The Economics of Cystic Fibrosis care in the East of England

James Austin Jarrett

Thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

Faculty of Health

School of Medicine, Health Policy and Practice

June 2010

James Austin Jarrett
June 2010
The Economics of Cystic Fibrosis Care in the East of England

Abstract

Cystic fibrosis (CF) is one of the most common genetic diseases in the UK. CF is a chronic, progressive, and life-shortening disease. As a multi-system disease, CF requires complex and constant care that can place a significant burden on constrained NHS resources. This thesis focuses on the economic impact on the NHS of paediatric CF in the East of England as well as investigating the cost-effectiveness of newborn screening for CF.

The literature indicates that there is a relatively large variation (US$4069 to $29849 per case per year) in the direct cost of care from a health system perspective. The only newborn screening study from a UK perspective is now relatively old (1998) and the emergence of new therapies and service delivery methods could alter the findings of that study.

First, a retrospective (1998, 2004-2007) longitudinal cost of illness study was carried out on paediatric CF patients in the east of England. Patients' outcome and resource use information was gathered from patient notes and computerised systems. The study takes into account direct costs to the NHS only, and uses NHS tariffs to attach a unit cost to resource use. Regression analysis was used to analyse the data for indicators of cost drivers, controlling for age, and diagnosis category.

Second, a cost-utility model was built to look at the long-term impact of newborn screening versus no screening (clinical diagnosis). Newborn screening for CF in the UK uses a blood-spot test which is tested using the immunoreactive trypsin test, a second confirmatory immunoreactive trypsin test and DNA analysis. The model was a probabilistic cohort simulation model that uses Markov processes to allow patient history over time. Once a patient had been diagnosed, they could enter one of four health states: Successful treatment, intermittent infection, chronic infection, and death.

The results of the study show that the total cost of care for paediatric CF patients in the East of England has increased over time (1998-2007) from approximately £1,040,087 to £1,062,008 and that the median cost per patient has increased from £3852 to £4249. The regression analysis shows that disease severity has a significant negative impact on costs and being in the youngest and oldest age group indicates having a higher cost. The results of the cost-effectiveness analysis indicate that no screening is associated with a cost of £110,238 and a QALY gain of 11.23 and screening is associated with a cost of £110,417 and a QALY gain of 11.51. Therefore, newborn screening has a cost per QALY of £641 when compared with clinical diagnosis.

This analysis has shown that newborn screening for CF is cost-effective, and that the government has made the correct decision by rolling out a national programme. Once identified by screening, treatment can begin immediately which may have a positive impact on future health for the patient. As shown by the regression analysis, keeping people healthier reduces the overall disease and cost burden on the NHS. The results of this thesis are most applicable to the UK context, however if treatment regimes and screening processes are similar in other countries, the results could be applicable in other contexts. The methods used in this thesis are applications of well known and utilised economic evaluation and statistical methods, and are appropriate for any diseases where prognosis and disease severity is variable.

Contents

Abstract 2
Contents 3
List of Boxes, Figures and Tables 6
Acknowledgements 8
List of abbreviations 9
Chapter 1: Introduction 10
Introduction 10
Cystic Fibrosis 11
Background and definitions 11
Diagnosis of CF 13
Incidence and Prevalence of CF 15
Funding arrangements 16
Treatment of CF 21
Economic Analysis 25
Introduction 25
Cost-of-illness studies 26
Cost-effectiveness analysis 28
Aims and Outline of Thesis 31
Chapter 2: Literature Review 33
Introduction 33
Economics of CF 34
Methods 34
Results 37
Discussion 52
Conclusion 55
Chapter 3: Methodology 68
Introduction 68
Types of costs 71
Concepts of cost 72
Study Design – Costing Issues 73
What costs to include? 73

Estimation of costs....74
Conclusions....84
Benefit measurement and valuation....85
Measuring Health....86
Valuing Health....87
Conclusions....111
Analysing Health Economic Data....112
Cost-of-illness methods....112
Conclusions....123
Economic Evaluation Methodologies....123
Conclusions....136
Chapter 4: Cost of Illness....138
Introduction....138
Methods....138
Study Population....138
Eastern Region Cystic Fibrosis Database (ERCFD)....138
Resource use....142
Costing....142
Statistical analysis....143
Model descriptions....145
Results....148
Study population....148
Data analysis....152
Discussion....187
Chapter 5: Cost-effectiveness analysis of newborn screening for cystic fibrosis....190
Introduction....190
Methods....191
Model design....191
Data analysis....204
Results....207
Discussion....218

Chapter 6: Discussion.....221
Overview of thesis.....221
Rationale and objectives.....221
Main findings.....222
Methodological considerations.....224
Methodological limitations.....224
Conclusions.....227
Policy implications.....229
Recommendations for future research.....231
Conclusion.....235
Appendix A: Medical Glossary *(193)*.....237
Appendix B: Literature searches.....240
Appendix C: Literature Review Pro-forma (adapted from (38, 39)).....242
Appendix D: Data collection sheet for the ERCFD.....243
Appendix E: Excel Workbook and Macros for cost-effectiveness model.....246
Reference List.....257

List of Boxes, Figures and Tables

Boxes

Chapter 1
Box 1.1: CF Trust Banding recommendations....20
Chapter 2
Box 2.1: Search strategies....35
Chapter 3
Box 3.1 Definitions of cost....73
Chapter 4
Box 4.1: Comparison of national CF database statistics to ERCFD in 2007....150

Figures

Chapter 1
Figure 1.1 UK Screening Algorithm....15
Chapter 2
Figure 2.1 Literature Review Algorithm....38
Chapter 3
Figure 3.1 Choices for costing over time....78
Figure 3.2 Visual Analogue Scale....89
Figure 3.3 Sample Decision Tree....132
Chapter 4
Figure 4.1 Map of hospitals participating in the ERCFD....141
Figure 4.2 Mean clinic costs over time....155
Figure 4.3 Mean inpatient hospital costs over time....155
Figure 4.4 Mean drug costs over time....156
Figure 4.5 Mean total costs over time....156
Figure 4.6 Trend in median total costs by age group....158
Figure 4.7 Trend of mean costs in screened versus not screened patients....159
Figure 4.8 Trend of median costs in screened versus no screened patients....160
Figure 4.9 Cumulative percentage plots total costs screened/not screened....160
Figure 4.10 Cumulative frequency plots of total cost screened/not screened 1998/2007....161
Figure 4.11 Box plots of total cost by diagnosis type and band....162
Figure 4.12 Box plots of total cost by diagnosis, age group and band....163
Figure 4.13 Comparison of total costs using different costing sources....164
Figure 4.14 Comparison of mean/median total costs using different costing sources....165
Figure 4.15 Comparison in trends between cost drivers....166
Figure 4.16 Changes the total costs of common drugs used in CF care....167
Figure 4.17 Trend in median costs by age group over time (Banding Tariffs)....168
Figure 4.18 Cumulative percentage of total cost between screened/not screened (banding tariffs)....169
Figure 4.19 Cumulative percentage of total cost between screened/not screened 1998/2007 (Banding tariffs)....170
Figure 4.20 Histograms of total costs by year....171
Figure 4.21 Histograms of logged total cost by year....172
Figure 4.22 Predicted versus actual values for models using total cost....186
Figure 4.23 Fitted versus actual values for logged total cost models....186
Chapter 5

Figure 5.1 Diagrammatic representation of CE/Markov Model 194
Figure 5.2 Total state costs over time by arm 211
Figure 5.3 Cost-utility plane 214
Figure 5.4 ICER distribution 215
Figure 5.5 Cost-effectiveness acceptability curves 216
Figure 5.6 Population EVPI 217
Figure 5.6 EVPPI 218

Tables

Chapter 1

Table 1.1 Phenotypic features consistent with a diagnosis of CF 14

Chapter 2

Table 2.1 Cost of Illness studies 56
Table 2.2 Cost descriptions 57
Table 2.3 Cost consequences studies 58
Table 2.4 Full and partial economic evaluations 64

Chapter 3

Table 3.1 HRQOL measures for cystic fibrosis 108

Chapter 4

Table 4.1 Study population characteristics 151
Table 4.2 Resource use and unit costs 152
Table 4.3 Mean and median resource use and costs 154
Table 4.4 Median costs over time by age group 157
Table 4.5 Results of skewness test on total cost and natural log of cost and fitted distributions 173
Table 4.6 Initial regression analysis on baseline data (total costs) 174
Table 4.7 Initial regression analysis on baseline data (natural logged costs) 178
Table 4.8 Pairwise correlation of total and logged total costs to band and age 179
Table 4.9 Model Fit Statistics 180
Table 4.10 GEE Regression results 1 181
Table 4.11 GEE Regression results 2 182
Table 4.12 GEE Regression results 3 185

Chapter 5

Table 5.1 Demographics at baseline (1998) 195
Table 5.2 Transition probabilities and Drichilet Beta distributions estimated from ERCFD 196
Table 5.3 Regression to calculate hazard function for relative risk of infection 199
Table 5.4 Covariance matrix, Cholesky decomposition and random variable for probabilistic analysis 200
Table 5.5 Starting cost and utility parameters 201
Table 5.6 Summary of model parameters 202
Table 5.7 Number of people in each health state by year 208
Table 5.8 Costs of health states by year 209
Table 5.9 Utilities gained in each health state by year 212
Table 5.10 Results of the probabilistic analysis 213
Table 5.11 Net Monetary Benefit 215

Acknowledgements

First and foremost I would like to thank my primary supervisor, Prof. Miranda Mugford, for her unwavering support and encouragement over the last 7 years. I would also like to thank the rest of my supervisory team, Dr. Richard Iles and Dr. Erika Sims. They have provided much needed input and enthusiasm at just the right time. I owe a great debt of gratitude to my colleagues at UEA who let me bounce ideas off them many times over. In particular I would like to thank Dr. Marcello Morciano, Ian Shemilt, and Ed Wilson. I would also like to thank my discussants and audience members for their helpful comments on various aspects of this project at the Manchester HESG meeting.

