order and associated memoranda to transfer Task Force functions to the Council on Competitiveness, letters and memoranda to form the Working Group, as well as the charge given to the Working Group, the membership list, the final report, and drafts of interim reports with handwritten comments and corrections.

⁵ The Lasagna report also recommended that the FDA delegate much of its drug application review work to external, private-sector contractors and that it expand the use of advisory committees to take up duties associated with monitoring the design and conduct of clinical trials performed under investigational new drug applications (INDs) and reviewing NDAs (U.S. House 1992, 258-9). The committee additionally recommended an expanded use of Institutional Review Boards (IRBs) by giving drug sponsors the option of submitting their INDs to an IRB instead of to the FDA (291). The Lasagna Report further suggested that IRBs should be permitted to review Phase I and II ‘noncommercial’ clinical research to ‘find new uses for marketed drugs in lieu of FDA review’ (261). Demonstration projects were suggested under which the FDA would implement these recommendations on a limited basis and review the results after a prescribed trial period. The Working Group rejected these demonstration projects (233), but the Quayle Council overruled the Working Group and announced a series of FDA reforms (dated 13 November 1991) including the use of external contractors for review and expanded roles for advisory committees and IRBs (200-209). The following year, President Bush lost the presidency to Bill Clinton, and many of the Quayle Council’s initiatives were consequently stillborn.

In the end, despite the hodge-podge patchwork of data, and despite uncertainty over the ultimate relationship between CD4 cell counts and clinical benefit, the committee voted to approve the drug — although the vote was split. The FDA did not immediately grant approval to the drug. Since internal deliberations within the FDA are not made public, it is difficult to know precisely the process taking place behind closed doors between July and October when the drug was finally approved. However, the approval ultimately granted by the FDA was conditioned on the sponsor making another appearance before the ADAC as soon as confirmatory data was available. An ADAC meeting was held in 1992 (FDA 1992c) to review this confirmatory evidence. The confirmatory study was ACTG 117, the study the FDA used as a supplement to the inconclusive data presented in the 1991 meeting for ddI. ACTG 117 was a study comparing the long-term single-agent administration of AZT to a therapeutic regimen on which patients began on AZT and then later crossed over to ddI. The study was randomized, double-blinded, and multi-centred, and used the primary endpoints of survival and previously undiagnosed AIDS-related events, as well as secondary endpoints which included CD4 cell counts. The managers of ACTG 117 conducted interim comparisons of the two treatment groups in February and August of 1991 (FDA 1992c).

Thus, we can suppose that the FDA delayed approval of ddI until later in 1991 out of a desire to see more interim data from ACTG 117 before making their decision. Judging from the ACTG 117 final data presented to the ADAC the following April, the interim data must still have been immature with respect to CD4 cell count analysis when the decision was made to approve ddI (because it was still inconclusive at the 1992 meeting). Nevertheless, the FDA approved the drug in October 1991, sending a strongly worded approval letter mandating another appearance before the ADAC to present confirmatory data of patient benefit with all due haste (FDA 1992c). The letter also informed Burroughs-Wellcome that if the final results of the study did not confirm the clinical benefit of ddI, the FDA would ‘expect an immediate withdrawal of the drug from the market’ (Gladwell 1991, A4)

Notably, the most updated version of labelling available for ddI lists ACTG 117 and 116a as the *pivotal studies supporting approval*, even though these studies were not part of the original ddI NDA.⁶ In no small measure, the basis of the ‘conditional’ approval of

⁶ <http://www.fda.gov/cder/foi/label/2006/020154s50,20155s39,20156s40,21183s16lbl.pdf> (accessed 25 January 2008). Look in the ‘monotherapy’ subsection of the ‘Clinical Studies’ section, p. 15.

ddI appears to have been the interim data from ACTG 117 requested by the FDA in July and (presumably) August rather than the farrago of data supplied by the drug sponsor in the NDA. In the case of ddI, the FDA had moved from the posture of facilitation that we saw in our discussion of cancer drugs to one of promotion.⁷

6.1.4 AZT and FDA Policy as Bolstering Use of Surrogate Endpoint

Clearly, impetus to approve ddI converged from many directions, not the least of which was political pressure. Actors in this drama were also well aware of the urgency of the disease itself and the need to get additional weapons into the therapeutic arsenal. Moreover, physicians who had used ddI in a clinical setting expressed optimism about the drug (James 1991).⁸ In addition to these factors, the argument can be made that the perceived success with AZT provided support for acceptance of the unvalidated endpoint for ddI. Although the first AZT pivotal trial relied on clinical endpoints, the use of CD4 counts as an indicator of drug action had been developed and used from the earliest clinical testing of AZT (Yarchoan 1998), and clinical data amassed since that time provided some support for the notion that CD4 counts were related to survival, at least for AZT and probably also for its analogues.⁹ Indeed, at the 1992 ADAC meeting to discuss the approval of ddC, Dr. Fred Valentine commented that all of the development work for understanding CD4 counts as a surrogate marker was done with AZT (FDA 1992c, 458). More than that, in a 1997 interview, Dr. Samuel Broder remarked:

⁷ Kessler later testified before a Congressional subcommittee, saying that for ddI the FDA ‘really stepped into the role — one of the first times in the agency’s history — of proponent of the drug’ (U.S. House 1992, 92).

⁸ In the ADAC meeting to approve the drug, one participant expressed frustration that he was certain the drug helped people, but that ‘this data doesn’t prove it’ (FDA 1991, 356-7).

⁹ Drugs with differing mechanisms of action can be associated with differing clinical significance for the same surrogate endpoints. Less variation would be expected within classes of drugs having similar mechanisms of action. However, these types of differences between drugs can make a substantial difference in the response of the surrogate marker and the interpretation of it. See FDA’s Dr. Susan Ellenberg’s comments on CD4 counts in FDA 1990c.

AZT is important both for what it did clinically—there is no question about it—but also for the principle that it established, AZT laid the groundwork for defining surrogate endpoints in other studies, for illustrating that anti-viral agents could work in patients, and for providing a template for moving quickly from a laboratory observation to a proof-of-concept clinical study, and from their [*sic.*] to a randomized prospective clinical trial (Broder 1997, 16).

Hence, AZT was not merely a conceptual resource for rule-writing and the establishment of regulatory policies and procedures (Chapter 4), but also an important scientific conceptual resource, influential in establishing future directions for research and in creating pathways for the use of surrogate markers in AIDS clinical studies.

Significantly, there is one more element typically overlooked in discussions of the impetus to use surrogate endpoints. When drug approval decisions were to be based on interim data taken from ongoing studies, Committee members were not allowed to see interim clinical data, *but only data from surrogate measures* (FDA 1992c, 231-235). The policy was deemed necessary because unfavourable interim clinical data was not necessarily indicative of a negative final outcome. However, in many cases clinical studies could not be completed because patients (or their physicians or possibly also the administrators of the study themselves) came to believe, sometimes erroneously, that the investigational therapy was ineffective on the basis of interim clinical data. While the advisory committee members themselves may have approached interim clinical data with appropriate caution, the 1972 *Federal Advisory Committee Act* mandated that advisory committee meetings should be open to the public. If negative clinical interim results were presented in an open public meeting, it was deemed likely that patients would begin to withdraw from the clinical studies before the results had matured, foreclosing on what could have turned out to be an efficacious treatment regimen, but one whose beneficial effects took some time to become apparent. For this reason, interim data based on clinical endpoints were not released to the public forum of the advisory committee meetings. Only surrogate endpoints were to be considered, with the expectation that the surrogate measures of response were more immediately manifested than related clinical benefits, and therefore more likely to be reliable on an interim basis. The hope, of course, was that the surrogate endpoints selected were indeed meaningful as an indicator of clinical benefit. But, as one

committee member complained, this policy effectively blinded the ADAC to data important for an informed approval decision (231-2).

Hence, even as surrogate endpoints were touted as an effective way to accelerate clinical development of drugs — why wait for mature clinical endpoints when people were dying? — the reality of the situation was that if one wanted to expedite a drug approval decision on the basis of ongoing but not-yet-complete clinical trials, one had little choice but to consider surrogate endpoints, since clinical endpoints were not made available on an interim basis. We have already seen that surrogate endpoints were nothing new in oncology trials. Indeed, they had been so common that in the mid 1980s the ODAC recommended a change of practice in requiring drug sponsors to present clinical endpoints to support NDAs (Chapter 5). Nevertheless, as shown in the last chapter, that recommendation was difficult to enforce in practice. Even the tenacious Charles Moertel, likely one of the proponents of the Committee's recommendation, found himself sometimes endorsing approval on thin but suggestive evidence. Now also we see that a major impetus towards using surrogate endpoints was the (unusual, but becoming more frequent) policy of using interim data for expedited approval, coupled with the mandated practice of open advisory committee meetings. While obviously through the person of David Kessler the regulatory approach to ddI was associated with the Lasagna report and the political agenda of Vice President Quayle, in the broader context of history the Lasagna report's call for the use of surrogate endpoints seems redundant — a response more to what was written in specific laws and regulations, and a reflection of contemporary politics, rather than a true-to-life assessment of what the FDA had been doing in practice.

6.2 The Re-Vision of Subpart E and Creation of Accelerated Approval

In considering the 'conditional' approval of ddI, we can ask: in what sense was the ddI experience consistent with the regulations and procedures already in place for expediting important drugs? Clearly there was very little resemblance between the approach taken to develop and approve ddI and the AZT-templated rules for Subpart E approval. This is not to suggest that there were not elements drawn from earlier experience. As I have already argued, earlier decision points in the development process, a lighter burden of evidence for approval, transference of some of the burden of evidence

to postmarket study, an accordingly modified view of risk-benefit assessment, etc., will all continue to be themes traced in this account. Here, however, I speak of a straightforward reading of the letter of the regulation, which specified early planning with FDA to create large multicentre Phase II trials capable of creating enough efficacy data to mitigate the need for Phase III. Of course, just because a rule was created did not mean that anyone *needed* to follow the AZT model to develop important drugs. The point is that, as far as I can tell, few or none of the drugs approved for cancer or AIDS in the years following AZT *did* follow this model.

This fact may help to explain the notably imprecise interpretation given to the Subpart E rules in the July 1991 ADAC meeting to discuss the NDA for ddI (FDA 1991c). In this meeting David Rosen, an FDA attorney, was on hand to provide the regulatory background ADAC members should use to consider the ddI drug application. Nowhere in his review of Subpart E did Rosen discuss the rule's specific procedures. Rather, he noted that the regulations were 'designed to expedite the development, evaluation and marketing of new therapies intended to treat persons with life-threatening and severely-debilitating illnesses, especially when no satisfactory alternative therapy exists' (187). He then gave the following interpretive summary of Subpart E:

The procedures outlined in Subpart E reflect the recognition that physicians and patients are generally willing to accept greater risk of side effects from products that treat life-threatening and severely debilitating illnesses than they would accept from products treating less serious diseases. These procedures also reflect the recognition that the benefits of the drug need to be evaluated in the light of the severity of the disease being treated (187).

Rosen later stated that the FDA chose to file the drug application for ddI, among other reasons, because the 'mandate of the Subpart E Regulations requires that we exercise the broadest flexibility in applying the statutory standards of safety and effectiveness in the evaluation of drugs used to treat persons with life-threatening and severely-debilitation [*sic.*] diseases' (191).

Here, then, consistent with a finitist view of knowledge, the concept of Subpart E was open-ended and revisable in application. As a result, another sort of conceptual appropriation and resignification was able to take place. The Subpart E rules were no

longer a specific set of procedures to collaborate with the sponsor to design *two* Phase II studies based on *clinical* endpoints, invite the sponsor to participate in treatment IND as the study became mature, and make a decision based on an assessment of the perceived benefit of the drug weighed against the risk of information deficit associated with phase II studies, possibly requiring postmarket studies to fill the information gaps. In his reinterpretation of the rule, Rosen selectively extracted elements to emphasize broad flexibility and a willingness to accept greater risk from certain drugs perceived to be beneficial for life-threatening diseases. This was a functional and philosophical restatement of the rule, expressing Subpart E as a posture towards drug sponsors and drug approval. This re-vision of Subpart E went down smoothly and without question from the ADAC. They were physicians, not attorneys or regulators. They looked to the FDA for interpretations of law and regulation. It is uncertain whether the FDA intended this interpretation of Subpart E to apply only to ddI or if this was the beginning of a consciously systematic redefinition of the rule. Whatever the case, this interpretation seems to have been embraced. A year later, in the April 1992 advisory committee meeting to discuss the approval of another AIDS drug, ddC (FDA 1992c), committee member Dr. Abrams applied Rosen's statement of the rule almost verbatim when he said that the Subpart E regulations 'suggest that patients and their providers are willing to take greater risks.' (547).

This was not the first time a similar view of Subpart E had been voiced. AIDS activist Jim Driscoll used this view of Subpart E to try to force the FDA prematurely to open an ongoing trial, ACTG 155, for data mining to support approval of ddC (FDA 1992c). Driscoll claimed that the FDA's refusal to do so was potentially illegal in the light of Subpart E. He claimed to have worked with others (including committee member Dr. Abrams) to file a petition with the FDA prior to the ddI approval, demanding that the Agency 'use the powers under Subpart E to call in the data from ddI and ddC and reach a decision on approval or at least review it by March of 1991' (FDA 1992c, 434). Driscoll promised close legal scrutiny if the evidence for ddC was judged insufficient, implying that Subpart E somehow obligated the FDA to raid ongoing studies for interim data to allow more rapid drug approval decisions. If one reads the subpart E provisions in procedural terms, it is baffling what 'powers' in the rule would authorize (let alone require) the FDA to take such action. However, if one reads Subpart E as Rosen did, as a mandate to exercise flexibility, then the threat seems more fitting.

Subpart E, along with AZT, had become a resource available for appropriation and redefinition for various purposes. It had been resignified through articulation-practice. The regulation based on AZT was beginning to be transformed in the hands of people — FDA regulators, AIDS activists, and probably others as well — who needed it to be something other, or something more, than what it originally was. It had become unmoored from the circumstances of its original production and converted into a flexible resource based on the extraction of meanings on a selective basis. Clearly, any realistic theory of regulation must take into consideration the fact that each application of a rule, indeed the meaning of the rule itself, is open-ended. Regulatory outcomes are ultimately individualized products of the contingent circumstances in which they are produced, not the predictable results of a mechanized regulatory process out of which outcomes fall like so many moulded plastic parts. With Subpart E not seeming to serve very well as a set of procedures, it began to be appropriated for other uses.

Indeed, although the Subpart E rule was touted as a response to Vice President Bush's 1988 call for accelerated approval regulations, the promulgation of Subpart E seems to have done very little to silence those voices calling for greater flexibility and expediency in drug development and approval. Meanwhile, it was clear that there were more nucleoside analogues in the pipeline. How would these be handled in the future? Perhaps it should therefore come as no surprise that in April of 1992, a half-year after the decision to approve ddI, the FDA published a new proposed rule for accelerated approval, Subpart H,¹⁰ this time modelled on the experience with ddI. FDA Commissioner David Kessler told the ADAC just a few days after the publication of the proposed rule: 'Your actions last year on ddI were the model for this new regulation' (FDA 1992b, 227).

According to the proposal published in the *Federal Register* (FDA 1992a), accelerated approval would be considered in two situations. The first situation was when approval can be reliably based on evidence of the drug's effect on a surrogate endpoint that reasonably suggests clinical benefit or on evidence of the drug's effect on a clinical endpoint other than survival or irreversible morbidity, pending completion of any necessary studies to establish and define the degree of clinical benefits to patients (13234).

¹⁰ Proposed 15 April 1992 (FDA 1992a) and finalized 11 December 1992 (FDA 1992b).

The second situation, expected to be a rare circumstance, was ‘when the FDA determines that a drug, effective for the treatment of a disease, can be used safely only if distribution or use is modified or restricted’ (13234). For the purposes of this thesis, we will focus primarily on the provisions and implications of the first circumstance, the use of surrogate endpoints for approval.

For accelerated approval, the FDA wrote that the procedures should only apply in circumstances in which a ‘serious’ or ‘life-threatening’ disease is targeted by the new therapy, and only ‘where a serious medical need is not met by currently available therapies’ (13235). Also if a new therapy represented a ‘clear improvement’ over existing therapies for a serious or life-threatening condition, it would be considered eligible for accelerated approval. In such instances, the FDA would be willing to consider as evidence ‘adequate and well-controlled trials establishing that the drug has an effect on a surrogate endpoint that is reasonably likely (based on epidemiologic, therapeutic, or other evidence) to predict clinical benefit’ (13235).

In a description notable for the way it incorporated quality-of-life measures along with more traditional measures of patient condition, the FDA defined a surrogate marker as a ‘laboratory measurement or a physical sign that is used in therapeutic trials as a substitute for a clinically meaningful endpoint that is a direct measure of how a patient feels, functions, or survives, and that is expected to predict the effect of the therapy’ (13235). Importantly, surrogate endpoints ‘can be established with different degrees of assurance’ (13235), since a disease symptom and a surrogate marker could be coincidental, caused by a common underlying factor, rather than causally connected to each other. The FDA used the example of respiratory impairment and fever in pneumonia: the fever does not cause the disease, and treating the fever will not improve the infection. Hence, in the case of pneumonia, fever would not be an appropriate surrogate marker to predict the elimination of infection and ultimately the cure of the patient. The FDA wrote that approval of a drug on the basis of a ‘well-documented’ surrogate endpoint ‘can allow a drug to be marketed much earlier’ than if a clinical endpoint had to be established — and in the proposed rule the FDA noted that the Agency had indeed made such approval decisions in the past in cases in which it was thought that the endpoint ‘was very likely to predict a clinical benefit’ (13235). However, since there can be a certain amount of uncertainty associated with the use of surrogate endpoints, and since the FDA might be willing to grant approval in situations ‘where there is some uncertainty of the relation of

that endpoint to clinical benefit' (13235), the FDA would require follow-up studies 'to ascertain the actual clinical benefit of the drug on such endpoints as survival, disease complications, or longer-term symptoms' (13236). It was anticipated in the proposal that 'the requirement for postmarketing studies would usually be met by studies already underway at the time of approval' (13236).¹¹ If postmarket studies were not conducted with 'due diligence', or if the drug failed to demonstrate a clinical benefit in postmarket study, or indeed, if any new information came to light establishing that the drug was not safe or effective for the indicated use, then the FDA would have recourse to a 'streamlined withdrawal process' (13238).¹²

Importantly, whereas the pursuit of postmarketing studies to fill information gaps was *possible* under Subpart E, to be requested at the FDA's discretion, here postmarketing studies are *mandatory*. It is also important to note that Subpart H is not an alternative to Subpart E, but in effect a regulatory overlay: it is possible to have a drug in development for which post-Phase I consultation with FDA has resulted in convincing Phase II evidence (with associated postmarket study possibly required as described in Subpart E) based on a plausible surrogate endpoint requiring validation in postmarket study (as described in Subpart H). In this way, the promulgation of Subpart H represents not only a continuation of the trend started with Subpart E toward postponing selected data-gathering until postmarketing, but a *compounding* of it. For drugs intended to treat serious or life-threatening diseases, the principle first formally articulated in Subpart E continues to be elaborated: the pre-market burden of proof continues to be lightened based on an expectation that the benefit of doing so will outweigh the risks, including the risk of deferred information-gathering.

¹¹ This is a feature of the Subpart H rule that some pharmaceutical representatives seem to have completely forgotten. See, for example, the complaints in a recent *Wall Street Journal* editorial piece by Dr. Richard Miller, president and CEO of Pharmacyclics and adjunct professor of oncology at Stanford University. Miller asserts that in 'recent years, the FDA has effectively regressed to a pre-AIDS mindset' by requiring that sponsors submit a protocol for confirmatory trials and that they begin enrolment before conditional approval is granted (Miller, Richard. 'Cancer Regression', *Wall Street Journal*, 1 August 2007, A15). The assertion seems disingenuous in light of the clear statement in both the proposed and final rule that confirmatory trials should already be underway. Moreover, in the model case, ddI, confirmatory trials were underway at the time of approval, as they were for the first 'official' case of accelerated approval, ddC. If anything, it appears that pharmaceutical companies have been dragging the FDA farther away from the original stated intent regarding the timing of confirmatory trials (indeed, many confirmatory trials have never been completed [FDA 2000]), and that the Agency now seeks to regain control of the process.

¹² A number of other conditions for withdrawal apply, including the discovery of false marketing claims by the sponsor. See FDA 1992a, 13238.

6.3 Subpart H and ddC

Just five days after this proposed Subpart H rule was published, the ADAC met to discuss the confirmatory data for ddI and the NDA for another nucleoside analogue to combat AIDS, ddC (FDA 1992b, FDA 1992c).¹³ The committee found the clinical data reasonably convincing to confirm the benefit of using ddI, however the data related to the CD4 counts had turned out to be more difficult to analyze than anticipated and was not ready at the time of this meeting; i.e., clinical benefit of using ddI was confirmed, but evidence of a meaningful relationship between the CD4 surrogate marker and clinical benefit was still lacking.

The timing of the publication of the proposed Subpart H rule so close to this meeting is unlikely to be coincidental. Clearly, the FDA was awaiting the confirmatory data to be completed even as they were writing the proposed rule, and they presumably reviewed the ddI confirmatory data to be presented to the ADAC before they published the proposed rule. We can reasonably suppose that if the clinical ddI results had not vindicated the decision for early ddI approval — indeed, if rather than touting their success, the FDA had needed to announce a lack of confirmation or even make an embarrassing (and potentially impossible) request to remove the drug from the market — the proposed Subpart H rule might never have been published, and most likely not at this time. However, with the completed ddI studies turning out to confirm clinical benefit, it was plainly useful to publish the proposed rule prior to the meeting to consider ddC.

David Kessler was on hand personally to explain the new accelerated approval rule to the committee. In his presentation to the ADAC, he assured the group that if they ‘approve a new drug on the basis of surrogate endpoints, we can assure that the necessary studies to confirm efficacy are, in fact, carried out’ (FDA 1992b, 227-8). ‘The commitment, the follow-up, is the key to the whole policy’ (229), he said. In this way, they can be assured ‘that our accelerated approval decisions, in the end, are indeed correct’ (230). In a nutshell: ‘The new regulations make it agency policy to accept less data for

¹³ These were two days of meetings held on 20 and 21 April 1992. The first day covered presentation of the ddI confirmatory data, committee discussion and voting, discussion of surrogate endpoints in AIDS, presentation of the new proposed rule for accelerated approval, and presentation of the data supporting the NDA for ddC. The second day covered the FDA assessment of the ddC data followed by committee discussion and voting, as well as some additional discussion of the function and applicability of accelerated approval. The transcripts of the two meetings are continuously paginated, making it tempting to treat them as one document. However, in the FDA records these are two separate volumes and must be requested separately. I therefore cite the transcripts separately, and for portions of this discussion I will jump between the two: FDA 1992b (pages 1-335) and FDA 1992c (pages 336 and upwards).

approval for a potentially life threatening disease than we might normally like for a drug for a less serious disease’ (230). Nevertheless, he insisted that this new regulation ‘does not represent a change in FDA’s traditional standard. Rather, it is an acknowledgement of what we have always stated, that we must be prepared to accept greater risks from a drug when greater benefits are possible’ (228-9). Kessler concluded his presentation by telling the committee that ‘if you decide that ddC meets the same criteria as dDI — namely, that an equally convincing surrogate endpoint exists to suggest efficacy, you should recommend invoking the accelerated approval policy’ (230).

Notably, however, since the CD4 surrogate endpoint was not yet validated in the confirmatory studies, many ADAC members still expressed misgivings about the underlying meaning or significance of CD4 cell counts, even while admitting that there was no choice but to continue using them — preferably in conjunction with other measures of patient response, and certainly not for any drugs other than nucleoside analogues. Dr. Paul Meier of the University of Chicago, who referred to surrogate endpoints as ‘treacherous’ in a previous ADAC meeting, expressed the opinion that CD4 counts alone could not be the basis for an approval decision (202). An increase in CD4 cell counts could simply mean that one is ‘chasing some cells out of reservoirs, lymphoid tissues, for a period, and might ultimately not be a good thing to have happen at all. So, its biological obviousness is really not there’ (203). Dr. Donald Abrams of San Francisco General Hospital expressed concern about ‘the lack of a real translation of the CD4 change into survival, because when I think of a surrogate marker, I am thinking that we are looking for something that is a surrogate for the ultimate endpoint which is survival’ (208-9). Indeed, some committee members expressed concern that a search for a surrogate endpoint for this condition could be misguided. Dr. Deborah Cotton of Harvard Medical School, who voted against the dDI approval decision the previous year, said: ‘while we often cite blood pressure, it is not clear to me that the majority of diseases — even chronic diseases — have surrogate markers. I think this might be a very nice thing when it happens, but not something that we should assume is there if we search hard enough and do enough theoretical analyses’ (205).¹⁴ Dr. Monto Ho of the

¹⁴ Blood pressure is the paradigm case example of a surrogate marker in which an easily measured physical parameter, blood pressure, is taken to be a meaningful stand-in for knowledge of a patient’s underlying clinical cardiovascular condition and a relatively reliable predictor of the likelihood of future events related to cardiovascular health. The assumption in the search for surrogate endpoints is that these types of stand-in measures exist to assess a patient’s condition for all disease types — an assumption clearly being questioned in this discussion.

University of Pittsburgh was more sanguine than Dr. Cotton on the question of the usefulness of CD4 counts as an indicator of ‘general trends’ (207), but nevertheless agreed that in ‘seeking an ideal surrogate marker, we should not be seeking the holy grail. Probably in this case there is no holy grail. The blood pressure situation is really quite unique and CD4 is definitely not like blood pressure as a surrogate marker for cardiovascular disease’ (206). Dr. Theodore Eickoff of the Presbyterian-St. Luke’s Medical Center summed it up: ‘I think I don’t have too much to add to this litany of discomfort that seems to be going around the table. Listening to the discussion, it seems somehow astonishing that this is actually the same group that voted last summer’ (215).

AIDS activists were taken aback at this continuing suspicion of surrogate endpoints, having apparently assumed that the vindication of the ddI decision had quelled any lingering misgivings about making decisions on the basis of CD4 counts. Mark Harrington, an AIDS activist participating in the committee sessions as an invited consultant,¹⁵ commented that ‘the committee is letting its misgivings about the risk it took last summer overshadow what should be really a rather enormous feeling of relief and gratification that if they made a mistake, they made the right mistake’ (220). Further, he added, the decision was not made solely on CD4 counts, but also on weight gain, reduction in the occurrence of opportunistic infections, and other indications of improved health. Others from the activism community were much less collegial in the next day’s open hearing session of the meetings, taking to the podium to rebuke and castigate the ADAC members for ‘backpedaling’ on the issue of CD4 counts (FDA 1992c).¹⁶ The fear, of course, was that the committee would back away from approval of the next antiviral drug to come before them, ddC — a drug that turned out to be relatively easy to

¹⁵ For the role of AIDS activists in drug development and approval at this time, and on Mark Harrington’s specific role as a leader in the AIDS advocacy movement, I again refer the reader to Epstein (1996). It is worth noting here, however, that Harrington’s performance in these meetings is impressive. Although he is not a physician, he often speaks with knowledge and authority virtually indistinguishable from the physicians on the panel. More than that, the panel members appear genuinely to like and respect him. Tellingly, at one point in these hearings, the chairman of the committee Dr. Henry Masur called on Harrington to come to the podium to make some remarks and mistakenly called him ‘Dr. Harrington.’ Masur then joked about his verbal misstep by saying that the panel was ‘looking for a university that would like to bestow an honorary doctorate on Mark’ (FDA 1992c, 567). I do not read any irony in the remark, but simply good-humoured fondness and respect.

¹⁶ Characteristically, Martin Delany’s comments are among the most ardently expressed. ‘Maybe you can wait for all the scientific “i’s” and “t’s” to be dotted and crossed, but I can’t. CDs mean a great deal and anyone living or actively treating this disease damn well know it. Why in God’s name can’t you people get it? That I am having to defend CD4s as a marker is an affront to PWAS [people with AIDS] and their physicians.... I don’t give a damn if you ever scientifically find linkage between CD4s and survivability. I look around me and at my own CD4 count and know that it is real’ (FDA 1992c, 348-9).

synthesize on an amateur basis, and had already been getting churned out of home laboratories and distributed through an underground AIDS patient network (Epstein 1996, 1997), a situation of which the committee was well aware.

The data supporting the NDA for ddC, while better than the ddI data set in terms of quality (as remarked upon by more than one observer), was nevertheless highly problematic. The drug sponsor, Hoffman-La Roche, was requesting approval for two indications, as a monotherapy for patients intolerant to or failing AZT, and also as a combination therapy of ddC with AZT. We will consider the ADAC deliberations for each indication below, followed by a discussion of the application of accelerated approval to ddC and the significance of these events for this thesis.

6.3.1 NDA: ddC as Monotherapy

To support the claim for monotherapy, Roche presented data from three studies, two conducted by the AIDS Clinical Trial Group (ACTG) and one which was an expanded access 'trial'. The first ACTG trial, called ACTG 114 (N3300), was a trial for patients having less than three months exposure to AZT. This was the *only* randomized, double-blind study included in the NDA. A total of 320 patients were randomized and received ddC monotherapy as compared to 315 who were treated with AZT (FDA 1992b, 253). Unfortunately, according to the FDA analysis, patients on the ddC arm had almost *twice* the risk of mortality as those on the AZT arm (FDA 1992c, 365; FDA 1992c, 442). Indeed, although it was planned to be a two-year study, a safety review board discontinued the trial after 12 months based on these statistically significant interim findings. Since nothing close to equivalence was demonstrated in this trial, the sponsor did a retrospective analysis using the placebo group from the original AZT trial (BW002) as a control, hoping to demonstrate at least that ddC was better than placebo. The FDA invoked various disclaimers about historical controls, but their analysis found no statistical difference between the ddC treatment group and placebo (FDA 1992c, 367). The sponsor also did a series of subgroup analyses in search of some segment of patients at some time who might have benefited from ddC, however the FDA reviewer demonstrated the futility of the exercise (365-7). In many cases the subgroups simply became too small to be meaningful.

The second study was ACTG 119 (N3492), a randomized, open-label (unblinded) trial intended for patients who had at least 12 months of prior AZT therapy. The trial was

designed to have 320 patients but only recruited 111. According to the sponsor, the difficulty with enrolment was that the ddI expanded access programme was siphoning off prospective patients (FDA 1992b, 258). (Very possibly the growing black market supply of ddC was also a factor.) With only 59 patients in the ddC arm and 52 in the AZT arm, the sponsor admitted that the study was not powerful enough to see any statistical difference between the two study arms (262). Likewise, according to the FDA analysis, the 95% confidence interval on the hazard ratio was inordinately wide: the risk of dying on the AZT arm was anywhere from half to four times the risk of dying on the ddC arm (FDA 1992c, 370).

In the expanded access protocol called N3544, patients were randomized on an unblinded basis to receive ddC on either a high-dose or a low-dose regimen (FDA 1992b, 263). The expanded access protocol was originally designed for patients who were either intolerant of or had failed AZT therapy. Patients intolerant to or failing ddI were subsequently also allowed onto the protocol. Over 5,000 patients were enrolled in this programme, but at the time of the cut-off date for analysis there were 3,479 patients who had had at least one follow-up examination and could therefore be included in the data set (263). There was no statistically significant difference between the high- and low-dose groups in terms of survival. In terms of CD4 counts, the high-dose group experienced a somewhat greater boost of lymphocytes. However, drug toxicity was a significant problem. The drop-out rate for patients on this protocol was exceptionally high, especially for the high-dose cohort: after 24 weeks, the number of patients on study had dropped from nearly 3,500 to 715 (FDA 1992c, 372).¹⁷

In his summary statement, Hoffman-LaRoche's Dr. Soo argued that the evidence for the monotherapy came from 'the totality of the evidence, not from just a single study' and insisted that 'the benefits of ddC monotherapy do vary among different subpopulations', so it was the committee's task to consider which population 'deserves' ddC monotherapy (FDA 1992b, 324-5). To be fair, there were additional Phase I/II studies not discussed here for brevity's sake, but included as part of the safety presentation, for which there was some suggestion of a dose response based on CD4 cell

¹⁷ Additionally according to the FDA's presentation, after 36 weeks there were only 190 patients remaining; and by week 48, there were only 16 patients remaining on study. The sponsor's Dr. Soo disputed the latter interpretation, however, saying that the low numbers of patients at weeks 36 and 48 were simply due to a lack of data for patients more recently enrolled in the protocol (FDA 1992c, 503). He did not dispute the dropout figure corresponding to 24 weeks, however.

counts.¹⁸ There was also a great deal of preclinical and pharmacokinetic data supporting drug activity. This was part of the ‘totality’ of the data to which Dr. Soo referred. Even so, the committee was not sanguine about the prospects for monotherapy. ACTG 114 not only demonstrated a survival *disadvantage* for ddC when used as a monotherapy in AZT-naïve patients, but showed that if ddC was used as initial therapy, the survival disadvantage grew and persisted even after the therapy was discontinued (FDA 1992c, 441). The only good news from this study was that the CD4 cell counts seemed to track consistently with the survival outcomes. However, weight gain did not. Disturbingly, patients having the most significant weight gain also had the highest mortality (441).

In any event, the requested indication for approval was not ddC as initial therapy, as studied in ACTG 114, but as a therapy for patients who were failing AZT or intolerant to it. ACTG 119 might have provided support for that indication, but was too small for any conclusions to be drawn. The expanded access protocol might also have been able to address the question, but because there was no control group, there was no context for evaluating the response of previously AZT-treated patients to ddC at the two dose levels. As the FDA’s Dr. Feigal noted in his summary of issues, one can randomize an early access program, but without other refinements in data collection early access programs ‘suffer from some of the worst features of clinical trials’: they are expensive, data is underreported, and it is difficult to extract any meaning from the results (441-2).

What was clear was the severe toxicity of ddC. Peripheral neuropathy was an especially severe adverse effect experienced by many patients. The condition when mild begins as numbness or tingling in the hands or feet, but in its more severe forms it is quite painful, can spread up the extremities, and even lead to a loss of reflexes and motor coordination. According to the FDA’s safety review (FDA 1992c, 396-403), at higher doses the drug was almost uniformly neuropathic and with ddC there was the problem of ‘coasting’ — the neuropathy could continue to worsen even after discontinuance of the therapy. Resolution of the symptoms could take weeks or months.

Some observers such as Mark Harrington argued that there was, in fact, a modest surge of CD4 counts observable in the first ACTG study presented (ACTG 114), lasting roughly 12 weeks — a ‘bump’ in CD4 not observable as a placebo effect in other studies, thus demonstrating drug activity (FDA 1992c, 496). For most other observers, however, to see drug ‘activity’ was not enough. As Dr. Eickoff said, ‘I am troubled by ACTG 114

¹⁸ See the FDA’s Dr. Gitterman’s safety assessment (FDA 1992c) beginning on p. 380.

because we did not see a major significant, sustained rise in CD4 counts of anywhere near the order of magnitude that we have seen with AZT. This is the group of patients in which you would anticipate being AZT naive that we ought to see a major, sustained rise in CD4 count' (512). Nevertheless, many struggled with the lack of therapeutic options for patients failing both AZT and ddI. Dr. Neil Schram, a consultant, suggested that the drug be made available only for patients who had failed both AZT and ddI. Dr. Fred Valentine suggested that monotherapy not be approved, but that ddC should still be made available under expanded access (513). Dr. Alan Novick of Yale University found another solution. Noting his 'obligation as a consumer advocate' he said that he could not see a circumstance in which this drug was potentially useful to patients as a monotherapy and he was 'not ready to recommend licensing of a drug that I see as substantially toxic' (532). However, he said, 'I personally found my loophole in being willing to affirm combination therapy' (532). (The data supporting combination therapy was presented with the monotherapy data in the meeting, but for the sake of clarity I have temporarily deferred discussion of that portion of the NDA.) The committee voted against the proposed indication 8-to-2 with 2 abstentions. They considered approving ddC as a monotherapy for patients who had failed both AZT and ddI, but this proposal also failed by a vote of 7-to-3 with one abstention (one member did not vote because he had to leave the meeting before its conclusion).

6.3.2 NDA: ddC in Combination with AZT

For evidence of ddC efficacy in combination with AZT, four studies were offered for the consideration of the committee, however only one of these appears to be part of the programme conducted by Hoffman-LaRoche to develop the drug. ACTG 106 (N3447) was a Phase I/II study, only partially randomized, with 6 arms and 56 patients (thus, only about 9 patients per arm), which compared various regimens of AZT and ddC treatment against one arm of AZT therapy alone. All the patients in this study were AZT-naïve. The clinical data were inconclusive, however the sponsor identified one combination arm in which the CD4 response appeared to be more durable than in the AZT-only arm (FDA 1992b, 279). For better 'context' (i.e., to get a control group larger than eight subjects), the sponsor also compared the response to the AZT-only arm of the first monotherapy trial (ACTG 114) as an historical control, claiming again that the response for the combination was enhanced in magnitude and durability in comparison to

AZT alone (279-80). Clearly, however, the size of the study prevented any statistically significant results.

Two more studies, ACTG 047 and 050, were described by the sponsor as ‘really not part of the overall program that we are presenting’ but nevertheless ‘interesting’ (301-2). Both were very small Phase I/II studies initiated by the ACTG to begin characterizing various alternating or intermittent combinations of AZT and ddC. ACTG 047 was for AZT-naïve patients and distributed 131 patients across seven treatment arms, including one continuous AZT arm. ACTG 050 was for AZT-intolerant patients and distributed 109 subjects across six treatment arms. According to the FDA review of these studies, in trial 047 there were indeed increases evident in CD4 cell counts having greater magnitude and duration than the AZT control arm. However, no differences were discernible in clinical outcomes. This did not mean that there were no differences; the study was so small, it lacked power to detect such differences (FDA 1992c, 415). However, these were AZT-naïve patients; this was ‘not the regimen for which the sponsor seeks approval’ and the sponsor’s conclusion was that there were ‘only two arms that appear to be candidates for further study’ (416). In trial 050, which was for AZT-intolerant patients, there was no control arm. There was also a very high incidence of dropout related to adverse effects, this time primarily because of AZT intolerance. The sponsor did not offer any conclusions. The FDA reviewer noted that the study may indicate less effectiveness of the combination therapy in AZT-exposed patients, but there are ‘caveats’ (417). These were not simply AZT pretreated patients, but hematologically intolerant ones. There might have been a ‘slight’ CD4 response, but certainly not of the magnitude seen in the 047 study of AZT-naïve patients (417). The FDA reviewer also noted that there were *no Phase II/III clinical trials* in the combination therapy portion of the NDA submission, but there were ongoing studies of previously AZT-treated patients (reviewed by the drug sponsor as part of the presentation) which were anticipated to be completed in less than two years (418).

One other study was presented which was not part of the Roche development programme but was supportive of combination therapy: an ongoing Burroughs Wellcome study of drug resistance, BW 34,225:02, which began less than a year beforehand (FDA 1992b, 296). One would be justified in wondering how interim data from a Burroughs Wellcome study came to be presented in an NDA meeting for Hoffman-LaRoche. The FDA requested it to supplement the Roche presentation (FDA

1992c, 563). This study had three arms, (AZT, AZT plus ddI, and AZT plus ddC) with a total of 150 patients, 92 of whom were relevant to this NDA (i.e., the AZT and AZT plus ddC arms). The primary endpoint of the study was related to drug resistance, but the secondary endpoints were surrogate markers such as CD4 and p24. While this interim data was based on a relatively short duration — a consideration the FDA reviewer was careful to state (FDA 1992c, 409) — the patients on the AZT arm experienced a boost in CD4 cell counts approximately half of that experienced by the patients on the AZT/ddC combination, a difference of roughly 45 cell counts (FDA 1992b, 298).

The committee was left to consider a series of Phase I/II trials, only one of which had been part of the sponsor's NDA. In his wrap-up comments to the committee, the FDA's Dr. Feigal commented that the 'data on combination is as sparse as any data that has been presented to the Committee. It has been presented for naive patients only. It has been presented with trials that have had inherent limitations, limitations that are not unusual for a Phase 1/2 trial' (445). He added that the 'Burroughs Wellcome resistance study, however, does confirm the effect and makes it highly likely, in my mind, that ddC plus AZT gives a larger surrogate risk than AZT alone. It approximately doubles the effect' (445-6).¹⁹ Most, but not all, of the committee leaned towards accelerated approval of the combination for varying reasons. Dr. Ho believed that there was a real effect on surrogate endpoints and, in any event, that the committee should remain consistent with the decision-making used for ddI (FDA 1992c, 557). Dr. Schram, a consultant, recommended approval because he in fact was already using the AZT/ddC combination in his practice for patients who failed prior therapy (552-3). Dr. Mark Smith, who had made a case for monotherapy partly on the ground that it was already being used by the underground AIDS community (514-15),²⁰ was unconvinced that the combination therapy

¹⁹ The phrasing is a bit confusing because of the use of the word 'risk', but clearly he means that the Burroughs study demonstrated improved CD4 counts for patients taking the ddC/AZT combination over those taking AZT alone. I take Feigal to be using the word 'risk' in the same way that the word 'hazard' is used in 'hazard ratio' — referring to a probability. See the discussion of hazard ratios in Section 5.1 of this thesis, especially note 13.

²⁰ Meaning the aforementioned black market in which AIDS patients manufactured and distributed the drug illegally. To clarify, Smith did not feel any of the data for the drug was particularly convincing. 'Is it [ddC] safe? Not particularly. Is it effective? I don't know. Not much, if it is. You know, I am not — I am frankly very underwhelmed by bumps and blips and 10 cells at 24 weeks. That doesn't move me much. I think the drug is active' (FDA 1992c, 514). Nevertheless, he was concerned about the 'clinical reality' of patients who could tolerate neither AZT nor ddI. He was also 'influenced by the clinical reality that there are lots of patients, who are taking this drug obtained in other ways' (515). Even as he expressed concern that FDA would get trapped into the dangerous precedent of legalizing a drug on 'basically the dare that, gee, we are doing this anyway, so you have to let us continue to do it', he ultimately concluded that

was useful. ‘I am not impressed with one arm of four patients compared to another arm of five patients. That is meaningless to me,’ he said (561). Nevertheless, he seemed encouraged by the fact that there were a number of studies in the pipeline which should clarify the efficacy of the drug combination. Those ongoing studies were definitely a positive factor in Dr. Abrams’ view of the situation as well (558-9).

More sanguine still was Dr. Fred Valentine, who considered this an ‘ideal’ scenario for accelerated approval because there was ‘a durable CD4 rise in the small study, which granted only had 9 patients per arm. But the BW [Burroughs Wellcome] study showed that the bump wasn’t wiped out with the addition of a decent dose of AZT. So we have preliminary data that is convincing’ and ongoing confirmatory studies (559-60). However, he warned that the confirmatory studies were crucial: ‘It is critical that these accelerated approvals not compromise our ability to get the definitive data because otherwise *we will be swimming in a sea of anecdotal medicine*. We must not compromise our ability to learn how to use these drugs properly and to get proper data.’ (560, emphasis added). Dr. Eickoff likewise commented that, based on Dr. Feigal’s assurance that there was statistical significance to the interim analysis of the BW study, he would ‘go on record as favoring an accelerated approval as combination therapy, echoing the same caveats that Fred Valentine mentioned about completing the existing and ongoing clinical trials’ (565). The Burroughs Wellcome study clearly provided assurance to a number of committee members that the combination produced a CD4 effect. Indeed, Mark Harrington made a point of thanking Burroughs Wellcome for the data, ‘which has, indeed, turned out to be pivotal for this indication in this preliminary look at their ongoing study’ (568). Once again, as for ddI, extra data obtained by request by the FDA was playing a decisive role in the Committee’s decision-making. Ultimately, the Committee voted 9-to-3 in favour of the combination therapy.²¹ The first formal case of accelerated approval had been decided.

As a postscript, we should note that stavudine (d4T) became the first drug made available under the new parallel track program in 1992. Two years later (in 1994) d4T was approved as an anti-retroviral therapy on the basis of CD4 cell counts (FDA 1994). By this time, however, enthusiasm for CD4 cell counts had been tempered by the previous

because of the circumstances surrounding this drug, effective clinical evaluation for the desired indications prior to licensing was going to be impossible (515).

