NBER WORKING PAPER SERIES

ADAPTIVE MINIMAX-REGRET TREATMENT CHOICE, WITH APPLICATION TO DRUG APPROVAL

Charles F. Manski

Working Paper 13312
http://www.nber.org/papers/w13312

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2007

This research was supported in part by National Science Foundation grant SES-0549544. I have benefitted from the comments of David Meltzer, John Pepper, Bruce Spencer, and Alex Tetenov. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Adaptive Minimax-Regret Treatment Choice, With Application To Drug Approval
Charles F. Manski
NBER Working Paper No. 13312
August 2007
JEL No. D63,I18

ABSTRACT

Suppose that there are two treatments for a condition. One is the status quo, whose properties are known from experience and the other is an innovation, whose properties are not known initially. A new cohort of persons presents itself each period and a planner must choose how to treat this cohort. When facing situations of this kind, it has become common to commission randomized trials of limited duration to learn about the innovation. Rather than wait for the outcomes of interest to unfold over time, surrogate outcomes that can be observed early on are used to judge the success of the innovation. A close approximation to this process is institutionalized in the drug approval protocol of the U. S. Food and Drug Administration. This paper brings welfare-economic and decision-theoretic thinking to bear on the problem of treatment choice, with application to drug approval. I introduce the adaptive minimax-regret (AMR) rule, which applies to each cohort the minimax-regret criterion using the knowledge of treatment response available at the time of treatment. The result is a fractional treatment allocation whenever the available knowledge does not suffice to determine which treatment is better. The rule is adaptive because, as knowledge of treatment response accumulates, successive cohorts are allocated differently across the two treatments. I use the AMR idea to suggest an adaptive drug approval process that permits partial marketing of new drugs while scientifically appropriate long-term clinical trials are underway. The stronger the evidence on health outcomes of interest, the more treatment would be permitted, with a definitive approval decision eventually made when sufficient evidence has accumulated.

Charles F. Manski
Department of Economics
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
cfmanski@northwestern.edu

1. Introduction

Suppose that there are two treatments for a condition. One is the *status quo*, whose properties are known from historical experience. The other is an *innovation*, whose properties are not known initially. A new cohort of persons presents itself each period and a planner must choose how to treat this cohort. The outcome of interest unfolds over multiple periods. For example, the treatments may be alternative cancer therapies and the outcome may be life span. Or the treatments may be alternative therapies for a chronic disease and the outcome of interest may be quality-adjusted life years.

When facing situations of this kind, it has become common to commission randomized trials of limited duration to learn about the innovation. The experimental sample receiving the innovation is typically a small fraction of the population, the size of this sample being determined by conventional calculations of statistical power. Rather than wait for the health outcomes of real interest to unfold over time, surrogate outcomes that can be observed early on are used to judge the success of the innovation. A statistical hypothesis test is used to make this judgment, the null hypothesis being that the innovation is no better than the status quo treatment and the alternative being that it is better. If the null hypothesis is not rejected, the status quo treatment continues in force and no one subsequently receives the innovation. If the null is rejected, the innovation replaces the status quo as the treatment of choice. A close approximation to this process is institutionalized in the drug approval protocol of the U. S. Food and Drug Administration (FDA). Other approximations are institutionalized in the decision processes of public and private health insurance entities about whether to cover the cost of new treatments.

The present FDA drug approval process is susceptible to two types of errors with long-term consequences. Type I errors occur when new drugs that actually are worse than status quo treatments in terms of health outcomes of interest are approved because they appear superior when evaluated using surrogate outcomes. Type II errors occur when new drugs that actually are better than status quo treatments are disapproved because they appear inferior when evaluated using surrogate outcomes. These are not finite-

sample statistical errors whose probabilities of occurrence can be reduced by recruiting larger samples of subjects into randomized trials. They are basic data errors that can be repaired only be improving measurement of the outcomes of interest, which requires longer trials than those performed at present.

Public health researchers have often called attention to the difficulty of extrapolating from surrogate outcomes to health outcomes of interest. Fleming and Demets (1996), who review the prevalent use of surrogate outcomes in FDA-required Phase 3 trials evaluating drug treatments for heart disease, cancer, HIV/AIDS, osteoporosis, and other diseases, write (p. 605):

> "Surrogate end points are rarely, if ever, adequate substitutes for the definitive clinical outcome in phase 3 trials."

Sculpher and Claxton (2005), who consider decisions about whether new pharmaceuticals are sufficiently cost-effective for reimbursement in collectively funded health-care systems, write (p. 441):

> "Arguably the biggest challenge that reimbursement agencies have to face in terms of the uncertainty surrounding existing evidence relates to costs and outcomes which have not been observed directly in trials. There are two frequent manifestations of this: linking intermediate outcomes to ultimate measures of health gain, and extrapolating costs and benefits over a longer-term time horizon."

From a scientific perspective, the obvious solution is to perform clinical trials of sufficient length to measure the health outcomes of real interest. However, this has been thought politically infeasible. Pasty *et al.* (1999) write (p. 789):

> "One systematic approach is a requirement that, prior to their approval, new drug therapies for cardiovascular risk factors should be evaluated in large, long-term clinical trials to assess their effects on major disease end points. The use of surrogate outcomes is avoided, and the major health outcomes are known prior to marketing. Such an approach would slow the time to drug approval and may meet with resistance from pharmaceutical manufacturers."

Indeed, pharmaceutical firms eager for returns on investments and patient groups wanting access to new drugs have often advocated shortening rather than lengthening the present time to approval.

This paper brings welfare-economic and decision-theoretic thinking to bear on the drug approval problem. Doing so suggests replacement of the present binary (up or down) approval process with an adaptive process that permits partial marketing and insurance coverage of new drugs while scientifically appropriate long-term clinical trials are underway. In the new approval process, the prevalence of treatment with a new drug would vary smoothly as empirical evidence accumulates. The stronger the evidence on health outcomes of real interest, the more treatment would be permitted. Eventually, a definitive decision would be made when sufficient evidence has accumulated on the safety and effectiveness of the new drug.

The adaptive process emerges from consideration of treatment choice from the minimax-regret perspective. The minimax-regret criterion, first suggested by Savage (1951), is a general principle for decision making with partial knowledge of relevant outcomes. Suppose that a decision maker must choose from a set of alternatives, and that he wants to make a choice that maximizes welfare. However, he does not fully know the outcomes that would be produced by each alternative. How should he behave?