There are many other non-academic people who have helped me along the way as well. I would like to thank my parents for instilling a love of learning. I'm also grateful for Sarah Cox and Jim Hunter for their friendship and hospitality. My eternal gratitude also goes out to Patricia Thomas for so many things, among them asking me to do some research on the economics of AIDS vaccines that led me to become a health economist.

Finally, this thesis would not have been possible without the support from the NIHR NCCRCD pre-doctoral fellowship. My thanks go out to them for trusting in me to deliver a project for which I am still passionate.

List of abbreviations

CBA – Cost benefit analysis
CEA – Cost Effectiveness Analysis
CEAC – Cost effectiveness acceptability curve
CF- Cystic Fibrosis
CFTR – Cystic Fibrosis Transmembrane conductance Regulator
CMA – Cost minimisation analysis
COI – Cost of Illness Analysis
CUA – Cost utility analysis
DH – Department of Health
ERCFD- Eastern Region Cystic Fibrosis Database
EVPI – Expected value of perfect information
EVPPI – Expected value of perfect parameter information
GEE – Generalised Equilibrium Equations
GLM – Generalised Linear Models
ICER – Incremental cost effectiveness ratio
IRT – Immunoreactive Trypsin
NBS – Newborn Screening
NCG – National Commissioning Group
NHS – National Health Service (of England and Wales)
NSCG – National Specialist Commissioning Group
OLS – Ordinary Least Squares
PA – Pseudomonas Aeruginosa
PCT- Primary Care Trust
PSA – Probabilistic Sensitivity Analysis
QALY – Quality Adjusted Life Year
SA – Staphylococcus aureus
SC – Shared Care
SCG – Specialist Commissioning Group
SHA – Strategic Health Authority
WHO- World Health Organisation

Chapter 1: Introduction

Introduction

This thesis focuses on the economic impact of paediatric Cystic Fibrosis (CF) in part of the Eastern Region (Cambridgeshire, Norfolk, and Suffolk) of England utilising data from the Eastern Region CF Database (ERCFD). CF is a chronic, progressive, and ultimately terminal disease. Like patients with other chronic diseases, patients with CF require constant care. Therefore, providing CF care can be very expensive. This thesis utilises two different health economic methodological approaches to illustrate the impact of CF on the English National Health Service (NHS).

Like every health service, the funds available to the NHS are not infinite, and therefore must be carefully managed. As the NHS carries out most of the care for paediatric CF patients within the Eastern Region, it is imperative that the clinical and economic impact of CF on the NHS is detailed. Knowledge of these impacts will enable decision makers to allocate funds more efficiently and potentially improve services for CF patients and their families.

Health economics offers various methods that can aid in answering these questions. One such method is a cost of illness analysis. This type of analysis can provide information on the burden of disease on the health-care system. This is particularly useful in gauging the scale of the problem. In this thesis, I will analyse the cost-of-illness to the NHS of caring for CF patients. The second methodological approach I will be using is a cost-effectiveness analysis (CEA). This method allows alternative forms of care to be compared in terms of their clinical and cost effectiveness, using a common outcome measure. This thesis will analyse the relationship between the

cost and clinical effectiveness of newborn screening for CF. Newborn screening for CF has now been rolled out across the UK despite there being little economic evidence that it is cost effective. This analysis will add to limited published evidence on newborn screening for CF in the UK context.

Cystic Fibrosis

Background and definitions

CF is an inherited, life-threatening disease that progressively worsens over time. CF most commonly occurs in people of Caucasian extraction. Approximately 1 in 2,500 babies born in the UK have CF. Newborn screening for CF was rolled out across the UK in 2007, although several areas (Scotland, the east and south west of England) have been screening for longer (3). CF causes severe respiratory problems and inadequate pancreatic function, due to the production of excess sticky mucus; patients can also have liver and kidney problems and most males with CF cannot conceive naturally. Disease severity varies between affected individuals: most patients die from respiratory failure, but approximately 15% die from other organ failure. Improvements in treatment have increased average life expectancy to over 30 years. Indeed, at least half of individuals born with CF since 1990 are now expected to live beyond 40, but there remains no known cure for the disease (4).

CF is an autosomal recessive genetic disorder[1]. Cloning of the gene for the CF transmembrane conductance regulator (CFTR) was finished in 1989. This means that both parents must carry one half of the faulty CF gene. One in 25 of the UK

[1] A glossary of medical terms relevant to Cystic Fibrosis can be found in Appendix A.

population carry one half of a faulty CF gene. Because the other half of the CF gene is correct, these people do not develop CF disease. In order to be affected, a child must inherit a faulty gene from both of its parents. The chances of two carrier parents conceiving an affected child are 1 in 4 for each pregnancy. There are more than 3000 known mutations to the CF gene, most of which are very rare. In the UK the mutation known as 'Delta F508' (ΔF508), which is a small deletion in the gene, accounts for around 70-75 percent of cases in the UK. (4) (5)

National CF databases

In order to perform audits and research into CF and the natural history of the disease, a national CF database was developed. Initially started in Scotland in 1992, the UK CF Database was extended across the UK in 1999, with customised software running on computers in Specialist CF Centres and in district hospitals that cared for CF patients. The database collected information from the annual reviews[2] of patients and therefore allowed a "snapshot" of CF patients' health outcomes each year. The project was initially funded by the Clinical Resource and Audit Group and the National Services Division of NHS (Scotland); the Database was then funded by the Cystic Fibrosis Trust[3] until 2007, when the database was closed. (6)

---

[2] A more in depth discussion of what clinical interaction takes place during an annual review comes later in this chapter. In general, all CF patients should have an annual review which includes seeing all members of a specialist team (consultant, specialist nurse, nutritionist, physiotherapist, psychologist) as well as various tests such as pulmonary function tests and BMI.

[3] The Cystic Fibrosis trust is a UK-based charity that is dedicated to all aspects of CF. The organisation is a key stakeholder in CF care in the UK, and has conducted research into many aspects of CF care,

The CF Trust developed a new, web-based CF Registry in 2007, which also uses customised software (PORT CF) which allows users to enter annual review data directly into the database. The CF Trust produces annual data reports which analyses the demographic and clinical outcome data which allows more accurate prevalence and incidence information as well as allows researchers to look for new trends in the health of people with CF in order to aid in creating national standards of care, the design of clinical trials, and improving care. (7)

Diagnosis of CF

The clinical diagnosis of CF is anything but straightforward. Though diagnostic tools have greatly improved over the past few decades, identifying patients for diagnostic testing is problematic. The World Health Organisation (WHO) recommended that "the diagnostic classification should be made on clinical rather than laboratory grounds, while acknowledging the importance of identifying CFTR mutations in those persons with clinical conditions such as pancreatitis and atresia of the vas deferens where some, but not all cases are CFTR-related."(2) A CF diagnosis is entertained due to a patient presenting with one or more of the symptoms listed in table 1.1, a sibling with CF, or has had a positive result from a newborn screening test. If any of these factors are found, a sweat test to ascertain the level of sodium chloride is undertaken (>60ng/ml is considered elevated). If a patient has been identified through newborn screening and two mutations for CFTR have been found, one sweat test is considered adequate. However, as there are inconsistencies in

---

as well as creating guidelines for care alongside other stakeholders in CF care such as the Royal College of
Paediatricians and Child Health. The Trust has published several consensus documents relating to the treatment of children and adults with CF.

the genotype-phenotype correlations associated with CF, sometimes a second confirmatory sweat chloride test is recommended.