²¹ It was technically 8-to-3. But Dr. Ho, who had to leave early, indicated that he would vote in favour.

year's revelations from two controversial but important studies suggesting that CD4 cell counts were not predictive of survival (Epstein 1997, 708-9). 'Cracks' were appearing 'in the consensus behind the mechanism that had been invoked to license ddI and ddC' (708). The FDA's Dr. Feigal's introductory remarks at the ADAC meeting to discuss the d4T application made the diminished status of CD4 counts clear. Feigal noted that while surrogate endpoints had been used as endpoints in trials, 'this is something that, as a reviewing division, we've discouraged in protocols. We think that the surrogate marker changes within trials are very useful, but we would prefer the trials to have clinical endpoints and use the surrogate changes for other purposes' (FDA 1994a, 18). Feigal then reminded the Committee that the d4T studies had been initiated some years beforehand, conveying the unspoken message that the study was designed and implemented at a time when CD4 cell counts were believed likely to predict clinical benefit. Despite this partial retreat from surrogate endpoints in general, and CD4 counts in particular, d4T was approved. While there may have been 'cracks' in the foundation, accelerated approval was on the regulatory books and the experiences with AZT and ddI would continue to reverberate in future rule-making and decision-making.

6.3.3 *ddC: First Formal Application of Accelerated Approval*

This was the first prospective instance of Subpart H accelerated approval faced by the ADAC. Notably, however, the FDA wanted feedback from the Committee to help define the applicability and purpose of the new rule. If they invoked accelerated approval for ddC, then FDA Commissioner Dr. Kessler wanted them to specify what additional data they wanted to see and when they wanted to see it (FDA 1992c, 474-5). If they chose to deny accelerated approval, then the FDA's Dr. Feigal instructed them to identify what additional evidence they would have wanted to consider this an appropriate application of accelerated approval (445, 450).²² Nevertheless, even as the FDA asked the ADAC to define the level of evidence required for accelerated approval, Committee members looked to the FDA for guidance — beginning with fundamental questions such as how the name of the rule should be interpreted. Dr. Ho asked that the FDA 'remind

²² Ultimately, the Committee did not formally provide this feedback, although in places the discussion touched on it. To do so would have required a third day of meetings, something none of the participants was willing or able to do on short notice. A range of issues therefore remained for the FDA to decide on its own, including what the specific indication for therapy should be, what the label information should say, which additional studies should be considered confirmatory, when they should be brought before the committee, and what results will suffice to be considered 'confirmatory' of benefit.

both the Committee and the public here what is conditional approval. I realize Dr. Kessler talked a lot about it, but the technical meaning of it is unclear' (475). Kessler responded using the example of dDI, saying that it was 'full approval, but the approval letter had certain conditions' (475). Dr. Ho then asked if it should really be called 'provisional' or 'conditional' approval rather than 'accelerated' (476). Kessler indicated that 'you should think of it in those terms' (477), however it is to be called 'accelerated approval' because the FDA did not want to send a message to third-party insurers that a drug approved under this process was in any way not a full FDA approval.²³

Clarification was also needed regarding the relationship of accelerated approval to other categories of drug distribution. Dr. Lane wanted to know whether accelerated approval was necessary when treatment IND and parallel track existed; or, conversely, if treatment IND and parallel track were necessary if drugs having essentially the same level of evidence required for expanded access could be approved on the basis of accelerated approval. Lane's question demonstrates how far the significance of treatment IND had migrated from its initial conception of a small-scale stop gap to allow access while a mature clinical study was being finalized for an NDA (see Chapter 4). Now, with parallel track, it was perceived as a large scale early distribution system easily equated with accelerated approval in terms of its scale and criteria for implementation. Notably, Kessler and Feigal did not fundamentally dispute this view of treatment IND. Rather, their response was that there are sometimes drugs which go into treatment IND before they are ready for approval and, in any event, treatment IND is more flexible and can be implemented more rapidly than accelerated approval (478).

In these questions, we see the kind of categorical clarification and redefinition likely to take place with the introduction of a new category of drug approval (Chapter 1). The function of expanded access needed to be set alongside accelerated approval and compared. Likewise, for the next application of the term 'accelerated approval' Kessler would point to dDI as the existing exemplar, and the Committee would search for salient characteristics of that experience. Indeed, in the discussion of accelerated approval, dDI is always present. Dr. Deborah Cotton expressed concern 'about having codification of what sounds to me like an experiment. I guess I am wondering why you chose to write

²³ Third-party insurers had a history of not compensating patients for therapies considered to be 'experimental'. This was one reason that pharmaceutical companies ended up bearing the cost of treatment IND: because insurers would not. So the FDA wanted to avoid giving the insurance providers a pretext for denying coverage for drugs approved under this process.

regulations instead of trying this out, with provisional approval of some drugs' (483). Kessler responded, 'I think we did. I mean, we did it with ddI' (484).

Earlier observations made in this thesis about the priming of a new category and the development of an informal consensus for a new rule with respect to Subpart E (Chapter 4) apply to this case equally as well. Dr. Eickoff noted that the accelerated approval rule had been published just the week before as a proposal, and wondered if there might be changes to the rule which might affect the decision of the Committee on ddC. Kessler, who had earlier referred to the proposed status of the rule as a 'technicality' (475), responded that 'we stood ready to act on ddI last year without the reg[ulation]. We stand ready to invoke accelerated approval if the Committee so recommends and the agency so believes it is an option, even in this interim period' (486-7). Again, if the goal were merely to accelerate approval of certain drugs on the basis of surrogate endpoints, the FDA could have proceeded to do so without writing new regulations. However, as Kessler told Dr. Cotton on the subject of using ddI as an 'experiment', 'I think we all feel pretty good about that and, again, I think there is benefit of having the regs' (484). Part of that benefit, as Kessler described it, was to provide a formalized framework to address the anxiety advisory committee members often felt about putting a drug on the market while lingering doubts remained. Said Kessler, 'The whole point of this is to relieve some of the hand-wringing, some of the angst that this is it because once you approve it, it is out and you will never have another shot at it' (482). But also, part of the benefit is reflected in Kessler's statement about everyone feeling good about ddI: the writing of the regulation helped to establish and legitimize the original action, or 'precedent', on which the regulation was based. I will develop this argument more fully in the final chapter of this thesis.

It is notable that although accelerated approval is a procedure for approving drugs on the basis of surrogate endpoints and does not specify the phase of study on which such judgments should be made, both ddI and ddC/AZT were approved on the basis of very early, Phase I/II studies. There is, of course, a relationship between the use of surrogate endpoints and clinical phases of evaluation. As we have seen in cancer trials, early clinical trials tended to use a 'shotgun' approach, testing a drug under multiple scenarios in small patient populations using surrogate endpoints. The studies seen in the ddC NDA were mostly similar in character, where an array of study arms were each populated with just a handful of patients in order to get an idea of the best drug

combinations or dosages to pursue in larger clinical studies. For the purposes of their original design, they would have served well. However it was precisely these sorts of studies that became available for consideration under Subpart H at the time of its creation. The use of such studies in new drug applications, in turn, necessarily lead to information deficits in addition to those of endpoint validation. In the ddC/AZT combination approved in this case, FDA and Committee observers lamented the lack of safety data for the combination in the target patient population, including data on the additive toxicity effects of AZT and ddC, and they struggled to understand which patient subgroup might stand to benefit the most from the combination. In this regard, Dr. Ho commented that ‘we have been under pressure both by the FDA and by the community to approve these drugs as soon as possible. I think the accelerated mechanism really implies that many of these questions are not solved’ (FDA 1992c, 575). But these were fundamental questions necessary to do the risk-benefit assessment called for by the regulation and required for writing the drug’s recommended labeling. In this way, the Subpart H rule effectively guaranteed that much more than endpoint validation would be necessary in postmarket study, at least for the foreseeable future — and the tension between data-gathering and early decision-making traced throughout this thesis was thereby exacerbated, with the trend continuing towards early decision-making.

By a different line of reasoning, Mark Harrington came to a similar conclusion in the meeting to discuss ddC. The failure of the initial AZT vs. ddC monotherapy study, and the sponsor’s premature use of it for the basis of an NDA, ‘suddenly puts a lot more focus on what really was not designed to be a pivotal efficacy study’ (569). Additional studies were needed to bolster the original claims. Therefore, he said, ‘I would urge those designing trials, whether inside the ACTG or in industry, to consider doing a control arm and randomizing earlier in the stage of their drug development, in what we sometimes call Phase 1/2 studies’ (569). The solution was to make earlier phase studies look more like later phase studies. The difficulty, of course, was how to do this without short-circuiting important safeguards for patients. Controlled studies mean larger studies. Indeed, in the particular example Harrington was using, investigators wanted to randomize a combination study for ddC/AZT (ACTG 106, briefly described earlier), but the FDA did not wish to allow it since the safety of the combination was unknown. An important categorization and sorting process would need to take place over the next years to work

out which characteristics of later phase trials should be applied to earlier phases of study for the sake of expedited drug development.

Harrington's comments return us to discussions of clinical phases from Chapters 3 and 4. Once again, the clinical phases potentially stand to be revised, this time because of the introduction of Subpart H. Changes in practice can take place organically, through technical improvements made to clinical development over time (as was the case between the 1963 IND rules and the new version proposed in 1979)(Chapter 3), or they can be induced by formalizing new practices through regulation, as for Subpart E (Chapter 4). In this case, the new regulation did not specifically articulate modifications to the clinical phases. However, as I have argued, the substance of Subpart H virtually guaranteed that, at least until adjustments could be made, accelerated approvals would be primarily approvals based on early phases of data, tipping the balance further away from data-gathering and more towards early decision-making. What Harrington's remarks make clear is that in order to restore some balance in the tension between data-gathering and early decision-making, some required information must somehow be developed earlier in the clinical trial process, rather than deferred to postmarket study.

6.4 Reactions to Accelerated Approval

The final version of the Subpart H rule was published in December 1992 (FDA 1992d). The comments received by the FDA on the proposed rule were, predictably, a mixture of those who believed the rule did not go far enough and those who feared it went too far. Patients' groups generally applauded the rule, while Dr. Sidney Wolfe of the consumer advocacy group Public Interest attacked the use of surrogate endpoints as not well enough understood to be useful in drug approval (Docket 91N-0278, Item no. C18). More than that, he said, the FDA has a 'grim' history with postmarket studies (3). FDA staff member Dr. Paul Leber submitted comments expressing concern that the 'proposed rule promulgates a standard for the evaluation of drug product effectiveness that seems inconsistent with the "substantial evidence" requirement' (Docket 91N-0278, Item no. C12, 1). Leber also suggested that the 'justifications offered for the adoption of the proposed rule reflect social and moral values that may or may not be widely shared by others in our society. Considering the importance of the proposed change, it may be more appropriate to bring about its adoption by statute rather than through notice and

comment rule making' (3). Another commenter expressed concern that the proposed rule 'may lead to the marketing of large numbers of clinically ineffective, but pharmacologically active, drugs' (FDA 1992d, 58944). However, physicians' groups like The American Medical Association (Docket 91N-0278, Item no. C14) and the American Society for Clinical Oncology (Item no. C25) were generally supportive of the rule. Their primary objection was related to the provisions restricting distribution of certain special case drugs to people or institutions with special training or facilities, saying these procedures exceeded the FDA's authority. Groups representing pharmacists also challenged the latter provision since it would effectively cut pharmacists out of the distribution loop. For its part, the FDA disagreed with all of these objections, asserting the authority to restrict distribution for safety purposes and the usefulness of surrogate endpoints in certain situations if used judiciously (FDA 1992d).

The Pharmaceutical Manufacturer's Association (PMA) and some large pharmaceutical companies argued that no new rules were necessary.²⁴ The PMA wrote that 'FDA already has procedures and regulatory flexibility for the expedited approval of important new drugs' (Docket 91N-0278, Item no. C47, 2). Whereas David Kessler had cited ddI as the precedent for new rule-writing, the PMA pointed to pre-Subpart H (as a final rule) approvals to argue that the existing mechanisms 'are operating effectively and are adequate to the task, as the recent example of dideoxycytidine (ddC) and dideoxyinosine (ddI) demonstrate. A completely new regulatory approach is not necessary' (2). The PMA also noted that the FDA had a 'long-standing policy of approving cancer drugs on surrogate endpoints, such as tumor shrinkage' and did not begin to modify that process until the 1980s (2). Bristol-Myers Squibb expressed similar opinions (Docket 91N-0278, Item no. C6). In response to these types of comments, the FDA argued that approval of ddC and ddI under existing procedures 'does not show that the rule is of no value' (FDA 1992d, 58944). Although studies having confirmatory value were almost complete when those drugs were approved, for future cases of accelerated approval, the 'provisions of the final rule will ensure that appropriate safeguards exist for timely generation of data on actual clinical benefit, for appropriate promotional

²⁴ For large pharmaceutical companies, new rules represent disruption to existing practices and correlative uncertainty regarding future FDA decision-making. Such changes therefore introduce elements of economic risk in a business sector already attempting to manage the substantial risk associated with new product development. For this reason, even if a new rule appears favourable to industry interests, if existing mechanisms for drug approval are deemed adequate to achieve the same goals, pharmaceutical companies will often oppose rule changes.

information about labeled indications, and for prompt withdrawal of the drug from the market if clinical benefit is not confirmed' (58944).

The Industrial Biotechnology Association (IBA), on the other hand, embraced many aspects of the proposed rule and appeared to believe it an opportunity to exploit the FDA in a moment of unexpected largesse. The IBA expressed concern that when there was no adequate surrogate endpoint available for clinical evaluation of an important drug, accelerated development and approval of the drug would be impossible. In such cases, it argued, the purpose of the Subpart H rule 'to approve drugs based on less information' would be undermined (Docket 91N-0278, Item no. C36, 2). Therefore, it suggested that the FDA should allow study based on a clinical endpoint with a less rigorous standard of statistical significance than normally required: 0.20 or 0.15 rather than 0.05.²⁵ 'Permitting such approvals would be no more risky than allowing approval based on the limited data that were sufficient to approve certain AIDS drugs' (2). The FDA rejected this suggestion bluntly, saying that the intent of the rule is 'to allow FDA to utilize a particular kind of evidence', not to 'place into the market drugs with little evidence of usefulness' (FDA 1992d, 58948). The IBA also wrote that while the language of the proposed rule seemed 'reasonable', the accompanying preamble was 'troubling' because the 'FDA should make it clear that one study could be the basis of approval' (Docket 91N-0278, Item no. C36, 2-3). To this comment, the FDA responded, 'FDA interprets the statute, and good science, as requiring at least two adequate and well-controlled studies to establish effectiveness', and reiterated that where single studies have been accepted, they have been 'of excellent design, showed a high degree of statistical significance, involved multiple study centers, and showed some evidence of internal replicability, e.g., similar effects in major study subsets' (FDA 1992d, 58948). Presumably, they were not thinking of ifosfamide when this rebuttal was written (Chapter 5).

One notable submission in the comments to FDA was that of Medicine in the Public Interest (MIPI). The president of this organization was Louis M. Lasagna, of 'Lasagna Report' fame. In many ways, the comments from MIPI serve as a measure of the extent to which the FDA's version of accelerated approval conformed to the Lasagna Committee's vision for accelerated approval. The first major area of concern for MIPI was the same one targeted by physicians' and pharmacists' trade organizations: the

²⁵ This refers to the p-value in statistical hypothesis testing. By convention, a p-value of 0.05 or less is considered statistically significant.

proposal that the FDA be empowered to restrict the distribution of a drug based on the determination that it ‘can be used safely only if distribution or use is modified or restricted’ (FDA 1992a, 13234). According to MIPI, this provision would create a mechanism for ‘improperly exerting regulatory control over the legitimate practice of medicine’ (91N-0278, Item no. C40, 3) and MIPI accused the FDA of ‘essentially seeking to establish various “classes” of medical prescribers’ (5). MIPI rejected as precedent the example of restricted distribution offered by the FDA in the proposal: ‘the fact that one such commitment was made in one instance in the past does not imply that the FDA has the authority to make it an absolute regulatory requirement’ (6).

The second major area of disagreement between MIPI and FDA on accelerated approval was the issue of *required* postmarket study as a condition of approval. MIPI argued that under the Federal Food Drug and Cosmetic Act (FDCA), a product ‘*shall* be approved if, among other things, it is determined to be safe and effective by the agency’ (8). If safety and efficacy are not sufficient to approve, then the FDA is not authorized by law to grant the approval; indeed, ‘approval must *not* be granted’ (8). However if the evidence is sufficient for approval, then the drug is safe and effective under the law, in which case it ‘must be approved without regard to any condition’ (8). Ironically, this reasoning could be used to argue that the accelerated approval rule should be scrapped and that no approvals should proceed on the basis of unvalidated surrogate endpoints. Indeed this was essentially part of the rationale behind the FDA’s legendary traditional conservatism. As we have seen, it was part of the reason that advisory committees were reluctant to grant approval in circumstances where there was medical need but also some doubt as to medical benefit; once the decision was made, the drug was ‘out there’ and could not be reeled back in except under the most obviously harmful circumstances. Nevertheless, MIPI was using this rationale to argue that accelerated approval ‘should not be legally conditioned on the requirement of costly post-marketing clinical trials’ (10-11).²⁶

While the comments from consumer protection advocacy group Public Citizen complained the accelerated withdrawal procedures specified under Subpart H were not fast enough (91N-0278, Item no. C18), MIPI objected to any abbreviated procedure for withdrawal. Their argument was that under the procedures for drug withdrawal specified by the FDCA for *any* marketed drug, due notice of the pending withdrawal must be given

²⁶ Although the group is called Medicine in the Public Interest, the insertion of the word ‘costly’ in this sentence suggests some alignment with drug companies, who would have to bear these costs.

and the opportunity for a hearing must be provided. While the Subpart H withdrawal process did allow for a hearing, the FDA sought to avoid the courts and instead resolve the case in an evidentiary hearing before an advisory committee — a shortcut to which MIPI objected. MIPI therefore reminded the FDA that it already had recourse to ‘imminent hazard’ provisions of the law which allowed immediate withdrawal of any product found to pose an imminent hazard to public health (91N-0278, Item no. C40, 12). Moreover the comment suggested that, as an alternative, ‘the agency, when confronted with a dangerous product on the market, could certainly request the sponsor to voluntarily withdraw the product’ (13). If a ‘legitimate hazard’ exists, argued MIPI, ‘most drug Manufacturer’s would readily comply’ (13). In response to this latter comment, the FDA reminded the reader of the DESI (Drug Efficacy Study Implementation) where, as we have seen in Chapter 3, drug-makers fought tenaciously in the courts against the withdrawal of drugs that were ineffective or unproven to have benefit (FDA 1992d, 58955). This experience made a powerful and durable impression in FDA collective memory. Drug companies would not consider unproven clinical benefit a ‘legitimate hazard’. Moreover, sponsors always had recourse to the courts, and the proposed rule made it clear that such review could be sought if the sponsor wished to contest the outcome of the investigative hearing.²⁷

Finally also MIPI objected to the stipulation in the rule that sponsors of accelerated approval products must submit all advertising and promotional materials to the FDA before the drug is approved. MIPI saw this as part of a pattern on the part of the FDA. ‘Increasingly over the last two years, the FDA has been criticized for trying to improperly and unduly restrict the flow of drug use information, especially in the context of Continuing Medical Education efforts and communications concerning valuable potential uses of cancer drugs’ (15).²⁸ For its part, the FDA argued that consideration of how the drug will be promoted was necessary as part of its assessment of the drug’s proposed labelling and indication — and as such, a contributor to the risk-benefit assessment. It was part of the FDA’s mandate under the law to assure that the drug’s labelling is not false or misleading (FDA 1992d, 58949).

²⁷ See the proposed CFR section §314.530(f) in FDA 1992a, 13241.

²⁸ According to Angell (2003), dissemination of ‘drug use information’ and ‘Continuing Medical Education’ are simply pharmaceutical company marketing and promotion of brand name drugs to physicians by another name.

At any rate, the FDA's version of the accelerated approval rule contained conditions, powers and restrictions not recommended by the Lasagna report and clearly not advocated by MIPI. Hence, although the Lasagna report embodied many of the demands of AIDS activists and was clearly politically influential, especially because of the centrality of position given to it by the Quayle Commission, the version of accelerated approval ultimately promulgated bore the distinctive marks of FDA safeguards and conditions, born out of previous experience.

All of these objections notwithstanding, FDA made no substantive changes between the proposed and final rules, and was confident enough that the proposed rule would stand as originally drafted that they approved ddC as a combination therapy under accelerated approval six months before the response to comments and final rule for accelerated approval were published.²⁹ Once again, we have to conclude not that the formal consensus process was unimportant, but that the most crucial, fundamentally formative processes of consensus had already taken place by the time the proposed rule was published. This primary consensus was achieved in part through a behind-doors political process and internal deliberation, but it was also importantly achieved through action: real decisions for real drugs in practice. Those practices and the tacit consensus which formed around them provided the basis for the new rules written: once again, formal rule-writing followed in the wake of practical action.

Importantly, just as we saw in the case of Subpart E rule-writing, there were elements of the experiences with ddI and ddC which were *not* written into the Subpart H rule. Perhaps most notably, in both cases the FDA was instrumental in obtaining supplemental interim data to support the NDA which ultimately provided the best evidence for approving the drug. In creating the new rule on the basis of existing exemplars, the FDA could have included a provision that such actions would be taken for all cases of accelerated approval. A demand for such a provision clearly existed, in view of AIDS activists' use of Subpart E to insist that the FDA had an obligation to raid ongoing studies not included in the NDA for ddC. Even so, the FDA chose not to include such a provision in Subpart H for reasons which could be guessed.³⁰

²⁹ The drug combination was approved on 19 June. (See listing under tradename 'Hivid' at <http://www.fda.gov/oashi/aids/virals.html>.) The final rule and response to comments were published on 11 December (FDA 1992d).

³⁰ Among other reasons, the FDA would no doubt argue that even though it went beyond the call of duty in these cases because of the extraordinary circumstances, and although it aims to be as cooperative as

The observation I wish to underscore here is that the experience with ddI (and with ddC when the rule was still a proposal) was constituted by a series of characteristics only some of which were deemed useful to codify for the sake of future practice. We again see a selective extraction of salient characteristics to serve as the definition of {I}. As careful as this selective definitional process presumably was, it was ultimately patterned on the experience with AIDS drugs. We will see in the next chapter that two early applications of Subpart H accelerated approval to cancer drugs (I₂) challenged the definition of {I}, calling for an adaptive approach to concept application.

possible to approve therapeutically important drugs, it is ultimately the responsibility of the drug sponsor to provide valid data as the basis for NDAs.

7. ACCELERATED APPROVAL AND CANCER

As we have seen, the concepts developed and codified in the Subpart H rules for accelerated approval were mainly developed to answer the question: what should happen when a new drug is likely to address an important unmet medical need, but the surrogate endpoints used to study the drug, while thought to be clinically meaningful, have not been proven so? This is the question confronted by the FDA in the late 1980s and early 1990s with respect to development of the nucleoside analogues and other drugs for AIDS. While certainly the Subpart H rules were forged as a result of numerous colliding forces, one cannot help but notice that the specific shape of the rules was framed in response to the specific problems the FDA faced at the time with AIDS drugs. As we will see in the sections which follow, as these rules for accelerated approval began to be applied to other serious conditions where the situation was unlike that encountered with the nucleoside analogues and AIDS, the scope and application of accelerated approval changed, even as FDA staff sought to maintain outward adherence to the frame provided by the rules.¹

7.1 Dexrazoxane: Misalignment of Rule and Application

7.1.1 *Cardioprotection and Tumour Protection*

Dexrazoxane (trade name Zinecard, also sometimes called DZR) was a drug found to mitigate a key form of toxicity associated with a cancer chemotherapy agent called

¹ To my knowledge, the cases described here are the first two cases of accelerated approval as applied to cancer. According to Bazell (1998), there was one earlier application of Subpart H to cancer drugs in the December 1992 approval of Taxol for second-line ovarian cancer. However the FDA does not classify this NDA as a Subpart H accelerated approval. See <http://www.accessdata.fda.gov/scripts/cder/onctools/summary.cfm?ID=123> (accessed 28 Feb 2008). See also the FDA's approval history for Taxol, where there is no indication of accelerated approval or conversion to full approval: <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm> (accessed 28 Feb 2008). It appears rather to have been a priority review, which means that the pending NDA was bumped to the head of the queue for consideration.

doxorubicin. The latter drug, also called adriamycin or sometimes simply ‘dox’, is a chemical cousin of an agent discussed in Chapter 5, daunorubicin. (Daunorubicin was part of the standard therapy for acute non-lymphocytic leukaemia and was the substance replaced by mitoxantrone in the investigational arm of the study.) Both chemicals are part of the anthracycline family of antibiotic drugs, both have been found effective to varying degrees against a number of forms of cancer, and both are associated with significant cardiotoxicity, meaning that the drugs tend to damage the heart muscle, leading to inefficient pumping of blood and adverse effects such as congestive heart failure. Doxorubicin’s toxicity is cumulative; the higher the total received dose, the greater the risk of heart damage. Used since the 1970s for a variety of cancers, the drug was often administered in chemical combinations, in part to reduce the dose of doxorubicin that would have to be given in therapy, hopefully preventing patients from reaching the cumulative dose threshold range at which cardiotoxic adverse effects became likely.

In late 1994, the Oncologic Drugs Advisory Committee (ODAC) met to discuss dexrazoxane (FDA 1994b), a drug purported to mitigate the harmful cardiotoxicity associated with doxorubicin therapy. Significantly, the ODAC had met once before, in 1992, to discuss this drug for this indication and had *denied* the sponsor’s application (17-18). Although the ODAC members agreed that the prevention of cardiotoxicity was clearly demonstrated in the studies presented to support the new drug application (NDA), in the largest of those trials there appeared also to be a *tumour protective* effect. In other words, in one pivotal trial and two confirmatory trials, the sponsor had tested the cardioprotective effect of dexrazoxane on breast cancer patients who were receiving the cardiotoxic doxorubicin as part of their treatment (the ‘FAC’ regimen, composed of 5-fluorouracil, adriamycin [doxorubicin], and cyclophosphamide). In each study, these cancer patients were divided into two groups. Patients in the investigational arm of the studies were given dexrazoxane with the FAC regimen; the control group of patients were given a placebo along with the FAC regimen. In each trial, the group receiving dexrazoxane evidenced significantly less cardiotoxicity than the placebo group (measured in terms of early evidence of trouble shown in heart scans, as well as clinical signs and symptoms of heart weakening). The cardioprotective effect was obvious enough that in 1991 a safety review committee recommended either discontinuing the studies (which had initiated enrolment in 1988) or allowing placebo patients to begin receiving dexrazoxane (20-21). Accordingly, the drug sponsor began administering dexrazoxane to the placebo

patients in January 1991 and used the data collected to that point to support the 1992 drug application. However, in one of the trials (the largest one), the patients receiving FAC with dexrazoxane relapsed *sooner* than those receiving FAC/placebo (measured as time-to-progression, or “TTP”), suggesting that the dexrazoxane was interfering with the therapeutic benefit of the FAC therapy for breast cancer. Under the circumstances, the ODAC voted to deny approval of the drug and the FDA accepted that judgement.

7.1.2 New NDA: ‘Protection from a Risk that Need Not be Assumed’

In 1994 the sponsor was once again before the ODAC with a new NDA for dexrazoxane as a cardioprotective agent. Significantly, the sponsor does *not* appear to have initiated any new clinical trials to support this new NDA, although they did not expend any effort to clarify this point in their presentation to the ODAC. They merely introduced and described two studies of cardioprotection in breast cancer patients, 88001 and 88006, studies which had already been included in the 1992 application and thereafter continued to produce data.² Following the 1991 protocol amendment to move placebo patients to dexrazoxane, those trials were structured in the following manner. One group received the FAC cancer therapy with dexrazoxane as before; the other group received FAC with placebo until the seventh course of treatment, whereupon the placebo patients began receiving dexrazoxane on an open-label (unblinded) basis. Each ‘course’ of FAC took 21 days and included 50 milligrams per square meter of doxorubicin; hence after the sixth course of treatment, the patient had received 300 mg/m² of doxorubicin, the quantity beyond which patients were thought to be at a much higher risk of cardiotoxicity.

Importantly, in this new NDA presented to the ODAC, the sponsor’s analysis was *not* based on comparison of the two trial groups described in the previous paragraph, as one might expect. Rather, in a somewhat complicated orchestration of the available data, the sponsor compared the *pre-1991* placebo group (those who had received placebo with FAC up until the protocol amendment in 1991), called the ‘PLA’ group, to the *post-1991* group which received placebo through the sixth course of treatment and then switched to

² When we examine the labelling for the drug as finally approved (which gives information on the studies used to support the application), we find a description of the three original studies, including the one demonstrating a tumour protective effect. The label specifies that two of these studies were subsequently amended to allow placebo patients to receive dexrazoxane, and then describes precisely the analysis of those two trials provided by the sponsor in the 1994 ODAC meeting. We can therefore conclude that for the second NDA the sponsor merely continued collecting data from two of the three original clinical trials. See <http://www.fda.gov/cder/foi/label/2005/020212s008lbl.pdf> (accessed 24 January 2008).

dexrazoxane, called the 'PLA/DZR' group. This comparison was said to demonstrate the effectiveness of dexrazoxane for cardioprotection when administered *after the sixth course of therapy* in breast cancer patients. The sponsor ignored the hundreds of patients who received DZR from the beginning of treatment, focussing instead on those who did not receive DZR until the seventh course of treatment with FAC. Thus, the two groups compared here are *not concurrent*. In effect, the sponsor was using the pre-1991 placebo group to act as an historical control for the post-1991 PLA/DZR group, which was originally a placebo control group but was now being transformed retrospectively into the investigational arm of the study. More than that, in a move at least one committee member considered highly unusual (44), the sponsor combined the PLA and PLA/DZR groups from two trials and presented the data to the ODAC as if they had come from one large trial. Obviously, the purpose of this data-pooling was to bolster the hazard ratios and p values obtained from the analysis and to double the apparent size of the compared groups.³

Consistent with the previous NDA submission, analysis of the data showed a clear cardioprotective effect for the PLA/DZR group which the advisory committee accepted as valid. Moreover, there was no statistically significant difference in TTP between the two groups. However, the committee approached this latter result with caution, expressing scepticism about the usefulness of FAC therapy beyond the sixth course. Many committee members insisted that any beneficial effects of FAC therapy on breast cancer are seen within the first courses. Nevertheless, in a practice sometimes called 'maintenance' therapy, many physicians continue administering the chemotherapy beyond this point in the hope of maintaining disease stability and forestalling disease progression. Referring to this practice, the Committee's Dr. Daniel Ihde of Washington University cited studies showing that protracted therapy, and even therapy beyond just three courses, did not appear to confer a benefit to breast cancer patients. For this reason, he wondered, 'at least in the trials that are being presented, if this is a protection from a risk that need not be assumed' (58).

Meanwhile, the sponsor was silent on the question of a tumour-protective effect of DZR in the earlier courses of therapy until finally Dr. Temple asked for an updated calculation from the follow-up of the original study. The hazard ratio for TTP in the

³ Dr. Bunn insisted on seeing separate analyses from each trial individually and was eventually satisfied that combining the data did not affect the overall conclusions of the studies (40-51).

original study was 0.6 with a p-value of 0.06, indicating a potentially significant tendency for the patients receiving DZR with FAC to experience disease progression sooner than the patients receiving FAC with placebo (74-5).⁴ The sponsor indicated that after continued follow-up the new hazard rate was 0.8 — still unfavourable, but not as much as before and, according to the sponsor, not a statistically significant finding.⁵ Still, this result hardly instilled confidence that early use of dexrazoxane would not interfere with the beneficial effects of the chemotherapy.

Curiously, the analysis for this new NDA demonstrated an unexpected survival advantage for the PLA/DZR group. This apparent advantage was not due to prevention of adverse cardiac events and could not be explained by the sponsor. The FDA presenter, Dr. Grant Williams, described this result as ‘very surprising . . . and not actually in a positive way’ (98). The apparent survival benefit could be an artefact of bias, and the same underlying phenomenon could be creating a false positive signal in the cardiac-related data as well. Although the sponsor’s analysis went to some lengths to account for any biases introduced by the comparison of non-concurrent groups of patients, Williams’ analysis revealed differences in patient entry criteria and tumour factors (e.g., whether measurable disease was required for entry) between the two groups compared in the analysis (90-1). At the same time, although Williams’ analysis confirmed that there was no difference in TTP between the PLA and PLA/DZR groups, he added a strong disclaimer: ‘The caution I have here is I don't think this is much data. This is a historical analysis of time to progression, and you know, how many times have we rejected historical analyses of time to progression?’ (98) With the small patient population (approximately 90 patients in each group) and differences in between-group prognostic factors, the data could not be relied upon to say what effect on TTP the drug really had. Williams concluded, ‘I think we have to assume that early on it did cause a difference in time to progression, and I think this will just have to wait for a later study’ (98).

The FDA had agreed that if the cardioprotective data were convincing, then accelerated approval could conceivably be granted on condition that the sponsor conduct

⁴ For an explanation of hazard ratio, see the discussion of mitoxantrone in Section 5.1 of this thesis. The p-value represents the probability that the result obtained can be ascribed to chance. By convention, a p-value of 0.05 or less is considered ‘statistically significant’.

⁵ In responding to Temple’s question, the sponsor’s spokesman at this point used the word ‘survival’ rather than ‘disease progression’. However, in other places in the transcript — e.g., the FDA review of the results from the original application on p. 86 of the transcript — it is made clear that the measured parameter was TTP.

an additional study to settle the question of whether the treatment was beneficial after the sixth course. The proposed postmarket study submitted by the sponsor would take 200 patients who had already received 300 mg/m² (based on the patient's body surface area) of doxorubicin and randomize half to doxorubicin with dexrazoxane while the other half of patients would receive no follow-on therapy whatsoever (157). As the FDA's Dr. Temple framed the issue: 'the reason we're talking about the accelerated approval is that we're not convinced that the decreased cardiotoxicity that is accompanying this particular regimen provides a net benefit to not treating at all' (133). He also stated that this situation fit the definition of accelerated approval because Subpart H authorizes approval not only on the basis of a surrogate endpoint in need of validation, but also on the basis of a clinical endpoint other than survival or irreversible morbidity for which the relationship to ultimate outcome is in question (158-9).⁶

Even so, the basis of this application seems unusual, to say the least. The drug had been previously rejected because of an apparent tumour-protective effect. This new NDA did not resolve the original question of tumour protection, but simply moved the administration of the dexrazoxane to a later, possibly unnecessary and inefficacious, phase of therapy. This new NDA merely served to reconfirm the cardioprotective function of dexrazoxane without addressing the original reason for rejection of the drug, tumour protection — nor, it seemed likely, would the follow-up study. In the effort to make the latter question irrelevant by delaying the administration of dexrazoxane until the seventh course of therapy, the sponsor only managed to create a new question regarding the medical relevance of FAC therapy beyond the sixth course — and by extension, the medical relevance of dexrazoxane in that setting. In a bizarre turn of events, therefore, the Phase IV study commitment for DZR was to determine whether there was a beneficial chemotherapeutic setting in which it could be used.

Some of the ODAC members expressed bafflement at the type of 'confirmation' reflected in this postmarket study. Dr. Ahmann complained about the questions posed by the FDA to the ODAC, saying, 'the question said, gosh, let's approve this in a hurry, and, number two, let's make them do a study to make sure that we approved it for the right reason or should have approved it' (163). For Ahmann, there was no need to rush. 'I don't know why we have to put our foot to the floor on an accelerator when, in fact, there

⁶ See Chapter 6. See also FDA 1992d, 58958, which gives the text for the *Codes of Federal Regulation* §314.510; also Dr. Temple's explanation in FDA 1994b, 158-9.

are that many uncertainties' (163) Dr. Temple's response was that accelerated approval was designed for circumstances in which one has 'pretty good reasons to believe that the drug will be useful and valuable' but there are also lingering uncertainties in need of clarification (164-5). For Temple, the evidence of cardioprotection was ample to justify accelerated approval, even if the question remaining was when and whether it was useful to use the drug for breast cancer therapy. Dr. Ihde expressed scepticism about protracted chemotherapy, saying that he asked some colleagues why they continue to treat metastatic breast cancer up until disease progression, and that the response was 'because the patient wants treatment' (145). Dr. Temple's response was, 'That sounds like a good thing to study' (145).

7.1.3 *Incongruity between Rule and Application*

The accelerated approval ultimately went forward, perhaps in part because of an awareness of the many possible applications of this agent not only within breast cancer, but for other cancers as well. Dr. Bunn noted that dexrazoxane had a cardioprotective effect in the treatment of small cell lung cancer (107). Dr. Temple recalled a 'dramatic' study in which children with lymphoblastic leukaemia experienced complete remissions on an anthracycline-based treatment but were 'nevertheless killed off by their therapy' (110). Why, he wondered, had this sort of disease setting not been chosen to study the protective effects of dexrazoxane? In a different context, Temple commented that 'we've always been puzzled . . . by the focus on breast cancer where benefit is, you know, somewhat less certain than in a number of other places, and therefore differences are hard to spot, but that is where the bulk of the data are' (130).

As one of the first two cases of accelerated approval considered by the ODAC, the medical necessity behind the use of the Subpart H rule in this case is unclear. While, granted, there was no cardioprotective agent to be used with anthracycline-based chemotherapy at this time (hence one could argue medical necessity), doxorubicin had been used for years in a variety of conditions and physicians had developed techniques for limiting cumulative doses of doxorubicin, monitoring patients closely, and managing any adverse effects. As Dr. Temple himself noted, 'The whole regimen is designed not to kill patients with congestive heart failure. That's why they get all of these scans. So you only get a death if you've screwed up' (72-3). Hence, he added, 'the main benefit, if indeed it is a benefit, is that you keep on giving the DOX' (73). According to Temple, 55 percent of

patients who stopped therapy did so because of the effects of doxorubicin. Thus, the main benefit of dexrazoxane appeared to be that you could give *more* doxorubicin. But, as we have seen, in this disease it was questionable whether giving more 'DOX' was really associated with any benefit and this became the issue to be resolved in postmarket study.

More significantly, the basis for granting conversion from accelerated to regular approval in October 2002 is also unclear.⁷ An advisory committee meeting on the confirmatory studies does not appear to have been convened, making it difficult to gain insight into the study used to support the standard approval. Normally, this type of information would be included in the clinical study section of the drug's label. However, the most current labelling of the drug (May 2005)⁸ only provides information on the original three studies from the rejected 1991 NDA and the two post-protocol amendment studies discussed in the 1994 ODAC meeting. Additionally, the label contains a warning that there is 'some evidence that the use of dexrazoxane concurrently with the initiation of fluorouracil, doxorubicin and cyclophosphamide (FAC) therapy interferes with the antitumor efficacy of the regimen, and this use is not recommended' (6). This warning section of the label reviews the results from the original studies rejected by the FDA⁹ and then concludes that 'ZINECARD should only be used in those patients who have received a cumulative doxorubicin dose of 300 mg/m² and are continuing with doxorubicin therapy' (6). More recent studies confirming this conclusion were not cited, leading one to wonder what evidence might have supported the conversion to regular approval. A series of searches in Medline did not unearth any published studies of the design proposed for postmarket study in the ODAC meeting, although two other similar studies expressed at best ambivalence towards the benefit of protracted therapy (Falkson,

⁷ The sponsor, Pharmacia, submitted to FDA a supplemental application on 28 December 2001 which, according to the October 2002 FDA acceptance letter, provided 'support' for conversion of dexrazoxane from accelerated to regular approval. <http://www.fda.gov/cder/foi/appletter/2002/20212se7-004,005,006ltr.pdf> (accessed 24 January 2008). Note that the approval history listing at the Drugs@FDA website erroneously identifies the 1995 approval as a standard approval and the 2002 approval as an accelerated approval. (The approval history can be found by searching for Zinecard at <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm>, accessed 24 January 2008). The linked materials, however, confirm reversing the dates. The October 2002 letter granted the conversion from accelerated to standard approval. The 1995 weblink leads to the approved oncology drugs page (<http://www.fda.gov/cder/cancer/druglistframe.htm>), where the 1995 approval is listed as an accelerated approval and indicates that clinical benefit was subsequently established, but does not say when or on what basis.

⁸ <http://www.fda.gov/cder/foi/label/2005/020212s008lbl.pdf> (accessed 24 January 2008).

⁹ 'In the largest of three breast cancer trials, patients who received dexrazoxane starting with their first cycle of FAC therapy had a lower response rate (48% vs 63%; p=0.007) and shorter time to progression than patients who did not receive dexrazoxane' (see p. 6 of the label).

et. al. 1998; Gregory, et. al. 1997) (and one publication — Swain, Whaley, and Ewer (2003) — merely reiterated the results of the original three studies, minus any mention of tumour protection). An FDA slide presentation on accelerated approval lists the ‘confirmation of benefit’ of Zinecard as ‘reduction of CHF’, meaning reduction in congestive heart failure.¹⁰ Perhaps most bizarre of all, a paper on accelerated approval published by FDA staff members indicates that the endpoints supporting accelerated approval for dexrazoxane were cardiotoxicity and response rate to assess potential tumour protection (Johnson, Williams and Padzur 2003, 1406). The tabulation in which this data was given implied that these were the parameters validated in confirmatory Phase IV study. However, as we have seen, the issue of tumour protection was actually side-stepped in the ODAC meeting while the labelling still contains warnings about the tumour protective effects of dexrazoxane. And, in any event, cardioprotection need not have been validated since this had already been well established in all of the studies presented to the FDA, including the ones supporting the rejected NDA. All of the available evidence suggests that if a postmarket confirmatory trial was done, it confirmed what was already proven — that dexrazoxane was effective for cardioprotection when administered with FAC — but did not address the question of the effectiveness of FAC therapy beyond the sixth course, much less the question of tumour protection.

It is tempting to speculate that a postmarket study on the effectiveness of FAC therapy in breast cancer beyond the sixth course may have yielded inferior results and that the FDA, frustrated with the fact that the drug would likely have been approved immediately had a different disease context been chosen to evaluate dexrazoxane, decided to grant full approval based on other evidence. Or, judging from the label information, the FDA may have subsequently decided that no additional evidence was necessary and therefore simply decided to approve on the basis of twice-proven cardioprotection and give a warning about early use in FAC therapy. Clearly, as we have seen, many patients experience cardiotoxicity during chemotherapy, not only from doxorubicin but from other agents as well. The ability to administer anthracycline-based therapies in high cumulative dosages may indeed extend the therapeutic options available to oncologists in a number of diseases and thereby represent a net benefit to patients. Ironically, as we saw with AZT (Chapter 4) and also with ifosfamide (Chapter 5), whereas off-label uses often

¹⁰ [http://www.fda.gov/ohrms/dockets/ac/05/slides/2005-4191S1_02_FDA-Dagher.ppt#272,13,Confirmation of Benefit](http://www.fda.gov/ohrms/dockets/ac/05/slides/2005-4191S1_02_FDA-Dagher.ppt#272,13,Confirmation%20of%20Benefit) (accessed 24 January 2008).

pose a significant concern for advisory committee approval decisions, here they may have contributed to the acceptance of the drug, despite the fact that consideration of off-label uses of any drug is putatively outside the realm of drug approval decision-making.

Whatever the ultimate basis of full approval, the point here is that this early application of accelerated approval to oncology looks very different from the AIDS model on which it was based. Dr. Temple indicated that accelerated approval applied in the case of dexrazoxane because this was an approval on the basis of an endpoint other than survival or irreversible morbidity, but this part of the rule does not fit neatly with the situation at hand. When the FDA wrote this particular provision of the Subpart H rules, they envisioned a therapy which had an apparently beneficial *secondary* effect but had not been proven to have the *primary* benefit sought, such as survival. The AIDS-modelled example given in the proposed rule was a hypothetical anti-HIV drug which augmented weight gain and reduced the incidence of opportunistic infections, but for which the effect on long-term survival was not yet known. ‘While the favorable findings would be a basis for approval’, the Agency wrote, ‘in some instances additional study may be necessary to clearly determine long-term effects’ (FDA 1992a, 13235-6). In the case of dexrazoxane, however, the endpoint — reduction of cardiotoxicity — was in fact the primary clinical benefit sought in testing the drug, and it had been well established. No additional validation of this outcome was required. But because the drug itself was secondary to the primary goal of cancer therapy, and because of the specific disease setting in which this drug was studied, other questions were raised relating to patient benefit. So, had the Phase IV studies been successfully accomplished as planned (or assuming they were), the ‘confirmatory’ data collected would have ironically had nothing to do with the *primary* benefit of dexrazoxane (cardioprotection), but would nevertheless have supported the drug’s conversion to regular approval. Emphatically, my purpose here is not to judge whether this application of the rule was ‘appropriate’. My point is that in this first application of Subpart H to the realm of cancer drugs, we already have a situation not precisely foreseen by the rule, requiring an adaptive approach. Put in the terminology used in Chapter 1, this addition to {I} of a new exemplar was noticeably different from existing exemplars, and demonstrates the open-ended nature of the process, particularly in the early stages of developing the set {I}.