The familiar Bayesian answer is that the decision maker should assert a subjective probability distribution over the unknown outcomes and maximize expected welfare. See Meltzer (2001) for applications to medical decision making. However, a subjective probability distribution is itself a form of knowledge, and the decision maker may feel that he has no good basis for asserting one. How then might he behave? The minimax-regret rule chooses an alternative that minimizes the maximum loss to welfare that results from not having complete knowledge.

Specifically, the *regret* of an alternative at a given value for the unknown outcomes is defined to be the difference between the maximum welfare that would be achievable given complete knowledge and the welfare that is achieved by this alternative. If one has complete knowledge, the best course of action obviously is to choose an alternative that minimizes regret, setting it equal to zero. In the absence of

complete knowledge, the minimax-regret rule chooses an alternative that minimizes maximum regret across all possible values of the unknown outcomes.

In previous work, I have studied minimax-regret treatment choice by a social planner who must choose treatments for the members of a population and who has only partial knowledge of treatment response; see Manski (2004, 2005, 2007a, 2007b). A general finding is that when there are two treatments and the available knowledge of outcomes does not suffice to determine which treatment is better, the minimax-regret rule does not assign all observationally identical persons to the same treatment. Instead, it fractionally allocates these persons across the two treatments, with the fraction receiving each treatment determined by the available knowledge. In this manner, use of the minimax-regret criterion enables a planner to socially diversify risks that are privately indivisible.

Choosing Treatments for X-Pox: A dramatic illustration of social diversification occurs in a hypothetical problem of treatment choice considered in Manski (2007b, Section 11.7). Suppose that a new viral disease called x-pox is sweeping the world. Researchers have proposed two mutually exclusive treatments, say a and b, which reflect alternative hypotheses, say H_a and H_b, about the nature of the virus. If H_a (H_b) is correct, all persons who receive treatment a (b) survive and all others die. A planner knows that one of the two hypotheses is correct, but does not know which one. The objective is to maximize the survival rate of the population.

In this setting, the risk of death is privately indivisible. An individual receives either treatment a or b, and this person either lives or dies. Yet society can diversify by having positive fractions of the population receive each treatment. Consider the rule in which a fraction $\delta \in [0, 1]$ of the population receives treatment b and the remaining $1 - \delta$ receives treatment a. Then the fraction who survive is either δ or $1 - \delta$. A planner who uses the minimax-regret criterion would set $\delta = 0.5$, implying that half of the population survives and half dies. □

Social diversification is a central qualitative feature of minimax-regret treatment choice. It is not a general feature of Bayesian decision making. For example, in the x-pox illustration, a Bayesian planner allocates the entire population to the treatment with the higher subjective probability of success.

The difference between my earlier work and the present paper is that the earlier work considered one-period planning problems where, as in the x-pox illustration, there is no opportunity for learning. Drug approval is a multiple-period planning problem, where performance of randomized trials with the innovation creates an opportunity for social learning. Indeed, randomized trials themselves implement fractional treatment rules, with random samples of the population assigned to alternative treatments. The objective is to learn the distribution of treatment response, in order to make better treatment decisions in the future.

I show here that an *adaptive minimax-regret* treatment rule achieves both social learning and diversification. This rule applies to each cohort the minimax-regret criterion using the knowledge of treatment response available at the time of treatment. The result is a fractional treatment allocation whenever the available knowledge does not suffice to determine which treatment is better. The rule is adaptive because knowledge of treatment response accumulates over time, so successive cohorts may receive different fractional allocations. Eventually, the planner may learn which treatment is better. From this point on, he assigns new cohorts entirely to the better treatment.

Section 2 formalizes the treatment choice problem, introduces the adaptive minimax-regret rule, and gives numerical illustrations. Section 3 proposes a drug approval process that incorporates important features of adaptive minimax-regret treatment choice. To keep the exposition simple, Sections 2 and 3 focus on the problem created by use of surrogate outcomes and abstract from other important issues that arise in the accumulation of knowledge for drug approval. The concluding Section 4 discusses various such issues.

While drug approval is a leading case in which surrogate outcomes have been used to evaluate innovations, the practice is widespread within medicine and in other realms. The ideas developed here should therefore have considerable application beyond drug approval.

Consider, for example, evaluation of educational interventions in early childhood. The outcomes of real interest may be years of schooling completed by adulthood and job performance in adulthood. Not wanting to wait for these outcomes to unfold over time, researchers have often used performance in the early grades of school to judge the success of innovations, with binary implementation decisions in mind. Here, as with drug approval, it may be better to institute an adaptive process in which the scale of implementation of an intervention varies as evidence accumulates.

2. Allocating a Sequence of Cohorts to a Status Quo Treatment and an Innovation

Section 2.1 sets out a formal problem of choice between a status quo treatment and an innovation. Section 2.2 introduces the adaptive minimax-regret rule. Section 2.3 gives numerical illustrations based on hypothetical treatment-choice problems.

To keep attention focused on the problem of surrogate outcomes, I make several simplifying assumptions here and in Section 3. First, I assume that the members of the population are observationally identical. In practice, persons may have observable covariates, and a planner may be able to differentially treat persons with different covariates. In such cases, the present analysis can be applied separately to each subpopulation of observationally identical persons.

Second, I assume that the randomized trial on the innovation is performed on a large enough random sample of persons that finite-sample statistical error is a negligible concern. Sample size sometimes is a significant issue in the assessment of trials, so I will discuss the matter in Section 4. However, the fundamental present concern, the use of surrogate outcomes in treatment choice, is an entirely distinct issue.

Third, I assume that the trial is a classical randomized experiment. In practice, trials used in the drug approval process often depart substantially from the classical ideal. First, trials typically are performed on

convenience samples of volunteers rather than on random samples of the patient population of interest. Analyses of findings usually presume, often without justification, that the distribution of treatment response among the volunteers who participate in a trial is the same as in the relevant patient population. Second, some of the volunteers who participate in trials may not comply with their assigned treatments or may leave the trial early, before their outcomes can be measured. To cope with these problems, analyses often assume that noncompliance and attrition are random or, alternatively, they apply intention-to-treat analysis. Third, trials are typically performed with blinded treatment assignment, even though treatments are observed in ordinary clinical practice. These features of present-day randomized trials for drug approval are problematic, so I will discuss them in Section 4.