Table 1.1: Phenotypic features consistent with a diagnosis of CF*

Chronic sinopulmonary disease manifested by
a. Persistent colonisation/infection with typical CF pathogens including *Staphylococcus aureus,* nontypeable *Haemophilus influenzae,* mucoid and non-mucoid *Pseudomonas aeruginosa,* and *Burkholderia cepacia.*
b. Chronic cough and sputum production
c. Persistent chest radiograph abnormalities (e.g. bronchiectasis, atelectasis, infiltrates, hyperinflation)
d. Airway obstruction manifested by wheezing and air trapping
e. Nasal polyps; radiographic or computed tomographic abnormalities of the paranasal sinuses
f. Digital clubbing
Gastrointestinal and nutritional abnormalities including
a. Intestinal: meconium ileus, distal intestinal obstruction syndrome, rectal prolapse
b. Pancreatic: pancreatic insufficiency, recurrent pancreatitis
c. Hepatic: chronic hepatic disease manifested by clinical or histologic evidence of focal biliary cirrhosis or multilobular cirrhosis
d. Nutritional: failure to thrive (protein-calorie malnutrition), hypoproteinemia and oedema, complications secondary to fat-soluble vitamin deficiency
Salt loss syndromes: acute salt depletion, chronic metabolic alkalosis
Male urogenital abnormalities: obstructive azoospermia

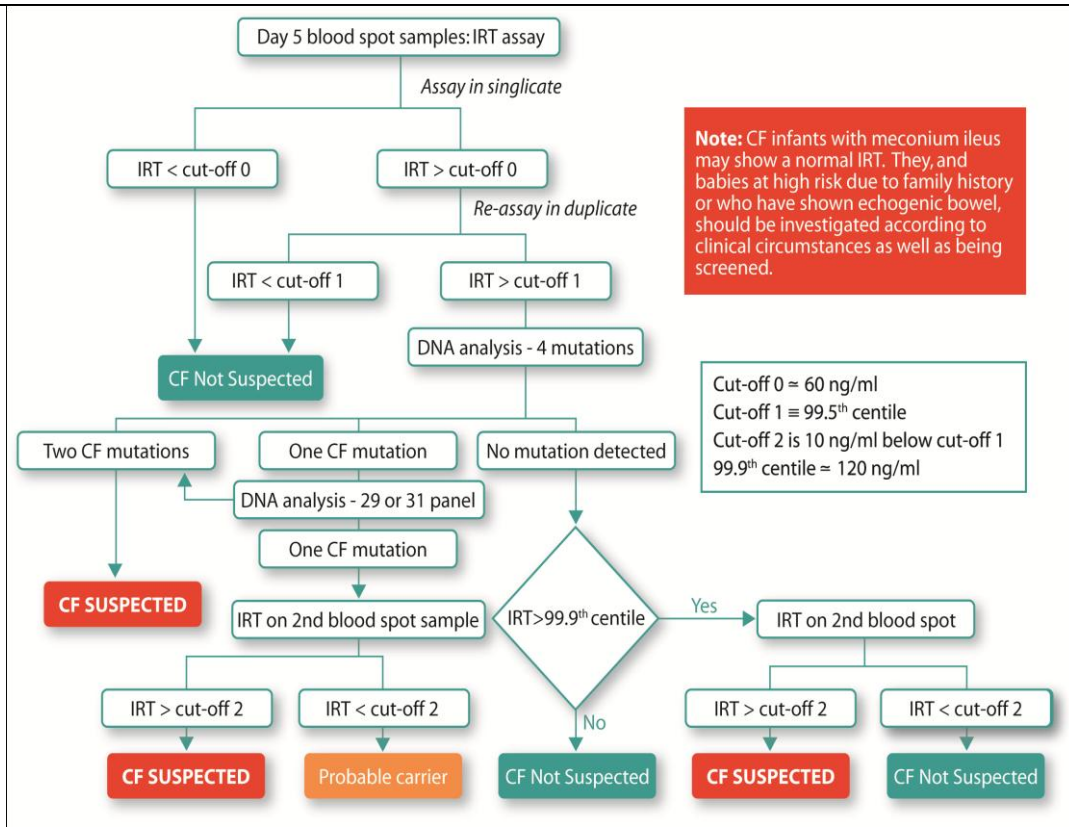
*symptoms are not mutually exclusive.
Table recreated from Rosenstein and Cutting (1998) (8)

Newborn Screening in the UK

In 2005, the UK National Screening Committee described the method of newborn screening for CF as testing the newborn blood spot for elevated immunoreactive trypsin (IRT). If the IRT is elevated, a second test is carried out. If the second test shows IRT to be greater than the 99.5th percentile, a follow-up (confirmatory) DNA test is undertaken. If no mutations or only one mutation is found, a second blood spot is taken and retested for IRT. If the IRT is above cut off 2 (10ng/ml below cut off

1), CF is suspected. (3) Figure 1.1 shows the screening algorithm that has been adopted by all UK screening laboratories.

Figure 1.1: UK National Screening Algorithm (3)



Incidence and Prevalence of CF

Incidence studies

Incidence is defined as the frequency of new cases of an event in a population within a given time period in a population, usually one year. Cystic Fibrosis incidence in the Caucasian population is thought to be 1 in 2500 live births per year. In 2002, McCormick et al (9) looked at the ethnicity of the registered CF population (n=5274 including adults) in the UK and found that 96.3% were Caucasian. Of the

remaining 196 patients registered, 99 were of mixed origin, 88 were from the Indian Subcontinent, 6 black Caribbean, and 3 black African. The authors calculate that one in every 27 UK CF patients has some non-Caucasian origin.

Prevalence studies

Prevalence is defined as the proportion of a population that is affected by a disease at a given time. There are generally 3 ways of calculating prevalence (point, period, and cumulative), but published prevalence estimates are usually point prevalence measures. In 2007, Dodge et al (5) estimated that there were approximately 8284 people living with CF in the UK, of which 4702 were children under 16 in 2003. The CF Registry Annual Data report in 2007 had 8080 patients registered on the registry, of which approximately 44% were under the age of 16. (10)

Funding arrangements

Policy context

While it is not the goal of this section to explain the funding systems of the NHS, it is helpful to understand how Cystic Fibrosis services are allocated money from the health budget. The NHS in England is funded by the Department of Health. Most funding for NHS services is allocated by a complex process of commissioning by the Primary Care Trust (PCT). The process includes assessing population needs, prioritising health outcomes, procuring products and services as well as managing service providers (hospitals, GPs, community services, etc.) Cystic fibrosis, however, is one of 36 specialist services that are covered by the Specialised Services National Definitions Set (3rd Edition). (11) Generally, specialised services are those with small numbers of patients that require relatively expensive care. The

definitions identify activities that are regarded as specialised and therefore need collaborative commissioning processes between groups of PCTs in order to make providing services and training specialist staff cost-effective.

Each Strategic Health Authority (SHA) has a Specialist Commissioning Group (SCG) which provides specialist commissioning for a population of around five million people. SCGs replaced the various ad hoc commissioning arrangements that were in place previous to 2007. The National Specialist Commissioning Group (NSCG) provides coordination of the commissioning done by specialist commissioners at the SHA level. The roles and responsibilities of the National Specialised Commissioning Group (NSCG) are:

- To provide oversight and coordination of commissioning undertaken by SCGs where the specialized service has a catchment/planning population which is bigger than that of a single SCG
- To support supra-SCG decision making. This is subject to endorsement by the 10 SHA Chief Executives and the implementation of the decision is returned to individual SCGs
- To facilitate and encourage collaborative working across and between SCGs, and between SCGs and the National Commissioning Group (NCG) through the initiation of joint projects, development of protocols and sharing of best practice
- To provide oversight of national commissioning of highly specialized services and agree the annual commissioning and management budget for the NCG (subject to endorsement by the 10 SHA Chief Executives)

- To work closely with the NCG to provide concerted advice to Ministers each year on which services should move into, or leave, the portfolio of services to be nationally commissioned, and on proposed designation of centres to provide these services
- To advise PCTs on the commissioning of specialized services

The NCG replaced the National Specialist Commissioning Advisory Group in April 2007 and commissions services on a national basis for a specific group of extremely rare conditions or very unusual treatments, including CF. The NCG is a standing committee of the NSCG and is supported by the National Commissioning Team, which has transferred from Department of Health (DH) to NHS London to undertake the day-to-day commissioning work and national support.

Most services commissioned by the NCG relate to a condition where the national caseload is less than 400 people. Examples include heart and lung transplantation and secure forensic mental health services for adolescents. The total annual budget for the NCG in 2007 was £346 million. This was previously a central DH budget but from 2007-8 funding has been returned to PCT baselines and transparently levied on a fair shares basis.

Ministers continue to have the final decision on the designation (and de-designation) of nationally commissioned specialized services, based on recommendations from the NCG and NSCG. Regulations were altered in March 2007 to allow SHAs to commission those specialized services listed on a schedule. An inter-authority agreement between the SHAs sets out delegated responsibility for NHS London to commission and contract for specialized services on behalf of the NHS in England.

NHS London is the host employer for the National Commissioning Team. However, while the funds are commissioned differently, Hospital Trusts and CF specialist centres are still using block contracts[4] for paediatrics and respiratory diseases to fund the care of CF patients at the time of writing, which ties into the current "payment by results" scheme that the NHS began using in 2007. The Department of Health states that "[t]he aim of Payment by Results is to provide a transparent, rules-based system for paying trusts....Payment will be linked to activity and adjusted for case mix."