7.2 Liposomal Doxorubicin for AIDS-Related KS

7.2.1 *Defining 'Response' in KS*

As we saw in the previous section, doxorubicin is active in many forms of cancer. In 1995, the ODAC met to discuss an NDA sponsored by Liposome Technology, Inc. (LTI) for the use of doxorubicin in patients having AIDS-related Kaposi's Sarcoma (KS) who had failed prior combination chemotherapy either due to progression of disease or unacceptable toxicity (i.e., second-line treatment) (FDA 1995). KS is a painful and disfiguring form of cancer with symptoms which call to mind the tribulation of the Biblical character Job. This cancer is expressed most visibly as abnormal tissue growth in patches under the skin. These raised reddish, purple or brown skin lesions appear especially on the face, legs and feet, and in the early years of the epidemic were a readily identifiable marker of AIDS-related disease, contributing to the pariah-like status experienced by many AIDS patients. Moreover, abnormal tissue growth can take place in the mucosal lining of the throat, nose, mouth, or anus, and can involve internal organs, especially the lungs, liver and gastrointestinal tract. These growths are associated with oedema (fluid retention and swelling) in the affected areas and pain. Thus lesions on the feet can make it excruciatingly painful for patients to walk; lesions in the mouth or throat make it painful to eat; involvement of the gut can lead to bleeding and blockages; and oedema in the lungs hinders breathing and if not resolved can lead to death. KS had been considered a relatively rare disease until the advent of AIDS, when it became a common complication associated with the immune system devastation characteristic of AIDS.

While doxorubicin was a familiar therapeutic agent to the participants in the meeting, the form of drug delivery being proposed was unique. In traditional intravenous doxorubicin administration, the drug circulates in the blood with only a fraction of the agent delivered to the tumour. In this new preparation, the doxorubicin was encased in liposomes — liquid-filled spheres made of phospholipids with a cholesterol membrane. To prevent premature breakdown by blood plasma, linear chains of polyethylene glycol molecules are attached to reactive sites on the exterior surface of the sphere. Effectively guarded with a polymeric shield, the spheres are able to evade attack within the blood plasma and continue to circulate until they eventually make their way into the tumour's vascular inlets, becoming chemical bombs embedded in the tumourous structures. Here, where the environment is more hostile than in the blood plasma, the liposome's membrane finally degrades allowing delivery of the doxorubicin payload directly to the tumourous tissue. According to the LTI's chief scientist, Dr. Frank Martin,

pharmacokinetic and tissue distribution studies had shown that liposomal preparations of doxorubicin acted very differently than the traditional mode of infusion, remaining in circulation in the body much longer and showing higher concentrations of the drug in KS skin tumours than in normal surrounding tissue (FDA 1995, 53). LTI assigned the trade name 'DOX-SL' to this liposomal preparation of doxorubicin.

Dr. Susan Krown, of Memorial Sloan Kettering Cancer Center (a sometime member of the advisory committee now speaking to the committee as part of the LTI presentation team) indicated that there was a clear need for second-line therapy because there were no therapeutic alternatives for patients who did not respond to first-line therapy. Indeed, there was a dearth of literature on refractory patients; Krown could only find three studies in the literature which even mentioned prior treatment, and these three studies together only represented a total of 28 patients (57-8). Moreover, there was no curative therapy for KS and all existing treatments were useful only for palliation of symptoms (59). 'In practical terms,' she said, 'I take this to mean that what we currently call a successful treatment for KS is a treatment that improves the functional or emotional well-being of the patient without causing excessive toxicity or exacerbating the underlying HIV infection or its consequences' (57).

In such a disease, how was one to measure response? One obvious choice was a reduction in cutaneous lesions (measured as the number or surface area of lesions). However, one could also include other factors such as the size, thickness, and colour of lesions (cosmetic factors); the quantity of associated oedema; the involvement of internal organs; and the reduction of pain. Improvements had been made in assessing total lesion bulk, but said Krown, instead of asking 'how are the lesions doing?', maybe we should have been asking, 'how is the patient doing?' (58). Less visible lesions, the ability to put on one's shoes and walk to the corner shop, and easier breathing or reduction of pain were all factors which would contribute to an overall improvement in the patient's mental attitude and sense of well being. However, there were no validated quality-of-life instruments designed specifically for AIDS-related KS. (58). For these reasons, the assessment of patient response would be a vexing issue for this NDA.

In the original protocol, patients were to be assessed using criteria similar to standards published by the AIDS Clinical Trial Group (ACTG) in 1989. This method of assessment called for an overall estimate of the lesion numbers and degree to which the lesions were raised, with effectiveness of any therapy to be judged according to a whole-

body assessment of patient improvement, including reduction of lesion size and bulk, involvement of viscera and internal organs, and extent of oedema (61). After the protocol was initiated, the ACTG modified their recommendations out of a fear that the original assessment method allowed for excessive subjectivity on the part of the examining physician. FDA expressed the same concern about assessment procedures in the protocol. Therefore, LTI retrospectively created another set of response criteria based on 'indicator lesions'. Five lesions were chosen and assessed for size, colour, and thickness. The condition of these lesions would be tracked throughout the study period. This indicator lesion method of assessment, although retroactive, was presented in the NDA as the primary endpoint of the study (134). An overall investigator assessment was also made involving an approximation of the total number of lesions, including the number of raised lesions, assessment of extent of oedema as well as lesion-associated pain (60-1). This 'investigator assessment' was presented in the NDA as a secondary endpoint (134). A partial response was defined as a 50% reduction in overall lesion bulk, or a flattening of 50% of the lesions, or if the calculated size of the indicator lesions had decreased by 50%. The measured response had to endure for at least two consecutive assessments made 21 days apart (62). As we will see below, neither system of assessment was unambiguous.

7.2.2 An Accidental Pivotal Trial

Another vexing issue would come from the nature of the study itself. The NDA was based on a single pivotal study (designated '30-12') which, as explained by Dr. Craig Carpenter, another sometime ODAC member, was not originally designed to be a pivotal trial (44). Originally two controlled trials with a third, backup 'trial' were planned (44, 66, 80). The first study, 30-10, was a randomized two-arm trial comparing DOX-SL to the chemotherapy combination adriamycin, vincristine, and bleomycin (ABV) (recall that adriamycin is doxorubicin without the liposomal delivery system). The second, 30-11, was another two-arm randomized trial comparing vincristine and bleomycin (BV) to DOX-SL. When patients on the ABV or BV arms of either of these randomized studies failed to respond to therapy, they were to be switched over to study 30-12 for second-line 'salvage' therapy with liposomal doxorubicin. Also, 30-12 would admit patients pretreated on other chemotherapeutic KS trials besides the ones planned by LTI (44). In effect, said Carpenter, this study was intended as a compassionate protocol for patients failing the first-line therapy (94). However, as word of mouth spread amongst AIDS patients and

their physicians that liposomal doxorubicin seemed to relieve KS symptoms, enrolment in 30-12 swelled and accrual to this group exceeded expectations long before the randomized studies comparing DOX-SL to BV and ABV were complete (44). Although there was a protocol for 30-12 specifying entry criteria for the patients to be admitted, study managers often waived certain entry criteria for patients who failed previous therapy (most often bypassing a waiting period of at least 28 days after treatment with radiation or chemotherapy prior to study entry) with the permission of IRBs on a compassionate basis (89-90, 95-100). Hence, the 'study' LTI brought before the ODAC was an unrandomized group of patients having no concurrent control group and inconsistent standards for admission.

The NDA submission contained a total of 383 patients, however close inspection of those patient records revealed a total of only 77 who received sufficient first-line therapy to be considered refractory to standard treatment or intolerant of it by protocol criteria. Based on the indicator lesion response criteria for these 77 patients, the drug appeared to be reasonably efficacious (67-71).¹¹ Moreover, cardiac assessment of five selected patients demonstrated that DOX-SL was no more cardiotoxic than non-liposomal doxorubicin (78-9). Therefore, concluded LTT's Dr. Mamelok, 'KS patients who have failed prior combination cytotoxic chemotherapy, due either to disease progression or toxicity, achieve a meaningful response rate and derive clinical benefit from Dox-SL that outweigh the risks of therapy' (79).

Perhaps so, but the Committee's questions made it clear that there were substantial flaws in the data presented. The ODAC's Dr. Paul Bunn wanted to know how many of the 77 patients used as the basis for the efficacy evaluation were actually eligible to enter the study according to the protocol (88-91). Surprisingly, the sponsor could not give a precise answer. 'In terms of meeting every entry criteria [*sic*] and every exclusion criteria, I think we will have to get back to you on that' (92). The sponsor admitted that 16 patients had prior therapy less than 28 days beforehand; 'several' patients had anaemia or neutropenia (92); one had thrombocytopenia. Approximately 25 patients altogether did

¹¹ Sixty-nine of 77 patients had at least one indicator lesion with unfavourable colour at baseline. Of these 69 patients, 55% experienced elimination of unfavourable colour. Twenty-eight of 77 patients had indicator-lesion oedema at baseline. Of these 28 patients, 77% experienced resolution of oedema. Seventy-two of 77 patients had at least one raised indicator lesion at baseline. Of these 72 patients, 50% experienced flattening of raised indicator lesions. Of 77 patients, 37 had moderate or severe pain at study outset. Fifty-one percent of these patients had a reduction of pain (to mild or absent) over at least one cycle of therapy.

not meet the entry criteria – but still no information was forthcoming on how many of the eligible patients responded to DOX-SL, a point Dr. Bunn pressed repeatedly (92, 93, 96-7, 99-100, 101), but the company had never done that analysis. The Committee Chair, Dr. Schiffer, also noted numerous gaps and omissions from the data, including how long beforehand refractory patients had been treated with adriamycin, and what the previous dose had been (98). He wanted to know which responding patients had previously received adriamycin, since it would be difficult to tell if the response was due to DOX-SL or adriamycin (108). He wanted to know how the indicator lesions had been selected (115). Committee members commented on the limited data collected on internal organ involvement of the disease (106), the lack of before-and-after photographs to document the lesion bulk and appearance of lesions (105), and the lack of Karnofsky performance scores to assess overall patient condition and mobility (113).¹² Said Dr. Abrams, ‘I just want to say that even in the wacky and weird world of nucleoside analogue therapies, we have adopted Karnofsky performance scores as a criteria [*sic.*] for evaluating responses and some drugs have been approved on the basis of Karnofsky performance status alone’ (113). He added that the absence of Karnofsky scores and before-and-after photographs in this application was ‘worrying’ (113). Schiffer nevertheless suggested that despite the faults of the study, the committee should seek to identify a subset of patients who benefited from DOX who had not previously benefited from adriamycin. If such a group could be identified, then the FDA would ‘temporarily approve this drug until the results of randomized trials are out’ (98).

The FDA analysis of the data by Dr. Murgu was even more illuminating of all the data flaws and gaps. For example, many physicians kept patients on the study after disease progression, hoping the patient might benefit from continued therapy. As a result, some patients were classified in the case reports as having ‘stable disease’ *after* disease progression had occurred. Likewise, since the indicator lesion method of assessment was retroactively applied (not part of the original protocol), patients who experienced disease progression according to this method of assessment would have remained on study as

¹² The Karnofsky performance scale measures the ability of patients to perform certain ordinary tasks such as walking, caring for themselves, etc. The scale ranges from zero to 100: a patient with a score of zero is dead; a patient with a score of 100 is fully functional and shows no signs of illness. Patients with scores below 40 are bedridden and severely diseased; scores between 50 and 70 indicate patients who are moderately diseased but able to live at home with some level of assistance. The scale and its application can be found in many places. See, for example, the table and references available at <http://www.hospicepatients.org/karnofsky.html>.

well (135). Moreover, many of the patients began the study having less than five indicator lesions, making the basis of accounting for partial responses inconsistent (136-7).¹³

One of the most serious problems, however, was the lack of control group for comparison. The FDA's Murgo pointed out that comparisons to other, previous studies were not possible because of the lack of historical studies on refractory patients (132). However in cases where the disease was serious or life-threatening and the therapeutic options were lacking, one could consider efficacy data against the background of a well documented disease trajectory (133). For this reason, said Murgo, when LTI told the FDA that they sought to use an open-label study as pivotal, the FDA informed LTI of the 'rationale and the importance of documenting pretreatment status' of each patient, since this pretreatment status would be used as the historical control (133). Despite this clear message, the case reports produced on the patients were inadequate to document their prior therapy (138). Alternatively, LTI provided to Dr. Murgo the individual medical charts for 71 of the 77 refractory patients on the study. Dr. Murgo produced a detailed analysis for the ODAC on the basis of those charts (138-53), examining previous patient therapy, reasons for going off previous therapy, various measurements of patient condition and response, entry criteria, response criteria, and other factors.

Ultimately, Murgo's assessment of the data was medical rather than statistical. He was trying, he said, 'to make some global assessment, using Dr. Krown's term "how the patient was doing" regardless of whether they were responding by the criteria that we just discussed, regardless of whether they were evaluable or nonevaluable' (153). Based on that broad medical assessment of individual patient charts, Murgo identified six patients for whom 'there was substantial evidence that the patient was benefitting from the Doxil treatment or probably benefitting from the Doxil treatment' (153-4). Then, 'at the risk of seeming anecdotal in front of this Committee' (155), Murgo proceeded to describe the individual case histories of each of those six patients (154 - 157). 'Here is the bottom line', he concluded: 'The one comment that I would like to make before the Committee has its discussion is to consider less the rate of response and more that the patients who did respond had very limited treatment options' (158).

7.2.3 A Question of Legal Standards

¹³ For example, a patient having only two indicator lesions would need for only one of those lesions to become reduced in size or thickness to be considered a partial response.

Confronted with these six cases, the Committee struggled to put a denominator to the numerator of six. Dr. Gelber asked if it was six out of 77 (159). Dr. Murgo replied that this ‘might’ be an interpretation of the results, but that the figure could not really be extrapolated to a larger population. ‘I really can’t even answer the question’ (160). Dr. Ingle asked about the criteria for judging benefit: ‘to have a clinical benefit they had to have an improvement in pain or clear decrease in symptoms. And if that is the case, what is your denominator? Is it 6 out of 40, or 6 out of 71 or —?’ (161-2). Murgo responded: ‘That is why I hoped that the rate of response would not be too much of an issue here. And it applies to that clinical benefit as well because I don’t know how to come to a proportion’ (162).

The ODAC’s Dr. Harwood noticed that four of Dr. Murgo’s six ‘responding’ patients did not meet eligibility requirements for the study. However, the FDA’s Dr. Temple had earlier noted that ‘not all failures to be eligible have the same implications’ (98) and that some patients, while ‘technically ineligible, are not useless’ (99). Following Murgo’s presentation, Temple took up the issue again, highlighting the fact that Murgo’s approach was only to exclude ineligible patients if the reason for the ineligibility could confound the assessment of therapeutic benefit (165-6). If, for example, an improved lesion had been treated the week before with radiation, then the benefit in that case could not be assessed. Murgo, said Temple, had taken an ‘extremely conservative view of clinical benefit’ (166). In fact, there were *at least* six patients who experienced benefit and there could be more (167). He mused that, ‘once again, one is, as usual, struck by the irony that in an attempt to do something in the most efficient way, the most compassionate way, you end up getting data that has to be interpreted so conservatively that the results, on the whole, are less impressive than they might well be’ (167).

By far the most serious struggle for the committee was whether the study met the legal standard for approval. Dr. Abrams posed the question to Dr. Murgo: ‘Can you explain to me how, in your impression or interpretation of the data, this is a controlled study?’ (170). Murgo reiterated the idea of using patient baseline status as an historical control. Dr. Temple added his own refinement to the concept, saying that the ‘baseline control idea is something of a misnomer (171). Rather, the control ‘is really what you think would have happened to that group had they been untreated. . . . So that expectation and knowledge becomes the control group’ (172). Despite this rather lenient interpretation of a controlled trial, when the advisory committee went to vote on the first

question presented to them — whether 30-12 was an ‘adequate and well-controlled investigation’ — many on the committee balked. Dr. Bunn said it was ‘barely’ adequate or well-controlled (173). Dr. Abrams said that this was more like an expanded access compassionate use protocol (173), adding: ‘I would say that if the Committee chooses to believe so, that there might be some evidence that this is controlled. But, in my opinion, it is not adequate or well controlled.’ (173). Dr. Ingle commented that the ‘adjectives here’ (meaning ‘adequate’ and ‘well-controlled’) were getting in the way. ‘Is it adequate and well controlled? No.’ But he said, it is ‘interpretable for its purpose’ (175). The study has ‘a lot of problems’, he said, but so does the disease (175). Dr. Temple intervened: ‘I just need to remind everybody that you can’t tell us that it is not a well-controlled study but that it should be approved. So if you are thinking of doing that, don’t bother’ (175). Dr. Ingle countered: ‘You define well-controlled, then’ (175). Again, Dr. Temple appealed to knowledge, saying that the regulations allow for ‘an historical control where you can say with assurance that you know what would have happened in the absence of treatment’ (175). He then added: ‘All of the drugs this Committee has recommended approval for, for refractory disease in cancer, have been based on historical controls. All is an overstatement, but most don’t have a randomized control group.’ (176). So, ‘the control group is what would have happened in the absence of therapy and the knowledge that the tumor would have gotten bigger and the patient would have done worse. That is the control group. And it has been considered reasonable where one knows the natural history of the disease with assurance’ (176). Thus, he said, it is up to the Committee to decide how well that natural history was known. Moreover, he added, the same could be said for a study describable as a series of cases (176-7) — in other words, in a series of cases involving refractory cancer in which the natural history of the disease is well understood, that series could be considered ‘well-controlled’. That conclusion is necessary, said Temple, ‘because it is the only basis for approval under law that it is a well controlled study. So we swallow hard, and in some of these cases, we can do that’ (177).

We have already seen in the previous chapter Dr. Temple’s frequent role as a voice of flexibility in drug approval decision-making, and this is perhaps the most extravagant example yet. In the case of Fludara, discussed in Chapter 5, he accepted the ‘patient baseline’ model of an historical control, saying that the FDA has ‘considered such [historically controlled] trials to include both those in which a historical experience is identified and those in which the patient’s baseline is used as his own control’ (FDA

1990d, 89). Here, he moved beyond the ‘patient baseline’ model of historical control to a definition in which the knowledge and expectations of qualified experts is what constitutes the ‘control’. More than that, he skirted tantalizingly close to articulating the idea that the law simply doesn’t contemplate situations like these and that it is therefore appropriate to ‘swallow hard’, bend the truth, and call a series of cases ‘well-controlled’ if, based on one’s personal expertise and experience, one believes that certain refractory cancer patients with no therapeutic options benefited from a therapy. In this definition of ‘control’, the individual ‘medical judgment’ therapeutic reformers sought to banish through systematized techniques (Marks 1997) is seen to be not only alive and well, but implanted in the definitional heart of the techniques designed to vanquish it.

The Committee was stumped. Although, as noted by Dr. Ozols, many community physicians were asking for approval because there was ‘a tremendous feeling . . . that this thing works, to some degree’ (186-7), there was little basis for that conclusion in this NDA. Chairman Schiffer put it bluntly, saying that while the disease was difficult to evaluate (a point made several times by others), ‘it is not entirely rocket science’ (187-8) and ‘[t]he fact of the matter here is a third were not eligible, a third were not evaluable, and Dr. Murgio had to dig really deep’ (188). In Schiffer’s opinion, it was poor data. Dr. Siu wondered if the question of adequate and well-controlled could be somehow ‘cloaked in saying that this was adequate and well controlled in the construct of accelerated approval and knowing that there would be post-marketing studies conducted’ (188). His question went unanswered. Dr. Abrams suggested putting off the vote on this question and discussing later questions. Dr. Temple reminded him that all the subsequent questions flow legally from this one (189).

The final vote on the question was an unusual tie: five votes in favour, five votes opposed, and one abstention. On this most fundamental question, the FDA did not have a decisive vote. Some of those voting against declaring the trial adequate and well-controlled indicated in various ways, however, that they nevertheless were favourably inclined towards the drug. Dr. Abrams was the most unambiguous on this point, saying, ‘my feeling is, having come out so strongly with my comments about adequate and well-controlled, I personally believe that this drug should be made available on accelerated approval. So I will vote no on the first question but hope that my adamancy will not influence the Committee to any great extent’ (184).

This attitude perhaps contributed to the reasoning behind Dr. Temple's next intervention. When the Committee began to deliberate the second question — whether the study demonstrated substantial evidence of efficacy — ODAC members still struggled with the ambiguity of the data. Dr. Bunn commented that he believed there was substantial evidence that there was a 30 percent partial response rate, but that he did not know how to interpret that finding. He said, 'to me, this really is a place where an accelerated approval might be worthwhile because, later on, in some randomized trial, you might actually figure out what that means' (193). Bunn explained that this was why he balked on calling this a controlled trial. 'In a controlled trial, you ought to know what that means. You see what I am saying — because of the natural history — but, anyway, I think that the lesions that were measured and gave you a 30 percent response rate. . . . What I don't know is what it means' (194). It seems clear that Bunn was talking about the lack of an adequate control against which to compare the result, given his reference to what the result means in terms of 'natural history'. But Temple interpreted Bunn's statement in terms of uncertainty associated with a surrogate endpoint: 'What I actually hear you saying is that you think the study is adequate to define some kind of response rate, but that you are not sure of the clinical endpoints, like pain, edema and all that sort thing, are well evaluated by this study' (194). Then, he applied this interpretation to the split vote on the first question. 'I have to tell you I will slightly understand what you are saying now to modify the first vote which is that you might find it a reasonable study or well-controlled study even on the question of tumor response but you are not persuaded it told you much about the clinical endpoints' (195). There were no objections from the committee. Moreover, with the assurance that substantial evidence of efficacy could mean substantial evidence of a partial response in six patients (197-8, 200), the Committee voted that the study provided the required substantial evidence (10 in favour with one abstention).

7.2.4 *The Meaning of 'Accelerated Approval' and 'Historical' Control*

The remaining questions for the Committee— whether the benefit of DOX-SL outweighed the risk, and whether the drug should be approved — were dispatched more quickly. What is notable here and elsewhere in the discussion is the Committee's attitude towards the purpose of accelerated approval. Earlier in the meeting, Dr. Temple clearly warned that if the committee thought the data were 'rubbish' and hoped that a better trial

was immanent, ‘that is not what accelerated approval was for’ (180). Moreover, as we have seen, Temple sought to frame the results of the first vote in terms of confirmation of a surrogate endpoint. Committee members accepted this framing and also seemed to accept that (again in Dr. Temple’s words), ‘accelerated approval was not meant to be a means by which inadequate data could be a basis for approval’ (179). Nevertheless, as the ODAC members discussed their thoughts and reasons for voting a certain way, accelerated approval was clearly perceived by some members as more than confirmation of a surrogate endpoint. We have already seen that Dr. Schiffer thought of this as a ‘temporary’ approval ‘until the results of randomized trials are out’ (98). Moreover, when considering the risk/benefit profile of the drug, Dr. Bunn said, ‘I have been equivocating and I am going to continue to equivocate. That is why I like to [*sic*] accelerated approval because there is some evidence, and whether it is sufficient or not is a tough one’ (203). Moreover, despite Dr. Abrams’ earlier insistence that the study was inadequate and uncontrolled, he remarked that ‘this drug is a perfect candidate for accelerated approval under the FDA guidelines’ (204), adding that he wanted to see the data on the 30-10 and 30-11 studies, as well as new trials on ‘earlier’ patients to assess overall efficacy. Embracing Dr. Temple’s interpretation of the first vote, Abrams added that ‘our approval on the basis of the indicator lesions is clearly a very good example of a surrogate marker’ (104), perhaps leading one to believe that he had repented of his earlier negative view of the pivotal trial. But then he remarked: ‘I do not believe that this was an adequate and well-controlled study let alone two of them, so I don’t believe that I could vote for unconditional approval’ (205). Similarly, when Chairman Schiffer called for a vote on accelerated approval, he also called for suggestions on the conduct of additional trials, ‘because, frankly, we have all been appalled by what we have heard today. There is no secret about that’ (209). Finally, after the unanimous vote for accelerated approval (with one abstention),¹⁴ Schiffer’s recommendations to the FDA included the suggestion that ‘this drug should not be given out liberally without the collection of adequate data’ (210-11) — as if the drug would not be distributed widely under accelerated approval. Many committee members clearly saw this as a ‘compassionate’ approval on the basis of a ‘compassionate’ trial having inadequate data pending controlled studies. Schiffer summed

¹⁴ The group voted eight to three that the benefit outweighed the risk (203). They voted unanimously that the drug should not be granted ‘full’ (standard) approval (207), and unanimously that the drug should receive accelerated approval (210). It is notable that the ODAC was unanimous on the vote for accelerated approval, despite dissenting votes on whether the benefit was worth the risk.

up the meeting by browbeating the sponsor for their poor quality data, saying, ‘My discomfort about some of these votes is that we acted on imperfect data but with enough of a hint of efficacy and enough of a burst of sympathy’ (211) — sympathy engendered in part, he said, by the stories of the patients who spoke at the beginning of the meeting (212). In a parting jab at LTI, Schiffer added: ‘And I think you all could have made it a lot easier for us’ (212).

Although the Subpart H regulations did not modify the need for adequate and well-controlled investigations as specified by statute, the advisory committee clearly viewed this decision as based on inadequate data and was guided instead by medical knowledge and ‘sympathy’. This medical knowledge on which the decision was based ultimately amounted to Dr. Temple’s definition of an ‘historical control’ as the expert’s educated assessment of the natural course of the disease apart from therapy. It is also significant that it was not the work of the drug sponsor which ultimately convinced the committee to grant the drug accelerated approval, but rather the work of the FDA — not only Dr. Temple’s shepherding of the committee members, but also Dr. Murgo’s careful and methodical medical review of the charts of all the refractory patients on the study. The Committee was clearly affected more by Murgo’s work than by LTI’s. The committee chair, Dr. Schiffer commented that if the drug were approved, LTI would not be allowed to take Dr. Murgo out for a celebration, ‘but he was the one who did the digging’ (188).

7.3 Discussion: Adaptations for Accelerated Approval

In this chapter, we examined the first two applications of the accelerated approval rules to cancer drugs. What we found, as could be expected from finitism, was only superficial resemblance to the first applications of the rules to AIDS drugs — and, indeed, only superficial resemblance to the Subpart H rule as created in the light of those experiences with AIDS drugs. In the case of dexrazoxane, the application of the rule clearly deviated from what would have been expected from a reading of the section of Subpart H authorizing accelerated approval based on an endpoint other than survival or irreversible morbidity. The primary benefit, cardioprotection, had already been proven. What was in need of confirmation following the initial rejection of the drug was an unanticipated adverse effect, tumour protection. In sidestepping the issue of tumour

protection and seeking approval for the drug after six courses of cancer therapy, the drug sponsor succeeded in creating an additional question for confirmation (the usefulness of FAC therapy for breast cancer beyond six courses) which had nothing to do with the primary benefit of dexrazoxane. The characteristics of this I_2 fit neither the rule as envisioned through articulation-practice nor the actual existing exemplars.

In the case of DOX-SL, the difficulty and variability of measuring ‘response’ in AIDS-related KS complicated the task of making a decision on the basis of a response rate. This variability, coupled with the lack of a control group, left the committee split over whether to call the data ‘adequate and well-controlled’. Dr. Temple managed to turn this split into a statement of the need for accelerated approval. In his interpretation, the problem wasn’t a lack of a control, but that the response data needed to be confirmed in subsequent study using clinical measures of benefit such as lack of oedema. Once Temple put the issue in these terms many committee members embraced this interpretation. Yet, ultimately, it is clear from the comments of the Committee members quoted above that many of them saw this data as inadequate, and viewed this accelerated approval not as proceeding on the basis of a surrogate endpoint with confirmation of clinical benefit, but as approval of a drug that might, on the basis of a handful of clinical cases described in detail by Dr. Murgo, have some benefit for patients suffering from a disfiguring and painful disease. For these committee members, what needed to be done on a postmarket basis were the adequate, randomized controlled trials that should have been completed in the first place. Again this is not, strictly speaking, the pattern of action envisioned by the Subpart H rule. This is clear from Dr. Temple’s comment that sometimes cases arise in which decision-makers must ‘swallow hard’.

Temple’s definition in this case of a historical control as essentially the expert’s knowledge is also notable. As we saw in Chapter 3, the 1969 description of the IND regulations in the *Federal Register* admitted ‘prior experience’ to serve as a form of historical control, but specifically referred to documented experience which could be used for quantitative comparisons. The ‘official’ version of the rules at this time, as found in the *Codes of Federal Regulation* (CFR) (which dispenses with the sort of preambles and explanatory passages characteristic of *Federal Register* publications) did not specifically mention historical controls. Rather, the regulations specified that specific case report data

be collected and maintained on all subjects identified as controls.¹⁵ Hence, even while these regulations made no specific distinction between historical and concurrent controls, nor did they specify a need for quantitative comparison, they did require specific documentation of a range of parameters for control subjects. Whichever definition one chooses to use, Temple's characterization of the concept as expert knowledge of disease history seems to stretch the concept to maximum advantage. This 'stretch' of meaning is not surprising when viewed in terms of finitist concept application.

All of the conclusions from Chapter 5 could be reiterated here. The theory of regulation should not view regulatory outcomes as simple reflections of laws on the books or Congressional dictates. The involved and active participation of individual regulatory decision-makers must be considered. While I have not done an exhaustive review of the advisory committee system and the way the FDA uses it, in the cases studied in this thesis the FDA clearly sets the tone of the meeting, frames how the data should be viewed, and sends many cues to the advisory committee to let a drug pass when desired.¹⁶ Repeatedly in this period, we have seen the FDA encouraging the approval of drugs, not only shepherding them through the process using flexible application of the rules and concepts, but also requesting more data to analyze, even from the sponsor's competitors, and doing additional analyses which turned out to be decisive in several of these cases. Nevertheless, we have also seen one case (Provenge, in Chapter 1) where advisory committee members were able to change the minds of FDA decision-makers to withhold approval of a drug. So the influence can work in both directions, and does not necessarily always work in favour of approval.

It is worth noting that dexrazoxane, and indeed most of the drugs discussed in this paper, were approved on a basis that was rejected during the DESI review. Let me immediately qualify this remark by saying that there are clear differences of context which cannot be ignored. As we saw in Chapter 3, during the 1960s the FDA was struggling to establish its authority to enforce the new standards of drug approval established in the 1962 drug amendments, and the drug companies were challenging that authority at every turn. At that point it was obviously important to establish a more 'scientific' standard and

¹⁵ See the 1986 version of the *Codes of Federal Regulation*, Title 21, Section 312.1 (a)(2)(10)(c) ('institutional review board', p. 63-4). This section specifies that for all control subjects, the age, sex, condition treated, dosage, frequency of dosage, results of clinical examinations and laboratory tests, and all previous therapies received needed to be documented.

¹⁶ Jasanoff (1990) has looked at the use of advisory committees, including FDA committees, and found that they were often used to ordain and justify the foregone conclusions of regulators.

adhere to it. More than that, the examples cited in Chapter 3 came from studies of combination antibiotics — a category of drugs deemed redundant and unnecessary by the FDA and National Academy of Sciences. Single antibiotic products were already on the market and very effective, hence there was no immediate medical need to be addressed. For these reasons, compelling evidence would be necessary for the FDA to change its position on the issue of combination antibiotics. By the 1980s and 1990s, the authority to set the standards had been upheld in the courts and the more pressing question for the FDA at this time was how to get important drugs to seriously ill patients more quickly. With these obvious differences of context acknowledged, we can still observe that as much as the FDA protests that the standards for drug approval have not changed, clearly through this period the burden of proof required for approval and the quantity of information-gathering deferred to post-market study shifted significantly for certain types of drugs.

It would be a mistake, I believe, to attribute this shift and the decision-making predicated upon it to simple ‘bias’ as Abraham (1996) might do. As I said in Chapter 5, what we see in this account is a case-by-case assessment of how much risk is acceptable, how much benefit should be expected, how much proof should be required. The advisory committees, a group of acknowledged experts, struggled with these questions in each meeting — in their microclimate of standards-setting. We can suppose that behind closed doors at the FDA, staff members also struggled with the same questions prior to coming before the advisory committees. While certainly political forces or commercial interests work to push that consensus one way or another in a less-than-disinterested matter, the result is no less a process of social census because of it. More than that, we have seen that for cancer drugs perceived to be therapeutically important, approvals have always tended to take place on the basis of surrogate endpoints and frequently have proceeded on the basis of less than ideal data. If anything, the regulatory changes made in the late 1980s and early 1990s were merely a reflection of already established practices.

8. REINVENTION AND MODERNIZATION

In this chapter we will turn our attention towards the forces which came together to create the *Food and Drug Administration Modernization Act of 1997* (FDAMA)(P.L. 105-115)¹ — the first major overhaul of the Food, Drug, and Cosmetic Act since 1962. As we will see, the FDAMA is notable for a number of reasons, not least of which is the way it incorporates existing FDA regulations (and concepts from those regulations). Despite the partial retreat we witnessed in the use of CD4 counts as a surrogate marker, the FDAMA further enshrined the practices prescribed in Subpart H accelerated approval. This was done even though the FDA earlier insisted that a legislative amendment was not necessary to implement accelerated approval procedures.² The manner in which this legislation borrowed from existing FDA regulation and informal practice would have consequences for those regulations and practices. Indeed, all of the drug approval practices and categories involved, including Fast Track itself, would need to undergo some definitional shifting relative to each other to adjust to the new law, consistent with what finitism would predict. We will also examine an instance of application of the new Fast Track provisions to a new drug application (NDA) bearing all of the hallmarks of concept application in the early stages of use. The topics of categorical shifting and concept application will be discussed towards the end of the chapter.

First, in the sections which follow, I will sketch the political context in which the provisions of the FDAMA were debated and created. Whereas Subpart E and H and the rules for expanded access were forged in a time largely dominated by the politics of AIDS (although, as we have seen, elements of these rules had already been in formation), the

¹ On the numbering system for congressional proposals and laws, see Chapter 3, note 16.

² In 1992, the FDA wrote that an amendment to the Food, Drug and Cosmetic Act was unnecessary to create effective accelerated approval and expanded access programs since the FDA was authorized by the Food, Drug, and Cosmetic (FDC) Act and the Public Health Service (PHS) Act to set out new regulations such as those in Subparts E and H (FDA 1992a; 1992b). Notably, Louis Lasagna, who was often a critic of the FDA, also expressed the opinion that major legislation was not needed to effect desirable policy and procedural changes at the FDA (U.S. House 1995c).

mid-1990s would resemble to a much greater extent the politics of the 1970s, when conservative (and not-so-conservative) criticism of the FDA dominated the debate over reform. However, a key difference exists between the two periods. In the late 1970s, Democrats controlled the Congress and also the Presidency. In the early 1990s, the Democrats again held such an advantage but lost it dramatically in the mid-1990s. This shift in power would turn out to be decisive to forge the first major amendment of the Food, Drug and Cosmetic Act since 1962. The detailed examination of congressional politics in this period will provide some perspective on interest group influence, the nature of congressional deal-brokering and consensus, and other issues pertinent to the models of regulation discussed in Chapter 1.

8.1 National Politics and New Directions for Reform

Bill Clinton was elected president of the U.S. in November 1992. Taking office in January 1993, Clinton quickly began implementing initiatives promised during his campaign. On 3 March 1993, he announced the creation of a 'National Performance Review', the goal of which, he said, was 'to make the entire Federal Government both less expensive and more efficient, and to change the culture of our national bureaucracy away from complacency and entitlement toward initiative and empowerment.' In sum, he asserted: 'We intend to redesign, to reinvent, to reinvigorate the entire National Government' (President 1993a, 350).

This concept of 'reinventing' government had been developed in collaboration with the Democratic Leadership Council, a group formed in 1985 to develop a policy agenda representing a 'third way' between the traditional 'liberal' wing of the Democratic party and conservative Republicanism (Borrelli 2001). In the context of the 'Reagan revolution' and the apparently associated shift-to-right of the American electorate, traditional Democratic policies were flagging, and Democratic candidates were all-too-easy to portray as poster children for a 'tax-and-spend' big-government agenda. One priority for the Clinton administration would be to assail this stereotype by downsizing the government bureaucracy and streamlining its operation. Clinton gave Vice President Al Gore responsibility for designing and carrying out the programme.

Gore's first report on the national performance review (NPR) was released six months later as directed by the president (Gore 1993). Replete with anecdotes of obsolete

and contradictory rules and regulations, of bureaucratic paternalism hampering the ability of government employees to do their jobs effectively, and of taxpaying citizens frustrated by ‘red tape’ in their efforts to obtain rightfully due benefits or information, the report made an extensive series of recommendations for reforming budget and procurement practices, reducing paperwork, improving customer service, and eliminating unnecessary regulation. The Clinton-Gore approach to the NPR was explicitly a collaborative one, where input would be sought from government employees, business leaders, and users of government services to identify problems and generate solutions. In the aftermath of this report, Gore oversaw implementation of the recommendations while Clinton issued a series of executive orders recommended by the report. Clinton began with an order for Federal departments and agencies to eliminate at least 50% of their internal management regulations (meaning regulations pertaining to personnel, organization, or management of the department or agency) (President 1993b).³ Another order (EO 12866) issued just two weeks later specified criteria for new rulemaking designed to assure that any new rules to be written were justifiably necessary ‘to interpret the law or are made necessary by compelling public need, such as material failures of private markets to protect or improve the health and safety of the public, the environment, or the well-being of the American people’ (President 1993c, 1925). The Order also stipulated an explicitly negotiated process (soliciting the input of affected parties) for the creation of new rules and mandated review of existing regulations, and the formation of a ‘Regulatory Working Group’ having representatives from all the major agencies in an effort to avoid contradictions and redundancies between closely related areas of regulatory jurisdiction.

In Gore’s initial review of government operations, specific comments on the Food and Drug Administration (FDA) were relatively few. Of course the FDA was subject to the same requirements for administrative streamlining as other agencies under the NPR, however the only mention of the FDA in the department-specific recommendations in the 1993 report was to recommend that FDA be permitted to collect user fees not only from drug sponsors seeking review of new drug applications (a practice initiated the year beforehand on a trial basis through the Prescription Drug User Fee Act, or PDUFA), but from any food, drug, or medical device manufacturer, processor or supplier using FDA

³ For a list of the executive orders and memoranda issued as part of the National Performance Review, see <http://govinfo.library.unt.edu/npr/library/direct.html> (accessed 8 November 2007).

services.⁴ This recommendation was designed to address a severe backlog of medical device applications (U.S. House 1993a). We can therefore assume that in their initial review the Gore team found the regulatory reform already underway at the FDA for accelerating drug approval consistent with the aims of the NPR and not in need of adjustment — or certainly, at least, not among their top priorities to address. Their main concern was addressing the log jam in the Office of Devices and Radiological Health. Many other agencies and departments did have lists of improvements to make, however, and the 1994 NPR status report to the President recommended myriad administrative improvements, reductions in the federal workforce (according to the report, Gore was largely targeting redundant layers of administrative management, not staffers), in the process claiming tens of millions of dollars in savings (Gore 1994).

In January 1995, Vice President Gore announced ‘Phase II’ of the National Performance Review, representing this new phase of review as an opportunity afforded by recent changes to the composition of Congress: ‘Recognizing the election of a new Congress in fall 1994 as an opportunity to further governmental reform, President Clinton asked Vice President Gore to launch a “second phase” of reinvention’.⁵ This characterization of the motivation for ‘Phase II’ was euphemistic, to say the least. The reality was that the Congressional elections of November 1994 resulted in far-reaching political change, tilting the balance of power in Congress heavily in favour of the Republicans.⁶ The results were most startling in the U.S. House, where the Republicans had not managed to gain a majority since the elections of 1952. Moreover, the gain of 52 seats achieved in 1994 was the largest net partisan swing in the House since 1948 (Jacobsen 1996).⁷

The Republicans revelled in their newfound Congressional control, promising to do nothing short of remaking Washington politically. Before even taking office in the new congressional session, House Republicans signed a ‘Contract with America’ with

⁴ See Gore (1993), Appendix A, recommendation number HHS10.

⁵ Kamensky (1999). See the beginning of the section titled ‘Phase II of the Performance Review – 1995’ at <http://govinfo.library.unt.edu/npr/whoweare/history2.html> (accessed 8 November 2007).

⁶ The specific reasons for this turn of political fortunes are beyond the scope of this thesis. For analysis of this event, see Abramowitz and Saunders (1998), Jacobsen (1996), Klinkner (1996), Stonecash and Mariani (1998).

⁷ See also the House election statistics provided by the Office of the Clerk of the U.S. House of Representatives: http://clerk.house.gov/member_info/electionInfo/1994/94Recapi.htm#RECAPITULATION (accessed 8 November 2007).

great fanfare, promising to bring to the floor of the House within the first 100 days of office a series of major legislative proposals. As breathlessly reported by a business writer for the conservative *Washington Times*, the ‘new vision’ represented by these proposals was: ‘Cut everything. Cut taxes, cut spending, cut government and cut the budget deficit’ (Nesbit 1995, B7). A social agenda was also clearly evident in the Contract, with proposals not only to balance the federal budget and reduce taxes, but also to promote ‘family-friendly’ tax reform (e.g., per-child tax credits) and punitive welfare reform to discourage promiscuity, among other provisions.⁸

For the purposes of this thesis, the most relevant aspect of the Republican agenda for reform came in their proposals for comprehensive regulatory reform, early forms of which were articulated in the Contract’s proposed *Job Creation and Wage Enhancement Act*. Senate majority leader Bob Dole announced in one of the first sessions of the 104th Congress that the Republicans would ‘roll back Federal programs, laws, and regulations from A to Z, from Amtrak to zoological studies, working our way through the alphabet soup of Government’.⁹ The *Washington Times* likewise noted a television appearance by the Speaker of the House Newt Gingrich in which the Speaker criticized the Federal Communications Commission (FCC), the FDA, and the Federal Transportation Commission (FTC), saying: ‘Had we had an FCC, FDA or an FTC in Silicon Valley, we’d be about 150,000 or 200,000 jobs short and we’d be back with mainframe computers because you’d still have bureaucrats studying whether or not to allow PCs to even exist’ (Nesbit 1995, B7). Thus concluded the *Washington Times* business writer triumphantly, ‘if you cut spending by going after as many programs as possible, and you cut the size of government by reducing the scope of several high-profile regulatory agencies, and then

⁸ In fairness, the reader should know that I was opposed to much of the Contract with America. The overall agenda appeared to be to cut taxes and increase military spending at the expense of social programmes. The proposed *Fiscal Responsibility Act* mandated a balanced Federal budget while capping the national debt at existing levels and setting limits on taxation. With these restraints in place, the proposed *Job Creation and Wage Enhancement Act* would have cut taxes on capital gains (which is *non-wage* investment income) by 50%, forcing commensurate reductions of outlays for Federal programmes while the proposed *National Security Restoration Act* increased Federal military spending. Meanwhile, the proposed *Personal Responsibility Act* denied welfare benefits to children born out of wedlock in cases where paternity could not be established or when the mother was under the age of 18 unless the mother and father chose to marry. Additionally, among other things, the *Taking Back Our Streets Act* proposed to divert money earmarked for social spending to build more prisons. This proposal would also have expedited executions of death row inmates by limiting their ability to make post-conviction appeals. See the complete Contract, including links to the text of proposed legislation, at <http://www.house.gov/house/Contract/CONTRACT.html> (accessed 8 November 2007).

⁹ *Congressional Record*, Senate, 04 January 1995.

you cut the deficit, all that's left is to cut taxes. Which Congress will do gleefully' (Ibid, B7).