A final simplification, maintained throughout this paper, is that I consider choice between a specified status quo treatment and innovation in isolation, without reference to how these alternatives were generated. It would undoubtedly be interesting to envisage an innovation process, with new alternatives appearing from time to time. The nature of the innovation process might be affected by the prevailing treatment rule. If so, treatment choice and the innovation process should be considered jointly.

2.1. The Setting

I present here a multi-period extension of a one-period planning problem previously studied in Manski (2005, 2007b) and elsewhere. A planner must choose treatments in each of the time periods $n = 0, 1, \ldots, N$. In each period, the set of feasible treatments is $T = \{a, b\}$, where treatment a is the status quo and treatment b is the innovation. The difference between the status quo treatment and the innovation is that only the former was available historically; that is, for $n < 0$. The latter first becomes available at $n = 0$.

A new cohort appears each period and requires treatment. Each member j of cohort n, denoted J_n, has a response function $y_j(\cdot): T \to Y$ mapping treatments $t \in T$ into outcomes $y_j(t) \in Y$. Subscripting $y_j(\cdot)$ by

j shows that treatment response may vary across the cohort.

Let P[y(·)] denote the distribution of treatment response across the cohort. Observe that I have not indexed P by n. Thus, all cohorts share the same distribution of treatment response. This assumption enables social learning. The planner can use observations of treatment outcomes in early cohorts to inform treatment choice for later cohorts.

Observability of Outcomes

Whereas the members of cohort n receive their treatments in period n, their outcomes unfold over the subsequent K periods. Although this is not necessary to the analysis, I assume for simplicity that $y_j(t)$ has the time-additive form

$$(1) \qquad y_j(t) = \sum_{k=1}^{K} y_{jk}(t),$$

where $y_{jk}(t)$ is the component of the outcome that is realized k periods after person j receives treatment.

To illustrate, consider the cancer treatment example given in the introduction. Here treatment t is a therapy and the outcome of interest is life span. Then a period may be a year, with K being a specified horizon of interest. Hence, $y_{jk}(t) = 1$ if person j would, in the event of receiving treatment t, be alive k years after treatment, and $y_{jk}(t) = 0$ otherwise.

The Treatment Choice Problem

The planner's problem is to allocate each cohort between the two treatments. A treatment rule is a vector $\delta \equiv (\delta_n, n = 0, \ldots, N)$ that randomly assigns a fraction δ_n of cohort n to treatment b and the remaining $1 - \delta_n$ to treatment a. The feasible treatment rules are the elements of the hyper-rectangle $[0, 1]^{(N+1)}$.

Let $u(t) \equiv u[y(t), t]$ denote the contribution to social welfare that occurs when a person receives

treatment t and realizes outcome y(t). I assume that the planner wants to choose a treatment rule that maximizes mean welfare summed across cohorts. Let $\alpha \equiv E[u(a)]$ and $\beta \equiv E[u(b)]$ be the mean welfare that would result if all members of a cohort were to receive treatment a or b respectively. The quantities α and β are not indexed by n because, by assumption, the distribution of treatment response is the same for all cohorts. The social welfare achieved by rule δ is

$$(2) \quad W(\delta) \equiv \sum_{n=0}^{N} \beta\delta_n + \alpha(1 - \delta_n) = \alpha + (\beta - \alpha) \sum_{n=0}^{N} \delta_n.$$

$W(\cdot)$ is an ordinary consequentialist social welfare function that aggregates individual contributions to welfare in an additive manner. A notable special case occurs when the function $u(\cdot)$ expresses private preferences. Then $W(\cdot)$ is the utilitarian social welfare function that weights all cohorts equally. A slightly broader definition of $W(\cdot)$ would permit the social welfare function to differentially weight cohorts that vary in size or to express time discounting. The present analysis extends easily to such cases.

The optimal treatment rule is obvious if (α, β) are known. The planner should choose $\delta_n = 1$ for all n if $\beta > \alpha$ and $\delta_n = 0$ if $\beta < \alpha$. All values of δ yield the same welfare if $\beta = \alpha$. The problem of interest is treatment choice when (α, β) is only partially known. In particular, I shall consider situations in which α is known but β is not.

It is often reasonable to suppose that α is known from historical experience. All members of cohorts $n < 0$ received the status quo treatment. Hence, a planner can learn α empirically if he is able to observe the outcomes experienced by cohort $-K$ or an earlier cohort.

The innovation having been introduced at period 0, empirical evidence cannot reveal the value of β before period K. When the planner treats a cohort $n < K$, he can only observe the components of the outcomes experienced to date by members of earlier cohorts who were assigned the innovation. Thus, at n

= 0, the planner has no empirical evidence. At n = 1 he can observe first-period outcomes for those members of cohort 0 who were assigned the innovation. At n = 2, he can observe second-period outcomes for members of cohort 0 who were assigned the innovation, as well as first-period outcomes for members of cohort 1. And so on.

2.2. The Adaptive Minimax-Regret Rule

The adaptive minimax-regret (AMR) rule extends the one-period minimax-regret (MR) rule to multi-period planning problems. I first review the one-period rule and then give the multi-period extension.

The One-Period Rule

Let N = 0. Then we have a one-period planning problem. If the planner assigns a fraction δ_0 of cohort 0 to the innovation and the remainder to the status quo, social welfare is

$$(3) \quad W(\delta_0) = \alpha + (\beta - \alpha)\delta_0.$$

The problem is to choose δ_0 in the absence of empirical evidence on β.

Application of the MR criterion requires only that the planner be able to place β within some bounded interval $[\beta_L, \beta_U]$. Such an interval always exists when the outcome is itself bounded. Then, if the planner knows nothing about the innovation, he can set β_L and β_U equal to the smallest and largest logically possible outcome values. Or $[\beta_L, \beta_U]$ may be a subset of the logically possible outcomes, excluding values the planner deems infeasible.

The MR rule is a function of α, β_L, and β_U. It assumes nothing about the position of β within the interval $[\beta_L, \beta_U]$. This contrasts with Bayesian planning, which requires assertion of a subjective probability

distribution on the interval of feasible values.