The CF Trust, however, has successfully argued to the Department of Health that a payment by result system is not fitting with CF, in particular for paediatrics, where a hefty amount of drugs are given to 'healthy' patients in order to keep them well. The CF Trust is instead recommending to the government that a more realistic approach would be for CF "Packages of Care" tariffs that would be associated with disease severity bands. (13) Currently, these recommended bands apply to all ages of CF care and are set out in Box 1.1. The banding system laid out in this box has been under review and in March 2009, the CF Trust submitted a document outlining a costing/banding approach to the Department of Health, which, if approved, would provide a mandatory national tariff based on an annual banded package of care

[4] A block contract is when a hospital receives a flat contract to care for a patient population regardless of the actual care given. This is an alternative to a cost-per-case (payment by results) contract which where a hospital is paid based on the cost of the medical services provided. 12. Chalkley M MD. Choice of contracts in the British National Health Service: An empirical study. Journal of Health Economics. 2008;27(5):1155-67.

starting in 2011. The costs associated with each band were derived from the adult study carried out by Robson et al (14) in 1992. (13)

Box 1.1: CF Trust banding recommendations

Diagnostic Year

Some patients may be admitted to hospital after they have been diagnosed whilst most will be cared for as outpatients. The outpatient input will be intense and they will spend a lot of time with doctors, CF nurses, physiotherapists and dieticians learning how to manage their CF.

Band 1

Patients who come only to outpatients, receive outpatient care in terms of input from physiotherapist, doctors, social workers, dieticians etc. They may receive nebulised antibiotics and courses of oral antibiotics from time to time and they receive regular pancreatic enzyme supplements and vitamin supplements as do 95% of CF patients

Band 2

Patients who receive the above and in addition receive outpatient intravenous antibiotics up to 3-4 times a year. They may occasionally be admitted. The input as outpatients may be more intense. Some of them will only be kept in Band 2 level by intensive and sometimes expensive drug treatment to keep Pseudomonas aeruginosa at bay.

Band 3

Similar to 1 and 2 but essentially intravenous antibiotics are received as an inpatient 3-4 times a year. They may also have Diabetes, require feeding gastrostomies, and require a higher input overall.

Band 4

Patients with severe disease, who will come into hospital at least 3-4 times a year for intravenous antibiotics, and have increasingly disease severity. They may have Diabetes and more resistant organisms. They may be under consideration for transplantation.

Band 5

These patients have usually been in Band 4 for at least a year and need to stay in hospital for 4-6 months throughout the year whilst awaiting transplantation or receiving palliative care. They are unable to go home because of oxygen dependence, nocturnal ventilation and feeding gastrostomies and need intravenous antibiotics every day, sometimes for 2-3 years, although on average, these patient's life expectancy is usually no more than a year to 18 months.

CF Trust Costing and Banding Document (2)

Treatment of CF

Setting:

Paediatric CF care in the UK is undertaken by multi-disciplinary teams in 24 CF Specialist Centres (SC) (usually based at a large teaching hospital) or at a CF clinic (usually located in a district general hospital) as recommended by the CF Trust. (4) The Specialist Centre provides a service to either children or adults (occasionally both), and usually serve between 50 and 100 patients for all their care. They also offer shared care facilities for those patients who receive most of their care at a CF clinic located in a general hospital local to the patient. However, a CF clinic should offer the same standard of care as the Specialist Centre. A CF team in a SC consists of at least a consultant paediatrician, clinical nurse specialist, physiotherapist, dietician, social worker, psychologist, pharmacist, clerk and secretary. The aim of this system is to provide the best care possible for the patient.

The CF Trust recommends that each CF patient has one annual review and at least one outpatient clinic visit at least every 3 months. For infants, the CF Trust advises that ill babies are seen every 2 weeks, and for those who are thriving, every month. In a clinic, it is recommended that height and weight are checked, a sputum culture or cough swab is taken, a consultation and physical exam with an expert doctor, spirometry (for patients over 5 years of age), a chest x-ray (if a noted fall in lung function), a review by the physiotherapist and dietician, as well as access to the specialist nurse, psychiatrist and social worker.

An annual review is carried out to ensure that optimal care is being provided for the patient. The review is generally provided by the specialist care team and covers all aspects of care. In addition to the care that is provided in clinics, the annual review carries out in-depth and detailed reviews and reports of progress of the patient by the entire CF team; this includes a full clinical work-up (spirometry, Swachman Score, chest x-ray, sputum culture or cough swab, blood count, electrolytes, liver function test, vitamin tests, Pseudomonas antibody levels, intestinal absorption tests, blood glucose test, urinalysis).

Of course, annual reviews, clinic visits, in-patient and outpatient visits are for the express purpose of treating and preventing illness in CF. As a multi-organ, multi-symptom disease, care is often difficult to manage and there are many different aspects of care to consider. Below is a brief overview of the main components of CF care that patients of all bands receive, although the intensity of the treatment can change as disease severity changes.

Nutrition:

Growth failure and weight loss are two effects commonly seen in patients with CF. Poor nutrition is often associated with diminished immune function, therefore making it more difficult to fight off infections. Until the late 70s, this was seen as an inevitable effect of the disease. However, trials in Toronto, Canada were able to show that a high calorie, high fat diet could stabilise or even improve growth patterns and increase survival. (15, 16) Today, most CF centres in the UK follow this high calorie diet, supplemented by pancreatic enzymes in all disease severity bands (17).

In patients where increased intake has not been successful for weight or growth gain, an alternative is to use pre-mixed supplements. In some cases (usually in band 3 or above), invasive nutritional interventions are necessary, such as nasogastric feeding with bolus, or more long-term therapies like gastrostomy or jejunostomy. These techniques have seen positive results for both growth velocity and overall nutritional status. (17)

Antibiotic treatment:

As nearly 90% of CF mortality is related to chronic pulmonary infection, treatment is usually directed at preventing or identifying and then eliminating bacterial infection from the lungs. The most common bacterial infections are from *Staphylococcus aureus*, *Haemophilus influenzae*, *Pseudomonas aeruginosa* and *Burkholderia cepacia*. Indeed, it has been shown that prevention of chronic *P. aeruginosa* infection will prevent or slow deterioration of the airways. (18) (19)

There is still debate over the best method to prevent and treat bacterial infection in CF. However, it is generally agreed that aggressive antibiotic therapy is the best method of fighting infections. Generally, the CF population is treated more often, for longer and on higher doses of antibiotics than the general population. As such, CF patients need intense monitoring of their drug regimens to prevent toxic side effects.

In the past two decades, there have been major advances in CF therapeutics. One of the most important advances in therapy was the discovery from clinical trials that chronic *P. aeruginosa* infection can be reduced, or at least delayed, by early treatment with nebulised colomycin and oral ciprofloxacin (20-22) or nebulised tobramycin. (23) (24)

Physiotherapy:

A key element of treating Cystic Fibrosis is physiotherapy. Various therapies have been developed by physiotherapists to help clear the airways of excess bronchial secretions that are common in CF patients. There are several airway clearance techniques (ACT) that are used by physiotherapists. Upon confirmed diagnosis, the recommendation is for twice daily postural drainage and percussion. This has been shown to possibly slow disease progression as well as improve long-term adherence to physiotherapy. (25) There is little scientific evidence to support the prescription of ACT, but it is still included in the care of CF patients. (26)

Mucolytics and emerging therapies

Mucolytic drugs are used to help clear excess mucus in the lungs by restoring the periciliary layer. Rogers (27) argues that mucolytic drugs should aim to change both viscosity and adhesiveness in order to effectively treat mucus hypersecretion. Currently in the UK, hypertonic saline is used to improve mucociliary clearance by restoring the periciliary layer. Recombinant human deoxyribonuclease I (rhDNase or dornase alfa) is a commonly used enzyme that helps to reduce sputum viscosity in children over 6 (and adults).

Cystic Fibrosis Related Diabetes Mellitus

CF related diabetes (CFRD) rises as age of survival increases. Reported prevalence of CFRD in children with CF under 10 years of age is low, however after 10 years of age, there is an age related increase in the prevalence of 5 percent per year. The CF Trust recommends that the routine use of the Oral Glucose Tolerance

Test (OGTT) and serial glucose monitoring is the most reliable method of screening for CFRD and should be clinical practice for patients over the age of twelve. Insulin treatment is indicated. (28)

This discussion of the treatment for CF is not exhaustive, but has attempted to highlight the main treatment regimes involved in most CF patients' care. It is clear that the treatment of CF is often difficult, and can vary between patients. Therefore, a concerted effort must be made by every member of the CF team, whether at a CF clinic in a local hospital or at the Specialist Centre, to undertake a coordinated, aggressive approach to treating CF in order to give patients a better quality of life.

Economic Analysis

Introduction

Economic analysis in health care has become widespread. This is not necessarily surprising as all health systems face tough choices about where to place limited resources to improve the health of the population. Often, there are various methods of achieving a certain health related outcome. Depending on the type of economic and/or health question, different types of economic techniques can be used to help decision makers compare interventions and hopefully make an optimal choice.

In the UK, CF is considered a specialist service, and is therefore competing with other specialist services for limited funds. It is important to establish the economic viability of the investment in CF services to the health service and to ensure that the patients are getting the best care for the money available. One aspect of care that is not covered by this thesis is the indirect costs of care (

In this thesis, I have attempted to describe the economic picture for CF in terms of the cost of care in the Eastern Region of the UK and to use the data gathered in this thesis and from the literature to assess whether or not newborn screening is cost effective. To do this, I used two particular health economic methodologies: cost of illness analysis, and cost-effectiveness analysis. These are described briefly below with a more detailed description in chapter 3.