At the start of the 104th Congress in January 1995, the Republicans wasted no time in pursuing their legislative agenda, introducing a series of measures related to the Contract with America. One of these bills, H.R. 450, the *Regulatory Transition Act*, called for a moratorium on all regulatory rule-writing pending the development of legislation to require a cost-benefit analysis for all new rules. The new legislation would also require each agency which made decisions based on assessments of risk to develop a *standardized approach* for risk assessment applicable to *all* decisions. H.R. 450 was introduced in the House on 9 January 1995 and passed in the House on 14 February 1995. A similar bill was introduced into the Senate on 12 January 1995 (S.219). Meanwhile, the House began holding hearings on how to legislate standardized risk assessment (U.S. House 1995a). In the American congressional system, whichever political party holds the majority in the House or Senate also takes leadership of the corresponding congressional committees. Thus, the 1994 elections not only gave the Republicans an advantage when voting on legislation in Congress, but also gave them an enhanced ability to control the hearing agendas of the congressional committees. What followed, among other things, was a series of hearings into the workings of the FDA.

In this context, we can infer that the motivation behind the second phase of the NPR was to move pre-emptively in the hope of maintaining some control of the reform agenda and influencing the direction taken by the new congressional Republican majority. The pre-emptive character of the NPR's Phase II is evident in a Vice Presidential Memorandum issued by Gore in January 1995 announcing the continuation of the review programme. The memorandum directed departments and agencies each to assemble a task force which should consider the question: 'If your agency were eliminated, how would the goals or programs of your agency be undertaken — by other agencies, by states or localities, by the private sector, or not at all?' The task forces were also to obtain feedback on how their customers would feel about such possible eliminations or changes.¹⁰ Clearly Clinton and Gore wanted feedback from the agencies and their customers, not only for the continued streamlining of government operations, but as a

¹⁰ Content of memorandum is posted at the NPR archive website:
<http://govinfo.library.unt.edu/npr/library/direct/memos/266e.html> (accessed 24 October 2007).

preparation for, and even a defence against, the kinds of regulatory reform being advocated by many Republicans.¹¹

Of course, the FDA had always been subject to oversight by Congressional subcommittees and the Congress had long demonstrated a keen interest in issues related to medicine, therapeutic innovation, and the public health, as is evident from the agenda of the previous (103rd) Congress. For example, while significant progress had been made in expediting the approval of drugs, the backlog of medical device product applications suggested that the regulatory process for medical devices had broken down. The 103rd Congress therefore held hearings on the causes of the breakdown (U.S. House 1993a; 1993b) and discussed instituting user fees for medical device applications (U.S. House 1994d) — an action, as noted above, recommended by Vice President Gore’s National Performance Review (Gore 1993). Another concern of the 103rd Congress had been assuring that adequate and appropriate informed consent was obtained from subjects in clinical investigations (U.S. House 1994b).¹² Many other hearings were also held.¹³ Overall, congressional committees under control of the Democrats seemed generally content with the FDA’s progress on reforming procedures for NDA review, and were

¹¹ Three days after this memorandum was circulated, the *Washington Times* business correspondent disingenuously wrote: ‘The Clinton administration has finally caught the vision as well. The Labor Department, which is in truth a huge regulatory agency, has pulled its political shock troops together for a specific task force designed to blunt any Republican offensives in Congress. Other agencies are doing the same thing already, though less overtly. The war has clearly begun’ (Nesbit 1995, B7).

¹² The Committee had found that Federal agencies (particularly the FDA and the Office of Protection from Research Risks, which is another arm of the Department of Health and Human Services) had conflicting policies for informed consent which tended to hamper the efforts of academic and industry researchers. More than that, the Committee wanted to look into claims that some institutional review boards, having been encouraged under Federal policy to develop their own informed consent guidelines, had instituted the practice of ‘deferred consent’, meaning that consent was not obtained from some subjects until after the study was already underway. Although from the FDA’s perspective this practice was illegitimate and inappropriate, some hospitals were using it, leading to ‘institutional review board shopping’ (U.S. House 1994b, 2) by unscrupulous researchers seeking an environment of lax human subject protection in which to work.

¹³ The 103rd Congress also conducted hearings to reauthorize the Orphan Drug Act (U.S. House 1994c); hearings into unfair prescription drug pricing practices (U.S. House 1994f; U.S. Senate 1994d); a hearing on the progress of research on a particular form of tumour radiation (U.S. Senate 1994b); an inquiry into the U.S. military using U.S. soldiers as research subjects without their knowledge or consent (U.S. Senate 1994a); a hearing on gene therapy, its potential for future use, and its policy implications (U.S. House 1994e), and a hearing on the negotiations to license the French ‘morning after pill’ RU-486 to an American non-profit foundation because the drug, in the words of committee Chairman Widen, ‘offers the best hope for American women to have a safe, effective alternative to surgical abortion before the end of this century’ (U.S. House 1994a). For this reason, he added, delays in the negotiation were ‘very troubling’ (1). The Congress furthermore authorized the National Institutes of Health to conduct research into whether some forms of alternative medicine are effective and held hearings to review the research to date (U.S. Senate 1993) and to discuss proposed legislation giving patients the right to demand alternative therapies from their health providers, if desired (U.S. Senate 1994c).

concerned to fix trouble spots associated with medical device approvals and conflicting regulatory policies. Their main prescription for the FDA was to find ways to increase funding, which would allow increases in staffing and other improvements for more efficient review of applications (especially device applications). This solution seemed to be working well for new drug applications thanks to the PDUFA of 1992 so, said congressional Democrats, why not extend the concept of user fees to all of the FDA's core missions?

Clearly the Contract with America and the related legislative agenda of the 104th Congress reflected a different set of priorities from that of the 103rd Congress. Just as in the previous decade, the AIDS crisis and AIDS activists served as a catalyst both to accelerate and to extrapolate already-existing reform efforts, in the mid-1990s the combination of the neoconservative reformist momentum and the Democratic attempts to channel, preempt and sometimes thwart that momentum, would be the spur towards comprehensive legislative reform of the FDA: the FDA Modernization Act (FDAMA).

8.2 Congressional Politics and the Footrace to Reform

8.2.1 *Some Brief Lessons from Medical Devices*

The first of the hearings held on FDA regulation took place on 30 March 1995 (U.S. House 1995b). The specific focus of this hearing was medical device approval. While the recent history of medical device approval is beyond the scope of this thesis, some aspects of the medical device hearings from this period (the 103rd and 104th Congressional sessions) are worth highlighting in this study. For one thing, it is clear that the rapid change and innovation taking place on the therapeutic drug side of the FDA was not mirrored in other aspects of its operation. As already noted, the previous Democratically controlled Congress had held hearings in 1993 and 1994 on medical devices (U.S. House 1993b; 1994d) to address problems with device application backlogs. Indeed, in March of 1993 the House Energy and Commerce Committee's Subcommittee on Oversight and Investigations published a detailed report on the management of the FDA's Center for Devices and Radiological Health (U.S. House 1993a). The document complained that a decade earlier the subcommittee had issued a report demonstrating the FDA's lack of compliance with the Medical Device Act of 1976 and recommending actions 'to instill order, logic, and fairness to the medical device approval process' (III).

According to the 1993 report, the FDA ‘took little effective or timely action’ (III) on these issues, and as a result structural and management problems remained and worsened. Clearly Congressional direction does not automatically lead to FDA action — another observation worth emphasis. Indeed, the FDA appears to be somewhat Balkanized, with each division acting according to its own rules and perceived mission, and responding to contemporary exigencies and political developments particular to that mission. This observation once again problematizes the public choice and neopluralist theories (Chapter 1) of regulation, since they treat regulatory agencies as black boxes through which the Congressional will is implemented. It additionally problematizes the public interest theory of regulation, which does consider the role of individual decision-makers, but does so in terms of individualistic motivations, failing to take into account the milieu of organizational culture and commitments in which the decision-makers are immersed. Even the civic republican theory, which is more concerned with processes of social consensus than any of the other perspectives, fails on this count because it ultimately is concerned with formalized processes of consensus, not informal ones.

The report issued by the subcommittee in March 1993 contained a series of recommendations, including: 1) reallocating personnel to augment the staff in the Center; 2) developing an ‘expedited approval track’ for ‘breakthrough products’ (86); 3) refusing to file incomplete applications (rather than filing the application and then engaging in a lengthy back-and-forth exchange over missing information); 4) reassessing the risk classification system used, and downgrading the risk classification of many devices which could be thus reviewed under a less rigorous procedure. Perhaps in part because of the presence of an activist FDA Commissioner, Dr. David Kessler, and perhaps especially because of Kessler’s recent experience with reform on the drugs side of the FDA, these recommendations were rapidly embraced. In a hearing later that year (U.S. House 1993b), the FDA announced a series of policy changes consistent with the Subcommittee’s recommendations, including the creation of a separate ‘fast-track’ queue for devices having superior ability to treat life-threatening or debilitating conditions and developing a risk-based system for review under which the lowest risk devices would receive only a ‘labeling review’ rather than a full ‘scientific’ review (46). Clearly the concept of fast-tracking for important or special applications was spreading. FDA also agreed to meeting performance measurement goals at this time (effectively, time limits on device application reviews)(53) and proposed to exempt certain low-risk devices from pre-market

notification requirements previously applied to devices deemed substantially equivalent to already-approved marketed products (59). Later that year, the FDA published a list of 148 categories of devices to be exempted in this way (FDA 1994c).

It is also worth noting that one major recommendation for reform in the NPR, the institution of user fees for device applications, was thwarted by congressional Republicans in the 103rd Congress. The medical device industry had long opposed the implementation of user fees, calling them ‘nothing short of an up-front tax on product development and innovation’ (U.S. House 1994d, 129). To break the log jam of device applications, however, some industry representatives finally agreed to a user fee bill pending in the 103rd Congress (H.R. 4728) (see U.S. House 1994d, 129-30 for a significant endorsement). In fact, three different versions of a device user fee bill were introduced by Democrats in the 103rd Congress (H.R. 4728, H.R. 4864 and S. 2276) but none of them were made into law. Despite industry agreement, Republican members of the committee dissented, asserting that no additional resources should be provided to the Center for Devices and Radiological Health without first implementing a complete overhaul of what they said was dysfunctional Center management (see the dissenting opinions in the Committee report to the House, U.S. House 1994g). The institution of medical device user fees did not actually take place until 2002.

Clearly when discussing ‘congressional intent’ or ‘congressional oversight’ in the context of regulatory theory, one must keep in mind that congressional motives and intention are rarely unitary. More than that, in the case of devices — and also in the case of drugs as described below— the main battle lines over specific proposals for reform tended to be drawn along the lines of political ideology. In no small way, the deal-brokering took place between the legislators *themselves*, above and beyond any legislator brokering of deals between opposing interest groups (as neopluralism views the process).

8.2.2 *The Fog of Reform*

Two weeks before the first hearing on the FDA was to take place, President Clinton characteristically held a news conference announcing, among other things, that the FDA would stop requiring costly and unnecessary environmental impact assessments on new drugs (one outcome of the NPR), would eliminate hundreds of pages of antiquated and unnecessary regulations associated with antibiotics and insulin, and had exempted 140 categories of low-risk devices from pre-market notification requirements,

with more exemptions on the way (President 1995, 430).¹⁴ In this way, he made more visible the actions already taken by the FDA and anticipated many of the themes to arise in subsequent hearings. In April 1995 the Clinton Administration also released a National Performance Review document called *Reinventing Regulation of Drugs and Medical Devices* (Clinton and Gore, 1995), summarizing many reforms to date (including exempting low-risk medical devices from premarket review, the use of advisory panels to expedite review of certain products, the use of user fees to meet FDA performance goals, etc.) and proposing additional reforms (including a new proposal for medical device user fees).¹⁵ The document also recommended reducing or eliminating the need for manufacturers of drug and biologic products to get approval for changes to production or processing facilities, allowing the makers of biological products to obtain licenses for pilot facilities instead of being required to build full-scale plants for production of still-investigational products, the elimination of outdated requirements for the manufacture of insulin and antibiotics, and other reforms subsequently discussed in hearings held in May and June of 2005 (U.S. House 1995c).

Despite the Administration's attempts to take the initiative on such issues, the new Congress tackled FDA reform with verve. As a Congressional Research Service report to Congress makes clear, 'Since the 104th Congress convened, FDA drug approval reform has been high on the agenda of the new leadership. House Speaker Newt Gingrich and Commerce Committee chairman Thomas Bliley have publicly expressed their own reservations about FDA's drug approval policies' (Randall 1996, 15). Put more bluntly by another critic, former FDA Commissioner Edwards, while testifying before a Senate committee: 'Certainly the FDA has always had its critics, but I think it is rare indeed to have the criticism of the Agency come from so many camps and to have so many different proposals for reform quoted at the same time' (U.S. Senate 1995, 3). While many of these complaints (e.g. the medical device application backlog) were valid and in need of attention, the congressional hearings which followed often exhibited more enthusiasm than comprehension.

To give the reader a sense of the quality of discourse often surrounding the proposals for reform at this time, take, for example, one of a pair of hearings held before

¹⁴ See the FDA's announcement for the exemption in FDA 1994c.

¹⁵ Note that a 1996 Congressional Research Service report (Randolph 1996) contains a section summarizing the contents of the 'Reinventing' document but does so inaccurately, presenting as proposals some items listed in the NPR document as already-implemented reforms.

the Subcommittee on Oversight and Investigations (a subcommittee of the House's Committee on Commerce) (U.S. House 1995c). In this hearing, held on 19 June 2005, Rep. Bliley, a Republican from Virginia (R-VA) who was the new Chairman of the Commerce Committee and a vocal critic of the FDA, exposed his poor understanding of the FDA and its workings in the following exchange with Ellen Stovall, Executive Director of the National Coalition for Cancer Survivorship:

BLILEY: Ms. Stovall, I would like to read to you from a letter to the New York Times of February 19, 1995 from Senator Connie Mack of Florida. The letter states that, and I quote: 'Before my brother died of cancer in 1979, I found it incredible that he was not allowed to take experimental pharmaceuticals because of concern over the side effects. It made no sense that while he was dying, faceless bureaucrats in Washington worried about the side effects of a pharmaceutical drug.' Is this an example of an FDA attitude that we must change?

STOVALL: You mean the fact that they wouldn't prescribe the drug because of side effects?

BLILEY: You've got a patient -- a patient is dying of cancer, and they wouldn't approve an experimental drug because of a potential side effect.

STOVALL: Well, the potential side effects are part of what the evaluation process is used to determine in terms of whether a product is safe or effective. All the products have side effects. I don't think it's in the interest of people with cancer, particularly those who are dying of cancer, to have to wait to see where every side effect of a product is proven or not. I think the patient needs to be much more involved in the process of determining what risks they're willing to take, particularly those whose treatment options have been exhausted.

In Senator Mack's place . . . I have talked to him about his brother's situation, and those were dire circumstances, and there is such a thing as compassionate use of products under consideration by the FDA that should have been exercised at that time.

BLILEY: Well, do you think the law should be changed to specifically allow this use of non-FDA approved drugs or biologics, if the patient has a fatal condition and has given his or her informed consent to the physician?

STOVALL: That's a very difficult one for me to answer. I think it's too general a question. I think each circumstance has to be evaluated on its own merit. I think, more importantly, the FDA approval process is so outmoded and so arcane that to speak to that specific situation would be very difficult.

I think that the example, you know, accelerated approval, helps these situations. I believe that patients need to be much more involved in the process, and I think therefore that would help the situation, along with their physician, there needs to be -- things that appear in the literature, for instance, about clinical trials that are also helpful. Sometimes before a drug has an opportunity to get labeling, there is information that appears in the literature that indicates its usefulness in these situations.

In this exchange, Rep. Bliley's question was based on a situation which took place *fifteen years prior* to the hearing.¹⁶ The letter clearly expressed frustration at not being able to obtain an *investigational* drug, however Bliley confused the letter to mean that the FDA 'wouldn't *approve* an experimental drug because of a potential side effect' (147, emphasis added). This left Stovall in a quandary as to whether Bliley was talking about access to investigational drugs or accelerated approval of drugs to treat life-threatening diseases — and as a result, she was left to ramble about 'side effects' in assessing drug safety, finally mentioning that 'compassionate use' procedures should have been available in 1979, almost as an afterthought, then veering into a confused discussion of accelerated approval. More importantly for a question about legislation to reform the FDA in 1995, the witness failed to inform the congressman that since 1979, the FDA had codified the treatment IND rules and established the parallel track programme for access of seriously ill patients to investigational drugs. Rep. Bliley's question about changing the law to 'allow this use of non-FDA approved drugs or biologics' makes clear that he was unaware of the existence of these programmes, and his preparation for the hearing apparently amounted to no more than the collection of decades-old anecdotes. The witness's inability to clarify *current* FDA procedures for pre-market access to investigational drugs cast doubt on her credibility as well.¹⁷ Moreover, what none of the witnesses pointed out to Bliley was that the FDA could *allow* premarket access, but they could not *compel* it; it is not unusual for the drug-makers themselves to deny access to investigational drugs on a pre-market basis. In such circumstances, there is nothing the FDA can do.¹⁸

¹⁶ Indeed, we could remark on the disingenuousness of Senator Mack in publishing a complaint in 1995 about something the FDA supposedly did – or failed to do -- in 1979. This is a tactic often used by critics of the FDA, whether through ignorance or guile: regurgitating outdated complaints and antiquated statistics without taking into account reforms which took place in the interim.

¹⁷ The last paragraph of her response is especially baffling. It is difficult to fathom what accelerated approval has to do with patients getting more involved, or getting more information.

¹⁸ One example is the drug-maker ImClone, which in 2001 was attempting to bring a drug to market under the Fast Track provisions of the FDAMA discussed in this chapter. The drug company had allowed limited access to the investigational drug for certain cancer patients. However, after publicizing promising preliminary results of clinical studies, the company was overwhelmed with requests for the drug.

Rep. Bliley, while not the only poorly prepared committee member, was the *chairman* of the committee, spearheading the movement for reform. Another key figure, the Chairman of the Subcommittee on Oversight and Investigations, Joe Barton of Texas (R-TX), revealed his basic misunderstanding of the underlying issues in an exchange which began by asking Dr. Louis Lasagna, Director of the Tufts Centre for the Study of Drug Development, if a legislative standard for efficacy should be written. Lasagna responded that such a standard would be difficult to write; rather, ‘expert opinion’ should be rendered in each situation regarding how much efficacy is needed to assure that a drug will do more good than harm (169) (in other words, he was advocating the kind of risk-benefit approach to drug approval decisions actually used by the FDA, although he did not specify for the chairman’s benefit that a risk-benefit based system of expert decision-making in the form of advisory committees was often already employed). The chairman pressed the question, asking, ‘Is there something within the FDA that says before a drug is approved it has to be shown to be 90 percent effective for 90 percent of the population or 100 percent effective for 100 percent of the population or 50 percent effective for 50 percent of the population or, again, is this a case each drug is measured on its merits and some judgment is made about the efficacy of that particular drug?’ (171). This time, Richard Moscicki, a vice president of Genzyme Corporation responded, indicating that it is ‘the latter’, i.e. a matter of judgment, based on results from clinical trials. Barton pressed on with his agenda:

Now the radical question: What if we eliminated by statute this whole question of efficacy? Make sure it’s safe, but then let the marketplace determine if it’s effective? Is that something we’d even want to do; and if you answer that no, that you want an efficacy standard, then I would postulate that we need to begin to try to formalize this whole issue of efficacy, because as long as we leave it in some warm and fuzzy judgemental situation, even with the best of intentions, if we insist on more and more certainty in our lives, efficacy is going to become more and more minute. I mean, it’s 95, 100, we’re going to get to the point of everything has to be 100 percent, which is impossible (172).

According to the head of the company, ImClone was unable to produce enough drug to meet the requests for treatment IND and decided to discontinue their treatment protocols in January 2001. An outcry went up throughout the patient community, but the company did not relent. Unfortunately for many patients, the drug was not approved until 2004, after the company had gathered additional data at the FDA’s request. See the hearings held on ImClone’s refusal to participate in treatment IND in U.S. House (2001b). See also Bazell’s (1998) account of Genentech’s resistance to initiating a treatment IND programme in the face of strong patient protest when testing the cancer drug Herceptin. Eventually the company relented, but only allowed 25 patients chosen by lottery to receive the drug each quarter.

Barton's ignorance of drug approval decision-making was profound. Even after being told that the FDA does not evaluate drug efficacy in terms of percentages, but uses a case-by-case risk-benefit approach, he continued to discuss efficacy as though the FDA could mandate '95' or '100 percent' efficacy. For him, the way to streamline drug development and approval was to eliminate the legal requirement to prove a drug effective prior to marketing. Let the market determine which drugs were better than others. Here was a bald-faced statement of the neoconservative agenda in the 104th Congress — an agenda consonant with the conclusions of the public choice theory of regulation and Peltzman's (1974) application of it to the FDA (Chapter 3). Lasagna explained that even prior to 1962, information on efficacy was needed to provide meaningful context in which to judge the relative safety of any drug. Missing or ignoring Lasagna's point (and contradicting his own assertion that the FDA was requiring more and more certainty), Barton noted, as if it were a retort, that the FDA had already lowered its standards of efficacy for cancer and AIDS drugs. Despite his defence of the idea, even on this panel of witnesses which was otherwise relatively hostile to the FDA, Barton found no support for elimination of the efficacy requirement. Dr. Arthur Kibbe, a pharmaceutical sciences professor, testified that it is 'too difficult for an individual or a physician to test every drug themselves to make sure it works, and it's too dangerous to let them do that with patients who might have serious conditions who could otherwise have been treated with something else, and just treated with this because it was new and it had a claim, without that claim having some substance to it' (173). Even Genzyme's Moscicki testified that although 'there are many individuals in the biotechnology industry who would find your bold idea very attractive and appropriate, I personally, as a physician, do still favor some evaluation of efficacy' (173). Dr. Lasagna said, 'I certainly don't want to see drugs marketed without some evidence of efficacy' (173), adding that the question was how much evidence to require. Stovall agreed with the other witnesses (174). Barton again asserted, 'My hypothesis is, the marketplace will determine effectiveness. As long as we guarantee that it's safe. If it doesn't work, the medical community over time, and a very short time if they're in it for profit, are not going to continue to prescribe it, nor to use it' (174). Lasagna responded: 'Doctors wouldn't even want to begin prescribing it if they weren't presented with evidence on efficacy' (174). Barton's weak response was: 'Well, the marketplace, it is just a hypothesis, it is not a fact' (174).¹⁹

¹⁹ Fortunately, a few others on the committee had a much better grasp of the FDA approval process and the

In all, during the two-year term of the 104th Congress, I count at least nineteen hearings having to do with FDA management, drug or device approval, patient access to investigational drugs, or off-label prescribing and information dissemination. Additional hearings were held on food safety and the FDA's process for designating food products and additives as 'generally recognized as safe' (GRAS), and issues with veterinary medicine and drugs. Congressional leaders additionally held numerous hearings on the regulatory process in general, the 'hidden cost' of government regulations to businesses, developing standardized risk assessment formulae for regulatory decision-making, etc. To be sure, some members of these committees did have in-depth knowledge of FDA workings (e.g., see note 19, this chapter). But now the agenda was being set by new committee leaders like Barton and Bliley, whose reform goals clearly differed from those of previous committee leaders, and (at least in the case of Barton and Bliley) whose ardent commitment to reform was unencumbered by serious study or preparation.

8.2.3 Reform Marathon

A 1996 Congressional Research Service (CRS) report noted that the 'U.S. pharmaceutical and biotechnology industries, recognizing opportunities for regulatory change, have made efforts to persuade Members of Congress that drug reform should be given high priority' (Randall 1996, 15). Put more bluntly we could say that, not surprisingly, many industry groups saw this new leadership of Congress as an opportunity to achieve the regulatory slackening which Democrats had denied them. Between the efforts of industry to influence the Republican agenda and the efforts of Democrats to retain some control of the agenda, there seems to have been a footrace to reform the FDA. Washington was awash with recommendations for FDA reform, many of which overlapped or borrowed from one another in confusing disarray. Generally speaking, most of the proposals attempted to address the hot-button issues of the day: accelerating and streamlining drug and device development and approval; revising rules associated with exports of drugs and devices (especially exports of drugs and devices not approved in the U.S., but allowed in other countries); outsourcing or privatizing some of the current functions of the FDA; making it easier to use data from foreign clinical studies in U.S.

issues at stake. Rep. Henry Waxman, a Democrat from California long involved in FDA oversight and reform, reminded Barton about Laetrile: it was perfectly safe, lots of cancer patients took it, and it did not work. He added that allowing patients to take such a drug freely, meanwhile foregoing other potentially helpful treatments in favour of what seems 'new' or 'innovative', is a 'real cruel hoax' (174).

new drug applications; rethinking the FDA's role in restricting promotional materials disseminated by drug-makers; and revising FDA's approach to regulating 'good manufacturing practices' in the production of drugs and biologics. Moreover, many of the proposals sought to accelerate drug development or approval simply by reverting to a version of the pre-1962 practice of notification rather than affirmative approval; i.e., the drug sponsor would notify the FDA (or a designated external surrogate for the FDA) that clinical studies were about to begin (in the case of an investigational new drug) or that a new drug was deemed ready to market (in the case of a new drug application). The notification would include data to support the claim that the drug should be studied or marketed, and if the FDA did not actively intervene to prevent the clinical studies or marketing within a prescribed time period, the proposed study or marketing would ensue automatically. For my purposes here, as it has been throughout the thesis, the focus will be primarily on proposals to accelerate drug development and approval, modify the standards of evidence for approval, and expand access to investigational drugs for seriously ill patients.

Two CRS Reports attempting to summarize the key issues at stake and proposals for FDA reform for the new Congress to consider were issued only a month apart, one in December 1995 (Vogt 1995) and the other in January 1996 (Randall 1996).²⁰ According to Vogt (1995, 43), concept papers and draft legislation for FDA reform were prepared by a number of industry groups and industry-friendly lobbying and research organizations:²¹

- The Biotechnology Industry Organization: *FDA Reform Act of 1995*
- The Health Industry Manufacturer's Association: *Forces Reshaping the Performance and Contribution of the U.S. Medical Device Industry*
- Hudson Institute: *The Human Costs of Regulation: The Case of Medical Devices and the FDA*
- National Electrical Manufacturing Devices Association (NEMA): *Re-Inventing the Regulation of Medical Devices: A Challenge for the Twenty-First Century*

²⁰ The 1995 report summarizes the results of a CRS-sponsored seminar held in September 1995 to address FDA reform questions from various perspectives. Thus, this latter report is deliberative in its approach to the subject. The 1996 report attempts to summarize and consolidate the major bills and position papers in circulation at the time. Both reports are useful to get a sense of the major issues and proposals in-play, however see note 15.

²¹ See Chapter 3, note 16. In the American system, proposed legislation is often named to reflect the title the bill would have if it became law.

- Pharmaceutical Research and Manufacturer's of America (PhRMA): *New Drug Regulatory Improvement Act of 1995*
- Price Waterhouse LLP: *Improving America's Health*
- The Progress and Freedom Foundation: *The Future of Medical Innovation*

Congressional proposals were almost as plentiful as those from private interests. Most relevant for our purposes was a bill introduced on 6 June 1995 by Rep. Ron Wyden (D-OR) called the *FDA Modernization Act of 1995* (H.R. 1742) into the House. Among other things, the bill mandated the review of Phase I clinical studies to be performed by Institutional Review Boards,²² eliminating the existing IND system and instituting a notification system (as opposed to an affirmative approval system) for investigational drugs. It also eased export restrictions and explicitly excluded peer reviewed journal articles from materials considered to be 'promotional'. The central feature of this legislation, however, was the introduction of a form of 'conditional approval'. Unlike the Subpart H rule, this version of conditional approval did not specify a particular endpoint or type of data to be accepted temporarily until confirmation could be obtained. Rather the proposed legislation simply dictated that, for a drug or device intended to treat a life threatening condition, approval could be granted 'on the basis of valid scientific evidence demonstrating a reasonable assurance of safety and effectiveness' (see H.R. 1742, Sec. 4.) with a timeframe set for expiration of the approval. If the drug sponsor failed to meet conditions specified by the FDA for the particular situation within the prescribed timeframe, the approval would be void. This concept of 'conditional' approval harkened back to the 'provisional approval' proposed in the *Drug Regulation Reform Act of 1978* (Chapter 3). Both proposals envisioned a predetermined sunset for the temporarily approved drug. However, whereas the 1978 legislation established a set, three-year period for all provisionally approved drugs and established criteria for the applicability of the procedure and for the burden of evidence required for approval ('significant evidence'), the 1995 proposal left to the discretion of the FDA the duration of conditional approval, the amount of evidence needed in a given situation, and the applicability to any situation. Other than the stipulation that the drug or device should treat a life-threatening disease,

²² Notwithstanding the earlier 'IRB shopping' scandal investigated by the previous Congress (note 12, this chapter), many observers sought to shift some of the FDA's responsibilities to IRBs. Rep. Wyden's proposal to have IRBs review proposed early-phase clinical studies is essentially the same as the proposals in the Lasagna Report for IRBs to review investigational new drug applications (see U.S. House 1992).

parameters for approval would be set on a case-by-case basis at the discretion of the FDA.²³

Shortly after the introduction of Wyden's bill, on 30 June 1995, Rep. Jon D. Fox (R-PA) introduced competing legislation. Fox's bill was called the *Life Extending and Life Saving Drug Act of 1995* (H.R. 1995). While having a different name than PhRMA's *New Drug Regulatory Improvement Act*, Fox's bill differed very little, if at all, from PhRMA's proposed legislation (Randall 1996).²⁴ H.R. 1995 sought to reverse many features of the 1962 legislation and roll back many of the existing requirements for drug approval. Among many features, the bill reduced the amount of information required for an investigational new drug (IND) application, and returned the IND system to a notification system. H.R. 1995 did not provide for a separate approval track for drugs intended to treat life-threatening diseases, as Wyden's did, but rather sought to accelerate the entire process for all drugs. The bill scaled back data and approval requirements for new drug applications (NDAs), mandated that many of the NDA review functions should be contracted out for external review, and eliminated affirmative approval procedures for new drug applications, returning to a pre-1962 format of notification of the FDA for drugs about to be marketed. It also facilitated the acceptance of drugs already approved in other countries, and required the FDA to pursue international harmonization of drug application requirements. Significantly, part of Fox's scheme for expediting approval of

²³ This aspect of the Wyden bill also suffered from the same flaw as the proposal for provisional approval did: once a drug has been released to physicians for prescribing, even on a conditional basis, it is very difficult to withdraw if even a modest minority of patients believe they have benefited from using it. One can easily imagine scenarios in which the drug sponsors had not met the conditions of full approval by the deadline, but physicians and patients desired continued access to the drug. As it stands, the same problem now exists with drugs approved under accelerated approval; the list of unfulfilled postmarketing commitments grows but none of those drugs have undergone 'accelerated withdrawal', although the mechanism is available to the FDA.

²⁴ According to Randall (1996), Fox's bill 'parallels' the PhRMA draft (18). Therefore, at the very least, we can say that the two documents matched in their key proposals. It is entirely likely, however, that 'parallels' is a euphemism for 'wholesale adoption' of the PhRMA proposals. In September 1995 Fox also introduced a *Life Extending and Life Saving Device Act* bill into the House (H.R. 2290). The CRS Report does not provide any information on the provenance of this proposed legislation, as the report focuses on reforms for drugs. However, it is not unreasonable to believe that the Congressman had significant 'assistance' from industry groups on this bill as well. This sort of activity is not necessarily unusual. Harris (1964) reported that in the process of legislative amendment to the proposed 1962 Kefauver-Harris bill, pharmaceutical company representatives created a series of slips of paper, each containing a desired amendment to the legislation, and passed those slips of paper to Republican allies in Congress, who willingly introduced motions for those amendments. Presumably, Democrats might also similarly serve as conduits to Congressional proceedings for the interests they represent, but I have not encountered specific examples of it in the course of research for this thesis.

all drugs rested in amending the Food, Drug and Cosmetic Act's definition of 'substantial evidence' to mean:

evidence consisting of scientifically sound data, *including data from one well-controlled clinical investigation and confirmatory evidence (obtained either before or after such investigation)* on the basis of which experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved could fairly and responsibly conclude, taking into account the entire knowledge base on the drug's effectiveness and safety, interpreted as a whole, that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling of the drug (H.R. 1995, Sec. 6; emphasis added).

This passage was notable in at least two respects. First, it specified that the 'entire knowledge base' about a drug was admissible as part of the review process for drugs. Under H.R. 1995, the data offered in the NDA potentially represented only a fraction of the information which could be considered for the sake of an NDA. While obviously the FDA and advisory committees have always brought their personal expert knowledge of drugs and diseases to bear on approval decisions, this provision opened the door to a world of data fishing. We have seen how for cancer and AIDS drugs, the FDA had already been doing a great deal of investigation and analysis beyond the contents of NDAs to evaluate drugs perceived to be therapeutically important (and having legal threats made as if Subpart E required them to do it) — often with the result that the studies presented in the NDA were virtually ignored due to their inadequacy while other sources of knowledge were sought. Under H.R. 1995 this cast-the-nets approach could potentially apply to *any* drug approval, not only to drugs for serious and life-threatening diseases.

Second, this passage appears to be the earliest legislative attempt to edit the 'substantial evidence' clause explicitly to allow single-study new drug applications. As we have seen, although the FDA approved AZT on the basis of a single Phase II study, the Agency specifically excluded this feature from the Subpart E rules, writing that 'the effectiveness of a drug should be supported by *more than one* well-controlled clinical trial and carried out by independent investigators' (FDA 1988a, 41521, emphasis added) and reiterating this position in its publication of the Final Rule for the Subpart H regulations (FDA 1992b, 58948). Nevertheless, after the highly visible example of AZT, the precedent for single-study trials seems to have been established in many minds.

According to Dr. Ellen Cooper, former Director of the FDA Antiviral Division, in the late 1980s ‘in general there was this sense that one controlled study would be enough’ (Cooper 2006).²⁵ Although H.R. 1995 never made it out of the Commerce Committee and other, yet-to-be-introduced versions of legislation finally became the FDAMA, the unique wording of this definition as ‘including data from one well-controlled clinical investigation and confirmatory evidence (obtained either before or after such investigation)’ appeared in the FDAMA almost verbatim. Hence the provenance of this provision appears to be traceable to PhRMA. This definition was cannily phrased such that the timing of confirmatory data collection was left open to question. In the context of a procedure like Subpart H accelerated approval, it could even be interpreted to mean that the confirmatory evidence could come after approval (although this is not the interpretation the FDA subsequently applied). This aspect of the definition would later cause some vexation for FDA regulators trying to work out a meaningful interpretation for daily practice (see Temple 2001).

In Fox’s bill we see two examples of how past practices or concepts originally intended to be used in extenuating circumstances can eventually become normalized for applicability to all drugs. The ‘whole knowledge’ approach to drug approval was not ultimately legislatively formalized, but the single-study standard for drug approval was. Ultimately, there is no conceptual firewall between procedures for therapeutically important drugs and all others.

Although Wyden did not explicitly provide for the acceptance of one clinical trial as the basis for drug approval in his legislation as Fox’s bill did, he subsequently embraced the idea — but only for ‘breakthrough’ products, meaning products representing therapeutically important advances — telling Dr. Lasagna in a hearing the following month (U.S. House 1995c) that he was calling for a ‘fast-track approval process, what amounts to a conditional approval . . . [for] breakthrough drug and device products. They could come forward on the basis of one well-designed clinical trial’ (148-9). Lasagna responded that the Agency had always had a ‘tradition’ of ‘tailoring the amount of information required for approval to the importance of the drug being proposed’ (149), noting AZT as an example. Then he added:

²⁵ This perception was pervasive enough that in a September 1990 meeting of the Antiviral Drugs Advisory Committee, Cooper was moved to declare: ‘Sponsors should stop planning to rely on a single Phase II study to demonstrate efficacy and then scramble to retrospectively turn Phase I or expanded access data into adequate and well-controlled studies’ to fabricate additional data, as needed (FDA 1990b, 28).

I think approving a drug on the basis of one trial for drugs that don't represent breakthrough advances is probably not desirable or defensible, but the more important the drug, the less already available to treat the condition under study, the more one should move in the direction you've just espoused, namely, requiring enough information to make everyone believe that if the drug is marketed and used as labeled, it will do a lot more good than harm, and then filling in the gaps later on (149).

This recommendation from Lasagna is notable, since the latter had been a consistent critic that FDA required too much evidence for drug approval.

In December of 1995 came proposed legislation for the *FDA Performance and Accountability Act of 1995* (S. 1477), introduced in the Senate by Senator Nancy Kassebaum (R-KS). This bill incorporated features of other, previously introduced bills, and sought to legislate many practices already in use at the FDA. For example, the bill allowed seriously ill patients who lack effective therapies to acquire investigational drugs thorough the following mechanisms, all of which were already practiced by the FDA: protocols for parallel track, treatment use (treatment IND), single-patient protocols, emergency use (a single-patient protocol in which the use precedes the protocol), and uncontrolled trials (e.g., large, open-protocol 'trials'). Like the earlier bills, this proposed legislation would return the IND process to a system of notification, with permission to begin clinical testing automatically granted after thirty days if the FDA did not intervene (S. 1477, Title III). This legislation also set time limits for review of NDAs, and required the FDA to contract with outside experts to review applications if the time limits were not met using internal review. Moreover, if time limits were not met on NDAs for products already approved in the UK or EU, then the product would be approved by default. Notably, Title VI of the legislation borrowed from Rep. Fox's bill the redefinition of 'substantial evidence' with one important difference of wording. The Kassebaum version states that substantial evidence 'may consist of data from one well-controlled clinical investigation (which may be waived by the Secretary) and confirmatory evidence (obtained either before or after such investigation)' (S. 1477, Title VI). According to the subcommittee report to the Senate recommending the bill for full Senate consideration (U.S. Senate 1996), this confusing wording (could even the single study be waived?) was intended to confirm existing FDA practice in using discretion to approve on the basis of one study or to require additional studies.²⁶

²⁶ See page 38 of the senate report: "The FDA usually interprets the requirement to demonstrate substantial evidence of effectiveness to require two adequate and well-controlled clinical studies, but has shown

More bills were to come. In the first quarter of 1996, two brief, single-issue bills were introduced into the House. H.R. 2932, introduced by Rep. Richard M. Burr (R-NC) in February 1996 was designed only to allow drug representatives to give reprints of peer reviewed articles to physicians if the articles were accompanied by a package insert and a disclaimer that the article had not been approved by FDA.²⁷ H.R. 3149, the *Terminally Ill Access to Medical Treatment Act*, was introduced into the House by E. Clay Shaw (R-FL) in March 1996. This similarly brief proposal would have mandated the FDA to approve drugs for limited distribution to terminally ill patients when the drug was not approved for non-terminally ill patients. The bill contained no specific information on what standard of evidence should be used for this limited approval (other than to say the drug should be deemed ‘safe’), how the approval could be withdrawn if the drug was found to be ineffective or unsafe, or how one could restrict approval only to terminally ill patients when off-label prescribing was legal and widely practiced.

Meanwhile, in March 1996 President Clinton announced a ‘National Cancer Initiative’ as part of the national performance review (President 1996; Anonymous 1996). The four main proposals of this initiative were, first, for the FDA to approve cancer drugs on the basis of *partial* tumour shrinkage when acceptable alternative therapies are lacking. Second, under this initiative the FDA would urge companies with cancer drugs under investigation to participate in the treatment IND programme. More than that, the FDA would contact companies with cancer therapies already approved in other countries but still considered investigational in the U.S. and invite them to make their products available to eligible patients through treatment IND when alternative therapies were unavailable. Third, the FDA would begin to reject supplemental applications submitted unnecessarily. Finally, the FDA would expand patient representation on their oncologic drugs advisory committee. Once again, these proposals either formalized procedures already in use by the FDA (e.g., approval on the basis of partial tumour shrinkage which, as we have seen in Chapters 5 and 7, was one of many practices employed in the surprisingly flexible standards of evidence used for some situations) or made administrative adjustments to

flexibility and approved some drugs on the basis of one adequate and well-controlled clinical study. The legislation confirms the current FDA interpretation that substantial evidence may, as appropriate, consist of data from one adequate and well-controlled clinical investigation and confirmatory evidence (obtained before or after the investigation)?

²⁷ Under existing law, the FDA was able to restrict such distribution of articles on off-label uses by considering the latter an act of promotion, and all materials used for promotion of a drug had to be reviewed by the FDA prior to dissemination.

existing FDA practices (e.g., ‘urging’ drug sponsors to participate in treatment IND, rather than ‘inviting’ them to do so as specified in Subpart E).

The rest of the year would see more substantial attempts at legislation in the Congress. Another, more comprehensive bill, H.R. 3199, introduced by Rep. Burr (R-NC) in March 1996, was called the *Drug and Biological Reform Act of 1996*. This bill sought to redefine ‘substantial evidence’ in multiple ways: not only in terms of data quantity (‘one or more clinical investigations’), but also in terms of the type of controls used in clinical trials and the type of disease studied: ‘a well-controlled investigation shall only include methods of control that are appropriate to the disease or condition for which a drug is intended as prescribed, recommended, or suggested in its labeling’ (Sec. 5). The legislation envisioned that in some cases, *no* controlled studies would be necessary: ‘The Secretary may waive the requirement to conduct any well-controlled investigation’ (Ibid.). Moreover, for drugs intended to treat life threatening conditions, the requirement for substantial evidence would be considered to be fulfilled ‘if it could be fairly and responsibly concluded by [experts] that there is a reasonable likelihood that the drug will be effective in a significant number of patients and that the risk from the drug is no greater than the risk from the condition’ (Ibid.). Additionally, this legislation sought to define any widely used off-label prescribing practices as ‘substantial evidence’ for the sake of a supplemental application: ‘if an application has been submitted for a new use of a previously approved new drug, such drug shall be considered to have substantial evidence by a demonstration that the new use is common among clinicians experienced in the field and represents reasonable clinical practice based upon reliable clinical experience and confirmatory information’ (Ibid.). The creative redefinitions of substantial evidence written into this bill are notable. Some of them, such as flexibility in application to drugs for life-threatening diseases or flexibility in the number of studies accepted, were already used in practice by the FDA, as we have seen in previous chapters. Widespread off-label use of physicians for a certain indication, however, while perhaps influential when clinical studies were also available, had not been considered a valid basis for allowing drug marketing since before 1962. The bill never made it out of committee. The 104th congress ended without any major legislation passed to reform drug and device approval.

8.3 FDAMA

8.3.1 *Legislation and the Art of Compromise*

The Republicans retained control of both the House and Senate in the November 1996 congressional elections, but momentum on the issue of reform had been lost. A *Washington Post* business section headline reported that ‘After the Election, it’s Oversight Rather than Overthrow’ (Skrzycki 1996): ‘Clearly, the thrill is gone. The last Congress’s capers of shutting down the government, trying to slash agency budgets and attempting to shutter whole departments are over — and widely viewed as a failure with voters’ (Skrzycki 1996, C11). Republicans were regrouping, reworking their reform agenda to proceed ‘piecemeal, if not wholesale’ (Waldmeir 1996, 4). Likewise for the industrial and business interests which two years beforehand saw an opportunity to write their own legislation. Skrzycki (1996) quoted a Washington business lobbyist as saying, ‘We’re going to operate with an agenda that is thoughtful, targeted and less sweeping’ (C11).

At the beginning of the 105th Congress, Rep. Fox reintroduced his bill for the *Life Extending and Life Saving Drug Act* to the House (now designated H.R. 1094, introduced on 18 March 1997). The legislation was once again referred to the House Commerce committee, where it appears to have died a quiet death. Rep. Burr also tried again, introducing a *Drug and Biological Products Modernization Act of 1997* on 23 April 1997 (H.R. 1411, Version 1).²⁸ Burr’s provisions for the determination of effectiveness were notably different from the previous year’s version (see H.R. 1411, Sec. 5) and the differences between the two bills are telling. Burr’s new bill retained the earlier definition of substantial evidence in terms of allowing flexibility for the type of study controls used and still sought to give the FDA latitude to waive the need for a clinical study altogether. Now, however, the earlier proposal to define the off-label prescribing practices of physicians as a form of substantial evidence was gone. Instead, for new applications of products already approved for another indication, this bill would require the FDA to ‘consider whether the new use is supported by adequate valid scientific information from clinical investigation reports in peer reviewed medical and scientific journals, patient registries or compendia, and other sources recognized by the Secretary’ (Sec. 5). Moreover, the earlier proposal to create a different (vague and very weak) definition for substantial evidence when applied to drugs intended to treat life threatening diseases was

²⁸ There are two versions of this bill: Version 1, which is the bill introduced into the house in April 1997, and Version 2, which is the version of the bill reported out of the House on 9 October 1997 for consideration by the Senate.

also gone. Instead, the proposed legislation required the FDA, when making a determination of effectiveness for a drug to treat a serious or life-threatening condition, to ‘consider whether the benefits of the drug outweigh the known and potential risks of the drug and the need to answer remaining questions about risks and benefits of the drug, taking into consideration the severity of the diseases and the absence of no [*sic*] comparable or satisfactory alternative therapy now’ (Sec. 5). In other words, this proposal would legislate the risk-benefit principle articulated in the Subpart E rule and already practiced by the FDA. Notably, the bill also borrowed some language from Fox’s definition of substantial evidence, which according to Burr’s bill, ‘may, where there is a high level of confidence in the scientific validity of the results of an adequate and well-controlled investigation, consist of data from an adequate and well-controlled investigation and adequate supportive scientific evidence (obtained before or after such investigation)’.