By definition, regret is the difference between the maximum achievable welfare and the welfare achieved with a specified treatment rule. The maximum achievable welfare is max (α, β). Hence, the regret of allocation δ_0 is max (α, β) − [α + (β − α)δ_0]. Regret depends on the unknown value of β. The MR rule computes maximum regret over all feasible values of β and chooses a treatment allocation to minimize maximum regret. Thus, the MR criterion is

$$(4) \quad \min_{\delta_0 \in [0,\,1]} \; \max_{\beta \in [\beta_L,\,\beta_U]} \; \max\,(\alpha, \beta) - [\alpha + (\beta - \alpha)\delta_0].$$

It is easy to see that the MR decision, denoted δ_{MR}, is $\delta_{MR} = 0$ if $\beta_U < \alpha$ and $\delta_{MR} = 1$ if $\beta_L > \alpha$. In the former (latter) case, the planner knows that the innovation is worse (better) than the status quo. Our concern is with situations where the planner does not know which treatment is better; that is, where $\beta_L \le \alpha \le \beta_U$. Manski (2007b, Section 11.3) shows that the MR decision then is

$$(5) \quad \delta_{MR} = (\beta_U - \alpha)/(\beta_U - \beta_L).$$

Proof: Maximum regret across the feasible values of β is

$$\max_{\beta \in [\beta_L,\,\beta_U]} \; (\alpha - \beta)\delta_0 \cdot 1[\beta < \alpha] + (\beta - \alpha)(1 - \delta_0) \cdot 1[\beta > \alpha] \;=\; \max\,[(\alpha - \beta_L)\delta_0,\, (\beta_U - \alpha)(1 - \delta_0)].$$

Thus, the MR rule solves the optimization problem

$$\min_{\delta_0 \in [0,\,1]} \; \max\,[(\alpha - \beta_L)\delta_0,\, (\beta_U - \alpha)(1 - \delta_0)].$$

The quantity $(\alpha - \beta_L)\delta_0$ is increasing in δ_0, whereas $(\beta_U - \alpha)(1 - \delta_0)$ is decreasing in δ_0. The MR allocation is obtained by choosing δ_0 to equalize these two quantities. This gives (5). □

Observe that the MR rule yields a fractional allocation when $\beta_L < \alpha < \beta_U$. The fraction of the cohort assigned to the innovation depends on the location of α within the interval $[\beta_L, \beta_U]$, with δ_{MR} increasing linearly from 0 to 1 as α decreases from β_U to β_L. This behavior is sensible. As α decreases from β_U to β_L, the potential gain from choosing the innovation rises and the potential loss falls.

Observe that leading alternatives to the MR rule, including the maximin rule and Bayes rules, generically do not deliver fractional treatment allocations when applied to this treatment-choice problem. The maximin criterion chooses an allocation that maximizes welfare when β take its lowest feasible value. Thus, the maximin criterion is

$$(6) \qquad \max_{\delta_0 \in [0,\,1]} \; \min_{\beta \in [\beta_L,\,\beta_U]} \; [\alpha + (\beta - \alpha)\delta_0].$$

Solution of this problem yields $\delta_0 = 0$ if $\beta_L < \alpha$ and $\delta_0 = 1$ if $\beta_L > \alpha$.

A Bayesian planner places a subjective probability distribution on the interval $[\beta_L, \beta_U]$, computes the subjective mean value of social welfare, and chooses a treatment allocation that maximizes this subjective mean. Thus, the planner solves the optimization problem

$$(7) \qquad \max_{\delta_0 \in [0,\,1]} \; \alpha + [E_\pi(\beta) - \alpha]\delta_0,$$

where π is the subjective distribution on β and $E_\pi(\beta) = \int \beta d\pi$ is its subjective mean. Solution of this problem

yields $\delta_0 = 0$ if $E_\pi(\beta) < \alpha$ and $\delta_0 = 1$ if $E_\pi(\beta) > \alpha$. [1]

The Multi-Period Extension

Now let $N > 0$. Extending the notation introduced above, suppose that at period n, the planner finds it credible to assert that β lies in a bounded interval $[\beta_{Ln}, \beta_{Un}]$. This interval may change over time, as empirical evidence accumulates on the outcomes experienced by members of earlier cohorts who were treated with the innovation. The interval $[\beta_{Ln}, \beta_{Un}]$ will shrink with n in many cases, but we do not need to presume this.

The adaptive minimax-regret rule applies the MR rule to each successive cohort, using the knowledge of β available at the time. Thus, the AMR decision at each n is

$$\begin{aligned} (8)\quad \delta_{AMR(n)} &= (\beta_{Un} - \alpha)/(\beta_{Un} - \beta_{Ln}) && \text{if } \beta_{Ln} \le \alpha \le \beta_{Un}, \\ &= 0 && \text{if } \beta_{Un} < \alpha, \\ &= 1 && \text{if } \beta_{Ln} > \alpha. \end{aligned}$$

The AMR rule achieves the dual objectives of social learning and diversification. Inspection of (8) shows that the necessary and sufficient condition for learning to occur is $\beta_{U0} > \alpha$. This condition is sufficient for learning because it implies that the planner assigns a positive fraction of cohort 0 to the innovation. Evidence then accumulates over the next K periods, after which he fully observes their outcomes and knows

---

[1] Manski and Tetenov (2007, Proposition 5) show that a Bayesian planner may make a fractional treatment allocation if the social welfare function is changed to $f[\alpha + (\beta - \alpha)\delta_0]$, where $f(\cdot)$ is monotone and continuously differentiable. Then the Bayes problem is

$$\max_{\delta_0 \in [0,\,1]} \int f[\alpha + (\beta - \alpha)\delta_0]\,d\pi.$$

Solutions to this problem are in the interior of the unit interval if $E_\pi(\beta) > \alpha$ and $\int f(\beta)d\pi < f(\alpha)$. In particular, this occurs if the function $f(\cdot)$ is sufficiently concave.

the value of β. Thus, performance of a randomized trial at N = 0 is an inherent consequence of the AMR rule when $\beta_{U0} > \alpha$.

The condition $\beta_{U0} > \alpha$ is necessary for learning because, if $\beta_{U0} \leq \alpha$, the planner assigns no one to the innovation and, hence, never learns its outcomes. The absence of learning has no consequence for welfare in this case. The planner knows from the beginning that the innovation cannot be better than the status quo treatment and, hence, there is no need for a randomized trial.