Cost-of-illness studies

Introduction

One of the main goals of this thesis is to describe the cost of paediatric CF care in the Eastern Region of the UK. To date, only one study has examined the paediatric costs of care in the UK context (29) (see Chapter 2).

Background

Cost-of-Illness (COI) studies are not full economic evaluations as they do not compare interventions. Rather, they are used to gather information and describe the total cost (economic burden) of caring for people with an illness. COIs are mainly used to highlight health problems by describing measures of occurrence, morbidity/mortality, and overall economic cost to society. A typical aim of COI studies is to influence the public health agenda and subsequent resource allocation. (30) COI studies also answer important questions such as:

1) What is the cost of disease on average?

2) In what populations is the disease particularly costly/less costly?

3) How much will it cost in a local vs. broad population?

4) How do costs behave with severity of illness, or over time?

COIs are also helpful for identifying a range of cost items for further cost-effectiveness analyses (31). However, COI studies provide no indication of how to achieve efficiency gains or health benefits (1, 32, 33).

Limitations of cost-of-illness studies

Because COIs generally only take account of the resources used, and not the consequences of using those resources, they have been criticised by welfare economists for not adequately aiding the decision making process (34, 35). In the past, COIs have also been criticised for using inconsistent methodologies. This makes comparisons between one disease and another very difficult and undermines their usefulness in "ranking" diseases. Drummond (1) suggests that researchers carrying out COI studies can use the following to improve the usefulness of the COI:

- reporting the direct and indirect costs separately and in aggregate;
- listing the separate components of direct costs so those associated with the most economic burden can be easily identified;
- predicting the economic burden of disease into the future to help health care planning; and
- investigating the impact of different treatment practices.

Cooper et al (31) also identifies problems in modelling the uncertainty around estimates of cost of illness, arguing that the costs be difficult to model using traditional methods due to strongly (right) skewed distributions and a significant number of cases with zero cost. However, the authors illustrated a Bayesian

approach to modelling costs, which they argue may alleviate these problems. A longer discussion of Bayesian methods is located in Chapter 3.

Conclusions

Well constructed, patient based COI studies can provide valuable data for decision makers[5] to use in decision analyses, for example. Although making decisions on a COI alone is not ideal. Patient-based studies are particularly useful as they can, if done well, point to areas of health care that could be refined to make more efficient use of resources. (36) (31)

Cost-effectiveness analysis

Introduction

Another goal of this thesis is to look at the cost-effectiveness of newborn screening for CF within the UK population. Cost-effectiveness is generally used to answer the question of what method of input is most efficient to achieve the desired output and what is the incremental cost of achieving an additional gain in health.

Background

Cost-effectiveness analysis (CEA) is a form of economic evaluation that examines both the costs and the consequences of a specific health technology. CEA is generally used to compare different interventions used to achieve the same specific (health) outcome. CEAs generally have a narrower viewpoint than cost-utility

[5] Segel 30. Segel J. Cost of Illness - A Primer: RTI International RTI-UNC Center of Excellence in Health Promotion Economics2006. gives examples from the USA of state smoking cost of illness estimates were used in state lawsuits against the tobacco industry to recoup Medicaid losses, and the CDC using cost of injury estimates to propose specialist injury centers. Segel goes on to quote from a study by the Bureau of Labor Statistics that found COI studies useful in educating employers and health professionals about prevention of workplace injuries and fatalities.

analysis and cost-benefit analysis, therefore CEAs rarely include indirect costs or analyse the impact on the wider economy of an intervention. Therefore, CEAs measure technical efficiency, in other words, identifying the most efficient way of achieving the outcome of interest (1, 32).

CEA relies on information collected prospectively (such as in a randomised clinical trial (RCT) or retrospectively. Ideally, information on resource use and costs are collected prospectively and on an individual basis, therefore improving data quality and minimizing bias. Data from a population perspective at individual patient level enables analysis of costs and the variability in costs, which greatly enhances the generalisability of the findings (1). This is a particularly important point in a disease area such as CF, as the prognosis is variable between patients.

Within health economics literature CEAs are often seen with quality adjusted life years (QALYs)[6] as an outcome measure. These studies are actually cost-utility analyses (CUA), and are a natural extension of CEA. However, the terms are often used interchangeably within the literature. CUAs measure the incremental cost of a programme against the incremental health outcomes expressed as a QALY or possibly a disability adjusted life year (DALY). Unlike CEAs, CUAs take a more multidimensional approach and incorporate the notion of value to the health outcome.

---

[6] A QALY is derived from life expectancy, and a measure of the quality of the remaining life years. The measure is used to assess benefits gained from interventions in terms of health related quality of life and the survival of the patient. For a more detailed description of how a QALY is derived and a discussion of utility based preference measures, see chapter 3.

As with any economic analysis, uncertainty exists around estimates of effect sizes as in most cases, we are faced with imperfect information. At times, the uncertainty can be large enough to change the findings if one extreme is used over another depending on the decision criteria. Therefore, it is necessary to undertake sensitivity analysis on models of cost-effectiveness (37). Sensitivity analysis allows a researcher to analyse the effect of varying different parameters in the model to assess the individual (or group) impact of those parameters. There are different methods of varying parameters which include probabilistic (when the parameters statistical distribution is known), extreme (when the statistical distribution is unknown), and threshold (when a decision cut-off point is known or needs to be defined) analysis.

Limitations of cost-effectiveness analysis

CEAs are designed to compare different interventions that have the same (health) outcome. Thus, if a health technology has more than one outcome that is of importance, CEA is of limited usefulness. Indeed, there are some that argue that CEA is not necessarily based in welfare economics as it does not consider the overall welfare being gained from different resources. (1, 32)

Another area of controversy emerges when gathering information for the costing of CEAs. Depending on the viewpoint or perspective, omitting certain costs within a CEA can make it difficult to draw economic welfare conclusions about the findings of societal economic efficiency.

Conclusion

Cost-effectiveness analysis can be a very useful tool for decision makers if they are trying to decide between technologies that achieve the same outcome. However,

CEA is of limited usefulness if decision makers are trying to decide on various interventions that have different impacts on different outcomes. In CF, CEAs can be very useful because most interventions are aimed at improving one area of health (e.g. lung function) or another. However, when planning an entire CF service, CEA is of limited usefulness as different interventions, such as pancreatic supplements and rhDNase do not aim to have an impact on the same aspect of the health of a CF patient. In that case, it would be more useful to carry out a cost utility analysis (CUA), which uses a generic measure of quality of life, to find the impact of a service on CF patients.

Ethics

This thesis made use of data that was collected from patient notes and hospital information systems. The ERCFD was granted ethics approval in 2004 to collect data from Cambridgeshire, Norfolk, and Suffolk and for secondary analysis of this data. As this thesis attempted to gather information from Hertfordshire, Bedfordshire, and Essex, further ethical approval was sought. Ethics approval for the data collection and secondary analysis was granted by the Cambridge 1 Research Ethics Committee in November 2008.

Aims and Outline of Thesis

In this thesis, I aim to illustrate the cost of providing care for paediatric patients at the tertiary hospital level in the east of England, as well as undertake a cost-effectiveness analysis of newborn screening versus no screening. The next chapter discusses a systematic review of the existing literature on the cost of paediatric CF as well as the cost-effectiveness of newborn screening. Chapter 3 conducts a theoretical and methodological review in order to inform best practice in conducting

the cost of illness and cost effectiveness analysis. The methods and results of these analyses are discussed in Chapters 4 and 5 respectively. Chapter 6 is a discussion of the overall conclusions of the thesis as well as suggestions of areas of future research.

Chapter 2: Literature Review

Introduction

In order to ensure that patients receive the best possible care, decision makers should have the most up-to-date knowledge of the economics of CF, both from a local, national and/or international viewpoint. First, it is necessary to have an idea of what the entire package of paediatric CF care costs the NHS in order to appropriately plan for the future. It is also important to understand whether the investments for technology, such as newborn screening, are good value for money. Ideally, the evidence on the costs and efficacy of CF care can be generalisable across borders. To help gather knowledge in this area, a systematic review has been carried out. A well constructed systematic review aims to:

- Specify its purpose
- Clarify whether the review is exploratory or testing a specific hypothesis
- Set criteria for inclusion or exclusion of literature
- Comprehensively search for appropriate material
- Identify which studies/literature should be included/excluded
- Discuss quality of methods used or evidence cited in papers included
- Discuss generalisability
- Identify important differences in study methodology

- Summarize/synthesise findings (if appropriate)

This systematic review was designed to be exploratory and has three main aims:

1. To critically analyse the literature for information on the cost (or cost-effectiveness) of technologies used in paediatric CF care and technologies used in the process.

2. To critically analyse the literature for information on the cost-effectiveness of newborn screening for CF.

3. To critically analyse the literature for economic information on the organisation of care.

The main outcomes of this literature review will be to inform an investigation into the cost of illness and cost-effectiveness of CF programmes in the Eastern Region of England.