Clearly, this was a more moderate version of the earlier bill, and many of the provisions which likely met resistance in Congressional committees the previous year had now been eliminated. The version of the bill ultimately reported out of the House (H.R. 1411, Version 2) dropped the section relating to the content and review of new drug applications and the determination of effectiveness (Sections 4 and 5 of Version 1). The final version added Fast Track provisions authorizing the FDA to approve a drug on the basis of a surrogate endpoint and to require postmarketing studies (Section 4 of Version 2), as well as authorizing expanded access to investigational drugs of needy patients lacking therapeutic alternatives (Section 5 of Version 2). The text of this new bill was incorporated into a sympathetic Senate bill, S. 830, introduced by Senator Jeffords (R-VT) on 5 June 1997. After much amendment in the Senate, and negotiation between the Senate and House over the final shape of the bill, S. 830 was the bill which ultimately became the FDAMA on 21 November 1997.²⁹

The provisions of the FDAMA clearly represent compromise, and a softening of the legislative agenda of the previous Congress. For example, while the law did

²⁹ The text of the FDAMA can be found in a number of places, including the FDA website: <http://www.fda.gov/cder/guidance/105-115.htm> (accessed 10 November 2007). Note that the FDAMA modifies the Food, Drug and Cosmetic Act (FDCA), so it must be kept in mind that the FDAMA text does not stand alone, but can only be fully understood as it is inserted as designated into the FDCA, which can be found in the U.S. Code, Volume 21, Sections 301 forward. Further, as will be discussed later in this chapter, the stated Congressional intention behind the FDAMA can be assessed in the report of the Senate committee which negotiated the bill and recommended it to the full Senate for a vote (U.S. Senate 1996).

incorporate a notification system to begin early clinical investigations (Sec. 117), it continued to require affirmative approval for new drug applications (Sec. 119). It did establish the use of ‘accredited persons’ (external to the FDA) for certain limited device review functions (Sec. 209), but disallowed their use for higher risk devices (Sec. 210) and did not mandate outsourcing of FDA functions to private contractors or external review panels for drug application reviews. It deferred the question of whether the FDA should accept data summaries instead of raw data in NDAs by requiring the FDA to study the issue and to publish within a year its conclusions about which types of data would be acceptable in summary form (Sec. 118). It reauthorized the PDUFA for another five years, providing for the collection of user fees for drug applications (with some exceptions) but not for devices (Sec. 101-107). It also addressed many issues that were never hot-button questions of reform in the previous Congress, but had been advocated in some hearings. For example, it encouraged drug-makers to conduct clinical studies of drugs for paediatric uses by offering extended patent protections for such products (Sec. 111). It also directed the FDA to work with industry to develop guidelines for including more women and racial minorities in clinical studies (Sec. 115).

These and many other examples demonstrate how the force of the previous neoconservative legislative agenda had been blunted in the FDAMA, leading to compromises on key issues. Indeed, we could consider the FDAMA a remnant or artefact of the attenuated Contract with America. This law would likely never have come about (certainly not in the form that it did) without the political changes of the 1994 elections. In the FDAMA, one can see reflections of the neoconservative agenda but they are fragmented, coexisting with elements of other, more moderate agendas. In this sense, the FDAMA represents the Democratic ideal: everyone got something but, in the end, no one vision appears to have dominated the rulemaking process.

8.3.2 *FDAMA and Existing Practice*

For our purposes here, of course, the most important features of the FDAMA are those relating to ‘Fast Track’ products, the definition of substantial evidence, and expanded access of investigational drugs. In these areas, the degree to which the FDAMA legislated practices already in use by the FDA is notable. This legislation directed the FDA to appoint expert advisory panels to assist in review and decision-making on new drug applications (Sec. 120). It required the FDA to meet with drug sponsors who

request meetings (Sec. 119). It required the FDA to allow access to investigational drugs for patients having serious or life-threatening diseases (Sec. 402). Additionally, for drugs ‘intended for the treatment of a serious or life-threatening condition’ and demonstrating ‘the potential to address unmet medical needs’, a drug sponsor could apply for ‘Fast Track’ designation coincident with, or any time following, submission of the IND application (i.e., from the beginning of Phase I testing). If a drug qualified for Fast Track designation, then the FDA was required to ‘take such actions as are appropriate to expedite the development and review of the application for approval of such product’ (Sec. 112). Under Fast Track, the FDA was granted specific authority to approve new drugs upon determination ‘that the product has an effect on a clinical endpoint or on a surrogate endpoint that is reasonably likely to predict clinical benefit’ with the option to require postmarket validation of the surrogate on the effect on a clinical endpoint, and with provisions for accelerated withdrawal of drugs approved under this mechanism (Sec. 112).

The Senate committee which negotiated this legislation and recommended it to the full Senate for consideration was aware that many of these provisions echoed existing practices. However, they appeared to believe that longstanding FDA policies and procedures had thrived in the gaps of the *Food, Drug and Cosmetic Act* (FDCA) like trees growing in the cracks of a cliff. They therefore wanted to amend the FDCA to provide legislative support for existing practices. For example, on the subject of expanded access, the committee wrote that for ‘many years, the need for patients to have access to unapproved therapies went unrecognized under the Federal Food, Drug and Cosmetic Act. The FDA established informal policies relating to compassionate use of investigational products shortly after enactment of the 1938 Act, but these policies remained informal and outside FDA regulations’ (U.S. Senate 1997, 16). The committee noted that the FDA had since established programmes for expanded access, ‘some of which are embodied in regulation’, and acknowledged that the programmes had been effective in expanding access to patients having HIV and cancer. However, the committee wrote that it wished to provide ‘statutory direction’ and to ‘emphasize that opportunities to participate in expanded access programs are available to every individual with a life-threatening or seriously debilitating illness for which there is not an effective, approved therapy’ (16). Indeed, the FDAMA extended the provisions for expanded access to both drugs and devices.

The committee had similar goals in creating the Fast Track provisions of the law. In their report, the committee acknowledged FDA existing practice to expedite drug approval for life-threatening conditions, but felt that the practical application had been too limited in scope: ‘For several years the FDA has allowed the expedited review and approval of drugs but such review has been largely confined to treatments for HIV/AIDS or cancer. This provision facilitates development and expedites approval of new drugs for the treatment of any serious or life-threatening diseases’ (3).³⁰ They therefore perceived a need for ‘a formal statutory mechanism for identifying breakthrough drugs early in product development that provides sponsors of such drugs a reasonable opportunity for early interaction with the agency [which] may help to further streamline the development and approval processes for such drugs’ (43).

On the redefinition of ‘substantial evidence’, the Senate committee observed that technical advances in the design and conduct of clinical trials over the previous 35 years made it possible to create ‘independent substantiation’ of a clinical outcome measurement within a large trial designed for such replication of results. Thus, the committee believed that ‘the structure of a particular clinical protocol and the quality of the data underlying a new drug application should guide FDA’s substantiation requirements’ (31). Noting that the ‘FDA usually interprets the requirement to demonstrate substantial evidence of effectiveness to require two adequate and well-controlled clinical studies, but has shown flexibility and approved some drugs on the basis of one adequate and well-controlled clinical study’ (30-1), the committee asserted that the legislation ‘confirms the current FDA interpretation that substantial evidence may, as appropriate, when the Secretary determines, based on relevant science, consist of data from one adequate and well-controlled clinical investigation and confirmatory evidence (obtained either before or after the investigation)’ (31). Despite this representation as essentially statutory codification of existing FDA practice, it is clear that this is a significant in-principle change, since the FDA’s ‘flexibility’ to approve drugs on the basis of one study was exercised in the review of drugs intended to treat serious and life-threatening diseases, while the law made it generally applicable.

³⁰ To be clear, both Subpart H and Subpart E were theoretically designed for *any* drug intended for life-threatening or serious conditions, not just AIDS and cancer. It just so happened that AIDS and cancer were the two primary life-threatening diseases most obviously in need of new therapies. Although, as we have seen, representatives of other diseases such as Alzheimer’s commented on the rules to assure that should a new therapy be developed for the condition they represented, it could in principle receive consideration under expedited approval.

Let us step back for a moment and look at the span of rule-making activity described in the thesis to this point. FDA policies and regulations were developed over time to fill in the gaps of the major legislation guiding their activities.³¹ As a result, the FDA created formal rules modelled on previous experience and existing practice in an effort to maintain parity between regulations and practice; then Congress created a law with provisions substantially based on FDA regulations and policies in an effort to provide ‘legislative support’ for FDA practice—in other words, at least part of the motivation was to maintain parity between laws and regulatory practice. In this way, two levels of formal rule-making have had to be revised to adapt to actual practice. Once again, rules followed practice. We have now seen this principle on both the regulatory and legislative levels.

Clearly also many of the provisions of the FDAMA sought to correct or improve FDA practice where it was felt to be inefficient or ineffective. Rather than reiterating and reinforcing FDA practice, this class of rules are reactionary in nature. They came about as a negative reaction to certain existing practices, and were intended to mitigate the perceived negative effects of those practices. Thus, even those elements of the FDAMA which do not significantly reiterate existing practices are nevertheless based on them, albeit in a negative sense; their inspiration and motivation lies in their critique of existing practice and they are developed in relation to that critique. Whether we wish to speak in a positive or negative sense, therefore, it is clear that the FDAMA was ‘based’ on FDA practice in some sense. Even seemingly ‘new’ rules can be said to follow practice when considered in this light.

What effect did this legislative intervention to create Fast Track have on subsequent regulatory practice? What happens when a new category of drug approval is created? In the next section, we will attempt to track the meanings of new and old categories in relation to one another over time.

³¹ Legislation is often purposely written in general terms to allow regulatory agencies the flexibility they require to address specific technical issues and changing circumstances. However, the level of specificity such laws should invoke has been a matter of debate among legal scholars (Harter 1982; see especially the introduction). In this case, clearly, legislators decided to intervene and create a higher level of detail than previously existed with the stated intention of giving ‘direction’ and hoping to refine and improve some of the process. This intervention was a remnant or effect of an especially turbulent period in Congressional politics, one in which FDA reform was a political football in an ideological struggle over the role of government in general and regulatory agencies in particular. We can suppose as a general rule that detail-specific legislative intervention is more likely to take place in such times of legislative conflict and competition.

8.4 Evolving Regulatory and Statutory Interpretations

8.4.1 *Reconciling Fast Track with Subpart H*

Under the terms of the FDAMA, the FDA was obligated to issue a guidance document within one year to ‘provide physicians, patient organizations, industry, and other appropriate persons a comprehensive description of the fast track provisions established under this legislation’ and to specify ‘the policies and procedures pertaining to fast-track drugs’ (U.S. Senate 1997, 44). The result, published in 1998, was *Guidance for Industry for Fast Track Programs* (FDA 1998b). The version of Fast Track presented in this guidance document forms an interesting and instructive contrast to the vision of Fast Track presented by the Senate committee in its report on FDAMA.

As noted previously, the fast track provisions of the FDAMA were an amalgam of Subparts E and H, with the heart of Fast Track virtually indistinguishable from Subpart H. Indeed, in the 1998 guidance document on fast track programmes, the FDA wrote that the FDAMA in effect ‘codifies in statute FDA’s Accelerated Approval Rule’ (FDA 1998b, 2), and again that it ‘essentially codifies in statute FDA’s accelerated approval regulations’ (15). That it did so, the Agency wrote, ‘affirms FDA’s authority to base marketing approval on data other than clinical efficacy data directly establishing an effect on the ultimate clinical outcome’ (2). In other words, the FDA saw the fast track provisions as a vindication of its decision to allow unvalidated surrogate markers as endpoints — a decision that created concern for some observers.³² Similarly, the agency wrote that the discretion granted by the FDAMA to approve drug applications based on only one clinical study ‘confirmed FDA’s interpretation of the statutory requirements for approval’ (FDA 1998b, 4).

While fast track may have been a vindication of FDA’s approach to accelerated approval, how was the FDA to implement it? Should the accelerated approval rule be effectively scrapped and replaced by fast track? This certainly would have been a feasible approach, and could even be seen as the congressional intent, although it is nowhere explicitly stated as such. This was not the FDA’s approach, however. Unwilling to amend or eliminate the established rule, the FDA treated the Fast Track law not as

³² See the FDA’s responses to comments published in the Final Rule for Subpart H regulations (U.S. FDA 1992). Many of the comments are critical of, and express concern over, the formal acceptance of unvalidated surrogate endpoints as allowing (in the words of one commenter) ‘belief rather than evidence to serve as the basis for a conclusion about the effectiveness of a new drug’ (p. 58943).

superseding any of its existing practices or policies, but as forming another layer atop of them. This view of the rules is clearly stated in the introduction to the guidance document on fast track, which declares that the FDAMA authorizes the FDA to facilitate development and expedite review of products intended to treat serious or life-threatening conditions, and that these ‘actions are not limited to those specified in the fast track provision but also encompass existing FDA programs’ to accomplish these goals (1). The document goes on to note that such programs include the Subpart E and Subpart H provisions, as well as priority review policy procedures and policies established by the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER).

Given this view of fast track as an overlay to existing practice, the Agency strove to distinguish between fast track and Subpart H. The FDA wrote that since by definition fast track drugs are intended for a serious or life-threatening disease and since they also must demonstrate the potential to address unmet medical needs, fast track therefore ‘does not apply to a product alone, but applies to a combination of the products and specific indication for which it is being studied’ (FDA 1998b, 3). Consequently, it is ‘the development program for a specific drug for a specific indication that will receive fast track designation’ (3) — not the drug itself, as in the case of Subpart H. However, given that the Subpart H rules also apply to drugs ‘for serious or life threatening illnesses’ in cases where the drug would ‘provide meaningful therapeutic benefit compared to existing treatment’ (FDA 1992d, 58942), it would seem that this ‘distinction’ is based on a false contrast. Subpart H is likewise designed to address a product and the indication for which that product will be used. More than that, the phrasing in the Senate committee report makes it clear that Congress was thinking of fast track designation as applying to a drug. (‘If the FDA determines that the drug should be designated as fast track, it will take appropriate action to expedite the development and review of the drug’)(U.S. Senate 1997, 43). Nevertheless, the FDA created this distinction and defined fast track such that once a drug development program received fast track designation, the drug sponsor would find a range of possible review options available, including priority review, ‘rolling’ review of applications, and accelerated (Subpart H) review (see FDA 1998b, 12-16). While there is some phrasing in the Senate committee’s explanation of the legislation which could be seen as supporting the notion of fast track approval as an umbrella programme for

expediting approval,³³ treating it in this manner inadvertently negates one of the key differences between fast track and Subpart H. Whereas Subpart H made postmarket study to confirm an endpoint or clinical outcome mandatory, fast track left postmarket commitments as *optional* at the discretion of the FDA (U.S. Senate 1997, 44). By maintaining Subpart H as one tool among many for fast track drug development, postmarket study is still mandatory for decisions based on surrogate endpoints, at least according to the letter of the regulation in Subpart H.

8.4.2 *Flexible, Invisible Subpart E*

As we have seen, the original, literal set of procedures specified by Subpart E was to collaborate with the sponsor at the end of Phase I to design *two* Phase II studies based on *clinical* endpoints, invite the sponsor to participate in treatment IND as the study became mature, and make a decision based on an assessment of the perceived benefit of the drug weighed against the risk associated with the information deficit associated with phase II studies, possibly requiring postmarket studies to fill the information gaps. In the 1998 fast track guidance document, the description of Subpart E emphasized early sponsor-FDA collaboration to produce effective study designs:

Under the Subpart E regulations for investigational new drugs . . . drug development is considered a continuum from early preclinical and clinical studies through submission of a marketing application. The regulations emphasize the critical nature of close early communication between the Agency and a sponsor, outline procedures such as pre-IND and end of phase 1 meetings as methods to improve the efficiency of preclinical and clinical development, and focus on efforts by the Agency and sponsor to reach early agreement on the design of the major clinical efficacy studies that will be needed to support approval (FDA 1998b, 1).

While this description of Subpart E is accurate, the themes the FDA chose to extract and those it chose to ignore from Subpart E are notable. The Subpart E regulations established the original formal procedures and rationale for compressing the clinical trial system; they articulated a new risk-benefit approach to drug approval in which approval would be based on less information than normally desired, and in which that lack of information was explicitly part of the ‘risk’ to be considered in the approval

³³ “This legislation is intended to clarify and coordinate some of FDA’s mechanisms for new drugs and biological products that are intended for the treatment of serious and life threatening conditions and that demonstrate the potential to address unmet medical needs for such conditions. It defines and clarifies the processes pursuant to which sponsors of these drugs may interact with the FDA and includes provisions that will ensure that these processes are well known and well understood” (U.S. Senate 1996, 43).

decision. In the 1998 guidance document, however, the main goal of Subpart E was to provide for early FDA-sponsor collaboration to improve efficiency and create successful clinical trial designs. Within this re-representation of Subpart E, clinical phases are effectively irrelevant; drug development is a ‘continuum’ in which Agency-sponsor collaboration is key. Significantly, this is the only mention of the relevance of Subpart E to fast track development programmes in this guidance document, and it seems tailor-made for the context. As I noted above, part of the intention of fast track drug development was to offer sponsors of drugs for life-threatening diseases early opportunities for collaboration with the FDA. Accordingly, this description of Subpart E fits that application.

We have seen this phenomenon before, when Subpart E was interpreted in terms of selected principles rather than specific procedures in the awkward application of Subpart E to ddI (see Section 6.2 of this thesis). In that representation, Subpart E was said to reflect a recognition that physicians and patients were willing to accept greater risk of debilitating adverse effects for more serious diseases, and that therefore the benefits of the drug needed to be evaluated in consideration of the severity of the disease. This earlier interpretation of Subpart E was also said to mean that the FDA needed to exercise the ‘broadest flexibility’ in applying the substantial evidence standards to drugs intended for life-threatening diseases. I also noted in Chapter 6 that in the aftermath of the approval of AZT, there seemed to be a scarcity of cancer and AIDS drugs developed and approved on the Subpart E model. The lack of real-world applications for the procedures in Subpart E seem to make it especially subject to what I have termed resignification. When applied to the situation with ddI (when the Subpart H rule was yet to be published), the emphasis was on conceptual features most salient to the ddI approval: risk-benefit evaluation and flexibility. When deployed as a general principle for a guidance document on fast track, conceptual elements relating to early sponsor-FDA collaboration became more applicable.

More recently, even more definitional shifting of drug approval categories is apparent. The FDA subsequently moved away from seeing the fast track law as a codification of Subpart H rules. The FDA webpage for fast track development programs³⁴ states that fast track is ‘independent’ of accelerated approval; that fast track ‘refers to a process for interacting with the FDA during drug development’ (note that this

³⁴ www.accessdata.fda.gov/scripts/cder/onctools/Accel.cfm (accessed 5 April 2004).

phrasing is reminiscent of the agency's earlier characterization of the Subpart E regulations) while Subpart H accelerated approval 'applies to the design and content of the studies used to support a marketing claim'. Moreover, at this webpage Subpart E is conspicuous by its absence: the page provides lists of approved drugs classified as either priority review or as Subpart H accelerated approval, but does not identify any drugs as approved under Subpart E (nor, indeed), nor does the webpage even mention Subpart E as a category. Although the FDA does meet with sponsors and offers consultation on clinical trial designs, for all practical purposes Subpart E has been erased as a set of specific procedures.

8.4.3 *Clinical Phases in Flux*

In earlier portions of this thesis, I demonstrated the care and detail with which FDA sought to define and discuss the clinical phases of drug evaluation. In Chapter 2, we saw that the first formulation of the clinical phases as part of the original IND rules (FDA 1963) specified two phases of clinical pharmacology followed by a third phase to determine efficacy. This last phase was the only one deemed a 'clinical trial'. According to the *Concept Document* (1979a), actual practice gradually diverged from this characterization such that a new description of the clinical phases needed to be written. Thus, both the *Concept Document* and the almost-concurrent *Clinical Considerations* guidance document produced by the FDA (FDA 1977) represented Phase II as the time when first evidence of effectiveness is generated while Phase III studies served as 'expanded' trials for more information. At the same time, I noted that pending FDA reform legislation contained provisions which also effectively promoted Phase II to the phase in which the first statutorily defensible evidence for effectiveness was generated. The clinical phase seemingly undergoing the most instability of meaning was Phase II, the function of which seemed to become increasingly important through this period in practice, even if the earlier version of them persisted in the *Codes of Federal Regulation* (see note 6, Chapter 4). The advent of AZT and Subpart E then heaped even greater responsibility on Phase II as an engine for effectiveness data, and provoked renewed consideration of the meaning and function of clinical phases.

The FDAMA would contribute to a continued re-evaluation of clinical phases, this time because of its redefinition of substantial evidence. The FDA, now dealing with a revised standard of evidence, published another guidance document in 1998 on providing

clinical evidence for the approval of new drugs and biologics (FDA 1998a). This document was designed not merely for fast track products or those intended to treat life-threatening diseases, but for all new drugs and biologics. Here, the FDA underscored sharply its longstanding policy that in rare instances there may be justification to approve a drug on the basis of a single study, however scientific evidence requires ‘*independent substantiation*’ (4, original emphasis) and therefore a ‘single clinical experimental finding of efficacy, unsupported by other independent evidence, has not usually been considered adequate scientific support for a conclusion of effectiveness’ (4). The document listed the various biases and irregularities common to clinical studies, stressing that the statistical and methodological safeguards developed to combat such biases ‘are often inadequate to address these problems in a single trial’ (5). Despite this ringing assertion that independent substantiation of clinical evidence will almost always be required for approval — and for good reason — the FDA nevertheless included in this guidance document a section on how to create ‘Evidence of Effectiveness from a Single Clinical Study’ (12). Not surprisingly, each of the suggestions for creating an acceptable single study tend towards the type of design features traditionally associated with Phase III studies: large studies conducted across multiple sites and institutional centres; ‘stratified’ studies, necessarily large because patients are randomized to subgroups based on disease severity, geographic residence, or demographic characteristics, effectively creating multiple smaller studies within one larger one; studies evaluating multiple endpoints and events, which again need to be relatively large to produce a statistically valid result for each endpoint measured; studies with ‘extreme p-value’, meaning a statistically powerful study, which typically requires a large population, especially if stratification or other design features dependent on the creation of analytic subgroups is used. Traditionally, studies with such designs would be considered Phase III. But what if a drug sponsor with a promising Phase I cancer drug obtained fast track designation and designed a large, multicentre, randomized trial to evaluate surrogate endpoints? Would such a study be considered Phase II or Phase III — or Phase II/III? Such hybrid designations are increasingly common, in part because of the technical advances made over the last decades in clinical trial design and statistical analysis. They are also more common because of the persistent demand to get potentially important drugs through clinical evaluation more quickly. We can identify another reason that the FDA’s guidance document on fast track drug programs sought to represent Subpart E as a symbol of the ‘continuum’ that is clinical

drug development. In the face of both political pressures and practice-based technical improvements in evaluating drugs for life-threatening diseases, once-discrete regulatory categories for clinical phases have been collapsing into one another.

8.5 Fast Track in Application: One Case

The Agency was charged with the task of setting out the rules for how fast track would work with other existing procedures even as it was obliged to begin accepting fast track applications. The FDA began accepting applications for fast track designation before it had even finalized the guidance document for industry, which had a deadline of November 1998 for completion.³⁵ In a congressional hearing on FDAMA implementation held in October 1998 (U.S. House 1998), FDA officials testified that 34 requests for fast track designation had been received and 19 had been granted. Unfortunately according to the FDA, information about fast track requests and associated FDA decision-making are not publicly reportable unless the drug sponsor already publicly released the information themselves (49). This prohibition hinders the task of examining FDA decision-making on this new statutory category of therapeutic product development and review.³⁶ One success story the FDA could discuss in this hearing was the breast cancer therapy Herceptin. This novel and important breakthrough monoclonal antibody therapy for cancer was granted fast track status in one day, given a ‘rolling review’ under fast track, and approved in four-and-a-half months (49). While this may sound like a dramatic success story for the FDA, obviously these studies were initiated prior to the passage of FDAMA, and in fact were completed before the FDAMA was made law.³⁷ Hence, in this case fast track designation did not facilitate ‘early’ identification of a

³⁵ A draft version was released in 1997 (FDA 1997) and the final version was not published until May 1998 (FDA 1998a).

³⁶ Even after a drug has been approved, it is not identified at the FDA website as having been a fast track product, which is bizarre, considering the level of information otherwise available for approval drugs (although, sadly, not for products whose NDAs have been rejected). More than that, for approved drugs the FDA now posts detailed summaries of their review of NDA data, copies of approval letters listing postmarket study commitments, links to advisory committee transcripts, etc. Look at any new chemical entity approved after about 1998 at the ‘Drugs@FDA’ website, <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm>.

³⁷ See the FDA clinical review summary document, <http://www.fda.gov/cder/biologics/review/trasgen092598r1p1.pdf> (accessed 12 November 2007) One pivotal trial was initiated in April 1995 and completed in June 1997; the other was initiated in June 1995 and completed in March 1997. FDAMA was made law in November 1997 (see legislative history in Appendix B).

breakthrough drug, nor did it foster FDA-sponsor collaboration on design of a development programme. By the time fast track status was granted, all that remained to be done was data evaluation and rolling review. It is likely that many of these very early fast track drugs were, like Herceptin, in advanced states of evaluation when fast track designation was granted.

Fortunately, we do have access to some detailed information about one case of fast track drug decision-making. In 2000, a small biotechnology firm called ImClone approached the FDA with preliminary data from a Phase II study, seeking to investigate the possibility of attaining fast track status. The investigational drug at issue, Erbitux (also sometimes called C-225) was another monoclonal antibody, a class of agents for which expectations had been growing because of the effectiveness of Herceptin in breast cancer (U.S. House 2002, 52). What happened from this point forward is a matter of some debate. At the heart of the issue is a biologics license application (BLA) so botched that the FDA took the unusual action of issuing a ‘refusal to file’ letter at the end of 2001, stating that ImClone’s Phase II study was ‘neither adequate nor well controlled’ (U.S. House 2002b, 20). Expectant patients who were denied investigational access to this highly anticipated drug (see note 18, this chapter) were stunned, and many complained to Congress. The resulting congressional hearings (U.S. House 2002) afford us a rare and valuable window onto a case of fast track decision-making and failed drug approval. For our purposes here, I will try to disentangle the threads of the story having to do with confusion over the meaning of fast track and the standards of evidence applicable for fast track products.³⁸

³⁸ This is a complicated case in which some company representatives (especially co-founder and CEO, Dr. Sam Waskal) may have not been ethically rigorous in their representations to the FDA and the public about Erbitux and the clinical data supporting it. It is consequently difficult to judge whether ‘misunderstandings’ between the FDA and drug sponsor were genuine miscommunications or whether, at times, ImClone wilfully ignored FDA instructions and purposely misled FDA reviewers. Fortunately, the hearing which followed (U.S. House 2002) was a full investigational hearing in which congressional staff and expert consultants conducted interviews and obtained internal memoranda, notes from meetings, and other documents which would be impossible to obtain without a congressional subpoena, allowing more insight than usually possible into a drug approval failure. By the time the hearings took place, Sam Waskal was being indicted for securities fraud and therefore exercised his constitutional right to refuse to testify against himself in the hearing. His brother, Dr. Harlan Waskal, the other co-founder, was never implicated in any of Sam’s misdeeds and the congressional investigation did not uncover any instances of apparent falsehood or misdirection on his part (as they had for Sam). Hence, in my opinion Harlan’s testimony (and he did testify) is much more credible than anything his brother might have said. Also, I should note that this investigation was taking place concurrently with investigation of Enron, a company whose name has become synonymous with egregious corporate greed and profound ethical poverty. This was one of a number of cases of corporate fraud in the news and the mood of the country was to spank selfish and greedy CEOs. Sam Waskal was all-too-easily painted with that brush. Hence, this

8.5.1 *The Meaning of 'Promising'*

Disagreement over the meaning and applicability of fast track took place almost immediately, not between the FDA and sponsor, but within the reviewing ranks at FDA itself. In that first exploratory meeting with ImClone in August 2000, the FDA primary reviewer discussed with ImClone representatives the eligibility criteria for fast track. By this time, the fast track guidance document had existed for almost two years (FDA 1998b). The document took some care to specify the conditions under which fast track status would be granted (see pp. 3-7). It specified criteria to define a 'serious' or 'life-threatening' disease; it gave a detailed account of factors to be considered in judging whether an unmet medical need existed; and it gave instructions on assessing a drug's potential. For situations where other therapies exist for a serious condition, an unmet medical need would be met if the investigational drug: 1) had a superior effect on disease outcomes as compared to existing therapies; or 2) had an effect on certain serious outcomes of the disease not treated by alternative therapies; or 3) was able to provide benefits to patients who cannot tolerate or are unresponsive to alternative therapies; or 4) provided benefits similar to those offered by alternative therapies but with substantially reduced toxicity; or 5) provided benefits similar to those offered by alternative therapies but with improvement in features such as patient compliance or convenience of use (5-6).

In this case, alternative therapies did exist, but not many and not very effective ones. From 1959 to 1996, the only chemotherapy regimens available for metastatic colorectal cancer (CRC) were based on 5-fluorouracil (5-FU). In the 1980s, a drug called leucovorin was found to modulate the metabolism of 5-FU and enhance response in some patients, however it was not until 1996 that another single agent was found to have activity against metastatic CRC (U.S. House 2002, 51). This agent was irinotecan (trade name Camptosar, also called CPT-11), and it had been approved initially in 1996 under accelerated approval with confirmation of clinical benefit and fulfilment of the postmarket study commitments in 1998.³⁹ The overall objective response rate for irinotecan was 15

congressional hearing also makes an interesting study in boundary-work (Geiryn 1983), and therefore at times I suspect the criticism against ImClone is not entirely fair.

³⁹ See the label and approval history for Camptosar at http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm?fuseaction=Search.Search_Drug_Name. Note that the row labels in this listing are incorrect: the link for the 1996 decision is called 'approval' while the one for the 1998 decision is called 'accelerated approval'. These are reversed, as becomes apparent by reading the label, letter, and review information provided at the links.

percent,⁴⁰ and for metastatic CRC this figure became the benchmark against which proposed fast track therapies would be measured.

ImClone was claiming to have achieved a response rate from the phase II study of 22.5 percent, so on the surface response rate did not appear to be an issue. However, there was one crucial problem with the ImClone study: it was not a trial of Erbitux as a single agent. Some limited data suggested that the drug was not itself active in metastatic CRC, but rather served to decrease tumour resistance to irinotecan. Therefore, ImClone designed the study such that patients who had stable disease or disease progression on irinotecan alone would receive a combination of Erbitux with irinotecan. This study design left the FDA with the question of whether the responses in the trial were due to Erbitux or the combination. It mattered a great deal, since irinotecan was a highly toxic chemotherapeutic agent having potentially severe adverse effects. If most of the responses were due to the much less toxic Erbitux, why continue to subject patients to irinotecan? No clinical studies had been done to establish the activity of Erbitux in metastatic colorectal tumours. At this time, ImClone could only cite some animal studies and one small renal cancer trial to support the hypothesis about the drug's single-agent activity.

In an internal FDA meeting to prepare for the August 2000 meeting with ImClone, the primary reviewer indicated in her notes that the benchmark of a 15 percent response rate could only be applied for the activity of a single agent. From this reviewer's perspective, since the single-agent activity of Erbitux was unknown, and since it was unknown whether patients needed to continue to be exposed to the highly toxic irinotecan, this particular study did not meet the standards for fast track.⁴¹ None of the other participants disagreed with her assessment (U.S. House 2002, 39). Nevertheless, in the 11 August meeting with ImClone, the most senior FDA medical officer told ImClone that 'the basic trial design is probably acceptable' (39), overruling the judgment of the

⁴⁰ The conversion to full approval was based on confirmatory studies yielding a 15% response rate while the accelerated approval had been based on a response rate of 12.8%. See the clinical studies summary in the 1998 version of the label at <http://www.fda.gov/cder/foi/label/1998/20571s8lbl.pdf> and the accelerated approval summary at <http://www.accessdata.fda.gov/scripts/cder/onctools/summary.cfm?ID=158> (accessed 12 December 2007).

⁴¹ The reviewer's notes read: '1) Is ORR [overall response rate] = 15% clinically meaningful for colorectal CPT-11 failure? Only if as a single agent. 2) CP02-9923 [the ImClone phase II study] meet accel. approval criteria and fast track? *No*' (U.S. House 2002, 39).

primary reviewer.⁴² According to an investigation report, the ‘senior FDA officer told Committee staff that her decision to accept the protocol was based on her belief that she should be flexible for a promising drug meeting an unmet medical need’ (39). She also said that she made the decision assuming that ImClone’s claims about the drug action with irinotecan were accurate (39).⁴³ When testifying at the congressional hearings, the senior reviewer said that ‘an approximately 20 percent response rate in patients with refractory disease was something that should be evaluated further’ (198). When pressed on this point, she added that ‘if 13 percent was sufficient to approve Irinotecan [*sic*], it is hard for me to believe that we should judge a much higher standard for Erbitux’ (198) (on response rates, see note 40 in this chapter). Ironically, after this controversy over whether Erbitux should have been granted fast track status and approved with a response rate of 20 percent, another competing drug (also a monoclonal antibody to treat cancers and a potential market competitor to Erbitux) called Iressa was later approved with a 10 percent response rate (219).

The FDA decision to grant fast track status to Erbitux foundered on the question of whether a preliminary result of roughly 20 percent response was ‘promising’ for this drug combination in this disease setting.⁴⁴ Notably, as detailed in the guidance document had been, it was written on the assumption that a single-agent response would be compared to another single-agent response to make a judgment of therapeutic promise. This act of fast track categorization therefore depended on a comparison for which there was no guidance. ‘Fast track’, as such, did not exist when previous therapies for colorectal cancer were approved, but judgments of ‘promising’ did. Unfortunately, the range of exemplars to choose from was limited: the only comparisons possible were with the

⁴² One highly significant element of the discussions in the congressional hearing is that by the time this meeting had taken place, the protocol had been modified to loosen the eligibility requirements. The FDA senior medical officer was looking at the original protocol when she made this remark. There is some disagreement as to whether FDA had both protocols at this time (ImClone says yes, FDA says no), and as to whether the ImClone representatives realized the medical officer was looking at the wrong protocol and, if so, why they did not correct her. But, in any event, the disagreement between the senior medical officer and the primary reviewer was not related to the difference in protocols, since they were both mistakenly working off of the original protocol.

⁴³ According to the medical officer, ImClone mentioned the single-agent clinical study without specifying that it was a study of renal cancer patients, not metastatic CRC patients. ImClone disputes that the issue of single-agent activity even came up at this meeting, although they agree it was discussed with the FDA in subsequent calls and meetings. Again, however, these disputes are not material to the disagreement between the senior medical officer and primary reviewer on fast track eligibility, since both actors were working with the same information and data set on which to make the decision.

⁴⁴ This estimate of the response rate turned out to be optimistic. When the data was closely scrutinized, the response rate fell to something comparable to that of irinotecan (see U.S. House 2002).

decades-old combination therapies based on 5-FU and with the relatively new standalone agent, irinotecan, which had been rushed to the market with a 12.8 percent response rate. In point of fact, fast track decision-making will always tend to suffer from an impoverishment of exemplars because it only applies to drugs for which existing therapeutic options are limited. Unless a ‘breakthrough’ significantly exceeds all expectations based on current standards of therapy, the definition of ‘promising’ will be problematic.

After reviewing the data provided by ImClone and discovering that the evidence for single-agent activity was lacking, the FDA requested a small additional study of the response rate of Erbitux alone in patients having metastatic CRC (40). ImClone completed the study by 12 October 2001 and this was the last bit of data submitted under the rolling BLA. The study showed a 10.5 percent response rate in 57 patients.⁴⁵ The phase II study was so small that, statistically speaking, it was possible that a patient on Erbitux alone could do just as well as a patient on the combination therapy. There was no way to discern a difference (41).

8.5.2 *Research vs. Therapy and the Meaning of ‘Protocol’*

The expert oncologist acting as an advisor to the congressional committee, Dr. Raymond Weiss, wrote a report detailing a series of flaws with the conduct of the clinical trial, including deviations from the protocol’s patient eligibility criteria, variations in the dose and administration schedule of irinotecan, lack of clarity as to whether patients on study were truly ‘refractory’ (due, in part, to the above-noted protocol change), and more uncertainty regarding which patients responded. These types of deficiencies were a central focus of the hearings, with committee members repeatedly asking how such a thing could have happened. How could the protocol have been well conceived when examples of deviations from it abound? For instance, some patients were enrolled in the trial based the oncologist’s ‘clinical judgment’ rather than on the basis of radiological scanning to demonstrate disease progression (210). Moreover, although the protocol disallowed patients having liver dysfunction, some patients having this condition were accepted into the clinical trials because, according to testimony, the physicians performing screening would ‘use their judgment to decide if an abnormal liver test put the patient at

⁴⁵ The oncology expert’s report to the congressional committee put the figure at a more conservative 8.7 percent (56).

any increased risk' (81). Seventeen patients did not receive the dosing stipulated by the protocol. And even the original interpretation of radiological scans by the participating clinics was in question; detailed re-examination by ImClone's soon-to-be parent company, Bristol Myers Squibb (BMS), resulted in significantly revised conclusions as to which patients responded to treatment and which did not.

The committee members found these deviations baffling, especially considering that some of the most prestigious cancer research centres and most respected names in oncology were involved in the trials. The fundamental problem of the hearing is: if the drug is good and FDA rejected the application, then the study must have been 'bad.' But how can the study have been bad when many of the most respected physicians and clinics in oncology are associated with it? In response to charges of 'bad science' in creating a flawed protocol, then-CEO Harlan Waskal defended both the protocol and the conduct of it, saying that '[w]hat is critical is that this study was not designed as a registration trial. It was a Phase II study early on in the development of this drug. It was only because of the unexpected results that we were able to go ahead and move it forward' (72). Waskal's fundamental argument was that there is a difference between the way a clinical trial protocol is designed and implemented for a Phase II clinical trial and the way it is done in a 'registration study' (an efficacy trial to support an NDA or BLA). The problem with this assertion is that the entire career of legal and regulatory reform over the previous dozen years had been based on the assumption that the knowledge produced by Phase II clinical studies is in principle compatible with the statutory requirements for drug approval. Indeed, as early as the late 1970s we have seen the claim by the FDA that Phase II studies produce the first, statutorily defensible evidence of effectiveness. So if this Phase II study was not 'flawed', as Waskal insisted, then how could it not have been adequate to support efficacy claims? Notably, no one in the hearings challenged Waskal's assertion, although he made it several times. Not only was this assertion not challenged, but the claim was echoed by FDA officials. The FDA's Dr. Keegan testified: 'I would say that we concur with the statements made by Dr. Waskal that that trial was not intended either by ImClone or ourselves to be a registration or a major efficacy trial' (189).

The reader will already know the answer to the question from Chapters 5 and 7, where we saw that Phase II cancer trial 'protocols' were often not much more than flexible rule-guided therapeutic regimens for a certain class of patients, revocable at the first signs of patient failure. Recall the comment of the noted cancer researcher Dr.

Einhorn, who said, ‘quite frankly, most of our Phase II trials are zero for 14 and we move on to the next Phase II drug’ (FDA 1988b, 37). Indeed, Einhorn’s research centre at the University of Indiana used the Phase II trials essentially as a last-chance repository for patients who failed standard therapy. Similarly, we saw in Chapter 7 that physicians broke the eligibility rules for the compassionate ‘trial’ of DOX-SL out of a desire to get patients suffering from Kaposi’s Sarcoma access to the drug. In practice, Phase II cancer clinical trials often combine therapy with research. For this reason, as we saw, data from such trials can be vexing to evaluate for an NDA when the results turn out to be better than expected.

Granted, Einhorn was in an academic institution while ImClone was a for-profit venture. Would not ImClone take a more careful, more systematic approach? One answer is that ImClone was the sponsor, but academic centres and therapeutic centres like the prestigious Sloan-Kettering Cancer Center were the ones administering the protocol. Thus, we should not be surprised that in an exchange over patient enrolment, when Rep. Fletcher asks why physicians would enrol patients having liver dysfunction if they knew that ‘it may possibly prevent this from being approved through the FDA’, Harlan Waskal responded: ‘I don’t think that was part of their consideration. It was their clinical judgment that *these patients were not being put at any type of risk* by enrolling them in the study...’ (81, emphasis added). In another exchange, Waskal defended a decision to change the dosing protocol partway through the study, saying that ‘these patients were treated by their doctors using what is the standard of care’ (89) and that ‘if a patient fails a cycle of treatment, a single cycle of treatment, with tumor enlargement or new tumors, it is unethical to continue to treat them’ with standard chemotherapy (89). From these passages and others in the testimony, it is apparent that prior to its attempted transition into a registration study for the purposes of approval, this Phase II clinical trial was in part used to treat patients, making judgments regarding their care sometimes at the expense of the protocol, while learning something about the drug at the same time. As Harlan Waskal testified, the favourable results were much better than expected. The answer to Rep. Fletcher’s question is that the physicians made all kinds of clinical judgments because they *had no idea* the study would *ever* be submitted as a BLA.

I noted in Chapter 3 the historical tension between therapy and research. According to Marks (1988), early-to-mid-20th century physicians resisted efforts to develop systematized clinical trials, in part because it restricted the clinicians’ decision-

making ability, turning them into mere technicians. As Marks noted, ‘centralizing the management of therapeutic studies did not eliminate the physician’s conflict between loyalty to the needs of the patients and to the objects of research. Someone would still be required to enforce the details of the experimental protocol’ (320). Marks’ observations clearly still hold true as recently as 2001 in oncology. In many cancer clinical trials, the expectation for a highly successful new drug is low enough that there is little incentive to establish elaborate procedures for screening and evaluating patients and to enforce them stringently. For clinical research oncologists, the goal is often practical compassion: try to give hope to otherwise hopeless patients while also learning something about a new investigational agent.

This view of what ‘protocol’ means is clearly divergent from the meaning of ‘protocol’ for the sake of a registration study. In both cases, the term ‘protocol’ refers to rules guiding a set of practices, however in each case the perception of the rules is quite different. Humans constantly make judgments about which rules are ‘bendable’ and which should be taken more literally. In the ImClone hearings, the phrase ‘clinical judgment’ often simply meant a judgment of which protocol rules could be suspended temporarily for the sake of perceived patient benefit. Thus, ultimately the ImClone BLA foundered on the practical difference between rule-following in an early-phase investigational setting for a cancer drug and rule-following in a clinical setting designed to produce data for a BLA. There was a difference.

8.5.3 *Social Considerations in the Judgment of Validity*

ImClone’s original claim was to have studied 127 patients and have obtained a 22.5 percent response. On review and re-examination of the case files of the patients on study, observers judged that there may have been as few as 89 ‘valid’ patients on study and a response rate as low as 13 percent (45, 54). Echoing a warning given to ImClone by its own consultant, the investigative report to the committee cautioned that ‘if indeed the denominator in 9923 [the phase II study] was below 100 (particularly if it were as low as 89, which the BMS independent radiologist appears to have indicated in the above e-mail), the entire study probably could no longer serve to support an accelerated approval application’ (45). Significantly, the ‘warning’ from ImClone’s consultant noted in the report came in January 2002, *after* FDA’s refusal to file, when ImClone was assessing whether the rejected study could still be salvaged (the investigative report quotes this

warning making it seem as though it came prior to the refusal to file). Moreover, it is clear that the FDA could be extraordinarily flexible in their analysis of data when it was deemed appropriate. In the most dramatic example we saw in Chapter 5, ifosfamide was approved on the basis of a single, uncontrolled study with a great deal of data gaps and only 59 patients to start with, whittled down to only a handful (at most 13) of evaluable, eligible patients after the FDA review. Granted, that was a therapy for a very rare form of cancer for which it was difficult to accrue patients to clinical trials. In Chapter 7 we also saw a single uncontrolled clinical study for DOX-SL to treat AIDS-related Kaposi's Sarcoma in which, after eliminating the data for patients who could not be considered refractory to first-line therapy, only 77 of 386 remained for analysis. Moreover, of those 77 patients, approximately 25 did not meet the protocol entry requirements and, to the frustration of ODAC member Dr. Bunn, the sponsor had not analyzed how many of the eligible patients responded to the therapy. Like the ImClone study, this 'trial' for DOX-SL was not originally designed to serve as a pivotal study for a drug application. Rapid accrual and promising therapeutic results thrust it into a role for which it was never intended where, under the intense scrutiny of FDA review, myriad flaws and data gaps were revealed. After detailed review, the FDA's Dr. Murgu identified only six patients who might have benefited from the treatment. And when asked what the denominator was (six out of how many?), Murgu said that he could not answer the question. In effect, there was no meaningful denominator.