The AMR rule diversifies each cohort's treatment allocation in a manner that reflects the available knowledge of treatment response. As empirical evidence accumulates and the interval $[\beta_{Ln}, \beta_{Un}]$ changes, the value of $\delta_{AMR(n)}$ varies accordingly. Eventually, observation of outcomes under the innovation reveals where β is larger or smaller than α. From that point on, diversification is no longer warranted and the AMR rule assigns all persons to the better treatment.

One caveat is warranted about the properties of the AMR rule. Although this treatment rule minimizes maximum regret for each cohort separately, given the evidence available at the time, it does not necessarily minimize maximum regret in terms of the multi-period objective function (2). Global minimization of maximum regret in multi-period decision problems is a subtle matter that requires joint consideration of all of the cohorts rather than sequential consideration of them one at a time. Determination of the global multi-period minimax-regret rule is an interesting subject for future research.

2.3. Numerical Illustrations

This section illustrates the AMR rule. I present two hypothetical treatment-choice problems. In each case the presumed outcome of interest unfolds over multiple periods. As empirical evidence accumulates, the AMR treatment allocation changes accordingly.

Treating a Life-Threatening Disease

When treating a life-threatening disease, the outcome of interest may be the number of years that a patient survives within some time horizon. For this illustration, I let the horizon be five years and I define y(t) to be the number of years that a patient lives during the five years following receipt of treatment t, where t is the status quo or the innovation. Thus, y(t) has the time-additive form (1), with $y_{jk}(t) = 1$ if patient j is alive k years after treatment, $y_{jk}(t) = 0$ otherwise, and K = 5.

The outcome gradually becomes observable as time passes. At the time of treatment, $y_j(t)$ can take any of the values [0, 1, 2, 3, 4, 5]. A year later, one can observe whether patient j is still alive and hence can determine whether $y_j(t) = 0$ or $y_j(t) \geq 1$. And so on until year five, when the outcome is fully observable.

Table 1 presents hypothetical data on annual death rates following treatment by the status quo and the innovation. The entries show that 20 (10) percent of the patients who receive the status quo (innovation) die within the first year after treatment. In each of the subsequent years, the death rates are 5 and 2 percent respectively. Overall, the entries imply that the mean numbers of years lived after treatment are $\alpha = 3.5$ and $\beta = 4.3$. The former value is known at the outset from historical experience. The latter gradually becomes observable.

Assume that the planner measures welfare by a patient's length of life; thus, u(t) = y(t). Also assume that the planner has no initial knowledge of β. That is, he does not know whether the innovation will be disastrous, with all patients dying in the first year following treatment, or entirely successful, with all patients living five years or more. Then the initial bound on β is $[\beta_{L0}, \beta_{U0}] = [0, 5]$. Applying equation (8), the initial AMR treatment allocation is $\delta_0 = 0.30$.

In year 1 the planner observes that, of the patients in cohort 0 assigned to the innovation, 10 percent died in the first year following treatment. This enables him to deduce that $P[y(b) \geq 1] = 0.90$. The planner uses this information to tighten the bound on β to $[\beta_{L1}, \beta_{U1}] = [0.90, 4.50]$. It follows that $\delta_1 = 0.28$.

In each subsequent year the planner observes another annual death rate, tightens the bound on β, and

recomputes the treatment allocation accordingly. The result is that $\delta_2 = 0.35$, $\delta_3 = 0.50$, and $\delta_4 = 0.98$. In year 5 he learns that the innovation is better than the status quo, and so sets $\delta_5 = 1$.

Treating a Chronic Disease of Aging

When treating a chronic disease of aging, the outcome of interest may be quality adjusted life years (QALY) within a specified time horizon. For this illustration, let the horizon be twenty years and let y(t) be the number of QALYs experienced during the twenty years following receipt of treatment t, where t is the status quo or the innovation. Thus, y(t) is time-additive, with $y_{jk}(t) \in [0, 1]$, and K = 20.

The welfare of a treatment is its benefit minus its cost. In this illustration, let the status quo be a no-program setting with zero cost and let the innovation cost \$5000 per person. The benefit of a treatment is the benefit of one QALY multiplied by the number of QALYs that a person experiences. I consider two values for the social benefit of one QALY, \$10,000 and \$20,000. Then $u(a) = v \cdot y(a)$ and $u(b) = v \cdot y(b) - 5000$, where v = 10,000 or 20,000. Moreover, $\alpha = v \cdot E[y(a)]$ and $\beta = v \cdot E[y(b)] - 5000$.

Table 2 presents hypothetical data on mean QALYs following each treatment. Two columns show $E[y_k(t)]$ for each k = 1, . . . , K and each value of t. The entries describe a situation in which the innovation never does harm relative to the status quo and raises the mean in in some years. Overall, the entries imply that the mean numbers of QALYs experienced during the twenty-year horizon are E[y(a)] = 13.28 and E[y(b)] = 13.57. It follows that if v = 10,000, then α = 132,800 and β = 130,700; hence, the status quo is the better treatment. If v = 20,000, then α = 265,600 and β = 266,400; hence, the innovation is the better treatment in this case.

The value of E[y(a)] is known at the outset from historical experience, but E[y(b)] becomes observable only gradually. To compute the AMR treatment allocation, I suppose the planner initially knows that treatment b never yields a result lower than $E[y_k(a)]$ and that it may raise it by a maximum of 0.10. Consider, for example, the first and twelfth years after treatment. The table shows that $E[y_1(a)] = 0.98$ and

$E[y_{12}(a)] = 0.80$. Hence, the corresponding values under the innovation are initially known to lie in the intervals [0.98, 1] and [0.80, 0.90]. These bounds generate the bounds on E[y(b)] shown in Table 2.

The final columns of Table 2 show the AMR treatment allocations. The patterns for the two values of v are quite different. If v = 10,000, then $\delta_n = 0.65$ in years 0 through 4, δ_n slowly decreases to 0.59 in year 10, then quickly rises to 0.74 at year 12, and finally falls to zero in year 18, when it becomes known that the status quo is better than the innovation. If v = 20,000, then δ_n stays close to 0.83 through year 10 and then rises to one in year 12, when it becomes known that the innovation is better than the status quo.