Economics of CF

Methods

A search strategy (Appendix B) was designed to be as wide in scope as possible in order to catch all the relevant papers. The search strategy used the following online databases: PubMed, Embase, National Health Service Economic Evaluation Database (NHS EED), and the Cochrane Library. These databases were chosen as

they cover both medical and economic journals and sources of evidence. Search terms had to be in the title or abstract and the search was limited to studies after 1990. Searches were limited to the last two decades in order to capture information that would be relevant to modern CF care. No limits were placed on language and MeSH terms were used.

Box 2.1: Search terms (all combined with Cystic Fibrosis)

PubMed and Cochrane Library	**Embase**
Cost$	Used NHSEED designed search strategy with limitation of paediatric/pediatric
Cost of illness	
Burden of illness	
Cost-effectiveness	**NHSEED**
Cost-utility	Newborn Screening
Cost-benefit	Shared Care
Newborn screening	Pediatric
Shared Care	Paediatric
Paediatric	
Pediatric	

To be included for data extraction, the title or abstract must contain the following:

1. Confirmation that the study is on paediatric CF patients

2. Some type (including reviews) of economic, cost and/or cost-effectiveness information on

a. Technologies used in treating CF

b. Newborn screening for CF

c. Organisational arrangements of care

Papers meeting the inclusion criteria were obtained and reviewed by JJ. The retrieved papers were classified using the criteria of the NHS EED and the BMJ working party on peer review of health economic literature (38, 39). The NHS EED database uses trained health economists to write abstracts assessing the quality, strengths and weaknesses of published health economic evaluation work. For economic evaluation papers, the NHS EED pro forma was used to extract data (See Appendix C), but JJ conducted the assessments 'blind' to the content of the found NHS EED abstracts. As a validation exercise, JJ compared the extracted data to the completed NHS EED abstracts (where available and appropriate).

Data from papers not classed as health economic evaluations (reviews, cost-descriptions, service organisation, economic methodology discussions) were discussed descriptively and any cost data was extracted and compared if appropriate (40). Where studies were not methodologically comparable, a simple descriptive list approach was carried out. All results were summarised in 2010 US $ using the OECD PPP and GDP deflator methods (41). If a price year was not stated, the year of publication was used to convert prices.

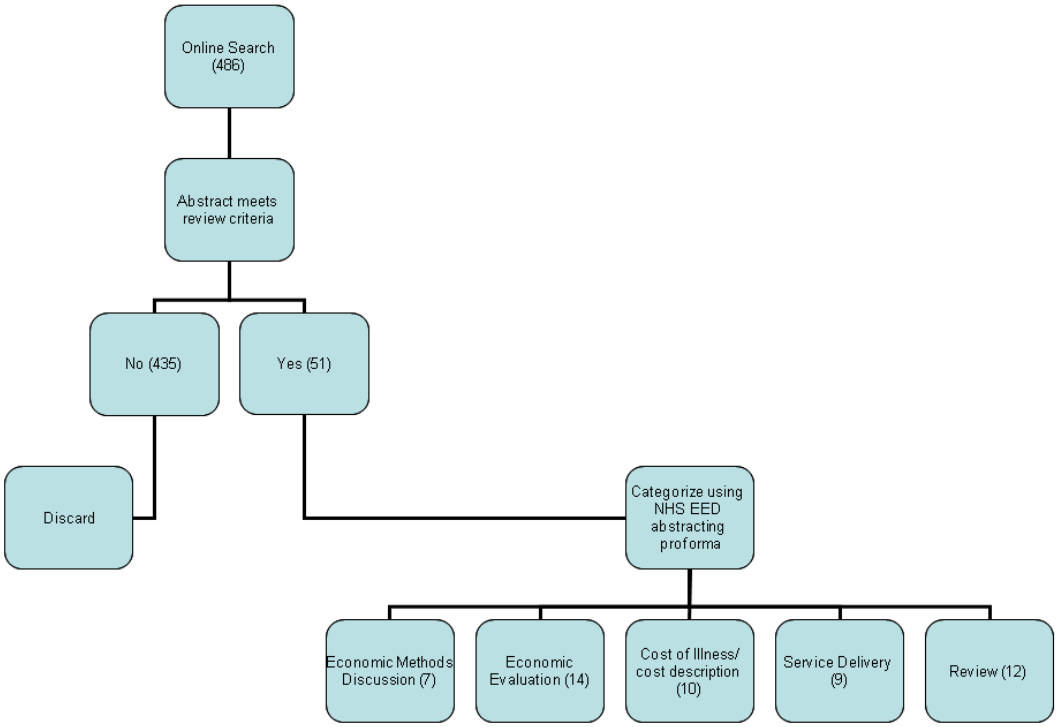
The references of papers meeting further criteria for review were hand-searched for grey literature sources that met the original inclusion criteria. The websites of UK and USA charitable organisations dealing with CF (i.e. The CF Trust in the UK and the CF Foundation in the USA) and the Department of Health were also searched for relevant literature. This was done in order to minimise publication bias as grey literature can often report negative findings that are not published (40). Citation searching for papers which cited the retrieved articles was not undertaken.

Results

The literature searches yielded a total of 486 unique references from PubMed, Embase and the Cochrane Library. Figure 2.1 illustrates the systematic process and indicates the numbers of papers at each stage. Fifty-one of those abstracts met both the inclusion criteria for further review. After retrieving and reading through the 51 papers, 12 of the studies were classified as reviews (42-53), 6 were cost-of-illness descriptions (29, 54-58), 4 cost-descriptions (55, 59-61), and 14 full or partial economic analyses (44, 62-74). There were 9 studies describing service delivery issues (13, 75-82) and 7 papers describing a health economic methodology with a CF example (83-89). The discarded papers usually did have the keywords in the title or abstract, but upon detailed reading of the abstract, the authors did not appear to actually undertake any kind of economic analysis. Checking the NHS EED for all CF related references and comparing bibliographies of included studies has minimised bias for economic evaluations, methodology, or costing studies as all but 2 references had entries on the database and there were no additions to the original list, indicating the search was thorough. Other studies, such as service delivery

studies, are prone to reviewer bias. There were no additional documents found on the charity or Department of Health websites.

Figure 2.1: Systematic review flow chart



Reviews

The twelve review papers did not contain any additional primary studies not identified by our initial search strategy. Krauth et al (47) carried out a review of the literature of cost of illness studies pertaining to CF. Seven reviews focused on the costs associated with screening for CF in newborns (43, 50) (42, 48, 49, 51, 53). Three reviews focused on the impacts of rhDNase (44-46). One review focused on issues surrounding quality of life estimates in CF children (52).

Study results

Table 2.1 summarises the papers dealing with cost of illness. Table 2.2 summarises the papers that were cost-descriptions of a technology. Table 2.3 summarises the

studies deemed cost-consequences. Table 2.4 summarises the partial or full economic evaluations. All costs were from the health service perspective unless otherwise stated. All JJ reviewed papers met the same conclusions as the NHS EED abstracts where available.

Study Population

Ten studies examined populations in North America (USA/Canada), 9 studies examined the UK population, and 6 studies looked at continental European populations (Netherlands, Germany, France, Italy). Demographic data was generally included in studies, or in some cases referred to a previously published article describing the population in detail.

All but two studies defined their population by using clinic-based or health maintenance organisation databases, otherwise authors used hypothetical cohorts based on parameters derived from the literature and expert opinion. Data based on clinic-based populations may benefit from being more precise as they are based on individual level data, but can suffer in terms of generalisability across settings, as the sample may not be representative of a population for whom services are commissioned (in the UK, this is currently mainly based on residential areas of service users). It was unclear in all papers but Sims et al (29) that this was the case, therefore caution is needed when interpreting the results. Another issue when looking at economic analysis in health care is that non-generalisability arises not necessarily from patient characteristics, but from clinical practice. However, when entire populations are under consideration (such as in universal newborn screening) and interventions are the same across settings, results could potentially be

generalisable. See chapter 6 for a discussion of the generalisability of results from this thesis.

Data collection - Resource use data

All of the papers used secondary data collection, which in this instance can be defined by using national or community databases, insurance claims or literature reviews to gather information on resource use. The studies on newborn screening for CF used data from the general population to make comparisons, while all others used CF specific secondary data.

Data collection – Unit cost data

Unit costs enable total economic costs to be estimated from health utilisation data. The literature reviewed here mostly used unit costs from health provider charges (i.e. health maintenance organisation tariffs or hospital costs), pharmacy prices and laboratory prices to find direct medical costs. Often charges could exceed costs, allowing for equipment replacement, future technologies, inflation, and pricing policies (1). Other sources of unit cost data were from national tariffs (such as the NHS Reference costs in the UK). No studies attempted to derive productivity costs such as time off work, or wages lost.

Data collection – Outcomes data

In those studies that collected and reported outcome data, information was collected by using patient databases and/or individual patient data. Newborn screening studies typically collected data on cases detected. However, one study also calculated life years gained (LYG) by assuming a gain of 40 years per CF death

averted by newborn screening (74). Another newborn screening study assigned utility values derived from the Quality of Well-Being Scale[7] to each health state in a model and multiplied survival time by the utility to produce quality adjusted life year (QALY) estimates. Those studies looking at therapies such as rhDNase and tobramycin collected data on hospitalisations and lung function (usually FEV1).