It may be that, as in the mid-1980s, the pendulum swing of standards for oncology trials was moving towards greater stringency; i.e., perhaps by 2001, after the breathless career of reform and accelerated approval decisions of the 1990s, the FDA was now in a reactionary phase of tightening. Another explanation can be offered in this case, however. By itself, the denominator was not the issue with the ImClone NDA. By themselves, the protocol violations and data flaws were not the issue. I would suggest that concerns like the study size or protocol violations, while not unimportant, took on greater significance than they otherwise might have for ImClone, in part, for reasons social in nature. This is not to imply that social relations were the predominant factor in this case. Clearly the technical problems with this application were substantial — the largest one of all being ImClone's insistence on testing the drug as a combination with irinotecan and assuming the single-agent activity of the drug in metastatic CRC. This was a fatal error. The FDA had been consistent in its insistence on discerning the individual contributions of each

agent in a combination prior to approving the combination. Nevertheless, significantly, the FDA *had been willing* to consider this drug combination for fast track approval. This fact by itself would tend to undermine the idea that the FDA was in a reactionary phase of rule-tightening. Rather, the FDA clearly sought to be flexible. If ImClone's claims regarding the drug's single-agent activity had been vindicated, the FDA might have been willing to work with some of the other technical flaws rather than rebuking ImClone with an RTF letter. But the single-agent data only served to undermine ImClone's position, both technically and in terms of social credibility.

In retrospect, all of ImClone's dealings with the FDA on the question of single-agent activity must have seemed disingenuous to say the least. Indeed, whether through misunderstandings, perhaps genuine naiveté in how to bring their first BLA through the process, or more nefarious motives, the congressional hearings are rife with examples of ImClone representatives appearing not to act as good faith participants in the FDA/sponsor collaborative process. If Dr. Einhorn's error-riddled data were given the benefit of the doubt because of his impeccable reputation as a cancer researcher, by contrast ImClone's error-riddled data were eventually consigned to the rubbish bin in part because of their seeming dissembling and not-completely-forthright approach to communications with the FDA.

This observation somewhat modifies Abraham's (1995) conclusion that industry tends to be given the benefit of the doubt when it comes to assessments of drug safety (and more generally in the licensing of drugs).⁴⁶ As we have seen, the FDA granted fast track status to Erbitux based on verbal representations made by company representatives regarding the available data; in this sense, certainly, they were given the benefit of the doubt. But 'benefit of the doubt' did not mean wilful blindness to the facts. FDA reviewers checked the data and found that ImClone's verbal representations had been misleading or incorrect. The FDA requested the additional, required data. And when that data was finally submitted and further put the lie to ImClone's verbal representations about the activity of the drug, the FDA reacted quickly and decisively.

⁴⁶ Abraham's article focuses on a case involving the analysis of adverse reaction reports to assess the safety of a drug marketed in the U.K. and U.S. However, his conclusions and policy prescriptions extend well beyond the issues of safety and product recall, noting the dependence of regulators on pharmaceutical company representations (with associated benefit of the doubt conferred) during premarket as well as postmarket.

Indeed, technical considerations aside, part of the assessment of ‘evidence’ can and (probably) typically does include an assessment of the *source* of that evidence. This assessment of the source takes place in at least two senses. Obviously the assessment makes allowances for clinical and regulatory context when the drug is considered potentially important for serious conditions. Does this data come from a major pivotal study designed to be decisive for the sake of a full approval? Or was this a case of ‘therapy-research’, in which a less formal phase II trial or loosely rule-bound congregation of ‘compassionate use’ patients turned out to produce noteworthy results adequate for the purposes of conditional approval, however flawed? Secondly, it also includes an evaluation (conscious or unconscious, as part of human interactions) of personal credibility. Whatever differences of technical detail and circumstance we can identify, one thing seems certain: whether consciously or not, a more stringent judgement of validity would be applied to data coming from Sam Waskal than to similar quality data from Lawrence Einhorn.

8.6 Conclusions

A number of important conclusions arise from this examination of lawmaking. This chapter problematizes the view of the relationship between congressional directive and FDA action. In many ways in this case, lawmaking retrospectively endorsed FDA practice rather than proactively directing it. We also saw that in the Balkanized culture of the FDA, some directives may be embraced more fully than others. Even where the FDA seeks earnestly to comply, the interpretation of a given directive is not a straightforward process of reading off the meaning of the new rule, but involves categorical jockeying to render the new rule sensible amid an array of existing related practices. Additionally, the practical application of such new rules can pose unanticipated challenges. In the case of ImClone, situationally variable meanings in the definitions of ‘protocol’ or ‘promise’ or ‘validity’ were exposed and created difficulties for rule-guided decision-making.

We can also observe that the congressional ‘will’ is rarely uniform. Attitudes towards the FDA and towards regulation more generally were readily identifiable along party lines, suggesting that the motivation for a single legislator to support one position over another has much to do with political ideology. For this reason, we can suppose that to the extent individual legislators are influenced by interest groups, influence tends to be

selective according to political orientation. Whatever brokering and compromise takes place between interest groups, it does not necessarily take place in a direct fashion, but is often mediated through the interests and actions of individual legislators, who negotiate with one another in committee deliberations and other venues.

Notably, although powerful pharmaceutical interests were put in a position of special access and influence after the ascendance of the neoconservatives in 1994, they were unable to achieve the comprehensive re-ordering of the system sought. Public choice theory would have predicted a decisive victory for pharmaceutical interests during the 104th Congress but, just as for the 1962 drug amendments (Chapter 3), the ultimate result was a political compromise. While the example of the 1962 amendments was perhaps unusual because the political compromise was forced to some extent by the unusual intervention of a drug-related crisis (thalidomide), no such excuse can be made for the FDAMA. In the latter case, the neoconservative Republicans were ultimately unable to build broad support for their agenda and Republican reformers discovered that it was more in their interest to cooperate and compromise rather than to adhere to a fruitless unilateralism.

Significantly, pharmaceutical ‘interests’ in determining the shape of the final legislation were also not uniform. We saw from congressional testimony that while some biotechnology firms might approve of, e.g., the elimination of requirements to prove effectiveness, the President of Genzyme opposed this move. We also saw that there could be disagreement between Republicans and their industry sponsors — as in, for example, the disagreement over whether to concede to the imposition of user fees to expedite the device application review process. Consumer interests are likewise divided.⁴⁷ Hence, when we talk of ‘public’ interest groups or ‘industry’ interest groups, we must be careful not to stereotype such interests as homogeneous. Indeed, as I will argue in more detail in the conclusions, the conflicted nature of legislative agendas, citizen support for legislators, and other participants in the various manifestations of principal-agent relationships (Chapter 1) tend to dilute the effectiveness of monitoring. We will see that monitoring is most effective only under certain conditions of consensus, and we will need

⁴⁷ Note, for instance, that the consumer interest group Public Interest opposed Subpart H and the use of surrogate endpoints, even as AIDS activists clamoured for them (Chapter 5). Moreover, AIDS advocacy groups became fractured over surrogate endpoints and issues related to the extent to which they identified with the scientists doing the studies (Epstein 1996; 1997).

to revise the way these concepts are constructed to make them empirically meaningful. These and other conclusions will be the subject of the next chapter.

9. CONCLUSIONS: FDA HISTORY, FINITISM, AND A ‘SOCIAL’ THEORY OF REGULATION

In the chapters up to this point, I have surveyed the history of Food and Drug Administration (FDA) rule-making, decision-making and reform with special emphasis on rules and procedures developed for patients having serious and life-threatening diseases. Along the way, I have sought to show how these events can be understood and analyzed using a number of social theoretic tools, especially meaning finitism, and have set markers for the reader where I believe these observations are meaningful for theories of regulation. In this chapter, I will now attempt to draw together these threads of the story into a theoretically cohesive fabric. First, I will review the lessons learned about FDA decision-making, seeking to build an interpretive framework which both validates the theoretical tools used and also extends their application and interpretation. In particular, I discuss some original concepts derived from these tools in an effort to underscore new insights arising from the relationship of regulation, tacit practice and consensus, and knowledge-making when seen in the lights of meaning finitism and other related perspectives.

I will then move into a discussion of the theory of regulation, showing how the insights of this thesis can both challenge and augment existing theories of regulation. In this discussion, I will not seek to provide a point-by-point rebuttal of existing theories. Instead, the goal is to show how most of these theories ultimately suffer from a theoretical perspective mired in the isolationism of individualistic social and economic theory. In so doing, I can demonstrate the superiority of a collectivist approach to group behaviour and can begin to sketch the outlines of a more empirically valid and vivid theory of regulation.

9.1 FDA Decision-Making and Meaning Finitism

9.1.1 *Rule-Making, Rule-Breaking, and Consensus*

In this account, we saw numerous examples of regulatory and legislative rule-making in which rule-writing formalized existing past practices. Features of the IND and NDA rewrites drew from existing practices, especially treatment IND, FDA-drug sponsor meeting policies, and the definition of clinical phases. The Subpart E rules were based on the experience with AZT (Chapter 4) and the Subpart H rules were based on the experience with ddI (Chapter 6). The Food and Drug Administration Modernization Act (FDAMA) enshrined existing FDA practices to expedite drug approval and expand access to investigational drugs, even as it also codified an informal practice withheld from the Subpart E rules as a basis for redefining ‘substantial evidence’ (Chapter 8). Cancer drug approvals from the late 1980s and early 1990s anticipated actions that would be taken in Subpart H and the FDAMA (Chapter 5). At a time when various political groups and appointed committees were calling for the acceptance of surrogate endpoints in drug approval, the FDA had been approving cancer drugs on the basis of response rates for years. Regulatory ‘exceptions’ were sometimes more like regulatory experiments for new rule-writing. (Recall Dr. Kessler’s response to Dr. Cotton when she expressed concern that the Subpart H procedures had not been given a trial run. He said that the approval decision for ddI was the trial. See Chapter 6.) And new rule-writing routinely incorporated select deviant characteristics of those exceptions to form the basis for the new rules — an important observation about the relationship between rule-‘breaking’ and rule-making.

Clearly, however, we also saw examples where rules were written not to embrace and formalize past practice, but to modify existing practices or to establish new ones. The 1962 drug amendments, 1963 IND rules, and 1970 regulatory definition of ‘substantial evidence’ fit under this latter description. These are examples of how rule-writing is typically conceived: a problem in the operation of a system is detected and a rule is created to correct the problem. A new set of procedures is created *solely* by performative enunciation. In this way, the rule provides a basis for coordinated future practice. Yet, for the period studied in this thesis, it is notable that rule-writing as formalization of existing practices seems to be at least as common as rule-writing to establish entirely new procedures, and arguably much more common.¹

¹ Subpart E, Subpart H, and treatment IND were clearly patterned on existing examples; the IND and NDA rewrites and the Food and Drug Administration Modernization Act (which contained fast track) were somewhat of a blend of existing practices with some revisions to current practice.

There exist other motivations for rule-writing beyond dictating new procedures. What do regulators do when confronted with an immediate problem in need of solution and the rules are maddeningly mute or inapplicable to the situation? In such a case, most responsible persons acting in a goal-oriented fashion would move to solve the problem in whatever way seemed appropriate, often times distorting the boundaries of existing rules in the process. We saw an example of this sort of distortion in the informal practice of treatment IND: pre-market distribution of the drug to needy patients for therapeutic use was conducted under the existing rules for investigational drugs as if the patient would be receiving the drug on an investigational basis (Chapters 3 and 4). Once a rule has been stretched or broken in this manner, does a new rule then need to be written? Not necessarily. A new rule is perceived to be useful only to coordinate *future* practice. Acting once or infrequently in such a manner can easily be dismissed as a ‘one off’ — an exception. But then, what if another situation arises requiring similar, goal-oriented intervention. With the addition of one or more examples of such action, the ‘exception’ begins to become a ‘practice’, and a new category of action begins to be *primed* (Barnes 1982). Once priming begins to happen, our hypothetical problem-solver is confronted with the prospect of more future circumstances requiring this new set of practices. The perceived need for a formalized process grows. The new rule will be written based on the past, exceptional practice; the rule-writer will select from the characteristics which made the action exceptional or deviant in the first place to create a new category of action. The fact that this is a selection process should be highlighted. The Subpart E rules did not contain a provision for single-study NDAs, although that was the basis of the approval for AZT. The Subpart H rules did not contain a provision for FDA collaboration in obtaining and analyzing sources of data other than those in the NDA, although this was done for ddI and ddC. In cases such as these, new rule-making is *constitutionally dependent* on rule-breaking. The former does not take place without the latter. And the latter provides for the former a pattern of exemplars with a menu of characteristics to serve as the basis for creation.

Writing rules under such circumstances tends not to be a matter of controversy since the underlying category of action is already primed. I have argued that in these cases, a *de facto rule* has come to exist through processes of consensus in practice. The perspective expressed here is consistent with assertions made by Keating and Cambrosio (2007) that clinical practice involves an array of informal but rule-like underlying

conventions which guide practice. They argue, for example, that before a clinical trial protocol is even developed, a ‘protocolized’ environment exists ‘in the sense that an archaeology of medical judgments would show that much of common practice flows from deeply embedded protocols whose “rules”, so to speak, have sunk out of sight’ (204). They refer to such conventions as constituting a type of ‘regulatory objectivity’ in biomedical practice (Cambrosio, Keating, et. al. 2006). Keating and Cambrosio draw this perspective from Canguilhem (1989), who concluded that norms follow normalization, which leads to the establishment of conventions and thus to the practices and methods which sustain those conventions (Cambrosio, Keating, et. al 2006, 192-3). While this perspective is clearly consonant with this discussion of rules, my focus is on how certain conventional or consensual practices emerge into the formal world of regulation — i.e., how some make the transition from endogenous conventions to seemingly exogenous rules. In this way I move beyond the world of tacit conventions to show that even in the highly formalized and supposedly rigidly rule-guided arena of the law, practices pre-exist rulemaking.

In cases where new rule-writing is derived substantially from existing practices, any formal consensus process such as notice-and-comment rule-making represents a secondary form of consensus. We can see the truth of this statement in, for example, the Subpart H rules. By the time the FDA promulgated Subpart H, regulators had in fact already worked with drug sponsors and clinical researchers in developing clinical trials which used surrogate endpoints; Kessler had already negotiated the terms of accelerated approval with politically powerful actors; and FDA had already allowed surrogate endpoints as the basis for drug approval with blessing from an advisory committee staffed with influential leaders of the medical community, many of whom act as consultants to pharmaceutical companies. Thus, many of the most authoritative community members had already reached a consensus on the practice of using surrogate endpoints for drug approval before any formal rules were published. This was why Kessler felt free to instruct the ADAC to consider ddC under Subpart H even though the proposal had just been published (Chapter 6): the primary consensus had already been achieved. We can likewise suppose that the FDA felt free to publish Subpart E as an interim rule (thus bypassing the proposal phase of rule-making) because it already had a strong mandate for action. Under these circumstances, those who submit their comments to FDA in response to proposed rules, regardless of their personal status or authority, are

participating in a secondary consensus process. In the period under study, although modifications were made to proposed rules in response to comments, what one mostly sees in FDA's response to comments is justification and clarification: the rule in question had already been implemented in practice, and the primary goal was now to police the boundaries of the new category to assure that it could function effectively as a guide to future action.

In cases such as this one, notice-and-comment rulemaking is therefore a process of fine-tuning or clarification, rather than a negotiation over whether a rule is needed at all and if so, what shape it should take. It is an exercise in articulation-practice (Chapter 4), in which potential scenarios of applicability are verbally rehearsed, with changes in meaning of the underlying concepts a possible result. However, in cases like the 1962 drug amendments and their aftermath, where a set of new procedures was established without a practice-based consensus, controversy was more likely to ensue.

In calling these practices 'new', I should underscore that the implication is not that these practices did not exist prior to the 1962 law. We have seen that the key components of the randomized controlled trial (RCT) were first used together in the 1930s and had been rising in ascendance through the 1950s (Chapter 3). Hence, drug sponsors were not being required to develop completely novel practices, but rather to implement a technology of clinical drug development which was available to them but had not previously been in widespread use. This is the sense in which I say that a practice-based consensus had not coalesced prior to new rule-writing. As a result, conflict ensued due to the lack of consensus between the FDA and the drug-makers on the type of evidence needed to prove the effectiveness of drugs (indeed, it seems clear that drug-makers were resentful that any new forms of evidence would be required at all).

From these observations we can say that although a performative enunciation created the new category of action and mandated its use, the actual practical definition of how to proceed was very much in question and open to debate. The performative enunciation was clearly a *contingent* one in which the interpretation of the rule would be challenged repeatedly. We can therefore conclude that in this regulatory context, and probably more generally, performative enunciation accomplishes a certain sort of priming, but by itself does not offer the same quality of priming evident when practice precedes the enunciation. The difference is in the degree to which existing exemplars exist, and in the degree to which a consensus has been achieved on the question of which, if any,

exemplars are appropriate for use in concept application. Performative enunciation can mandate the use of a new category of action, but it cannot mandate consensus, nor can it create exemplars from nothing. In the last instance, perhaps we can say that all performative enunciation is contingent, with the degree of contingency varying contextually across a spectrum from ‘negligible’ to ‘contested’.

Notably, however, whether new rule-writing is a result solely of performative enunciation or whether it is a result of practice-based consensus and de facto rules, new rule-writing never *ends* the question of the practice to follow, but rather *begins* it. Each new instance of rule application will be, in effect, new — a matter of negotiation. In the 1960s, with consensus lacking over broad questions of application, years of conflict and court challenges would follow. By contrast in the 1980s and 1990s, consensus on the new rules was virtually a *fait accompli*, and the question of how to follow the rule was much more a matter of technical questions than a controversy over the rule itself. Nevertheless for both types of rules, we witnessed case-by-case working out of the application of the rules, as would be expected by finitism. It was here that, in Barnes’ (1982) terms, ‘proper usage’ of communally accepted concepts was developed ‘in processes involving successions of on-the-spot judgments’ (30). It was here that each use of a concept was ‘in the last analysis. . . accounted for separately, by reference to specific, local, contingent determinants’ (30).

As we saw, subsequent applications of Subpart E and of the AZT experience varied widely. Indeed, the experience with AZT was used to support various, and sometimes contradictory, interpretations of the way to go forward (Chapter 4). We also saw that Subpart E itself, having an apparent dearth of actual applications to drug approval (probably due to the advance planning aspects of the rule), ended up being interpreted and reinterpreted, first in its application to ddI (Chapter 6), then in the aftermath of the FDAMA, as the FDA sought to reconcile existing categories of drug approval with the new law (Chapter 8). In each case, the definition of Subpart E changed to suit the contingent circumstances of the moment. We likewise saw that the first applications of Subpart H to cancer deviated significantly from the regulatory model of accelerated approval based on AIDS (Chapter 7) and that lack of internal FDA consensus over the proper application of the concept ‘therapeutically promising enough to qualify for fast track’ contributed to a failed fast track drug application (Chapter 8).

In Chapter 1, based on meaning finitism, I predicted that when a new category of rules is created, we would likely see a cascading dithering effect through related categories — an instability of salient characteristics defining related concepts. This type of dithering was apparent, for example, in the shifts of meaning for Subpart E, ‘fast track’, and Subpart H relative to one another after the passage of the FDAMA (Chapter 8). It could also be seen in the revision of the concepts of ‘risk’ and ‘benefit’ according to changes in regulatory practice. Decision-making on the basis of less information than desired with subsequent confirmation led to a new articulation of the risk-benefit assessment. Now ‘risk’ included not only toxicity of the drug for the patient, but the risk associated with *information deficits* at the time of approval. This new definition of risk, first introduced in Subpart E, was transmitted into Subpart H and the FDAMA, and with each step was given a potentially heavier information deficit to bear. Likewise, whereas once the idea of the ‘benefit’ of a drug referred to the clinical benefit — meaning the ability of the drug either to extend survival or to alleviate the clinical manifestations of a disease — with the increased use of surrogate markers and the establishment of accelerated approval on the basis of those markers, the concept of benefit began to change as well. Now ‘benefit’ did not necessarily mean a positive clinical outcome, but could mean a *theoretical possibility* of such outcomes based on medical reasoning from various laboratory measurements. This half of the risk-benefit assessment became a projection or prediction of benefit, a probabilistic estimate, rather than a view of the thing itself. In this way, expedited approval involved a judgement of a double probability based on an incomplete view of the data — like guessing the final shape of a jigsaw puzzle where some of the pieces are missing and others have shape but no colour. This judgment clearly varied according to individual assessment based on each expert’s personal experience, training, and professional orientation and commitments; in the context of FDA or advisory committee deliberations, it was a matter of negotiation. The standards for drug approval were changing on a case-by-case basis; the advisory committee proceedings we have studied could be seen as microclimates of standards-making and consensus-making on the ‘promise’ of any given drug.

We also witnessed a fluidity in the already porous boundaries between the definitions of clinical phases. Part of this fluidity was related to technical changes or improvements in the investigatory process itself. However, part of it can be traced to changing regulatory practices. Indeed, although application of rules and concepts was

ultimately a case-by-case affair, the rules and practices studied in this thesis changed over time along a retrospectively definable trajectory. We can identify an increasing tension between data-gathering and early decision-making, with movement of key decision points to progressively earlier moments in the drug development process with a correlative reduction in the burden of evidence required for each decision. Such shifting was evident in decisions for when to hold FDA/sponsor meetings, when to allow expanded access, when to discontinue trials, and when to approve drugs. Increasing amounts of information were correlatively shifted to postmarket study. These shifts necessarily involved a related slippage in the meaning of scientific concepts like ‘validity’, ‘replicability’, and ‘proof’ in practical application. And these shifts to earlier decision-making also necessarily involved related shifts in the definitions of the underlying clinical phases, with Phase II (and in some ways even Phase I) taking on a progressively increased burden for producing evidence of efficacy. All of these phenomena, along with the redefinition of risk and benefit, are traceable to the contingent, changing balance between information-gathering and early decision-making.

I have argued that these contingent shifts in standards-making do not necessarily constitute ‘bias’ in Abraham’s (1995) sense of deviating from communally accepted standards specifically to favour pharmaceutical interests. In the first place, as I argued in the conclusions to Chapters 5 and 7, if one’s theoretical apparatus is designed only to detect bias, then bias is the only social influence one will be able to detect when standards fluctuate. However, a fully engaged sociology of scientific knowledge must recognize the full range of interests motivating decision-making, including professional, technical, and social commitments (Barnes 1977). More than that, although Abraham implicitly uses a relativist standard of assessing ‘bias’ by evaluating it in light of accepted standards of good practice (Chapter 1), he nonetheless tends to treat these as rigid and absolute, even though communal standards and judgments can change over time — indeed, in clinical drug development and therapeutic studies, certain issues of scientific practice and standards-making can vary according to discipline or research centre, with agreement being far from uniform (as in, for example, the implementation of the clinical phases). Granted, the specific examples Abraham chooses to study often represent the most egregious examples one can imagine of ‘deviation’, in which, for example, dead laboratory animals are negotiated into being considered un-dead. Such examples would fit the re-conception of ‘bias’ I proposed at the end of Chapter 5. So my argument is not with his empirical

research, which does often demonstrate instances of bias in the narrow sense. My critique is only directed at the theoretical interpretive frame he has defined since it would tend to attribute bias in cases where, as in this thesis, a fragile, sometimes non-uniform, changing medical consensus underlies the decisions being made. Ultimately, the question of how ‘non-uniform’ concurrence must be to be disqualified as a ‘consensus’ is itself a matter of judgment for which full concurrence would be elusive.

There can be no doubt that for the regulatory context studied in this thesis, Barsian finitism and Bloorian rule scepticism (Chapter 1) fit the description of what people actually do with rules. All of the rules studied in this account have the character of being socially sustained through processes of consensus. Even for those rules initially mandated solely by (contingent) performative enunciation, where a higher level of initial conflict might be evident, the authority of the Congress to make performative enunciation in the first place was a matter of general consensus, and this was enough to begin the priming process for a specific rule. However, even though Congress has the authority to dictate the rules, the meaning of the rule is not fixed in the language of the rule itself, but must be worked out in practice. The reader might wish to counter that Congress intentionally leaves rules open-ended so that regulatory agencies might specify the technicalities. In this way, the reader might argue that the meaning of the rule was therefore intentionally not fixed by Congress, and rule scepticism is therefore irrelevant. My response would be that while Congress often does leave the details of implementation to the regulatory agency, we have seen that even after the FDA subsequently provided technical definitions of the rules (i.e., the 1963 IND rules, the 1970 rules for substantial evidence, etc.), the meanings of the rules were not fixed by the rules themselves, but determined in daily practice.

What I have developed here is what Bloor (1997) would call a ‘collectivist’ view of rules. He contrasted this type of account to a traditional, individualistic account in which the compulsion to follow the rule comes from within the rule itself, and each individual responds to that compulsion individually. Bloor termed this view of rule-following as *meaning determinism* (Bloor 3). Proponents of the latter position claim that normativity — the judgment of correct or incorrect usage of the rule — is fixed by the rule itself. Once the meaning of the rule is fixed, following it is unproblematic.

This view of rule-following is contradicted at every turn in the historical account provided here. More than 40 years after the passage of the 1962 drug amendments, the

specific standards used for individual cases of drug approval are contingent and situationally specific. Moreover, the view of normativity provided by the individualistic account seems impossible. Even if we permit some social processes to occur during the ‘fixing’ stage of rule use (say, learning the rule from someone else), these resources vanish once the rule has been learned. At this point, we are left with nothing but the rule and the putative user of that rule to establish the ‘correct’ application. In that moment, we must assume that the rule itself is immutable, regardless of the circumstances of the application or the motives of the putative user. We must also assume that its meaning is perfectly accessible to the user: for the user to make a judgment of the correctness of use, her knowledge of the rule must be complete, absolute, and unchanging. Only under these circumstances can she perceive the rule’s judgment of the actions she takes under its authority. If such conditions were truly in force, no one would *ever* apply a rule incorrectly, except due to wilful disobedience or woeful incompetence. The rule takes on a deity-like perfection while the putative user is granted (by the rule itself) omniscience in all things related to the rule. This version of normativity could only be plausible if there were an absolute source and measure of knowledge to which we had perfect access for each rule we use. This is not merely meaning determinism, but *meaning absolutism*. The idea that the rule itself provides the sole basis for normative judgments is a clever but ultimately futile attempt to adjust and defend a fundamentally flawed individualistic theoretical position.

From this study of drug regulation, it could be suggested that changes in the meaning and application of rules is not done wilfully or arbitrarily, but because people working in a goal oriented fashion encounter a situation in which the rule was not adequate for the circumstances. In the realm of mathematics, it would be difficult to imagine a situation in which a result of ‘4’ was seen to be inadequate or inappropriate for the rule ‘2+2’ (difficult but not impossible. See Chapter 1, note 22). However, situations do arise in which the existing rules and practices in mathematics are not adequate for the desired application, and in these cases the most useful solution is not to break existing rules, but to create new ones. For instance, if algebraic techniques were found to be inadequate and cumbersome for precise calculation of areas under curves or volumes of solid bodies, the necessary solution to the problem would not involve ‘breaking’ the rules of algebra, which would accomplish nothing, but creating a new approach which would be better designed for such applications, such as integral calculus. In this way, perhaps

mathematical rules can be seen as having a character more like the 1962 drug amendments, where ‘bending’ existing rules simply would not accomplish the purpose of requiring substantial evidence of effectiveness prior to permitting a new drug on the market. A new set of procedures needed to be established for that purpose.

Mathematics aside, I have made the case for regulation, a venue typically seen as rigidly rule-guided. We can now examine how rule-making and rule-following are related to knowledge-making in drug development.

9.1.2 *Knowledge Praxis*

In this process of creating new rules, we can see an intellectual pattern implicated generally in the creation of new knowledge: traditional practices and past experiences were employed as the basis for an action, decision, or new rule; then that action, along with its immediate outcomes, was added to the stock of experience (consisting of selectable conceptual elements, individually or in patterned assemblages) with which the next set of decisions was made or actions taken. We could refer in general to this process as ‘knowledge praxis’ — the cyclical, cumulative experience and selective application of knowledge. Importantly, this is not a simple additive process. The perceptions of those original experiences or procedures became *modified* through application to the new situation and in combination with the resulting new experiences. Therefore, we can say that this process is fundamentally *transformative*. In Chapter 4, I used the example of AZT to make this argument, noting that AZT began as a meaningless physical substance (or more accurately, a substance having the meaning ‘failed cancer drug’), and was then transformed through resignification, i.e., it was re-identified as a research object, characterized with respect to its activity against AIDS, and assigned values in terms of ‘safety’ and ‘efficacy’ with correlative associations such as ‘promising’ or ‘successful’. In Barnsian language, we could call it a selective reassessment and resorting of salient characteristics associated with the term ‘AZT’ — a resorting which takes place in the course of practice and for which there is a relatively high degree of consensus among those doing the practice. Resignification can be seen as a fundamental operation of knowledge praxis. AZT then became intelligible as a conceptually and culturally significant object and, as such, one which could be appropriated as a resource implicated in the development of certain attitudes, decisions, and actions, both scientifically and

politically. I have referred to this latter process as transformative resource creation — a frequent, but not inevitable, result of resignification.

At the end of Chapter 4 I argued that transformative resource creation differed from concepts like that of ‘translation’ in actor-network theory, (Callon 1986; Latour 1987; 2005), and also Pickering’s (1999) ‘mangle of practice’ because the idea of ‘transformative resource creation’ emphasizes the social-consensual derivation of knowledge in a way that the ‘mangle’ does not and ‘translation’ cannot. Although meanings are indeed generated in the ‘actor-network’, not prior to it, only humans can create and sustain meaning. And while physical objects impose themselves on our being in ways not of our own choosing, such as the emergence and devastation of the AIDS virus itself, outside of human social context even the AIDS virus has no *meaning*. What is considered to be knowledge of the AIDS virus is derived and agreed upon through social processes (Kusch 2002). Transformation of such objects into important cognitive and social resources for humankind can therefore only be accomplished by humans in a collective knowledge-making enterprise.

What do these concepts of knowledge-praxis have to do with the regulation of therapeutic drugs? In the previous section I discussed a movement of scientific concepts in tandem with regulatory ones, in which attempts to regulate earlier decision-making in drug development led to changes in the subsequent use of categories like ‘risk’, ‘benefit’, ‘replication’, ‘validity’, etc. This movement points to an imbrication of science and science-based regulation in which each has the opportunity to influence the other as both are applied, intertwined, to a given situation. In this way, new concepts and sometimes even new knowledge arises. For example, the requirement to conduct randomized controlled studies for clinical drug evaluation led to the articulation of new techniques and changing conceptions of clinical phases, which then fed back into the regulations.²

² The temptation here is to argue that the regulation of scientific endeavours is distinct from the regulation of activities which do not have at their core a physical underlying referent. The regulation of, say, commissions structures of mutual funds, while certainly having very real economic effects, ultimately does not refer to any object which could ‘push back’ in the manner that Pickering’s (1999) subatomic particles or Callon’s (1986) scallops or my example of AZT actually do in their interactions with humans because of the physical qualities they bear. Or does it? Once the machinery for using and responding to commissions structures or interest rates is established, is there not a certain solidity to the underlying referents, even if they are ultimately S-terms, not N-terms? It is not crucial to the main argument of this thesis, however I suspect that whatever differences exist between scientific and non-scientific regulation probably amount to the differences between S-terms and N-terms. So long as group of people choose to treat S-terms as ‘real’, then they effectively are.

We have seen that in each new situation encountered, no decision is inevitable; there are multiple directions that could be taken depending on the elements one chooses to emphasize within the stock of experience. The same principles apply to articulation-practice. In this verbal rehearsal of concept application the meaning of the concept can change with the application. The alarmed reaction of the AIDS Action Council to the FDA's subtle shift of meaning of the treatment IND rule when applied to Subpart E is a clear indication that such semantic reconceptualization has practical consequences, especially in rule-making. These rehearsals in concept application can themselves become a type of exemplar forming the basis for future action – indeed, in rule-making they are held out as exemplars as a model for future action. In this sense, articulation-practice is a form of modeling — often an intentionally performative one. But the performativity need not be intentional. Just as, say, financial models can performatively shape the behaviour they seek to model (MacKenzie 2006), these exercises in concept application can tend to shape future practice, even when they do not explicitly appear in new rules but are only rehearsed in, say, public speaking engagements,³ journalistic interviews, or draft guidance documents. Articulation-practice is therefore a chief means by which resignification can take place in a regulatory context.

Since concept application is open-ended, so also is knowledge praxis. Moreover, the choice of which aspects to emphasize within the stock of experience is one key way that interests enter the decision-making process. Often times, talk of 'interests' guiding choices is misconstrued to mean partisan political manipulation or petty selfish interests. While this account clearly includes some of those types of incentives, failing to look beyond such factors would represent a sadly impoverished picture of human motivation and judgement. Very often, as noted above, the interests of concern are narrowly technical goal oriented concerns, professional interests, institutional priorities or policies, or more broadly political interests (Barnes 1977). At the end of Chapter 5 I argued that many such interests are apparent in advisory committee decision-making for drug approval. More than that, we must think of interests not (necessarily) as a rigid agenda in

³ And, notably, a recent psychological study has shown what we might have intuitively guessed: when speakers address audiences, they tend to shape the facts of their presentation to comport with what is known of the audience's predominating pre-existing beliefs. What's more, the greater the speaker's desire to connect with the audience in this manner (i.e., a genuine attempt at connection, as opposed to a cynical reworking of the facts to score points), the greater the likelihood that the speaker will subsequently come to believe the version of the facts presented to the audience. See Echterhoff, Higgins, et. al. (2008). This is clearly a form of resignification.

operation, but as principled priorities or commitments deployed to guide practical decision-making, which can be re-prioritized or re-defined to meet the contingencies of daily practice. For example, as we have seen, the working definition of ‘treatment IND’ was repeatedly remade in the period under study (Chapters 4 and 6). On a formal level, part of that remaking did involve conventional (and sometimes petty) politics, with a struggle over who should have the authority to decide when an unapproved drug should be distributed to patients (U.S. House 1987). But we have also seen that on-the-fly modifications to treatment IND as a living, working tool proceeded very much as a response to immediate situations in need of solutions. In these situations, the technical question ‘how much data is enough to justify early access?’ fades in and out of prominence depending on the moment, and is often supplanted with a seesawing competition for priority between the physician’s commitment to help patients and the researcher’s commitment to collect valid data so that the physician might help patients more effectively. The relative restructuring of these priorities in any given moment must influence which elements from the stock of experience seem most valid for the situation, and actors choose accordingly.

As already noted, rule-making in this story decidedly came in the aftermath of practice, as a *reflection* of it; in this account, there is no way that rules can be seen as directing behaviour unless one appeals to vague notions regarding the ‘spirit’ of the law. Of course the FDA *does* seek to follow the intention of the law in a goal oriented fashion, and in fact, Congress has given the Agency broad latitude to do so. However, the many court cases challenging the FDA’s view of the intention of the law stand as testimony not only to the contingency of the FDA’s authority, but also to the various and multiple interpretations arising from the same statutes. Ultimately, concept application in general, and in particular regulatory concept application, can be considered a trial and error process. The ‘trial’ takes place when the observer is confronted with a new situation, makes judgements of similarity and difference to other known cases from a stock of experience, and thereby makes a decision. The judgement of truth or error becomes a retrospective task of the relevant community of observers. Any difference between the general case and the specifically regulatory context likely rests in the fact that, while all putative appliers of concepts and followers of rules seek to create retrospective justifications for actions taken (Mulkay 1980; 1991), regulators have an unusually good tool for the job: the act of rule-writing itself can be seen as justification for actions already

taken, as the ultimate tool for legitimating prior decisions. Indeed, this was exactly how the FDA viewed the Congress' 'legislative support' for existing practices in the FDAMA (Chapter 8). Moreover, this legitimation often follows an ironic two-step manoeuvre in which 'exceptional' previous decisions are cited as justification for the new rule, even as new rule-writing creates a sense of legitimacy for previous decisions. Thus, as we have seen, AZT was one of the 'exceptions' cited as precedent in the Subpart E rule, even as the Subpart E rule codified the approach taken with AZT. Similarly, the Working Group's recommendations for accelerated approval cited dDI as an example of previous practice to legitimate the recommendation, even as Commissioner Kessler claimed that the Working Group's recommendations facilitated the approval of dDI. In this way, we begin to understand the concept of 'precedent' not as a one-way vector pointing to a new practice, but as more circular in nature: precedent is ultimately self-referential and self-justifying.

Of course, there are many reasons for writing new rules, not the least of which is that rules serve as an organizing principle for 'correct' action in a community. If the FDA did not bother to codify frequently occurring informal practices, it would contribute substantially to confusion for drug-makers working to get drugs to market, physicians seeking pre-market access to drugs for their very ill patients, and others. Hence, the above discussion should emphatically not be construed as saying that the FDA writes rules frivolously. Another way of approaching the same argument would be to assert that a chief difference in the application of knowledge between the regulatory context and an individual context is that the regulatory body carries *authority* in the community which the individual user of a concept usually lacks — authority to make performative enunciations, albeit contingent ones.

This contingency is one reason why justification of past decisions is especially important for regulators, and it is where the concept of precedent is most usefully employed. Indeed, it is reasonable to say that the concept of precedent, like the performative enunciation, is a useful tool for priming new categories. Here it is important to clarify our working definition of 'precedent'. Regardless of what our dictionaries might say, we must recognize that in sociological terms 'precedent' does not refer in any simple or direct way to a past action or decision. Rather, the concept of 'precedent' is a thoroughly social construction referring to the *perception* of a given action or decision as a basis for future action and a *collective agreement* that this is so. This perception is an

after-the-fact construction of meaning for a past action. Consequently, just as not all declarative statements are performative but must be made under certain conditions to have force, so not all 'exceptional' actions are seen as precedent. For an action or decision to become precedent, one of two conditions must pertain. A person having authority to create precedent (such as a judge) must take action citing the original deed as a legitimate justification (in effect, precedent-making by performative enunciation). Alternatively, we could envision a more gradual process whereby, over time, a general consensus develops regarding the original, exceptional, action. (Here we might think of gradually changing moral standards, where practices once seen as scandalous eventually become widespread and generally accepted.) In the regulatory and legal contexts, of course, precedent is typically formally instituted and will tend to proceed by performative enunciation. This quality, along with the aforementioned self-referential quality of precedent, leads to the realization that we can understand precedent simply as a social kind term in Barnes' (1983) sense (Chapter 1). The declaration of a certain bit of metal to be 'money', for example, is functionally identical to a declaration that a past action can stand as precedent and similar conditions apply. In both cases, the perception of the referent (the bit of metal or the past action) is changed by the performative enunciation (the metal becomes 'money'; the action becomes 'precedent'), and this altered perception becomes an accepted basis for future action.

If the discussion of precedent to this point is valid, then we can also make two additional general statements which are applicable for all social kind terms. First, in precedent-making, the specific action taken is underdetermined by the underlying exemplar. Just as in the example of the creation of Subpart E based on the experience with AZT, a variable range of characteristics from the underlying exemplar might be chosen to form the basis for the newly perceived precedent. Second, application of the precedent is open-ended and revisable. For instance, the apparent precedent set by the application of 'treatment IND' to AZT was modified in practice when applied to ddI. Significantly, we can now see that the establishment of a precedent is a special case of resignification: key elements of an event are extracted and redefined as a model to follow; following the model then requires case-by-case concept application as described above. In the case of a legal precedent, however, the resignification becomes a formal declaration.

We are now in a position to propose a revised vision of the response to AIDS, and of regulatory processes more generally. The concepts and practices used to tackle the new

challenge of AIDS in the 1980s had their roots in previous decades. In application to AIDS drugs, in the ensuing knowledge praxis, elements of the existing repertoire of knowledge were applied and modified selectively on a case-by-case basis. In the process were modified concepts like ‘risk-benefit’, treatment IND, indeed even ‘substantial evidence’ itself, the legally mandated definition of evidence fundamentally underlying all drug approval decisions. This is ultimately where regulatory innovation took place — in the trenches, so to speak, where regulators ‘stretched’ the rules or made ‘exceptions’ to achieve a goal, in the process extrapolating the application of certain concepts along identifiable trajectories. ‘Precedents’ perceived as successful or useful for future application became the basis for new rule-writing, even as the new rules served to justify the previous actions. Nevertheless, characteristic of conceptual resources transformatively created through knowledge praxis, the perceived blueprint(s) for future action did not constitute an integrated whole. Rather, various elements of the experience were available for appropriation on a selective basis, allowing for multiple and sometimes contradictory views of what the blueprint should look like. Ultimately, the retroactive justification for new rules was made possible in no small way because these ‘new’ procedures were already established through the actions of authoritative members of the community. All that remained was selective formalization of the practices through performative enunciation.

Considering how often the FDA is criticized as a lumbering bureaucracy, their agility in this period is significant. Throughout this time, the existing rules were applied in a way that was remarkably flexible and adaptive. The reader may wish to believe that this flexibility was only a function of the extreme political context of the time. Certainly, the influence of politics cannot be dismissed. Consider, however, that one of the reasons for comprehensive regulatory revision cited in the 1979 *Concept Document* was a mismatch between rule and practice: ‘While the regulations have remained essentially constant . . . the new drug approval process has continued to evolve and become more complex’ (FDA 1979, 4). Politics or no, the challenge of daily practice is persistent. If contemporary society suffers from a proliferation of rules and ‘juridification’,⁴ it is because even if rules are necessary for organizing correct action, they are ultimately inadequate to do so, forever lagging behind the continuous, moment-to-moment creative activity of human agency in knowledge praxis.

⁴ A phenomenon whereby organizations or individuals become subject to regulation to such a degree that they become overwhelmed. For an overview and alternative perspective, see Haines and Sutton (2003).

9.2 Toward A Finitist-Collectivist Theory of Regulation

In Chapter 1 I noted that according to Croley (1998), government action is plagued by the costs associated with collective action and with the principal-agent slack which inevitably flows from the delegation of authority. He argued that the ‘textbook’ view (3) of regulation as mitigating ‘market failures’ was invalid, since the ‘mere presence of market failures and the scarcity of congressional resources hardly justify agency authority. An important question remains whether agencies can fill statutory gaps at a cost worth the benefits, a contentious issue’ (4). He therefore argued that, in sum, ‘any serious theory must somehow consider that collective goods, including monitoring, will in the absence of some catalyst tend to be underproduced’ (24-5).

This formulation suffers in a number of ways. First, the phrase ‘market failure’ obscures the true picture of the cost-benefit trade off in regulation because it implies that when the market is functioning normally, consumers can and should be protected. In fact the basic purpose of the market is to exchange goods and services for a profit — a goal typically incompatible with consumer protection. What is really at stake is the *inability* of the market to protect consumers, not a failure to do so. The usual counterargument is that a company hurting consumers through its products or services would not long stay in business (recall Congressman Barton’s statement of this formula when advocating for the elimination of drug efficacy requirements in Section 8.3.3 of this thesis). While perhaps generally true, it is an extremely weak argument. The idea is to prevent unreasonable exploitation or safety hazards to consumers *before* they happen, not to let it happen and correct it afterwards. Consider the current (as of 2007) subprime mortgage crisis in the U.S.⁵ Unsound mortgage lending practices have resulted in an unprecedented level of defaults on loans, with cascading effects through the financial industry and U.S. economy, with global repercussions. While the mortgage and banking industries have been economically punished consistent with the idea of the market correcting itself, was it worth the devastation to the economy and so many people’s lives to let the market

⁵ The media is replete with reports of the current crisis. See the account of one county official from Ohio (Rokakis 2007), where some of the most egregious practices took place as well as some of the most appalling human consequences. See also the perspective of Elliot Spitzer (2008) former New York attorney general and now governor of New York, who claims that the Bush administration blocked attempts to rein in predatory mortgage lenders. See also Ehrenreich (2008), who puts the mortgage crisis into a wider perspective, arguing that while wealthier people prospered during the ‘boom years’, lower income Americans slid further into the economic shadows.

operate on its own logic? Would not it have been better to prevent the predatory lending practices to begin with? (See also note 29 in Chapter 1.)