3. Revising the Drug Approval Process

3.1. The Present FDA Process

The present FDA process for drug approval begins with preclinical laboratory and animal testing of new compounds by pharmaceutical firms. Those that seem promising then go through three phases of trials of increasing size and varying objectives. Phase 1 trials, which typically take about a year and are performed with twenty to eighty healthy volunteers, aim to determine the basic pharmacological action of the drug and the safety of different doses. Phase 2 trials, which usually take about two years and are performed with several hundred volunteers who are ill with a specific disease, give preliminary evidence on the effectiveness and short-term side effects of the drug. Phase 3 trials, which usually take about three years and are performed with several hundred to several thousand volunteers ill with the disease, give further evidence on effectiveness and side effects. Following completion of Phase 3, the firm files a New Drug Application and the FDA either approves or disapproves the drug for prescription by physicians.[2]

[2] See www.fda.gov/cder/handbook/develop.htm.

Hypothesis tests are used to compare the innovation with the status quo treatment, with the status quo given the privileged position of the null hypothesis. Approval of a new drug normally requires one-sided rejection of the null hypothesis of zero average treatment effect $\{H_0: \beta = \alpha\}$ in two independent trials (Fisher and Moyé, 1999). This sets a high bar for approval, requiring that pharmaceutical firms demonstrate "substantial evidence of effect" for their products (Gould, 2002).

Two features of the FDA process stand out from the perspective of this paper. First, although the FDA protocol has multiple stages, it essentially expresses a binary rather than adaptive approach to treatment. Prior to approval, a new drug is used to treat only those patients who volunteer to participate in a trial and who are randomized into the appropriate treatment group. Even in Phase 3 trials, this group is typically quite small relative to the patient population in a country as populous as the United States. Following approval, a new drug may be used to treat the entire patient population. Thus, the fraction of the patient population who receive the treatment is close to zero prior to approval and may approach one following approval.

Second, the FDA regularly makes its approval decision using data on surrogate outcomes rather than evidence on outcomes of interest. As documented by Fleming and Demets (1996) and others, Phase 3 trials commonly are too short in duration to measure the health outcomes of real concern. The FDA has attempted to compensate for the use of surrogate outcomes in drug approval by encouraging longer term scrutiny of approved drugs through a process of "post-market surveillance." However, at present, the FDA post-market surveillance program only aims to detect adverse side effects of approved drugs, not to better measure their effectiveness in treatment. [3]

3.2. A Drug Approval Process Blending the FDA Protocol and AMR Treatment Choice

A drug approval process implementing AMR treatment choice would differ in important ways from

[3] See www.fda.gov/cder/regulatory/applications/postmarketing/surveillancepost.htm.

the present FDA process. Approval would be an adaptive process rather than a binary one, with the prevalence of treatment with a new drug varying smoothly as empirical evidence accumulates. The empirical evidence used to determine treatment allocations would mainly be data on outcomes of real interest rather than surrogate outcomes. A process implementing the AMR rule would eliminate the current sharp distinction between approval of a new drug and post-market surveillance. Instead, the planner would continually monitor the available evidence and adjust the treatment allocation as appropriate. In this way, society would achieve both social learning and diversification.

Implementation of a close approximation to the AMR rule could be feasible in a society with a socialized or single-payer health care system. It seems unrealistic in the present American context, where private pharmaceutical firms develop and market new drugs, and where treatment decisions arise from the decentralized interaction of physicians, health insurance organizations, and patients. However, an approval process that revises FDA drug approval to incorporate important features of the AMR rule may be realistic in the American context.

What I have in mind is a process that begins, as at present, with a pharmaceutical firm performing preclinical testing followed by Phase 1 and 2 trials. It seems prudent to continue these preliminary stages of the approval process without substantial revision. The changes would appear in the subsequent Phase 3 trials and in the FDA decision process. First, the duration of Phase 3 trials would be lengthened sufficiently to measure health outcomes of real interest. The specification of these outcomes would be decided by the FDA, with input from relevant parties. Second, the present binary approval decision following a Phase 3 trial would be replaced by an adaptive process that monitors the trial while in progress and that periodically makes limited-term partial approval decisions.

By "limited-term partial approval," I mean that while a Phase 3 trial is underway, the FDA would give a renewable license to the firm to market no more than a specified quantity of the new drug over a specified time period. The duration of the license would depend on the agreed schedule for reporting new findings in

the trial. For example, if the firm reports updated outcome data to the FDA annually, then the licensing decision would be updated annually as well.

On each iteration of the licensing decision, the maximum quantity of drug that the firm is permitted to market would be set by the FDA with the assistance of an expert advisory board, similar to those now used in drug approval. This is where the AMR rule comes in. To give the licensing decision transparency and coherence, the FDA could be mandated to compute the AMR treatment allocation with a specified social welfare function. The role of the expert advisory board would be to set lower and upper bounds $[\beta_{Ln}, \beta_{Un}]$ for treatment effectiveness that are scientifically appropriate given the available empirical evidence.

Finally, when the Phase 3 trial is complete and the specified outcomes of interest have been observed, the FDA would make a long-term approval decision as at present. If the drug is deemed safe and effective, the firm would be permitted to market it with no quantity restriction. Further marketing would be prohibited otherwise.

The revised drug approval process differs from AMR treatment choice mainly in the method used to allocate the population between the innovation and the status quo while the Phase 3 trial is underway. The AMR rule calls for randomized allocation of the entire patient population, thus ensuring that the distribution of treatment response is the same in both treatment groups. Here, randomized allocation would occur only within the sample of persons who participate in a Phase 3 trial. Otherwise, the pharmaceutical firm would market its permitted quantity in the ordinary manner, presumably setting price to maximize profit subject to the FDA-specified upper bound on sales.[4] The adaptive approval process would achieve social learning and diversification, with part of the population receiving the innovation and the remainder receiving the status quo

[4] This discussion presumes a relatively simple setting in which the members of the patient population are observationally identical and the innovation has use only for treatment of the specific disease being studied in the Phase 3 trial. If patients are observationally heterogeneous, the AMR rule may call for different treatment allocations within different subpopulations. If the innovation has uses for treatment of multiple diseases, the distribution of the permitted sales across patients with different diseases may be a concern. I call attention to these issues here but do not attempt to resolve them.

while the Phase 3 trial is in progress. However, it would not ensure that the distribution of treatment response is the same in both treatment groups.