Cost of Illness:

All the included studies used individual patient level data to derive estimates of resource use and relevant national or local tariffs to estimate costs. Sims et al is the only study to have taken place in the UK, and is also the most recent (2007) estimate of the cost of caring for a paediatric CF case. A study in the United States by Ireys et al (56) found that the mean direct cost of caring for paediatric CF patients was $18802 per year. Johnson et al (92) found that in Canada, the mean direct cost of caring for children with CF was $4969 per patient.

In 1991, Wildhagen et al (58) in the Netherlands found that on average, direct hospital-based costs (hospital days, consultation, diagnostic tests) of caring for

---

[7] The Quality of Well-Being Scale is a preference-based, quality of life instrument which utilises a questionnaire to assign values according to an individuals' functioning and symptoms 90. Kaplan RM, Anderson JP, Wu AW, Mathews WMC, Kozin F, Orenstein D. The Quality of Well-Being Scale: Applications in AIDS, Cystic Fibrosis, and Arthritis. Med Care. 1989 March;27(3):Supplement.. Subsequent to the Cystic Fibrosis Quality of Life scale 91. Quittner AL, Schechter MS, Rasouliyan L, Haselkorn T, Pasta DJ, Wagener JS. Impact of socioeconomic status, race, and ethnicity on quality of life in patients with cystic fibrosis in the United States. Chest. 2010;137(3):642-50., the Quality of Well-Being Scale provided the only health-related quality of life measure for Cystic Fibrosis patients. A further discussion of preference-based measurement techniques is conducted in Chapter 3.

paediatric CF were $14985, but that total costs (including general practitioner visits, help from relatives, travel expenses, and special aids) were $17874 per patient per annum. It was unclear, however, how the non-hospital resource use and costs were collected. Nevertheless, this indicates that non-hospital based costs are significant in the overall cost of caring for a CF patient. Wildhagen et al (57) also carried out a questionnaire-based study on the same patients that found the mean non-hospital care costs for paediatric patients with CF annually was $11429.

Information on the cost of paediatric cystic fibrosis care in the UK is very limited. In this review, only 1 paper discussed the costs of paediatric care in the UK. Sims, et al (29) discuss the differences in costs between the screened and non-screened paediatric CF population in the UK. The authors used a retrospective cohort study design using the UK CF Database, excluding those patients who were diagnosed by meconium ileus or family history, as these patients were likely to present early irrespective of newborn screening. The authors found significant differences in the cost of therapy between patients who were diagnosed by newborn screening and those diagnosed clinically. The authors report that newborn screened patients had a mean cost per annum of $8194 versus $13613 for clinically diagnosed patients. The authors also estimated that drug cost savings for the UK could have off set the estimated cost of adding CF to the UK's national newborn screening service.

Cost descriptions

Four studies found in the literature review described the costs and/or the outcomes but did not perform an economic evaluation. All studies (55, 59-61) described the

costs of screening programmes. Three studies were based on US data (all from the state of Wisconsin), and one on Irish data.

Gregg et al (59) describes the costs of two methods of newborn screening in Wisconsin. One was using IRT with confirmatory sweat test and the other using a confirmatory genetic test. The analysis was carried out as part of a major clinical trial and therefore based on individual patient data. Costs are described in terms of cost per case detected and only take into account the cost of the tests. The authors found that it cost an estimated $13917 per CF patient detected using IRT/Sweat test, and between $15539 and $14678 using IRT with the confirmatory DNA test depending on the test used.

A 2003 study from the Wisconsin trial was carried out by Lee et al (60). The authors looked at the costs arising from a screening programme using the IRT/Sweat test methodology. They estimated that if all costs arise only from the screening programme alone, rolling out the programme nationwide would cost $1184551, or $11750 per diagnosis, or $2.91 per birth. Rosenberg et al (61) also analysed the data from the Wisconsin trial and looked at costs arising from a screening programme using IRT/DNA (DeltaF508 only) assay versus IRT/DNA (CFTR) assay. The authors found that screening for only the DeltaF508 mutation only yielded a cost of $10210 per diagnosis, or $3.01 per birth. Screening using a DNA test that screens the entire CFTR for mutations costs an estimated $13009 per diagnosis or $4.53 per birth. These studies were all carried out by similar teams working on the Wisconsin trial. Therefore, it is not necessarily surprising that their results are similar for the cost of a screening programme.

In Ireland, Farrell et al (55) found that the cost of using a sweat test for a CF screening programme was approximately $4221 per CF diagnosis. The authors did discuss that because Ireland has the highest incidence of CF in Europe, and that the test is already routinely used in practice meant that their cost estimate was lower than most estimates of screening programmes for CF in the United States, where CF screening is not always routinely done, and therefore potentially have high set-up and initial running costs.

Cost-consequences:

Cost-consequences studies give a summary measure of costs and of outcomes, but do not combine the results to get a cost-effectiveness ratio or perform any marginal analysis, and therefore are only partial economic evaluations. This review found seven cost-consequences analyses in the literature. Four of the studies focused on the usage of the mucolytic drug rhDNase (also known as Pulmozyme or DNase), two studies looked at the impact of tobramycin, and 1 study looked at four different types of nebulizers to deliver bronchodilators and antibiotics to CF patients. All the studies took the perspective of the healthcare system.

Menzin, et al (67) carried out a multinational analysis of the use of rhDNAse in France, Germany, Italy and the UK. The authors used clinical trial data (85) and linked costs to resources used to fight respiratory tract infections. While the authors collected data on lung function, the analysis did not include this information. The cost of rhDNase was not known at the time of the trial, and was added to the analysis after it was on the market. The authors found that the mean cost savings of using rhDNase for respiratory tract infection related care was $1462 in France,

$1275 in Germany, $1180 in Italy, and $942 in the UK. The authors did acknowledge that the differences in the costs could be attributed to the differences in the health care systems. Oster et al (68) used the same clinical trial data and found that the mean cost for respiratory tract infection related care for the control group was $7436, $5671 for the group using rhDNase once daily, and $6731 for the group using rhDNase twice daily. The authors concluded that compared to normal therapy, rhDNase reduced average cost of RTI-related care by $1112-2298. However, these estimates did not include the cost of rhDNase. Neither Menzin et al or Oster et al included any discussion of uncertainty around their estimates. No sensitivity analyses were carried out. Neither study discussed issues around generalisability.

Suri et al (73) also carried out a cost-consequences analysis of rhDNase use in CF patients. The authors compared the use of alternate day rhDNase use, daily use of rhDNase, and the use of hypertonic saline (HS). The team found that the total health service cost was $10650 for the daily use of rhDNase compared to $8015 for HS. In comparing daily versus alternate day rhDNase, the authors found the mean costs were $3272 and $1604 respectively. The authors conclude that because of the cost of administering daily rhDNase was high, that alternate day therapy was better, and that either rhDNase option was better than HS.

In 2003, Iles et al (64) used patient records to carry out a retrospective study on the use of Tobramycin (TOBI) inhalers on the paediatric patients in the east of the England. The authors described the costs of hospitalisations and drugs from multiple centres. However, an explicit comparator was not given. The mean cost of therapy for patients was $51257 when TOBI was used and $44266 for the year

preceding the TOBI therapy. This represented a $6991 ($6737-$19735 95% CI) mean difference. The authors found that, while TOBI was more expensive overall, the use of the drug did lead to a reduction in hospital stays and the use of other drugs. Lelorier et al (66) found that in two provinces (Quebec and Ontario) in Canada, the mean weighted costs savings per patient for using TOBI was between $1984-$7967 in Quebec and $1362-6848 in Ontario. The comparator was a placebo. No price year was stated. Neither study carried out sensitivity analysis. As Iles et al and Lelorier et al did not use the same comparator, some caution should be taken in comparing the results of the studies. However, both papers found that the use of TOBI reduced the number of hospitalisations and a reduction in the number of other drugs taken, both of which increase quality of life for CF patients and could perhaps result in a savings if the reductions were large enough.

Cost-effectiveness analyses:

The search yielded two published studies analysing the cost-effectiveness of different newborn screening strategies for CF and four analyses of different methods of giving rhDNAse to children. One screening study was carried out in the UK, and the other in the Netherlands. Four treatment studies were carried out in the UK, and one in Canada.

Both of the screening studies used hypothetical patient cohorts to model the cost and effectiveness of different methods of newborn screening. In 1998, Simpson, et

al (70) used a decision tree with Markov processes[8] to model a hypothetical UK health authority newborn population and compared costs and outcomes of using two stage immunoreactive trypsin test (IRT) screening with confirmatory genetic testing versus no newborn screening. The authors modelled the disease progression by starting all cases as pre-symptomatic, with a given chance of moving into a symptomatic disease state, a severe lung disease state, and finally death. The authors excluded those cases diagnosed by meconium ileus or family history as these cases would've received the same prognosis under both strategies. In the 'no screening' strategy, the patients would be diagnosed clinically by the presentation of symptoms (similarly for false negatives in the screening strategy). The authors assumed that the difference in the annual transition probability of remaining pre-symptomatic in the screening arm was 69%, and 59% for those diagnosed clinically (resulting in a delay of the emergence of symptoms of 6 months).