Another objection comes in the very idea of conducting a fair cost-benefit analysis of regulation. In Chapter 3 I noted President Reagan's requirement to calculate a 'net cost to society' of each proposed regulation and I argued that such a calculation was necessarily based on an incommensurable combination of heterogeneous elements, some quantifiable and others not, leaving a great deal of interpretive latitude. I would elaborate here that in any cost-benefit evaluation of regulation, neither the true costs nor the true benefits are fully definable. Future woes do not arrange themselves according to a schedule discernable from past woes, so cost-benefit assessments based on past events have extremely limited utility. Our history of errors and successes in drug approval says little about the safety of future drugs, which will involve entirely new classes of compounds, new forms of production, and accordingly revised development and approval practices. Such future knowledge is inaccessible.⁶ Likewise, a cost-benefit analysis based on past events cannot say what costs *might* have been incurred had there not been certain laws and regulations in force. Would there have been a disastrous roller coaster accident, factory explosion, or outbreak of maladies from environmental chemical exposures? Alternate timelines are the stuff of science fiction, not policy analysis. This is not to say that certain reasonable assumptions cannot sometimes be made for the sake of an analysis. However, these types of regulatory cost-benefit analyses often admit such enormous latitude of assumptions, they can be readily tailored to fit whichever conclusions are desired.

A third objection more specifically relevant to the material in this thesis is that, as I will argue more fully below, contemporary views of regulation tend both to overstate and to misconstrue the nature of principal-agent slack. Ultimately, this is so because these theorists tend to fall into Olsonian assumptions about the nature of collective action, completely missing the importance of between-actor consensus and communication in the

⁶ Also, as the example of subprime mortgages shows, even when some people see the danger, where profit is involved it is difficult to make policy changes based on the *possibility* of future consequences. Proponents of global warming have faced the same resistance in the U.S. Unfortunately, in the social networks which develop around profit-making activities, greed and denial of future consequences can become characteristics which, rather than being sanctioned, are encouraged and fostered as if they were public goods. Such networks cannot be readily dissolved since the economic system depends on them. The only other solution is to modify their behaviour through a form of social coercion. As I argued in Chapter 1 for the abuser of water restrictions during drought, members of greed-fostering networks must be integrated into a larger network in which their activities can be monitored and socially sanctioned. In the case of the water user, public exposure was sufficient to modify behaviour. In the case of entire networks centred around profit-making activities, the only answer continues to be regulatory oversight with public exposure of wrongdoing, along with whatever penalties may apply.

overlapping social networks in which humans live and work. I have argued in earlier sections of this chapter that such consensus processes are essential for the operation of knowledge praxis and even for processes like ‘legal precedent’, which are created by performative enunciation but can only be sustained through collective belief and consensus. The same community which sustains such precedents can just as readily overturn them through the same type of goal-oriented adaptive practices which led to the creation of the rules for treatment IND or Subpart E. For these reasons, I will argue that any theory of regulation which does not accommodate tacit processes of consensus and daily practice is impoverished.

9.2.1 *Centrality of Regulatory Decision-Making and Tacit Practice to Regulatory and Legislative Outcomes*

Let us return to the theories of regulation discussed in Chapter 1. As we saw, public choice theory is an approach based on treating regulatory outcomes as an analogue to a market exchange for regulatory goods. In this model, interest group efforts to obtain regulatory goods were focussed on the legislators; the citizen-legislator relationship was given primacy over other relationships in the regulatory ‘economy’. So also for neopluralist theory, an updated pluralistic model of interest group competition designed to challenge public choice theory, and which tended to assume that regulatory behaviour would be a straightforward response to whichever interests prevailed in competitive interest group jockeying. In these schemes, individual goals and motivations of the regulators themselves are ignored.

This was one of Croley’s (1998) key criticisms of these two major approaches to regulatory theory, and in this respect his overall criticism accords well with the material in this thesis. In the first place, argued Croley, incumbent legislators are relatively safe once they have been elected, so the emphasis on help with re-election as ‘payment’ for regulatory services rendered is overstated. More than that, Croley pointed to studies suggesting that public office-holders are motivated more by ideology than by electoral considerations (42-43) — an observation consistent with the empirical narrative of this thesis, in which for both the 1962 drug amendments and the events leading to the 1997 Food and Drug Administration Modernization Act, the chief allies of the pharmaceutical industry were Republicans while the chief promoters of authority for the Food and Drug

Administration were the Democrats.⁷ More than that, the regulatory process is black-boxed in the public choice version of events, with legislator-regulator slack essentially assumed away. The assumption is that the administrators carry out the will of the legislators unproblematically, and therefore any special interest pressure exerted on legislatures is transmitted to administrative outcomes like an electric pulse in a telephone wire (26):

legislative inputs → regulatory outcomes

However, Croley observed, regulators are not necessarily subject to the same pressures as legislators, and may have very different interests to them. If so, special interests could be negated through the administrative actions of the regulators. In short, his criticism is that this model of regulatory action should look more like (27):

legislative inputs → regulatory process → regulatory outcomes outputs

Croley went on to develop an improved model of regulation (29) in which constituent demands feed into legislative goals, which are combined with congressional organizational and procedural inputs to create legislative directives which then feed into regulator goals. The regulator's goals then combine with regulatory organizational and procedural inputs to create preliminary regulatory outputs. These preliminary outputs then undergo executive, legislative, and judicial review prior to being finalized as outcomes, which then feed back into constituent demands.

While certainly a vast improvement over the oversimplified models of administrative procedure proposed by public choice and neopluralism, Croley's model can nevertheless benefit from the empirical content of this thesis. In Croley's model, influence cascades downwards through the layers of bureaucracy *unidirectionally*, not allowing for any feedback from administrative practice. However, as we have seen, regulators have a great deal of discretionary power in their day-to-day decision-making. The tremendous flexibility regulatory agencies can use in practice to address each extenuating circumstance and each instance of something 'new' tends to lead to de facto rule changes over time. This is a phenomenon overlooked even by civic republicanism, which was the perspective most sympathetic to interactionism and finitism in arguing that

⁷ See Chapters 3 and 8, especially Section 3.1 and Chapter 8 note 24.

attitudes to proposed regulatory actions are not predetermined but arise from a process of interaction and negotiation. As in the example of the FDAMA, rather than directing agency action in a proactive or anticipatory way, the legislature may find itself following in the wake of agency practice, seeking to provide ‘legislative support’ for existing practice. While we could certainly envision a legislature seeking to rein in such flexible regulatory action and discretionary practice as well, rather than seeking to endorse and solidify it, either way the legislature would be reacting to existing practice, not merely directing it proactively. Not only is the administrative process not amenable to black-boxing in any viable discussion of regulation, but it can and does serve as the inspiration for legislative action. Hence (using a pared down version for the sake of simplicity), Croley should insert feedback loops for knowledge praxis (Figure 9.1).

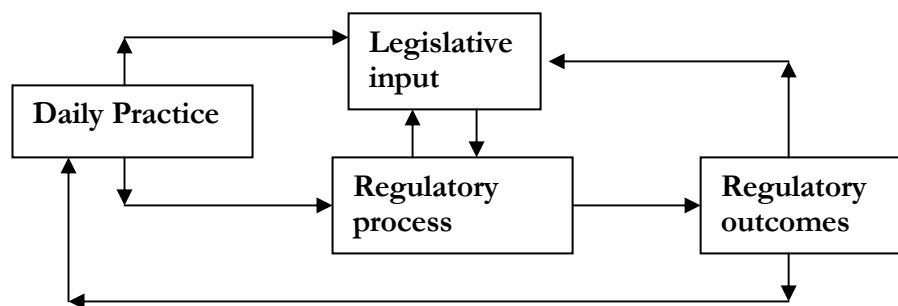


Figure 9.1
Regulatory Process and Knowledge Praxis

This kind of feedback was a predominant feature of the final version of the FDAMA. It was also a significant factor in the development of the PDUFA: Drug sponsors did not ask legislators to pay user fees to the FDA; rather this was a concession to the FDA and their congressional allies in the hopes of obtaining quicker review of their applications. This latter observation also makes clear that the ‘constituency’ providing inputs to the legislative process could very well include the FDA itself.

These knowledge praxis feedback loops are nourished by the shifting applications of rules, procedures, and standards characteristic of finitist concept application as demonstrated throughout this thesis. Although the degree to which this feedback takes place may vary from agency to agency, and likely takes different forms depending on the agency’s function and daily activities, the example of the FDA makes clear that not only

regulatory outputs, but also legislative inputs to regulation, are influenced by regulatory daily decision making and underlying practices, formal and informal.

9.2.2 Nature of 'Monitoring' and 'Slack'

An objection we can make to the theories of regulation summarized in Chapter 1 is that all of them tend to treat legislative intent as singular and unproblematic. Yet it is clear that the goals of legislatures are neither unambiguous nor univocal. Members of congressional oversight committees argue over whether the FDA's performance is adequate and over the appropriate actions needed to achieve more desirable outcomes. House and Senate committee reports on proposed legislation contain not only descriptions of legislative intent, but also often descriptions of committee member objections to certain aspects of the legislation — objections which do not dissolve once the proposal is made into law and may be re-expressed depending on the constitution of committee memberships and the political party in charge. Moreover, to the degree that legislators are susceptible to lobbying groups, we have seen that legislators tend to align themselves with selected interests on an ideological basis. Hence, a great deal of 'slack' is built into the legislator-regulator relationship simply by virtue of the fact that the effectiveness of monitoring is necessarily reduced when legislators themselves disagree on how agencies should execute their duties.

The importance of this observation is that since, in effect, a source of that slack can be the disagreement of the legislators themselves, the slack is not necessarily tightened simply through increased monitoring. In the principal-agent relationship between Congress and regulators, monitoring (as mostly performed through congressional oversight hearings) is by definition inefficient except when bipartisan consensus is close to unanimous, in which case the desires and directives of congressional committees are much more focused.

The same can be said for the principal-agent relationship between citizens and elected representatives. Recall (from Chapter 1) that a chief prediction for public interest theory was that 'hot' issues under high public scrutiny will tend to be resolved in favour of the public, since regulators will be operating in a low-slack environment. However, even for such closely monitored issues, elected officials need only fear the slack reducing effects of monitoring where there is a high degree of consensus on what their actions should be. The public interest theorists seem to assume that intense interest in an issue equates with a

high level of public agreement over what the outcome should be. But there is no guarantee that the electorate will speak with one voice simply because it feels strongly. Indeed, in the U.S., on the hot issue of abortion the electorate is both keenly aware of political developments and sharply divided on them. More than that, on less hotly contested issues where active monitoring is less likely, voters tend to vote by political party, not by issue. For this reason, we can say that on most issues regardless of the level of public interest, short of blatant malfeasance there will typically always be some segment of the constituency which supports the actions of a specific legislator, whatever those actions are. Under such circumstances, the effectiveness of monitoring to prevent slack is diluted considerably.

Hence, under normal circumstances slack is not solely the product of a lack of monitoring, but also a lack of *consensus* among those doing the monitoring. Citizens will produce a divided vote. So will legislatures. Oversight committees can be divided and argue among themselves over the proper role or action of a regulatory agency. In this way, we see that the notions of ‘slack’ and ‘monitoring’ are weakened considerably as analytical categories when the realistic disunity of the principals is taken into account. This is not to say that these categories can be discarded completely, but they must be modified. Monitoring is most important and slack is at its lowest when consensus is high. The less consensus exists, the less influential monitoring will be, even under conditions of very low slack.

I should also point out that, since the advent of AIDS activism, we have increasingly seen instances where dissatisfied constituents, such as the Provenge protestors (Chapter 1), take their grievances directly to the FDA in the form of protests, angry internet web pages or, as in the case of the Abigail Alliance (discussed below), through formal petitions lodged in FDA dockets and law suits. Hence, it is a mistake to consider only citizen-legislator and legislator-regulator relationships in this discussion of principal-agent slack. It is unclear to what extent such actions are influential upon the FDA. To date, the Provenge activists have gone unsatisfied. So also has the Abigail Alliance, which complained in its civil suit against the FDA that a citizen’s petition filed with the Agency went unanswered.⁸ AIDS activists had more success (Epstein 1996), however their multi-pronged attack involved much more than simply protesting. Their

⁸ The suit was filed in 2003 in the District Court of the District of Columbia. The document used to file the complaint can be found at the Abigail Alliance website, <http://www.abigail-alliance.org/> (accessed 15 February 2007).

agenda was not merely to secure approval of a certain drug, but ultimately to change the way therapeutic research and drug development was being done. Many medical scientists and regulators sought to defend their professional interest in the conduct of ‘sound’ science in the face of activist demands to cease placebo-controlled studies, use surrogate endpoints, approve investigational drugs more quickly, etc. To combat this resistance, many AIDS activists gained an expert understanding of AIDS clinical research and trial design, earning credibility with key figures in the FDA and the NIH, and gaining access to internal discussions of clinical trial designs and strategies. Hence, the victories of AIDS activism were gained as much through consensus-building and dialogue as through coercive monitoring and slack-reduction. Again, the notions of ‘monitoring’ and ‘slack’ are insufficient to predict the outcome of direct constituent lobbying of the FDA without the additional concepts of communication and consensus — this time, consensus between the principal and agent.

To be meaningful, it seems that the concepts of monitoring and slack need to be detached from the individualistic framework in which they have been embedded and should be understood in collectivist terms. Recall that in Chapter 1 I argued that the reason Mancur Olson’s individualistic market based model of group behaviour seemed generally to work was that Olson unwittingly introduced a form of mutual susceptibility into his model through individual monitoring of the price of group goods (i.e. in small groups, members would be able to detect ‘free riding’ because of a price increase to all). I further argued that this individualistic approach to group behaviour was ultimately inadequate to explain why Olson’s model was diverging with empirical reality over time; to explain the latter phenomenon we needed to shift to a collectivist view of group behaviour in which mutual susceptibility proceeded through *communication*, which could then create the conditions for group consensus. And, as Barnes (1995) has argued, once consensus has been achieved, the possibility of coordinated action follows. The cracks in Olson’s foundation begin with the difference between an individualist and collectivist account.

With this critique in mind, let us turn to monitoring and slack. How can monitoring lead to a decrease in slack? Imagine an individual principal monitoring the activities of a given agent — say, a citizen logging onto Thomas.gov to see how her congressional representatives voted on a particular bill (an activity, by the way, which belies the frequent claims that monitoring is expensive. Often times, it is not. What takes

more effort and cost is converting the potential energy of monitoring into the kinetic energy of political action). Let us assume that our citizen is displeased with the outcome. Without an alignment of other, like-minded monitors, no action can follow. More than that, although there may be a coincidental alignment of the attitudes of thousands of other monitors, if those monitors are atomized and unaware of each other, no action can follow. Monitoring in this sense is politically impotent. The only way political action can take place is through conscious alignment of attitudes between monitoring principals. And the only way a conscious recognition of an alignment can take place is through communication, regardless of group size. But once a conscious alignment has been established in a mutually susceptible network, action becomes a possibility, although it comes with a cost. This is often the greater hurdle to be overcome to reduce slack, since consciously aligned individuals must be convinced to write letters, participate in rallies or other political events, give money to organizers to fund awareness-raising efforts, etc. This is the primary mechanism by which monitoring can lead to reduced slack. Without communication and consensus, monitoring is impotent. More than that, between-monitor communication can induce changes in posture; it can foster alignment and realignment, building a larger base of consensus than would have existed otherwise.

These observations modify the basic theoretical premises not only of public choice, but also of neopluralism, and the public interest models of regulation, all of which rely on these concepts of slack and monitoring to anticipate and describe regulatory outcomes. The Olsonian picture of collective action is impoverished because it ignores the nature of social networks and the crucial importance of communication for consensus. Likewise, these regulatory theories miscalculate the conditions under which monitoring is effective and whether slack will be induced because they neglect the heterogeneous and divided nature of the principals and ignore the importance of communication for overcoming those divisions.

9.2.3 *The Rationality of Consensus*

In light of the content of this thesis, we also should question what kind of decision-making is truly included in 'rational' decision-making. We saw in Chapter 1 that Olson's model of collective action was such that it was rational to free ride whenever possible and this was why group goods were always underproduced, whether in large or small groups. But is it? In Olson's model each group member is a potential free-rider,

constrained only by the likelihood of being detected when the group realizes that the price of the group good is rising. Why should the potential free-rider fear detection? Olson would respond that in a small group the free-rider runs the risk of being ejected and therefore denied the group good. However, if a group has achieved a sufficient level of communication to form a consensus that a certain good is worth pursuit and worth whatever individual cost is incurred, and if the likely penalty for free-riding is to be cut off from that group good, then it seems to me that it is generally *not* in the interest of individual members to free ride — not unless some additional (or other) benefit could be obtained through loss of the group good.

For the latter case take, for example, a hypothetical situation contemplated by Olson in his first chapter. Imagine a cartel in which individual members (firms) agree to fix the prices of their products. One free rider is found to be undercutting the other's prices and cornering the market, thereby reaping greater profits than the others. Under these circumstances, ejecting the free rider will not correct the problem, since the free rider will then be free to continue undercutting her competitors. If the free rider cannot somehow be compelled to comply, the cartel will dissolve and the individual members will be forced to return to a competitive situation in which the outcome is uncertain. Olson examined such situations to illustrate how competition and free riding work. However, while free riding is obviously a theoretical possibility in this situation, it seems to me that free riding is *not* actually in the interest of any individual member. Most firms would prefer a reliable source of ample profits over an unpredictable competitive marketplace environment — and this principle can be clearly seen in the history of early twentieth century U.S. and the ultimate need to create anti-trust legislation (see Chapter 1, note 33). In the absence of government coercion, the natural tendency of firms in competition appears to be to collude. Even with anti-trust laws in place, competition tends to lead to mergers, fewer larger firms, and ultimately (if unchecked by government intervention) monopolies (Smith 1997). Thus, although free-riding would result in the freedom to undercut the cartel, this freedom would not last long. Cartels have found common ground in the abiding desire for consistency of profit over the uncertain prospect of more (or less) profits. Hence, even in this case it is generally not in the interest of the group member to free ride.

The assumption of Olson's model is that each individual's interest ultimately conflicts with that of the other individuals in the group, since each member has more

interest in free-riding while receiving the group good than in making a contribution to receive a group good. Hence, the relationships of the group are fundamentally competitive and antagonistic, not cooperative. This view is not entirely consistent with the example of cartels just described. While the members of a cartel are nominally in competition with one another, in the context of price-fixing the group members have achieved a consensus and their interests are clearly aligned. Deviation from that alignment would inevitably lead to a result not favoured by any of the members. Of course, the larger the group the more difficult it is to sustain effective communication for mutual susceptibility. In such cases, we must concede to Olson an increasing tendency to free ride, simply because consensus, mutual encouragement and sanctioning will all tend to become weaker under such circumstances. However, the overall point is that where consensus exists that a group good is worth pursuit, the rational interest of the individual actor often lies in cooperation rather than deviation, especially where mutual susceptibility is strongly sustained through communication.

Clearly, Olson's view of free riding precludes any group having members who might contribute out of a sense of personal loyalty to other group members, or out of an ideological commitment to the 'cause' represented by the group. Olson tended to banish such sentiments to the realm of philanthropic and religious organizations which he said have a different purpose than the economic organizations he analyzed, and therefore operate differently.⁹ It may have been easier to make that argument in 1965, when Olson was writing, before a great many political lobbying organizations grew from the rich soil of religious and social concerns. The political landscape is now cluttered with much more than the trade organizations and labour unions populating much of Olson's analysis, which he sought to keep to groups having a 'significant economic aspect' (Olson 1971, 6). Now we have everything from Greenpeace to the Abigail Alliance padding through the hallways of Capitol Hill in an effort to get their agenda heard. Furthermore, many of these policy agendas have clear economic dimensions. Anti-abortion groups lobby to block expenditure of state funds on abortion clinics; environmental groups lobby for land set-asides, protective measures for certain species of plants and animals, curbing industry emissions, developing alternative energy sources, and other initiatives which require public funding and can have direct economic impacts. These groups also often receive

⁹ See Olson (1971) page 6, note 6. Although Olson made this distinction, he nevertheless suggested that clubs and religious organizations would be analyzable in principle by the same techniques he used for groups having economic interests at the core of their activities, but he believed it to be a fruitless pursuit.

substantial voluntary contributions and employ professional staffs to manage budgets and promote the policy aims of the group — something not fully envisioned by Olson's theory of 'byproduct' groups. Such organizations cannot be classified as primarily philanthropic and cannot be dismissed as primarily social clubs or a provider of services with politics as a byproduct.

These groups become much more comprehensible if we analyze them not as loosely aligned 'latent' groups who coincidentally share a particular interest and which somehow became organized, as Olson does, but if we see them as composed of networks of people who consciously share a strong consensus over a particular issue. Seen in this light, voluntary contributions, voluntary service, and commitment out of loyalty to fellow members or to the 'cause' become more sensible. Of course, some people who sympathize with the cause may not volunteer time or money or other resources to the cause. This fact does not detract from the underlying principle at stake: for those who feel most strongly about the public good to be provided, and for those most fully integrated into mutually susceptible networks which promote consensus on the need for that public good, self interest is perceived in terms of achieving that public good. The actor in question believes that she will improve her own world, as well as the world of her children and grandchildren. The actor believes voluntary cooperation is therefore not only merited, but is worth the price to achieve the goal sought. Other motivations may also be at work. But in such cases, self interest is clearly seen to correspond with the general public interest.

In other words, whether we are discussing the behaviour of firms or the behaviour of activists in an interest group, our model of collective action will have greater empirical validity if we admit that self interest often calls for connection, collaboration, cooperation, and consensus, rather than competition. The problem with Olson's view of group behaviour is not that he assumes people will always act in a self-interested manner. Rather, the difficulty is in seeing all 'rational' behaviour in terms of competition and antagonism. This view is a direct result of attempting to analyze groups in terms of individualism; when one strips away meaningful, mutually susceptible social connections, the only rational behaviour is that which undermines others for the sake of self. But because we live and work in layers of interconnected social networks, communication-seeking and consensus-seeking *are* rational self-interested behaviours.

We saw in Chapter 8 how in the 1994 congressional elections, the Republicans won a sweeping victory. Assuming a public ‘mandate’ for action, and with the alliance of powerful special interests of the kind predicted to cohere by Olson’s logic and anticipated to prevail by public choice theorists, the newly empowered congressional Republicans marched into Congress armed with a broad agenda for reform in the shape of the Contract with America — and ended the 104th Congress two years later having achieved virtually nothing in terms of reforming pharmaceutical regulation. In the 105th Congress, Republicans achieved more moderate adjustments to the system through a process of compromise and consensus-building. Neopluralist theorists would see this as a case of the Republican agenda succeeding up to the point that it met resistance from other interests, and there is validity to that argument. But another lesson can be drawn from these events: the unilateralist approach did not work. It was in the interest of the Republicans to communicate with their opponents and find common ground for compromise. The alternative was to remain deadlocked in unilateral isolation and achieve nothing. To be sure, there are times when congressional representatives refuse collaboration and refuse compromise on ideological grounds. To do so may serve the purpose of mobilizing one’s ideological supporters — one’s political base — and may therefore help an elected official to achieve re-election. However, if one’s goal is to pursue legislative accomplishments, self-interest lies in consensus and communication, not antagonism and competition.

9.2.4 *The View from Here*

From our discussion of regulatory theory both in this chapter and in Chapter 1, we can see that public choice theory has been extremely influential in two ways. First, it informed many of the early influential criticisms of the FDA and clearly continues to live in, for example, the neoconservative agenda of the 104th Congress. Second, it has clearly become the theory that all the other theories seek to correct, counter, or improve. Obviously I have tended to side with the chorus of voices which tend to find fault with public choice. Not only does it grossly underestimate the importance of the regulatory process, but Croley (1998) is correct in saying that its conclusion that regulation should be eliminated in favour of market forces does not comport with its own analysis. More than that, I would add that it makes a fatal error in the logic it uses to arrive at this conclusion. Public choice theorists argue that narrow interests not aligned with the general public

interest typically win in the competition to secure regulatory goods, therefore, regulation is inadequate to protect the public interest and it should be left to the markets. The flaw in this formulation is that it assumes interest group influence will not affect markets with anywhere near the potency that they do regulation. In fact, eliminating regulation and allowing the markets to reign merely guarantees that only one type of narrow interest will prevail: commercial interests. The only defences for the consumer will be caveat emptor and the withholding of purchase — neither a viable prospect for consumers having serious diseases and requiring pharmaceuticals and other medical products.¹⁰

In any event, for the most part here I have not attempted a point-by-point rebuttal or response to any particular theory of regulation. Instead, I have sought to demonstrate how all the theories fall short because of their reliance on individualistic analytical categories. Even where the importance of regulatory decision-making is acknowledged, as in Croley's (1998) corrective, the use of traditional individualistic categories of analysis leads to oversight of an important feedback loop in regulation and legislation founded in tacit practice and de facto consensus about those practices. Individualistic notions of regulatory process therefore disregard knowledge praxis and the mutually influential relationship between science and the regulation of science. This traditional viewpoint similarly led to a misconception of how groups cohere, and therefore also resulted in misfires in the analysis of how, and under which conditions, monitoring can reduce slack. It also resulted in a misconceived view of principal-agent relationships. Finally, I have argued that this individualistic perspective leads to a fundamental misconception of the constitution of self interest. I am not arguing that because people live and work in social networks, they always behave in altruistic ways. To the contrary, I accept that people generally behave in ways calibrated to enhance their own self interest, whether in economic transactions or not. Indeed, I have argued that some social networks can promote greed or other behaviours which are less than civic minded. I have suggested that the solution for such networks is to integrate them into other mutually susceptible networks providing consensus, support, and sanctioning for civically responsible behaviour. My central argument is that the analysis of self interested rationality needs to

¹⁰ Stigler (Cohen and Stigler 1971) has argued that regulation has dulled the defensive reflexes of consumers, making them little more than babes in the woods when it comes to marketplace transactions. If they were not so reliant on government for protection, they would redevelop those skill and be less likely to be fleeced.

be reconsidered in the light of the social networks in which we live — that in such a context it is rational for individuals to seek communication, consensus, and cooperation.

A brief word needs to be said about the one theory of regulation which seemed more consistent with the viewpoint expressed here, civic republican theory. This view of regulation seeks consciously to foster a negotiated approach to regulatory outcomes, taking the view that equilibrium of interest groups in such situations is not a result of conflict, competition, and struggle, as for the other theories, but represents a collective consensus forged in the process of communication and negotiation. While, as I have argued (in Chapter 3), this perspective ignores important tacit processes of consensus in favour of more formal ones, this approach is clearly the most attuned than any of the other perspectives to the way in which meaning is produced in interaction. In some ways it was inappropriate for Croley to include civic republicanism in his analysis of regulatory ‘theories’, since this perspective is designed not to model the regulatory process as it exists, but to advocate for a regulatory process which consciously seeks negotiation and consensus over adversarial conflict. (This is likely one reason civic republicanism passes over tacit processes of consensus; it is not as if one can *plan* to create new tacit practices and de facto rules.) The conclusions of this thesis are entirely sympathetic to that prescription, since it is my contention that doing so would likely be in the self-motivated interest of the negotiants.¹¹ Ironically, even though civic republicanism has at its core a view of decision-making sympathetic with interactionism, Croley (1998) tends to judge it in terms of the individualistic categories of analysis arising from Olson’s *Logic*. At one point, for example, Croley asked how the various participants in regulatory negotiations overcome Olsonian barriers to organization, complaining that civic republicanism ‘provides no account of how parties emerge to engage in regulatory deliberation’ (81). We can see now that this complaint would be valid only if civic republicanism subscribed to the individualistic categories suggested by Olson. However, at least implicitly, it does not seem to. If the critique I have provided in this thesis is valid, then there is an alternative basis for argument on which civic republicanism could stand. It seems clear that these conclusions can represent the beginning of a dialogue in the development of a

¹¹ If one participant was extremely powerful and influential, it could theoretically be in that participant’s interest to remain aloof from the process, using other means to influence the outcome. In practical terms, however, that actor would have to be very influential and confident indeed to boycott negotiations in which he has an obvious stake in the outcome.

theoretically enhanced form of civic republicanism, informed in part by the analysis and perspective provided in this thesis.

9.3 Afterward

To conclude this thesis, let us return to where we began: the controversy over the delayed approval of the prostate cancer drug Provenge. An *ABC News* story (ABC News 2007b) on the controversy reported, ‘The campaign for Provenge has been seized on by other groups who want to make it easier for patients to get experimental therapies’. The report then quoted Frank Burroughs of the Abigail Alliance, a patient advocacy group which is suing the FDA in federal court for the *right* to expanded access of investigational drugs.¹² Burroughs was quoted as saying that ‘when you have a drug like Provenge, you should let people have access to it who have run out of options’.¹³

To hear such comments, one would think that the 1980s and 1990s never occurred. Indeed, in addition to the lawsuit, Burroughs’ group has been influential in the creation of proposed legislation which continued to push the boundaries of the tension between data-gathering and early decision-making traced thorough this thesis. The bill, S. 1956, introduced by Senator Sam Brownback in November 2005, would among other things *prohibit* the use of placebo-controls and no-treatment concurrent control groups in clinical trials for life-threatening conditions where positive controls can be used effectively (Sec. 4). It would also legislate the creation of a programme to encourage the development of surrogate endpoints and biomarkers for serious or life-threatening conditions and initiate another programme (to be conducted by the Institute of Medicine) to evaluate existing biomarkers and surrogate endpoints for validity (Sec. 5) . It would require the Oncologic Drugs Advisory Committee (ODAC) to admit at least two patient

¹² This lawsuit rather naively alleges that the sole hindrance to patients hoping to receive investigational drugs under treatment IND is FDA’s policy preventing drug-makers from charging more than a cost recovery fee for drugs distributed for compassionate use prior to approval. On this basis, the suit claims that the FDA’s policy hinders patients’ abilities to make appropriate therapeutic choices, thereby infringing their rights to privacy and liberty under the Constitution; and that the FDA policy effectively ‘operates as a death sentence’ for seriously ill patients lacking therapeutic alternatives, therefore violating the individual’s Constitutional guarantee against ‘deprivation of life without due process’. Notably, two of the four examples of such infringement of rights cited in the complaint involved attempts to get access to ImClone’s Erbitux (Chapter 8, see especially note 19). See a copy of the original suit posted at the Abigail Alliance website, http://www.abigail-alliance.org/WLF_FDA_Lawsuit.pdf. (accessed 15 February 2007). Quoted material is from pp. 10-11.

¹³ Patients with the specific form of prostate cancer studied in the Provenge trials have limited treatment options to be sure, but options are available: hormonal therapy (the effectiveness of which is not proven) and docetaxel (which has been clinically shown to be effective, but is toxic (FDA 2007, 70).

representatives and to give those representatives voting rights (Sec. 7). More significantly, the legislation would replace fast track drug approval with a three-tiered system (Sec. 3). Tier I drugs would be approved conditionally only on the basis of preclinical studies, pharmacological safety studies, case reports and other preliminary clinical information, *not* on statistically evaluable data.¹⁴ Tier II drugs would be conditionally approved on the basis of early clinical trials using surrogate endpoints or biomarkers or clinical endpoints. Tier III would represent what we think of as traditional clinical evidence for approval. The bill would require the FDA to establish a ‘new’ programme for expanded access, obliging the Agency to publish (yet more) guidance on mechanisms available for investigational access to drugs, encouraging Tier I and II conditional approvals, and facilitating early access to investigational drugs.

This legislation never made it out of the Senate committee to which it was referred, but the proposed legislation would have mandated decision points even earlier in the drug approval process, with a form of limited approval taking place mainly on the basis of animal studies and preliminary toxicity data — and would have done so even as diabetes patients and their physicians try to decide whether to ignore new revelations on the apparently cardiac-dangerous Avandia or to switch back to the older renal-toxic Rezulin.

In this way, the very recent history of the FDA looks very similar to the history of the FDA from ten, twenty, and thirty years ago. However, perhaps what the case of Provenge illustrates is a bit of a counterrevolution in which the FDA and medical experts are beginning to push back against calls for ever-earlier decision-making and ever-weaker bodies of evidence for those decisions — even for approval of drugs intended for life-threatening diseases. As examples of drugs with serious adverse effects like Avandia become more common, as seems likely, we will have reaped what was sown by accepting greater risks of information deficit for the sake of earlier decision-making.¹⁵ The question is not whether the FDA, Congress, activists, public choice theorists, and others acted

¹⁴ Tier I essentially represents a form of ‘compassionate use’ under which the drug can be sold for profit. Since the drug is ‘approved’, there are no restrictions on sale. However, the legislation seeks to limit distribution of the drug to very ill patients lacking therapeutic options, in part by requiring patients to sign an informed consent waiver as well as an affidavit waiving their rights to sue the drug-maker or the prescribing physician in the event of serious or fatal adverse effects arising from the use of the product.

¹⁵ The FDA has denied that withdrawals from the market are indicative of a fundamental problem with the approval process (Friedman, et. al. 1999). However, since this article was written the number of examples of marketed drugs having unrecognized serious adverse effects has multiplied, making it increasingly difficult to argue that any appearance of a trend is merely coincidental.

rightly in creating this history. Rather, the question is whether we, the public, are willing to accept the risk of information deficit in exchange for earlier access to drugs now that we see the specific shape that risk can take. The question of how much data is sufficient for decision-making is ultimately not a scientific one, but a social one. It is to be determined, as in the past, not only through political struggle, but also through the practical actions taken by regulators and others on a day-to-day basis to solve problems and create knowledge.

WORKS CITED

- ABC News (2007a), 'Diabetes Drug Linked to Heart Attacks', *ABC World News* [Television broadcast], 21 May 2007, 7:00 pm, EST. Transcription by author from video available at <http://abcnews.go.com/health> (accessed 22 May 2007).
- (2007b), "'We Need to Move Now": Cancer Patients Urge FDA Approval of Provenge, As Their Survival Options Dwindle', ABC News Online, <http://abcnews.go.com/WN/MedicineCuttingEdge/story?id=3246469&page=1> (accessed 5 June 2007).
- Abraham, John (1993), 'Scientific Standards and Institutional Interests: Carcinogenic Risk Assessment of Benoxaprofen in the UK and US', *Social Studies of Science* 23: 387-444.
- (1994), 'Distributing the Benefit of the Doubt: Scientists, Regulators, and Drug Safety', *Science, Technology, & Human Values* 19 (4): 493-522.
- (1995a), 'The Production and Reception of Scientific Papers in the Academic-Industrial Complex: The Clinical Evaluation of a New Medicine', *The British Journal of Sociology* 46 (2): 167-190.
- (1995b), *Science, Politics, and the Pharmaceutical Industry: Controversy and Bias in Drug Regulation* (New York: St. Martin's Press).
- Abraham, John and Courtney Davis (2007), 'Interpellative Sociology of Pharmaceuticals: Problems and Challenges for Innovation and Regulation in the 21st Century', *Technology Analysis & Strategic Management* 19(3): 387 - 402
- Abraham, John and Julie Sheppard (1997), 'Democracy, Technocracy, and the Secret State of Medicines Control: Expert and Nonexpert Perspectives', *Science, Technology, & Human Values* 22 (2): 139-167.
- (1999), 'Complacent and Conflicting Scientific Expertise in British and American Drug Regulation: Clinical Risk Assessment of Triazolam', *Social Studies of Science* 29 (6): 803-843.
- Abramowitz, Alan I. and Kyle L Saunders (1998), 'Ideological Realignment in the U.S. Electorate', *The Journal of Politics* 60 (3): 634-52.
- Ahart, Gregory (1979), 'The Food and Drug Administration Approval Process', Statement Before the Committee on Science and Technology, Subcommittee on Science, Research and Technology, 19 June 1979 (Washington, DC: U.S. General Accounting Office) [Available at <http://www.gao.gov/docsearch/repandtest.html> (accessed 22 June 2005)].
- Angell, Marcia (2000a), 'Is Academic Medicine for Sale?', *New England Journal of Medicine* 342: 1516-1518.

- (2000b), ‘The Pharmaceutical Industry: To Whom is it Accountable?’, *New England Journal of Medicine* 342: 1902-1904.
- (2004), *The Truth About the Drug Companies: How They Deceive Us and What to Do About It* (New York: Random House).
- Anonymous (1979), ‘Dubious Tactics in the War on Cancer’, *Business Week* (27 August 1979), 46D.
- Anonymous (1991), ‘CAST Deaths Revealed’, *Lancet* 337 (8747) (20 April 1991): 969.
- Anonymous (1996), ‘White House Announces Cancer-Drug Initiatives’, *American Journal of Health-System Pharmacy* 53 (10)(15 May 1996), 1110.
- Anspach, R. (1987), ‘Prognostic Conflict in Life-and Death Decisions: The Organization as an Ecology of Knowledge’, *Journal of Health and Social Behavior* 28 (3): 215-231.
- (1993), *Deciding Who Lives: Fateful Choices in the Intensive Care Nursery* (Berkeley: University of California Press).
- Austin, J. L. (1962), *How to do Things with Words* (Oxford: Clarendon Press).
- Barnes, Barry (1977), *Interests and the Growth of Knowledge* (London: Routledge and Kegan Paul Ltd.).
- (1982), *T.S. Kuhn and Social Science* (London and Basingstoke: Macmillan).
- (1983), ‘Social Life as Bootstrapped Induction’, *Sociology* 17 (4): 524-545.
- (1995), *The Elements of Social Theory* (Princeton: Princeton University Press).
- Barnes, Barry and David Bloor (1982), ‘Relativism, Rationalism and the Sociology of Knowledge’ in *Rationality and Relativism*, eds. Martin Hollis and Steven Lukes (London: Basil Blackwell), 21-47.
- Barnes, Barry, David Bloor and John Henry (1996), *Scientific Knowledge: A Sociological Analysis* (Chicago: University of Chicago Press).
- Bazell, Robert (1998), *Her-2: The Making of Herceptin, A Revolutionary Treatment for Breast Cancer* (New York: Random House).
- Bean, Melissa Marie (2003), ‘Fatal Flaws in the Food and Drug Administration's Drug-Approval Formula’, *Utah Law Review* 881 – 913.
- Becker, Gary S. (1979), ‘Altruism, Egotism and Genetic Fitness: Economics and Sociobiology’, *Journal of Economic Literature* 14(3): 817-826.
- (1981), ‘Altruism in the Family’, *Economica* 48: 1-15.
- (1983), ‘A Theory of Competition Among Pressure Groups for Political Influence’, *Quarterly Journal of Economics* (98)(3): 371-400.
- (2002), ‘Get the FDA Out of the Way, and Drug Prices Will Drop’, *Business Week* (16 September 2002): 16.

- Berg, Marc (1992), 'The Construction of Medical Disposables. Medical Sociology and Medical Problem Solving in Clinical Practice', *Sociology of Health and Illness* 14: 151-80.
- (1997), *Rationalizing Medical Work: Decision-Support Techniques and Medical Practices* (Cambridge, MA: MIT Press).
- Berg, Marc and Annemarie Mol, eds. (1998), *Differences in Medicine. Unraveling Practices, Techniques and Bodies* (Duke University Press).
- Blaikie, Norman (2000), *Designing Social Research: The Logic of Anticipation* (Oxford, UK and Malden, MA: Blackwell).
- Blakeslee, Alton (1977), 'Group Attacks FDA Authority over Anti-Cancer Drug Testing', *Washington Post* (20 January 1977), A3.
- Bloor, David (1983), *Wittgenstein: A Social Theory of Knowledge* (London: Macmillan).
- (1991), *Knowledge and Social Imagery* (Chicago and London: University of Chicago Press).
- (1992), 'Left and Right Wittgensteinians', in *Science and Practice as Culture*, ed. Andrew Pickering (Chicago: University of Chicago Press): 266-82.
- (1996), 'Ideology and the Sociology of Knowledge', *Social Studies of Science* 26 (4): 839-856.
- (1997), *Wittgenstein, Rules and Institutions* (London and New York: Routledge).
- (2007), 'Epistemic Grace: Antirelativism as Theology in Disguise', *Common Knowledge* 13 (2/3): 250-280.
- Bodenheimer, Thomas (2000), 'Uneasy Alliance: Clinical Investigators and the Pharmaceutical Industry', *New England Journal of Medicine* 342: 1539-1544.
- Borchardt, John K. (2000), 'Is the FDA's Drug Review Process Breaking Down?', *The Scientist* 1 (1) (20 December): 1220-02.
- Borrelli, Stephen A. (2001), 'Finding the Third Way: Bill Clinton, the DLC, and the Democratic Platform of 1992', *Journal of Policy History* 13 (4): 429-462.
- Box-Steffensmeier, Janet (2004), *Event History Modeling: A Guide for Social Scientists*. (Cambridge: Cambridge University Press).
- Breyer, Stephen (1982), *Regulation and Its Reform* (Cambridge, MA and London: Harvard University Press).
- Brickman R. and Rip A. (1979), 'Science Policy Advisory Councils in France, the Netherlands and the United States, 1957-77: A Comparative Analysis', *Social Studies of Science* 9 (2): 167-198.
- Brock N., J. Stekar and J. Pohl (1979), 'Antidote against the urotoxic effects of the oxazaphorine derivatives cyclophosphamide, ifosfamide and trofosfamide', [German] *Naturwissenschaften* 66(1): 60-1.