Although it seems unrealistic in the American context to mandate randomized treatment of the entire patient population while Phase 3 trials are in progress, I would note that mandated randomization does have precedent in American society. Examples include random drug testing, calls for jury service, and the Vietnam draft lottery. In the medical arena, the FDA recently approved conduct of a sequence of studies in which critically ill patients are to be randomized into treatment without consent; see Stein (2007). Nevertheless, it would be a major departure from current norms to broadly mandate randomized health care.

3.3. Discussion

Even if the proposed drug approval process must, for social acceptability, deviate to some extent from AMR treatment choice, I believe that it would substantially improve the present FDA process. As pointed out in the Introduction, the present process makes an up or down decision using data on surrogate outcomes and, hence, is susceptible to errors with long-term consequences. Type I errors occur when new drugs that actually are worse than the status quo are approved because they appear superior when evaluated using surrogate outcomes, and Type II errors occur when new drugs that actually are better than the status quo are disapproved because they appear inferior when evaluated using surrogate outcomes.

These long-term errors would not occur under the proposed process, which would perform Phase 3 trials of sufficient length to reveal health outcomes of real interest. In the present system, performance of longer Phase 3 trials has been resisted because this would delay the FDA's binary approval decisions. By permitting partial marketing of new drugs while Phase 3 trials are underway, the proposed process should make it acceptable to increase the length of Phase 3 trials as appropriate.

It should be particularly appealing to use the AMR rule to adjust the permitted scale of marketing as

evidence accumulates. The minimax-regret criterion has a firm welfare-economic foundation—no other decision rule yields a treatment allocation that is uniformly better across all feasible values of β. The MR criterion is arguably more objective than Bayesian treatment choice, where the allocation depends on the subjective probability distribution placed on β. Adaptive implementation of the MR criterion treats each cohort as well as possible given the available knowledge—it does not ask the members of one cohort to sacrifice its own welfare for the benefit of other cohorts. Finally, the AMR treatment allocation has the remarkably simple form (8), which should be easy to explain heuristically, if not technically, to the general public.

Although the AMR treatment allocation is simple in form, serious normative thought and empirical analysis will be necessary to determine appropriate values for α and $[\beta_{Ln}, \beta_{Un}]$. The need for normative thought arises when society, acting through the FDA, determines what social welfare function to use when comparing treatments. In the numerical illustrations of Section 2.3, I used life span in one case and a benefit-cost form based on QALYs in the other. The latter case requires decisions on how to measure QALYs and their social benefits.

With the social welfare function chosen, empirical analysis is needed to determine α and $[\beta_{Ln}, \beta_{Un}]$. The numerical illustrations showed how to do this in situations where the outcomes of interest are time-additive and the empirical evidence gradually reveals these outcomes. Other cases may be more complex.

In particular, the available empirical evidence may include data on surrogate outcomes as well as on the outcomes of interest. Then the FDA will have to decide what the surrogate outcomes reveal about the value of β. This decision may have considerable impact on treatment allocations in the early periods, before much is known about the outcomes of interest. Of course appraisal of surrogate outcomes is not a new task—the present approval process already requires extrapolation from surrogate outcomes to outcomes of interest. What will be new is appraisal of how data on surrogate outcomes should affect the interval $[\beta_{Ln}, \beta_{Un}]$ that the FDA uses to express partial knowledge of β.

4. Treatment Choice and Drug Approval with Imperfect Data

Sections 2 and 3 focused on the problem created by use of surrogate outcomes and abstracted from other issues that arise in the accumulation of knowledge for drug approval. In particular, I assumed that empirical evidence on outcomes of interest is obtained from a classical randomized experiment performed on a large random sample of the population. This concluding section discusses treatment choice and drug approval when these assumptions do not hold.

Section 4.1 supposes that the available data are from a classical experiment with a finite sample of subjects. Section 4.2 considers settings in which the available empirical evidence does not attain the classical experimental ideal.

Another issue, not considered here in detail, is that one may have only partial knowledge of outcomes under the status quo treatment. Minimax-regret treatment choice continues to be applicable in such situations. Manski (2007b, Complement 11A) gives the general result when it is known that $\alpha \in [\alpha_L, \alpha_U]$ and $\beta \in [\beta_L, \beta_U]$. The treatment allocation is fractional whenever the intervals $[\alpha_L, \alpha_U]$ and $[\beta_L, \beta_U]$ overlap.

4.1. Experiments with Finite Samples of Subjects

As mentioned in Section 3, FDA approval of a new drug presently requires one-sided rejection of the null hypothesis of zero average treatment effect $\{H_0: \beta = \alpha\}$. Hypothesis testing is difficult to motivate from the perspective of treatment choice. First, there is no decision-theoretic rationale for the standard practice of handling the null and alternative hypotheses asymmetrically, fixing the probability of a type I statistical error and seeking to minimize the probability of a type II error. One should instead handle the two errors symmetrically. Second, error probabilities only measure the chance of choosing a sub-optimal rule.

They do not measure the loss in welfare resulting from a sub-optimal choice.

From the perspective of this paper, an appropriate way to choose treatments with finite-sample data is to apply the finite-sample version of the minimax-regret criterion, studied in Manski (2004, 2005), Manski and Tetenov (2007), Stoye (2006), and Schlag (2007). Manski and Tetenov (2007) is particularly relevant to the present discussion. Considering choice between an innovation and status quo treatment when outcomes are binary, they compare the regret function of the finite-sample MR rule with that obtained when a hypothesis test is used to choose treatments. The MR rule is closely approximated by the "empirical success" rule, which uses the sample success rate with the innovation to estimate β and then chooses treatments accordingly. In contrast, a standard hypothesis test is more likely to choose the status quo over the innovation.

The articles cited above all concern situations in which one has complete outcome data on the sampled subjects. In the context of this paper, this occurs K periods after initiation of the experiment. Application of the finite-sample AMR rule prior to period K requires determination of the finite-sample MR rule when outcomes are not yet fully observed but have been observed to lie in certain intervals. This is a subject for future research. A useful starting point may be Stoye (2007), who has studied finite-sample MR treatment choice when outcomes are either fully observed or completely missing.