- Brock, N (1980), 'The development of mesna for the inhibition of urotoxic side effects of cyclophosphamide, ifosfamide, and other oxazaphosphorine cytostatics', *Recent Results in Cancer Research* (74): 270-8.
- Broder, Samuel (1997), Interview by Dr. Victoria Harden (Director, NIH Historical Office) and Dr. Caroline Hannaway (NIH Historical Contractor) (Bethesda, MD, 2 Feb 1997). <http://aidshistory.nih.gov/transcripts/transcripts/Broder97.pdf> (accessed 3 October 2006).
- Bud, Robert (2005), 'Antibiotics, Big Business, and Consumers: The Context of Government Investigations into the Postwar American Drug Industry', *Technology and Culture* 46 (April 2005): 329-349.
- Burgess, Robert G. (1984), *In the Field: An Introduction to Field Research* (London and New York: Routledge).
- Buyse, Marc.E., Maurice J. Staquet and Richard J. Sylvester (eds.) (1984), *Cancer Clinical Trials: Methods and Practice* (Oxford, New York, and Toronto: Oxford University Press).
- Callon, Michel (1986), 'Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay' in *The Science Studies Reader*, ed. Mario Biagioli (New York and London: Routledge), 67-83.
- Cambrosio, Alberto, Peter Keating, Thomas Schlich, and George Weisz (2006), 'Regulatory Objectivity and the Generation and Management of Evidence in Medicine', *Social Science and Medicine* 63: 189-199.
- Canguilhem, G. (1989), *The Normal and the Pathological* (New York: Zone Books).
- Carpenter, Daniel (2003), 'Groups, the Media, Agency Waiting Costs, and FDA Drug Approval', *American Journal of Political Science* 46 (3): 490-505.
- Casper, Monica and Adele Clarke (1998), 'Making the Pap Smear into the "Right Tool" for the Job: Cervical Cancer Screening in the USA, circa 1940-95', *Social Studies of Science* 28 (2): pp. 255-290.
- CAST Investigators (1989), 'Preliminary Report: Effect of Encainide and Flecainide on Mortality in a Randomized Trial of Arrhythmia Suppression after Myocardial Infarction', *New England Journal of Medicine* 321: 406-12.
- Christakis, Nicholas A. and James H. Fowler (2007), 'The Spread of Obesity in Large Networks over 32 Years', *New England Journal of Medicine* 357 (26 July): 370-79.
- Clark, J., D. Potter, and J. McKinlay (1991), "Bringing Social Structure Back into Clinical Decision Making." *Social Science and Medicine* 32: 853-66.
- Clinton, Bill and Al Gore (1995), *Reinventing Regulation of Drugs and Medical Devices*, National Performance Review (April 1995).
- Cohen, Manuel F. and George J. Stigler (1971), *Can Regulatory Agencies Protect Consumers?* (Washington D.C., American Enterprise Institute for Public Policy Research).
- Cole, Simon A. (1998), 'Witnessing Identification: Latent Fingerprinting Evidence and Expert

- Knowledge', *Social Studies of Science* 28 (5/6): 687-712.
- Collins, Harry M. (1975), 'The Seven Sexes: A Study in the Sociology of a Phenomenon or the Replication of Experiments in Physics', *Sociology*, 9: 205-24.
- Collins, Harry M. and Trevor Pinch (1979), 'The Construction of the Paranormal: Nothing Unscientific is Happening', *The Sociological Review Monograph* 27: 237-270.
- (1982), *Frames of Meaning: The Social Construction of Extraordinary Science* (London: Routledge and Kegan Paul).
- (1994), *The Golem: What Everyone Should Know about Science* (Cambridge: Cambridge University Press)(Canto Edition).
- (2002), *The Golem at Large: What You Should Know about Technology* (Cambridge: Cambridge University Press)(Canto Edition).
- Columbia University Teachers College (2004), 'Study Of Obscure Amazon Tribe Sheds New Light On How Language Affects Perception', *ScienceDaily*, 20 August 2004, <http://www.sciencedaily.com/releases/2004/08/040820083420.htm> (accessed 7 Feb. 2008).
- Cooper, Ellen (2006), Interview by Author (digital recording, telephone interview), 10 August and 14 August 2006.
- Cox-Maksimov, Desiree (1998), *The Making of the Clinical Trial in Britain, 1910-1945: Expertise, the State and the Public* (PhD Dissertation, University of Cambridge).
- Croley, Steven P. (1998), 'Theories of Regulation: Incorporating the Administrative Process', *Columbia Law Review* 98 (1): 1-168.
- Daemmrich, Arthur (2004), *Pharmacopolitics : Drug Regulation in the United States and Germany*. Chapel Hill, NC: University of North Carolina Press.
- (2002), *Therapeutic Cultures: Pharmaceutical Regulation and Medical Politics in the United States and Germany*, (PhD Dissertation, Cornell University).
- (1998), 'The Evidence Does Not Speak for Itself: Expert Witnesses and the Organization of DNA-Typing Companies', *Social Studies of Science* 28 (5): 741-72.
- DeVita, Vincent T., Samuel Hellman and Stephen A. Rosenberg (1993), *Cancer: Principles and Practice of Oncology* (Philadelphia: J.B. Lippincott Company).
- Dranove, David and David Meltzer (1994), 'Do Important Drugs Reach the Market Sooner?', *The RAND Journal of Economics* 25 (3): 402-423.
- Echterhoff, Gerald, E. Tory Higgins, René Kopietz, and Stephan Groll (2008), 'How Communication Goals Determine When Audience Tuning Biases Memory', *Journal of Experimental Psychology: General* 137 (1): 3-21.
- Edgar, Harold and David J. Rothman (1990), 'New Rules for New Drugs: The Challenge of AIDS to the Regulatory Process', *The Millbank Quarterly* 68/sup. 1: 111-142.

- Edmond, Gary and David Mercer (1996/97), 'Recognising Daubert: What Judges Should Know About Falsificationism', *Expert Evidence* 5/6: 28-40.
- (1998), 'Trashing "Junk Science"', *Stanford Technology Law Review* 3 [Electronic Journal]. Available at: http://stlr.stanford.edu/STLR/Articles/98_STLR_3/.
- (2000), 'Litigation Life: "A Phenomenology of Toxic Torts"', *Social Studies of Science* 30: 265-316.
- Ehrenreich, Barbara (2008), 'The Boom Was a Bust For Ordinary People', *Washington Post* (3 February 2008): B01.
- Einhorn, Lawrence H. (1990), 'Treatment of Testicular Cancer: A New and Improved Model', *Journal of Clinical Oncology* (8): 1777-1781.
- Epstein, Steven (1995), 'The Construction of Lay Expertise: AIDS Activism and the Forging of Credibility in the Reform of Clinical Trials', *Science, Technology, & Human Values* 20 (4): 408-437.
- (1996), *Impure Science: AIDS, Activism, and the Politics of Knowledge* (Berkeley, Los Angeles, and London: University of California Press).
- (1997), 'Activism, Drug Regulation, and the Politics of Therapeutic Evaluation in the AIDS Era: A Case Study of ddC and the "Surrogate Markers" Debate', *Social Studies of Science* 27 (5): 691-726.
- Ezzell, Carol (1988), 'US Looking for Shortcuts to Speed Drug Approval', *Nature* 334 (18 August 1988): 553.
- Fisher, Morris A. (2003), 'Physicians and the Pharmaceutical Industry: A Dysfunctional Relationship', *Perspectives in Biology and Medicine* 46 (2): 254-272.
- Food and Drug Administration (FDA) (1962), 'New Drugs for Investigational Use, Proposed Exemptions', *Federal Register* 27 (10 August 1962): 7990-7992.
- (1963), 'Procedural and Interpretive Regulations; Investigational Use', *Federal Register* 28 (8 January 1963): 179-182.
- (1969a), 'Novobiocin-Tetracycline Combination Drugs; Calcium Novobiocin – Sulfamethizole Tablets', *Federal Register* 34/93 (15 May 1969): 7687.
- (1969b), 'Novobiocin-Tetracycline Combination Drugs; Calcium Novobiocin – Sulfamethizole Tablets', *Federal Register* 34/152 (9 August 1969): 12958 – 12968.
- (1969c), 'Hearing Procedure for Refusal or Withdrawal of Approval of New Drug Applications and for Issuance, Amendment, or Repeal of Antibiotic Drug Regulations; Interpretive Description of Adequate and Well-Controlled Clinical Investigations', *Federal Register* 34/180 (19 September 1969): 14596 - 14598.
- (1970a), 'Hearing Requests on Refusal or Withdrawal of New Drug Applications and Issuance, Amendment, or Repeal of Antibiotic Drug Regulations and Describing Scientific Content of Adequate and Well-Controlled Clinical Investigations', *Federal Register* 35/33 (17 February 1970): 3073 – 3074.

- (1970b), ‘Hearing Regulations and Regulations Describing Adequate and Well-Controlled Clinical Investigations’, *Federal Register* 35/90 (8 May 1970): 7250–7253.
- (1974), ‘Requirements of Notice of Opportunity for Hearing, Request for Hearing, and Grant or Denial of Hearing’, *Federal Register* 39 (50) (13 March 1974): 9750-9766.
- (1977), *General Considerations for the Clinical Evaluation of Drugs* (September 1977), HEW (FDA) 77-3040.
- (1978), ‘Obligations of Clinical Investigators of Regulated Articles’, *Federal Register* 43/153 (8 August 1978): 35210- 35236.
- (1979a), *Investigational and New Drug Regulation Revisions: Concept Document* (October 1979), FDA Docket 79N-0388.
- (1979b), IND/NDA Rewrite Public Meeting, 7 November 1979, Washington, D.C., FDA Docket 79N-0388. [Transcript: Public Meeting.]
- (1981), *Guidelines for the Clinical Evaluation of Antineoplastic Drugs* (9 June 1981), FDA Docket 81D-0150.
- (1982), ‘New Drug and Antibiotic Regulations. Proposed Rule’, *Federal Register* 47/202 (19 October 1982): 46622-46666.
- (1983), ‘New Drug, Antibiotic, and Biologic Drug Product Regulations. Proposed Rule’, *Federal Register* 48/112 (9 June 1983): 26720-26749.
- (1987a), Anti-Infective Drugs Advisory Committee Meeting, 16 January 1987, Bethesda, MD [Transcript: New Drug Application for AZT].
- (1987b), ‘Investigational New Drug, Antibiotic, and Biological Drug Product Regulations; Treatment Use and Sale. Reproposed Rule’, *Federal Register* 52/53 (19 March 1987): 8850-8855.
- (1987c), ‘Investigational New Drug, Antibiotic, and Biological Drug Product Regulations; Treatment Use and Sale. Final Rule’, *Federal Register* 52/99 (22 May 1987): 19466-19477.
- (1987d), Oncologic Drugs Advisory Committee Meeting, 7 December 1987, Rockville, MD [Transcript: Meeting on mitoxantrone to treat acute non-lymphocytic leukaemia and breast cancer.]
- (1988a), ‘Investigational New Drug, Antibiotic, and Biologic Drug Product Regulations; Procedures for Drugs Intended to Treat Life-Threatening and Severely-Debilitating Illnesses. Interim rule’, *Federal Register* 53/204 (21 October 1988): 41516-41524.
- (1988b), Oncologic Drugs Advisory Committee Meeting, 19 April 1988, Rockville, MD [Transcript: Meeting on ifosfamide for refractory testicular cancer.]
- (1990a), Antiviral Drugs Advisory Committee Meeting, Volume II, 30 January 1990, Bethesda, MD [Transcript: Update on Expanded Access: the dDI Experience].

- (1990b), Antiviral Drugs Advisory Committee Meeting, 30 March 1990, Gaithersburg, MD [Transcript: AZT for Pediatric AIDS].
- (1990c), Antiviral Drugs Advisory Committee Meeting, Volume II, 18 September, Bethesda, MD [Transcript: Design of Pivotal Efficacy Study to Support NDA Approval of New Antiretroviral Agents and Immunomodulators for the Treatment of Individuals with HIV Infection].
- (1990d), Oncologic Drugs Advisory Committee, Volume II, 11 September, Gaithersburg, MD [Transcript: Fludara (Fludarabine) for the Palliative Treatment of Chronic Lymphocytic Leukemia Refractory to Other Therapy].
- (1990e), ‘Investigational New Drug, Antibiotic, and Biological Drug Product Applications; Clinical Hold and Termination; Amendment’, *Federal Register* 55/98 (21 May 1990): 20802 – 20805 [Announcement of Parallel Track].
- (1991a), Antiviral Drugs Advisory Committee Meeting, Volume I, 18 July 1991, Bethesda, MD [Transcript: New Drug Application for ddI].
- (1991b), Antiviral Drugs Advisory Committee Meeting, Volume II, 19 July 1991, Bethesda, MD [Transcript: New Drug Application for ddI continued].
- (1992a), ‘New Drug, Antibiotic, and Biologic Drug Product Regulations; Accelerated Approval. Proposed Rule’, *Federal Register* 57/73 (15 April 1992): 13234-13242.
- (1992b) Antiviral Drugs Advisory Committee Meeting, 20 April 1992, Bethesda, MD [Transcript: Presentation of Confirmatory Data for Previous Year’s Approval of ddI and presentation of data for ddC New Drug Application].
- (1992c) Antiviral Drugs Advisory Committee Meeting, 21 April 1992, Bethesda, MD [Transcript: New Drug Application for ddC].
- (1992d), ‘New Drug, Antibiotic, and Biologic drug Product Regulations; Accelerated Approval. Final rule’, *Federal Register* 57/239 (11 December 1992): 58942-58960.
- (1994a), Antiviral Drugs Advisory Committee Meeting, 20 May 1994, Rockville, MD [Transcript: New Drug Application for d4T].
- (1994b), Oncologic Drugs Advisory Committee Meeting, 12 December 1994, Rockville, MD [Transcript: New Drug Application for Zinecard and Ethyol].
- (1994c), ‘Medical Devices: Exemptions from Premarket Notification for Certain Classified Devices’, *Federal Register* 59/234 (December 7, 1994): 63005-63015
- (1995), Oncologic Drugs Advisory Committee Meeting, 14 February 1995, Rockville, MD. [Transcript: New Drug Application for DOX-SL and for Zoladex.]
- (1997), *Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products*, DRAFT GUIDANCE (March 1997) [Docket No. 97D-0100].
- (1998a), *Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* (May 1998) [Docket No. 97D-0100].

- (1998b), *Guidance for Industry: Fast Track Drug Development Programs – Designation, Development, and Application Review* (September 1998) [Docket No. 98N-0813].
- (2000), Oncologic Drugs Advisory Committee Meeting, 13-14 December 2000, Bethesda, MD [Transcript].
- (2007), Oncologic Drugs Advisory Committee Meeting, 29 March 2007, Gaithersburg, MD [Transcript, NDA for Provenge].
- Friedman, Michael A., Janet Woodcock, Murray M. Lumpkin, J. Shuren, A. Hass, and L. Thompson (1999), ‘The Safety of Newly Approved Medicines: Do Recent Market Removals Mean There Is a Problem?’, *Journal of the American Medical Association* 281 (18): 1728-1734.
- Frohock, Fred M. (1986), *Special Care: Medical Decisions at the Beginning of Life* (Chicago: University of Chicago Press).
- Fujimura, Joan (1987), ‘Constructing Do-Able Problems in Cancer Research: Articulating Alignment’, *Social Studies of Science* 17: 257-93.
- GAO (1980), ‘FDA Approval – A Lengthy Process that Delays the Approval of Important New Drugs’, Report by the Comptroller General of the United States to the Subcommittee on Science, Research, and Technology, House Committee on Science and Technology, 18 May, HRD-80-64 (Washington, DC: United States General Accounting Office).
- Garrety, Karin (1997), ‘Social Worlds, Actor-Networks and Controversy: The Case of Cholesterol, Dietary Fat and Heart Disease’, *Social Studies of Science* 27 (5): 727-773.
- Gelijns, Antonine (1990), ‘Appendix A: Comparing the Development of Drugs, Devices, and Clinical Procedures’, In *Modern Methods of Clinical Investigation*, ed. Antonine C. Gelijns (Washington DC: National Academies Press).
- Gieryn, Thomas (1983), ‘Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists’, *American Sociological Review*, 48 (6): 781-95.
- (1999), *Cultural Boundaries of Science: Credibility on the Line* (Chicago: Chicago University Press).
- Gladwell, Malcolm (1991), ‘Second AIDS Drug Given Conditional Approval’, *Washington Post* (10 October 1991), A4.
- Gore, Al (1993), *From Red Tape to Results: Creating a Government that Works Better and Costs Less* (First Report of the National Performance Review). Available at <http://govinfo.library.unt.edu/npr/library/review.html> (accessed 8 November 2007).
- (1994), *Creating a Government that Works Better and Costs Less* (Second Report of the National Performance Review). Available at <http://govinfo.library.unt.edu/npr/library/review.html> (accessed 8 November 2007).

- (1997), *Businesslike Government: Lessons Learned from America's Best Companies* (Final Report of the National Performance Review). Available at <http://govinfo.library.unt.edu/npr/library/review.html> (accessed 8 November 2007).
- Gouldner, Alvin W. (1962), 'Anti-Minotaur: The Myth of a Value-Free Sociology', *Social Problems* 9: 199-213.
- Grabowski, H. (1982), 'Public Policy and Innovation: The Case of Pharmaceuticals', *Technovation* 1 (3): 157-189.
- Grant, Murray (1981), 'Progress and Problems in FDA Drug Approval Process', Address Presented to the 25th Annual Educational Conference by The Food and Drug Law Institute, 15-16 December 1981, Washington, DC. Available at <http://www.gao.gov/docsearch/repandtest.html> (accessed 22 June 2005).
- Harris, Richard (1964), *The Real Voice* (New York: MacMillan).
- Harter, Philip J. (1982), 'Negotiating Regulations: A Cure for Malaise', *The Georgetown Law Journal* 71(1): 1-118.
- Health, Education and Welfare, Department of (HEW) (1977a), *Interim Reports, Volume I: Review Panel on New Drug Regulation*, (Original Publication 15 November 1976) (Washington D.C.).
- (1977b), *Interim Reports, Volume II: Review Panel on New Drug Regulation*, (Original Publication 28 February 1977) (Washington D.C.).
- (1977c), *Interim Reports, Volume III: Review Panel on New Drug Regulation*, (Original Publication 25 April 1977) (Washington D.C.).
- (1977d), *Final Report: Review Panel on New Drug Regulation*, (31 May 1977) (Washington D.C.).
- (1978), *The Administration Proposal: Drug Regulation Reform Act of 1978, Section-by-Section Analysis* [H.R. 11611 and s.2755 95th Congress, Second Session] (Washington, DC).
- Henderson, Diedtre (2007), 'Prostate Therapy's Delay Sets Off a Fight – Drug's Believers Pressure Regulators', *Boston Globe* (18 June 2007), E1.
- Hilts, Philip J. (1986), 'Experimental Drug AZT was Designed for Tumors; Skill, Luck Led to Promising Tests on AIDS', *Washington Post* (19 September 1986): A11.
- (2003), *Protecting America's Health: The FDA, Business and One Hundred Years of Regulation* (New York: Alfred A. Knopf).
- Howarth, Ken (1998), *Oral History: A Handbook* (Thrupp, Stroud, Gloucestershire UK: Sutton).
- Howell, Martha and Walter Prevenier (2001), *From Reliable Sources: An Introduction to Historical Methods* (Ithaca and London: Cornell University Press).
- Hyman, Dick (1976), *The Trenton Pickle Ordinance and Other Boneheaded Legislation* (Brattleboro, VT: Stephen Greene Press).

- Jacobsen, Cary C. (1996), 'The 1994 House Elections in Perspective', *Political Science Quarterly* 111 (2): 203-223.
- Johnson, John R., Grant Williams and Richard Pazdur (2003), 'End Points and United States Food and Drug Administration Approval of Oncology Drugs', *Journal of Clinical Oncology* 21 (7) (April 1): 1404-1411.
- Johnson, John R. and Robert Temple (1984), 'Food and Drug Administration Requirements for Approval of New Anticancer Drugs', *Cancer Treatment Reports* 69 (10): 1155-1159.
- James, John S. (1991), '(ATN) ddI and ddC: The Call for Early Approval', *AIDS Treatment News* 112 (5 October 1991). Available at <http://www.aegis.com/pubs/atn/1990/ATN11201.html> (accessed 20 March 2006).
- Jasanoff, Sheila (1985), 'Peer Review in the Regulatory Process', *Science, Technology, & Human Values* 10 (3): 20-32.
- (1990), *The Fifth Branch: Science Advisors as Policymakers* (Cambridge, MA and London: Harvard University Press).
- (1995), *Science at the Bar: Law, Science, and Technology in America* (Cambridge, MA and London: Harvard University Press).
- (1998), 'The Eye of Everyman: Witnessing DNA in the Simpson Trial', *Social Studies of Science* 28 (5): 713-740.
- (2005), *Designs on Nature: Science and Democracy in Europe and the United States* (Princeton and Oxford: Princeton University Press).
- Jones, David S. (2000), 'Visions of a Cure: Visualization, Clinical Trials, and Controversies in Cardiac Therapies, 1968-1998', *Isis* 91 (3): 504-541.
- Josephson, Matthew (1995 [1934]), *The Robber Barons: The Great American Capitalists, 1861-1901* (San Diego: Harcourt).
- Kamensky, John (1999), *National Partnership for Reinventing Government (formerly the National Performance Review), A Brief History* (January 1999) (Online report, <http://govinfo.library.unt.edu/npr/whoweare/history2.html>).
- Keating, Peter and Alberto Cambrosio (2002), 'From Screening to Clinical Research: The Cure of Leukemia and the Early Development of the Cooperative Oncology Groups, 1955-1966', *Bulletin of the History of Medicine* 76: 299-334.
- (2007), 'Cancer Clinical Trials: The Emergence and Development of a New Style of Practice', *Bulletin of the History of Medicine* 81: 197-223
- Kelman, Mark (1988), 'On Democracy-Bashing: A Skeptical Look at the Theoretical and "Empirical" Practice of the Public Choice Movement', *Virginia Law Review* 74 (2)(March): 199-273.
- Kelman, Steven (1987), '“Public Choice” and Public Spirit', *Public Interest* 87 (Spring): 80 – 94.
- Klinkner, Philip, A., ed. (1996), *Midterm: The Elections of 1994 in Context* (Boulder, CO: Westview Press).

- Kripke, Saul (1982), *Wittgenstein on Rules and Private Language* (Oxford: Blackwell).
- Kurtz, Howard (1981), 'Warning Pledged on Cancer Drug Effects', *Washington Post* (7 November 1981), A11.
- Kusch, Martin (1999), *Psychological Knowledge: A Social History and Philosophy* (London and New York: Routledge).
- (2002), *Knowledge by Agreement: The Programme of Communitarian Epistemology* (Oxford: Oxford University Press).
- (2004), 'Rule-Skepticism and the Sociology of Scientific Knowledge: The Bloor-Lynch Debate Revisited', *Social Studies of Science* 34 (4): 571-591.
- Lasagna, Louis (1989), 'Congress, the FDA, and New Drug Development: Before and After 1962', *Perspectives in Biology and Medicine* 32 (3): 322 – 343.
- Latour, Bruno (1987), *Science in Action: How to Follow Scientists and Engineers Through Society* (Milton Keynes: Open University Press).
- (2005), *Reassembling the Social: An Introduction to Actor-Network-Theory* (Oxford: Oxford University Press).
- Laudan, Larry (1977), *Progress and Its Problems* (London: Routledge and Kegan Paul).
- Levine, Michael E. and Jennifer L. Forrence (1990), 'Regulatory Capture, Public Interest, and the Public Agenda: Toward a Synthesis', *Journal of Law, Economics and Organization* 6: 167 – 198.
- Lichtenberg Frank R. and Joel Waldfogel (2003), 'Does Misery Love Company? Evidence from Pharmaceutical Markets Before and After the Orphan Drug Act', National Bureau of Economic Research, Working Paper 9750 (June 2003) (Cambridge, MA: National Bureau of Economic Research). Available at <http://www.nber.org/papers/w9750> (accessed 13 September 2006).
- Lilienfeld, Abraham, M. (1982), 'The Fielding H. Garrison Lecture: *Ceteris Paribus*: The Evolution of the Clinical Trial', *Bulletin of the History of Medicine*, 56(1): 1-18.
- Lipton, J. and Hershaft, A. (1985), 'On the Widespread Acceptance of Dubious Medical Findings', *Journal of Health and Social Behavior* 26 (4): pp. 336-351.
- Löwy, I. (1996), *Between Bench and Bedside: Science, Healing, and Interleukin-2 in a Cancer Ward* (Cambridge, MA and London: Harvard University Press).
- Lynch, Michael (1992), 'Extending Wittgenstein: The Pivotal Move from Epistemology to the Sociology of Science' in *Science as Practice and Culture*, ed. Andrew Pickering (Chicago: University of Chicago Press): 215-65.
- MacKenzie, Donald (2006), *Like an Engine Not a Camera: How Models Shape Financial Markets* (Cambridge, MA and London: MIT Press).
- Marchione, Marilyn (2007), 'AP Analysis: How a Drug's Risks Emerge', <http://abcnews.go.com/Health/wireStory?id=3205382> (accessed 1 Feb 2008).

- Markle, Gerald. and Peterson, J. (eds.) (1980), *Politics, Science, and Cancer: The Laetrile Phenomenon* (American Association for the Advancement of Science Selected Symposium 46. Boulder, CO: Westview Press).
- Marks, Harry (1988), 'Notes from the Underground: The Social Organization of Therapeutic Research' in *Grand Rounds: One Hundred Years of Internal Medicine*, eds. Russell C. Maulitz and Diana E. Long (University of Pennsylvania Press, Philadelphia, PA), pp. 297-336.
- (1997), *The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900-1990* (Cambridge: Cambridge University Press).
- McCrea, Frances B. and Gerald E. Markle (1984), 'The Estrogen Replacement Controversy in the USA and UK: Different Answers to the Same Question?', *Social Studies of Science* 14 (1): 1-26.
- McDowell, W.H. (2002), *Historical Research: A Guide* (London, New York: Longman).
- Meldrum, M.L. (2000), 'A Brief History of the Controlled Trial: From Oranges and Lemons to the Gold Standard', *Hematology/Oncology Clinics of North America* 14 (4) (August 2000): 745-760.
- (1994), '*Departures from the Design: The Randomized Clinical Trial in Historical Context, 1946-1970*'. Dissertation (PhD), State University of New York at Stony Brook.
- Merrill, Richard A. (1994), 'Regulation of Drugs and Devices: An Evolution', *Health Affairs* (Summer 1994): 48-69.
- Messner, Donna A. (2006), 'Rule-Making under Fire: Accelerated Drug Approval as a Case Study of Crisis-Mode Regulatory Change', 2006 Conference of the International Society for Policy and Social Studies of Science and Technology, Washington D.C., 22-23 April 2006.
- Mintz, Morton (1962), '“Heroine” of FDA Keeps Bad Drug Off of Market', *Washington Post*, 15 July, A1+.
- (1965), *By Prescription Only* (Boston: Houghton Mifflin).
- Mitchell, William C. and Michael C. Munger (1991), 'Economic Models of Interest Groups: An Introductory Survey', *American Journal of Political Science* 35 (2): 512-46.
- Morrison, Alan B. (1986), 'OMB Interference with Agency rulemaking: The Wrong Way to Write a Regulation', *Harvard Law Review* 99(5) (March 1986): 1059-1074.
- Mulkay, M. (1980), 'Interpretation and the Use of Rules: The Case of the Norms of Science', *Transactions of the New York Academy of Sciences* 39: 111-125.
- (1991), *Sociology of Science* (London: George Allen and Unwin).
- Mullins, Nicholas C. (1981), 'Power, Social Structure, and Advice in American Science: The United States National Advisory System, 1950-1972', *Science, Technology, & Human Values* 7 (37): 4-19.
- Nesbit, Jeff (1995), 'Republicans Poised to Cut Government Down to Their Size', *The Washington Times*, Jan 6, B7.

- Nightingale, Stuart L. (1981), 'Drug Regulation and Policy Formulation', *The Milbank Memorial Fund Quarterly. Health and Society* 59(3): 412-444.
- Nissen, Steven E. and Kathy Wolski (2007), 'Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes', *New England Journal of Medicine* 356(24)(14 June 2007): 2457-2471.
- Olson, Mancur (1971 [1965]), *The Logic of Collective Action* (Cambridge, MA and London: Harvard University Press).
- Orlando, Vivian I. (1999), 'The FDA's Accelerated Approval Process: Does the Pharmaceutical Industry Have Adequate Incentives for Self-Regulation?', *American Journal of Law & Medicine* 25: 543-68.
- Oudshoorn, Nelly (1993), 'United We Stand: The Pharmaceutical Industry, Laboratory, and Clinic in the Development of Sex Hormones into Scientific Drugs, 1920-1940', *Science, Technology, & Human Values* 18 (1): 5-24.
- (1999), 'On Masculinities, Technologies, and Pain: The Testing of Male Contraceptives in the Clinic and the Media', *Science, Technology, & Human Values* 24 (2): 265-289.
- Peltzman, Sam (1973), 'An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments', *Journal of Political Economy* 81 (September-October): 1046-91.
- (1976), 'Toward a More General Theory of Regulation', *Journal of Law and Economics* 19 (2): 211-240.
- (1998), *Political Participation and Government Regulation* (Chicago and London: University of Chicago Press).
- Pickering, Andrew (1981), 'Constraints on Controversy: The Case of the Magnetic Monopole', *Social Studies of Science* 11 (1): 63-94.
- (1999), 'The Mangle of Practice: Agency and Emergence in the Sociology of Science' in *The Science Studies Reader*, ed. Mario Biagioli (New York and London: Routledge), 372-393.
- Pinch, Trevor and Wiebe E. Bijker (1987), 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other' in *The Social Construction of Technological Systems*, ed. Wiebe E. Bijker, Thomas P. Hughes and Trevor Pinch (Cambridge, MA and London: MIT Press), 17-50.
- Posner, Richard A. (1974), 'Theories of Economic Regulation', *The Bell Journal of Economics and Management Science* 5(2)(Autumn, 1974): 335-358.
- President (1971) [Nixon], 'The Budget Message. The President's Message to the Congress Transmitting the Budget for Fiscal Year 1972. January 29, 1971,' *Weekly Compilation of Presidential Documents* 7(5) (1 Feb 1971), 130 – 139.
- President (1978a) [Carter], 'State of the Union: Annual Message to the Congress. January 19, 1978', *Weekly Compilation of Presidential Documents* 14 (3) (19 January 1978) : 98-126.
- President (1978b) [Carter], 'Executive Order 12044: Improving Government Regulations',

- Federal Register* 43/58 (24 March 1978): 12661-12670.
- President (1981) [Reagan], 'Executive Order 12291: Federal Regulation', *Weekly Compilation of Presidential Documents* 17(8) (17 February 1981): 124- 130. [Also in *Federal Register* 46/33 (17 February 1981): 13193-13198.]
- President (1993a) [Clinton], 'Remarks Announcing the Initiative to Streamline Government', *Weekly Compilation of Presidential Documents*, 29 (3) (3 March 1993): 350-352.
- President (1993b) [Clinton], 'Executive Order 12861: Elimination of One-Half of Executive Branch Internal Regulations', *Weekly Compilation of Presidential Documents* 29 (37) (11 September 1993): 1737.
- President (1993c) [Clinton], 'Executive Order 12866: Regulatory Planning and Review', *Weekly Compilation of Presidential Documents* 29 (39) (30 September 1993): 1925-1933.
- President (1995) [Clinton], 'Remarks on Regulatory Reform in Arlington, Virginia', *Weekly Compilation of Presidential Documents* 31 (10) (March 16, 1995), 426-31.
- President (1996) [Clinton], 'Remarks on the Anticancer Initiative, March 29, 1996', *Compilation of Presidential Documents*, Administration of William J. Clinton/March 1996, Book I, Published papers, 523-24.
- Randall, Blanchard (1996), *Drug Regulation: Historical Overview and Current Reform Proposals*, Congressional Research Service Report to Congress (Updated 17 January 1996)(Washington, DC: Congressional Research Service, Library of Congress).
- Ravetz, Jerome (1979 [1971]), *Scientific Knowledge and its Social Problems* (Oxford and New York: Oxford University Press).
- Reich, Robert B. (1985), 'Public Administration and Public Deliberation: An Interpretive Essay', *The Yale Law Journal* 94 (7)(Jun 1985): 1617-1641.
- Richards, Evelleen (1988), 'The Politics of Therapeutic Evaluation: The Vitamin C and Cancer Controversy', *Social Studies of Science* 18 (4): 653-701.
- (1991), *Vitamin C and Cancer: Medicine or Politics?* (New York: St. Martin).
- Ritchie, Donald A. (2003), *Doing Oral History: A Practical Guide 2nd Ed.* (Oxford and New York: Oxford University Press).
- Rokakis, Jim (2007), 'The Shadow of Debt. Slavic Village Is Fast Becoming a Ghost Town. It's Not Alone', *Washington Post* (30 September 2007), B01.
- Rothman, David J. (1991), *Strangers at the Bedside: A History of How Law and Bioethics Transformed Medical Decision Making* (New York: Basic Books).
- Rothman, David J. and Harold Edgar (1992), 'Scientific Rigor and Medical Realities', in *AIDS: The Making of A Chronic Disease*, eds. Elizabeth Fee and Daniel M. Fox (Berkeley, Los Angeles, Oxford: University of California Press): 194-206.
- Rubin, H.J (1995), *Qualitative Interviewing: The Art of Hearing Data* (London: Sage).

- Siegel, Joanna E. and Mark J. Roberts (1991), 'Reforming FDA Policy: Lessons from the AIDS Experience', *Regulation* 14(4) (Fall 1991), 71+. Available at <http://www.cato.org/pubs/regulation/regv14n4/reg14n4-siegel.html> (accessed 13 June 2006).
- Silverman, David (1993), *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction* (London, Thousand Oaks, CA, and New Delhi: Sage).
- Silverman, William A. (1985), *Human Experimentation: A Guided Step into the Unknown* (Oxford, New York, and Tokyo: Oxford University Press).
- Sjöström, H. and Nilsson, R. (1972), *Thalidomide and the Power of the Drug Companies* (Harmondsworth, Middlesex, England: Penguin).
- Skrzycki, Cindy (1996), 'After the Election, It's Oversight Rather Than Overthrow', *Washington Post* (8 November 1996): C11.
- Smith, Paul (1997), *Millennial Dreams: Contemporary Culture and Capital in the North* (London and New York: Verso).
- Solomon, S. and Hackett, E. (1996), 'Setting Boundaries between Science and Law: Lessons from Daubert v. Merrell Dow Pharmaceuticals, Inc.', *Science, Technology, & Human Values* 21 (2): 131-156.
- Spitzer, Eliot (2008), 'Predatory Lenders' Partner in Crime: How the Bush Administration Stopped the States From Stepping In to Help Consumers', *Washington Post* (14 February 2008): A25.
- Spruance, Spotswood L., Julia E. Reed, Michael Grace and Matthew Samore (2004), 'Hazard Ratio in Clinical Trials', *Antimicrobial Agents and Chemotherapy* 48 (8): 2787-2792.
- Stein, Rob (2007), 'FDA Delay in Cancer Therapy is Attacked', *Washington Post* (6 July 2007), A01.
- Sternberg, Steve (2007), 'Diabetes Drug Called Heart Death Risk', *USA Today* (22 May 2007), 1A.
- Stewart, Richard B. (1975), 'Reformation of American Administrative Law', *Harvard Law Review* 88 (1667 – 1813).
- Stigler, George J. (1971), 'The Theory of Economic Regulation', *Bell Journal of Economics and Management Science* 2 (1): 3-21.
- (1974), 'Free Riders and Collective Action: An Appendix to Theories of Economic Regulation', *Bell Journal of Economics and Management Science* 5(2): 335-365.
- Stone, Brad (1987), 'Fast-Tracking the First AIDS Drug – "Retrovir"', Interview with Ellen Cooper and James Bilstad, *FDA Consumer* (October 1987). Available online: http://findarticles.com/p/articles/mi_m1370/is_v21/ai_5273117 (accessed 26 June 2006).
- Stonecash, Jeffrey M. and Mack D. Mariani (1998), 'Republican Gains in the House in the 1994 Elections: Class Polarization in American Politics', *Political Science Quarterly* 115 (1): 93-

- Strauss, Anselm, Shizuko Fagerhaugh, Barbara Suczek and Carolyn Wiener (1985), *The Social Organization of Medical Work* (Chicago: University of Chicago Press).
- Streiner, David L. (2007), 'Alternatives to Placebo-Controlled Trials', *Canadian Journal of Neurological Sciences* 34 (Supp 1): S37-41.
- Stygall, Gail (2001), 'A Different Class of Witnesses: Experts in the Courtroom', *Discourse Studies* 3 (3): 327-349.
- Sunstein, Cass R. (1990), 'Political Economy and Administrative Law: A Comment', *Journal of Law, Economics, and Organization* 6: 299-306.
- Swain, Sandra M., Fredrick S. Whaley and Michael S. Ewer (2003), 'Congestive Heart Failure in Patients Treated with Doxorubicin, A Restrospective Analysis of Three Trials', *Cancer* 97 (11): 2869-2879.
- Temin, Peter (1979), 'Technology, Regulation, and Market Structure in the Modern Pharmaceutical Industry', *The Bell Journal of Economics* 10 (2): 429-446.
- (1980), *Taking Your Medicine: Drug Regulation in the United States* (Cambridge, MA and London: Harvard University Press).
- (1985), 'Government Actions in Times of Crisis: Lessons from the History of Drug Regulation', *Journal of Social History* 18 (3) (Spring 1985): 433-8.
- Temple, Robert (2001), 'Experiences with Single Trials Leading to Section 115 of FDAMA and May 1998 Evidence Document', U.S. Food and Drug Administration, Center for Drug Evaluation and Research. Slide Presentation, 23 January [date assumed based on path name of file]. Available at: <http://www.fda.gov/cder/present/dia1-23-2001/temple-dia/> (accessed 5 May 2004).
- (2005), Interview by Author (tape recording, Rockville, MD), 31 May 2005.
- Thomas, Lacy G. (1990), 'Regulation and Firm Size: FDA Impacts on Innovation', *The RAND Journal of Economics* 21 (4): 497-517.
- Timmermans, Stefan and Allison Angell (2001), 'Evidence-Based Medicine, Clinical Uncertainty, and Learning to Doctor', *Journal of Health and Social Behavior* 42 (4): 342-359.
- Timmermans, Stefan and Marc Berg (2003), *The Gold Standard: The Challenge of Evidence-Based Medicine and Standardization in Health Care* (Philadelphia: Temple University Press).
- U.S. House (1964), Hearings Before the Subcommittee of the Committee on Government Operations, *Drug Safety* (24, 25 March, 8 April, and 3 June 1964) (Washington DC: U.S. Government Printing Office).
- (1987), Hearing before the Human Resources and Intergovernmental Relations Subcommittee of the Committee on Government Operations, *FDA Proposals to Ease Restrictions on the Use and Sale of Experimental Drugs* (29 April 1987) (Washington, DC: U.S. Government Printing Office).

- (1992), Hearing before the Human Resources and Intergovernmental Relations Subcommittee of the Committee on Government Operations, *Council on Competitiveness and FDA Plans to Alter the Drug Approval Process at FDA*. (19 March 1992) (Washington, DC: U.S. Government Printing Office).
- (1993a), *Less than the Sum of Its Parts: Reforms Needed in the Organization, Management and Resources of the Food and Drug Administration's Center for Devices and Radiological Health*, Report by the Subcommittee on Oversight and Investigations of the Committee on Energy and Commerce (May 1993), Committee Print 103-N (Washington, DC: U.S. Government Printing Office).
- (1993b), Hearing before the Committee on Science, Space and Technology, Subcommittee on Investigations and Oversight, *Regulating Innovation: Regulating FDA's Device Approval Process* (17 September 1993), No. 69 (Washington, DC: U.S. Government Printing Office).
- (1994a), Hearing before the Committee on Small Business, Subcommittee on Regulation, Business Opportunities, and Technology, *RU-486: Status of the U.S. Commercialization Project, Transfer of Anti-Progestin Technology to the United States* (16 May 1994), Serial No. 103-80 (Washington, DC: U.S. Government Printing Office).
- (1994b), Hearing before the Committee on Small Business, Subcommittee on Regulation, Business Opportunities, and Technology, *Problems of Securing Informed Consent of Subjects in Experimental Trials of Unapproved Drugs and Devices* (23 May 1994), Serial No. 103-85 (Washington, DC: U.S. Government Printing Office).
- (1994c), Hearing before the Committee on Energy and Commerce, Subcommittee on Health and the Environment, *Orphan Drug Reauthorization* (discussion of H.R. 4160)(16 June 1994), Serial No. 103-113 (Washington, DC: U.S. Government Printing Office).
- (1994d), Hearing before the Committee on Energy and Commerce, Subcommittee on Health and the Environment, *Medical Device User Fees* (discussion of H.R. 4728)(14 July 1994), Serial No. 103-139 (Washington, DC: U.S. Government Printing Office).
- (1994e), Hearing before the Committee on Science, Space, and Technology, *Gene Therapy: Status, Prospects for the Future, and Government Policy Implications* (28 September 1994), No. 168 (Washington, DC: U.S. Government Printing Office).
- (1994f), Hearing before the Committee on Small Business, Subcommittee on Regulation, Business Opportunities, and Technology, *Questionable Sales Practices in the Drug Industry* (12 October 1994), Serial No. 103-107 (Washington, DC: U.S. Government Printing Office).
- (1994g), *Medical Device User Fee Act of 1994* (26 September 1994), House Report 103-751 (Washington, DC: U.S. Government Printing Office).
- (1995a), Joint Hearings before the Committee on Health and Environment and the Subcommittee on Commerce, Trade, and Hazardous Materials, *Risk Assessment and Cost/Benefit Analysis for New Regulations* (1 and 2 February 1995), Serial No. 104-3 (Washington, DC: U.S. Government Printing Office).

- (1995b), Hearing before the Committee on Energy and Commerce, Subcommittee on Oversight and Investigations, *A Consumer's Perspective on Medical Devices* (30 March 1995), Serial No. 104-18, (Washington, DC: U.S. Government Printing Office).
- (1995c), Hearing before the Committee on Energy and Commerce, Subcommittee on Oversight and Investigations, *Drugs and Biologics* (25 May 1995; 19 June 1995), Serial No. 104-23 (Washington, DC: U.S. Government Printing Office).
- (1998), Hearing before the Committee on Energy and Commerce, *Implementation of the FDA Modernization Act of 1997* (7 October 1998). Serial No. 105-136. (Washington, DC: U.S. Government Printing Office).
- (2002), Hearing before the Subcommittee on Oversight and Investigations of the Committee on Energy and Commerce, *An Inquiry into the ImClone Cancer Drug Story* (13 June and 10 October 2002). Serial No. 107-142. (Washington, DC: U.S. Government Printing Office).
- (2007), Hearing of the House Committee on Oversight and Government Reform, 'FDA's Role in Evaluating the Safety of Avandia' (6 June 2007). (Washington, DC: U.S. Government Printing Office).
- U.S. Senate (1974), Joint Hearings before the Subcommittee on Health of the Committee on Labor and Public Welfare and the Subcommittee on Administrative Practice and Procedure of the Committee on the Judiciary, *Regulation of New Drug R&D by the Food and Drug Administration, 1974* (25 and 27 September 1974) (Washington DC: U.S. Government Printing Office).
- (1977), Hearings before the Subcommittee on Health and Scientific Research of the Committee on Human Resources, *Food, Drug and Cosmetic Act Amendments of 1977* [S.1831] (26 and 27 July 1977) (Washington DC: U.S. Government Printing Office).
- (1993), Special Hearing before a Subcommittee of the Appropriations Committee, *Alternative Medicine* (24 June 1993), Senate Hearing 103-146 (Washington DC: U.S. Government Printing Office).
- (1994a), Hearing before the Committee on Veterans' Affairs, *Is Military Research Hazardous to Veterans' Health? Lessons from World War II, the Persian Gulf, and Today* (6 May 1994), Senate Hearing 103-984 (Washington, DC: U.S. Government Printing Office).
- (1994b), Hearing before the Committee on Energy and Natural Resources, *Boron Neutron Capture Therapy* (3 May 1994), Senate Hearing 103-664 (Washington DC: U.S. Government Printing Office).
- (1994c), Hearing before the Committee on Labor and Human Resources, *Access to Medical Treatment Act* (on S.2140) (22 July 1994), Senate Hearing 103-707 (Washington DC: U.S. Government Printing Office).
- (1994d), Hearing before the Committee on Governmental Affairs, *Drug Pricing: Poor Prescription for Consumers and Taxpayers?* (27 July 1994), Senate Hearing 103-1045 (Washington, DC: U.S. Government Printing Office).
- (1995), Hearing before the Committee on Labor and Human Resources, *The FDA and the Future of American Biomedical and Food Industries* (5 and 6 April 1995), Senate Hearing

- 104-45 (Washington DC: U.S. Government Printing Office).
- (1996), *Food and Drug Administration Performance and Accountability Act of 1995* (20 June 1996), Senate Report 104-284.
- (1997), *Food and Drug Administration Performance and Accountability Act of 1997* (1 July 1997), Senate Report 105-43.
- Vogt, Donna U. (1995), *Food and Drug Administration: Current Reform Initiatives*, Congressional Research Service Report to Congress (8 December 1995)(Washington, DC: Congressional Research Service, Library of Congress).
- Wainwright, P., Clare Williams, Mike Michael, Bobbie Farsides, and Alan Cribb (2006), 'Ethical Boundary-Work in the Embryonic Stem Cell Laboratory', *Sociology of Health & Illness* 28 (6): 732–748.
- Waldmeir, Patti (1996), 'FDA Reprived but Not Left Alone: Patti Waldmeir Reports on Post-Election Reform of the U.S. Food and Drug Administration'. *Financial Times* (London) (23 December): 4.
- Wardell, William M. (1973), 'Introduction of New Therapeutic Drugs in the United States and Great Britain: An International Comparison', *Clinical Pharmacy and Therapeutics* 14 (September-October): 773-790.
- (ed.)(1978), *Controlling the Use of Therapeutic Drugs: An International Comparison* (Washington, D.C.: American Enterprise Institute for Public Policy Research).
- Yarchoan, Robert (1998), Interview by Dr. Victoria Harden (Director, NIH Historical Office) and Dr. Caroline Hannaway (NIH Historical Contractor), 30 April 1998, Bethesda, MD Available at <http://aidshistory.nih.gov/transcripts/transcripts/Yarchoan98.pdf> (accessed 3 October 2006).
- Yin, Robert K. (2003), *Case Study Research: Design and Methods*, 3rd ed. (Applied Social Research Methods Series, Vol. 5, London, Thousand Oaks, CA, and New Delhi: Sage).
- Yoshioka, Alan (1998), Streptomycin, 1946: British Central Administration of Supplies of a New Drug of American Origin with Special Reference to Clinical Trials in Tuberculosis (PhD Thesis, Imperial College of Science, Technology, and Medicine).