4.2. Non-Classical Experiments

In Section 2 I cited several reasons why the trials used in the drug approval process often depart from the classical ideal. First, trials typically are performed on convenience samples of volunteers rather than on random samples of the patient population of interest. Second, some of the volunteers who participate in trials may not comply with their assigned treatments or may leave early, before their outcomes can be measured. Third, trials are typically performed with blinded treatment assignment, even though treatments are observed in clinical practice.

AMR treatment choice is applicable in such situations. Whatever the data problem may be, it will manifest itself in a bound $[\beta_{Ln}, \beta_{Un}]$ that the planner finds credible to assert in period n. The AMR decision in period n continues to be given by (8). Thus, imperfect data generated by non-classical experiments creates no conceptual problem. It only requires that the planner determine the bound appropriate to the setting.

Indeed, AMR treatment choice is applicable when the only available data are surrogate outcomes. The planner needs to determine the conclusions about β that he can draw by combining the available surrogate-outcome data with credible assumptions that link these data to outcomes of interest. This done, the AMR rule may be applied.

Data from non-classical experiments do create a practical issue for drug approval. The discussion of Section 3 presumed that the FDA would eventually be able to make an up or down decision on a new drug, certainly after K periods and perhaps earlier. If the available data do not meet the classical ideal, the FDA's expert advisory board might never reach the point where it feels able to shrink the bound $[\beta_{Ln}, \beta_{Un}]$ enough to determine whether the innovation is better than the status quo. In such a scenario, AMR treatment choice calls for permanent fractional treatment assignment. This is perfectly sensible in principle, but pharmaceutical firms and patients may not be willing to accept a non-zero permanent limit on the degree to which a drug can be marketed. At some point, an up or down decision may be necessary in practice.

References

Fisher, L. and L. Moyé (1999), "Carvedilol and the Food and Drug Administration Approval Process: An Introduction," *Controlled Clinical Trials*, 20, 1-15.

Fleming, T. and D. Demets (1996), "Surrogate End Points in Clinical Trials: Are We Being Misled?" *Annals of Internal Medicine,* 125, 605-613.

Gould, A. (2002), "Substantial Evidence of Effect," *Journal of Biopharmaceutical Statistics*, 12, 53-77.

Manski, C. (2004), "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72, 1221-1246.

Manski, C. (2005), *Social Choice with Partial Knowledge of Treatment Response*, Princeton: Princeton University Press.

Manski, C. (2007a), "Minimax-Regret Treatment Choice with Missing Outcome Data," *Journal of Econometrics*, 139, 105-115.

Manski, C. (2007b), *Identification for Prediction and Decision*, Cambridge, Mass.: Harvard University Press.

Manski, C. and A. Tetenov (2007), "Admissible Treatment Rules for a Risk-Averse Planner with Experimental Data on an Innovation," *Journal of Statistical Planning and Inference*, 137, 1998-2010.

Meltzer, D. (2001), "Addressing Uncertainty in Medical Cost-Effectiveness: Implications of Expected Utility Maximization for Methods to Perform Sensitivity Analysis and the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research," *Journal of Health Economics,* 20, 109-129.

Psaty, B. et al. (1999), "Surrogate End Points, Health Outcomes, and the Drug-Approval Process for the Treatment of Risk Factors for Cardiovascular Disease," *Journal of the American Medical Association*, 282, 786-790.

Savage, L. (1951), "The Theory of Statistical Decision," *Journal of the American Statistical Association*, 46, 55–67.

Sculpher, M. and K. Claxton (2005), "Establishing the Cost-Effectiveness of New Pharmaceuticals under Conditions of Uncertainty—When Is There Sufficient Evidence?" *Value in Health*, 8, 433-446.

Stein, R. (2007), "Critical Care without Consent," *Washington Post*, May 27, A01.

Schlag, K. (2006), "Eleven," Technical Report, European University Institute.

Stoye, J. (2006), "Minimax-Regret Treatment Choice with Finite Samples," Department of Economics, New York University.

Stoye, J. (2007), "Minimax-Regret Treatment Choice with Finite Samples and Missing Outcome Data," Proceedings of the 5th International Symposium on Imprecise Probability, Prague, Czech Republic.

Table 1: Treating a Life-Threatening Disease				
year (n or k)	death rate in k^{th} year after treatment		bound on β in year n	AMR allocation in year n
	Status Quo	Innovation		
0			[0, 5]	0.30
1	0.20	0.10	[0.90, 4.50]	0.28
2	0.05	0.02	[1.78, 4.42]	0.35
3	0.05	0.02	[2.64, 4.36]	0.50
4	0.05	0.02	[3.48, 4.32]	0.98
5	0.05	0.02	[4.30, 4.30]	1

Table 2: Treating a Chronic Disease of Aging					
year (n or k)	mean QALY in k^{th} year after treatment		bound on E[y(b)] in year n	AMR allocation in year n, by social benefit of one QALY	
	Status Quo	Innovation		$10,000	$20,000
0			[13.28, 14.70]	0.65	0.82
1	0.98	0.99	[13.29, 14.69]	0.65	0.83
2	0.98	0.99	[13.30, 14.68]	0.65	0.83
3	0.98	0.99	[13.31, 14.67]	0.65	0.84
4	0.98	0.98	[13.31, 14.65]	0.65	0.84
5	0.98	0.98	[13.31, 14.63]	0.64	0.83
6	0.98	0.98	[13.31, 14.61]	0.64	0.83
7	0.95	0.96	[13.32, 14.57]	0.63	0.83
8	0.95	0.96	[13.33, 14.53]	0.62	0.83
9	0.90	0.92	[13.35, 14.45]	0.61	0.84
10	0.90	0.92	[13.37, 14.37]	0.59	0.84
11	0.80	0.90	[13.47, 14.37]	0.66	0.93
12	0.80	0.90	[13.57, 14.37]	0.74	1
13	0.50	0.50	[13.57, 14.27]	0.70	1
14	0.50	0.50	[13.57, 14.17]	0.65	1
15	0.40	0.40	[13.57, 14.07]	0.58	1
16	0.40	0.40	[13.57, 13.97]	0.48	1
17	0.10	0.10	[13.57, 13.87]	0.30	1
18	0.10	0.10	[13.57, 13.77]	0	1
19	0.05	0.05	[13.57, 13.67]	0	1
20	0.05	0.05	[13.57, 13.57]	0	1