To populate the model parameters (probability, cost data, and quality of life estimates for the three health states), the authors used evidence gathered from the literature and assumptions. For the population incidence of CF, the authors used information from Dodge et al (93) (from 1997). The probability of transitioning into each state was based on UK age-specific survival rates as well as hazard rates derived from Dodge et al (93). The model was set up to allow each person to

[8] A decision tree is a method that can be used to model decisions and their possible impacts. The structure of the tree depends on chance events (where the tree branches), with the resource costs and utility associated with that chance. A Markov process allows a decision tree model to take into account time. Decision trees and Markov processes are discussed in more detail in Chapter 3.

transition through each state before eventually dying of CF-related respiratory symptoms. Evidence on the sensitivity and specificity of the screening test (both IRT and DNA) the authors extracted data from Robson et al (14). Cost information (counselling time, IRT, DNA analysis, sweat chloride test) was taken from various local and national sources (e.g. Annual financial returns for NHS Trusts). Costs of taking the blood spot itself were assumed to be 'sunk costs' in the National Screening Programme where other routine newborn screening already requires this procedure. Disease state costs were based on the findings of Robson et al (14) from a population of 161 adult patients at a UK CF unit. The authors excluded non-CF-related costs, and discounted future costs at 6% per year. Quality of life information was derived from a study on an adult CF population by Congleton et al (94) which used the Quality of Well Being scale to derive utilities. The authors ran three different scenarios based on survival information (conservative, balanced, and optimistic) and found that for a balanced CF patient survival rate in the UK, the cost per diagnosis was $10192 ($3.47 per screen); cost per case diagnosed clinically was $1771. They also state that newborn screening produces an incremental cost per QALY of $13446.

In 2006, van den Akker-van Marle et al (74) analysed four different methods of newborn screening in the Netherlands. The authors compared 1) IRT screening with a follow-up IRT test for confirmation, 2) IRT with DNA confirmatory testing, 3) IRT with DNA and IRT confirmatory tests, 4) IRT with DNA and denaturing gradient gel electrophoresis analysis (DGGE) as confirmatory tests, and 5) no screening. The authors built a decision analysis model, although unlike Simpson et al, this model

appears not to use Markov processes to accumulate costs and outcomes over a lifetime for different states of health.

The model was based on a hypothetical Netherlands birth cohort of 200000 neonates. The authors assumed a gain of 40 life-years per CF death prevented by newborn screening. The authors appear to have based the gain in life years by screening on an RCT (95) which indicated a reduction of at least 50% in mortality for children with CF diagnosed by newborn screening versus clinical diagnosis. However, it is not explicitly stated how they derived the life years gained in the paper. The authors only included the direct costs from the health care perspective (although a societal approach was stated in the paper) including the costs of IRT, DNA, DGGE and sweat tests, as well as genetic counselling and the cost of adding CF screening to the national screening programme (something that Simpson et al assumed was sunk). Lifetime costs were derived from Wildhagen et al (58). However, the authors state that costs incurred during life-years gained due to CF screening were excluded from the analysis. An annual discount rate of 3% was used.

Both univariate and multivariate sensitivity analyses were carried out. The findings showed that the IRT/IRT method had the most favourable cost-effectiveness ratio of $30501 per life year gained. The IRT/DNA/DGGE strategy had a ratio of $40585 per life year gained, IRT/DNA strategy $47105 per life year gained, and finally IRT/DNA/IRT strategy had a ratio of $48950 per life year gained. The authors concluded that if future reproductive decisions were taken into account, the savings from screening could amount to $2.2 million annually. They also stressed that newborn screening has proven health benefits to the newborn.

Christopher et al (44) carried out a cost-effectiveness analysis of rhDNase therapy in patients with mild-moderate lung disease. Effectiveness, resource use, and cost data were collected from the literature. The authors appraised the literature using the critical appraisal skills programme (CASP) checklist (96). The authors first estimated the life years gained by assuming that annual risk of death exceeded 50% once FEV1 fell below 30% of predicted and that rhDNase slows lung function decline (based on Kerem et al (18)). The model assumed FEV1 declines at a rate of 4.3% per year from birth until age 13, then the rate of decline slows to 2.77% per year, that treatment is started for those patients whose lung function falls below 61.1% predicted, and rhDNase would give patients an FEV1 5-8% higher than if they were not on it, and that death will occur in the year FEV1 falls below 28% of predicted. Life years gained and direct costs (costs of rhDNase and hospital stays) were discounted annually at 6%. Quantities and costs were not reported separately. Some costs were taken from the BNF and from UK NHS Tariffs for the south and west region. The authors found that for all patients, cost per LYG was $101049 (95% CI 39576-109711), and for those patients with FEV1 lower than 70%, cost per LYG was $30796.

In 2003, Grieve et al (62) carried out a study in the UK based on 47 patients to compare the use of daily rhDNAse with alternate day use of rhDNAse with hypertonic saline. The perspective is likely to be the UK NHS, but it was not specifically stated. Generalisability was not discussed. The authors looked at lung function (FEV% predicted) as an outcome measure. Costs and outcomes were linked to individual patient records. They found that the mean incremental cost effectiveness ratio (ICER) of daily rhDNase compared with HS was $204. The ICER

of alternate-day rhDNase compared with HS was $164. The ICER of daily compared with alternate-day rhDNase was $369. The authors state that if the decision makers have a Willingness to pay of $369 per 1% gain in FEV1, the probability of taking rhDNase daily being cost-effective versus hypertonic saline was 91%, and that the probability of alternate-day being cost effective compared with hypertonic saline was 88%. Therefore, they state that for a given willingness to pay ($344) either rhDNase strategy is more likely to be cost-effective and that alternate day therapy may be more cost effective.

This conclusion is supported by two studies by Suri et al (71, 72) . Both studies had similar designs. Effectiveness and resource use data was collected from patients between 1999 and 2000 at two London hospitals. Quality of life was measured by the Quality of Well-Being scale (90). Cost information was gathered from hospital finance departments and national sources. All costs appropriate for the perspective were gathered. The authors compared the use of alternate day rhDNase use, daily use of rhDNase, and the use of hypertonic saline (HS). The team found that the total health service cost was $10551 for the daily use of rhDNase compared to $7941 for HS, the mean difference was $2611 (95% CI: $814-4297). In comparing daily versus alternate day rhDNase, the authors found the mean costs were $10584 and $9634 respectively, with a mean difference of $951 (95% CI: $913-2925). The authors conclude similar results to that of their cost-consequence study. There was little difference in daily versus alternate day rhDNase in terms of costs or benefits, but that both were more effective than HS, and therefore likely to be cost effective.

However, in 1994, Perras et al (69) found that the costs of using rhDNase in Canada far outweighed the benefits (measured as a reduction in future hospital costs) unless

the price of the drug nearly halved. The authors carried out a cost-effectiveness analysis looking at rhDNase versus no rhDNase. The perspective was the Canadian health care system. The outcome measure was the hospitalisation rate and evidence was derived from the literature and one RCT (97) carried out in the USA. The source for resource use was from a clinical database which had data from a cohort of patients from two Canadian CF centres in Vancouver and Winnipeg at 6 months, 12 months and 15.9 months. The authors found that the cost of using rhDNase to reduce hospitalisation by 1 visit was $38876 at 6 months, $15478 at 12 months, and $7211 at 15.9 months. The authors carried out univariate sensitivity analysis to find the threshold value at which implementing the use of the drug would be cost-effective. They found the price of the drug needed to be $16 per dose (half the actual price at that time) and to reduce the number of hospitalisations by 30% in a year to be cost-neutral. The study did not discuss generalisability or uncertainty of their results.

Discussion

It is evident that there are relatively few studies on either the cost of paediatric cystic fibrosis or cost-effectiveness analysis of medicines or technologies used in care, especially from a UK perspective. Overall, the literature indicates that the direct cost of care to the health care system ranges from $4969 to $29849 case per annum at 2010 prices and that newborn screening has been shown to be cost-effective when compared to no screening. There also seems to be limited evidence or agreement on the costs of various treatments (e.g. TOBI and rhDNase) in the literature, and therefore more research on TOBI and rhDNase and the other drugs used in CF care may need to be carried out in order to help decision makers. Similarly, the only

screening study from a UK perspective is now relatively old (1998) and the emergence of new therapies and service delivery methods could alter the findings.

It is apparent that individual study results vary greatly. Reasons for this may include the use of the arithmetic mean which may not represent the most common cost per case due to skewed distributions often associated with cost data[9], different disease severity of patients involved in the study, structural differences between geographical locations (i.e. different health systems), and base year of the study (changes in health care practice, medications, etc.).

In their review, Krauth et al (47) rightly point out that many of these papers omit key cost variables, or restrict their analysis to one variable of interest (i.e. intravenous antibiotics) and therefore are likely to underestimate the cost of caring for CF patients. For example, none of the studies found looked at the cost of providing physiotherapy to CF patients, yet almost all CF patients receive physiotherapy as part of their regular interactions with the health service. In particular, the omission of key modern therapies, such as rhDNAse (first introduced in the UK in 1994 (98)), will heavily impact the cost of care. Only Sims et al (29) and Johnson et al (92) include this cost item in their studies. All of the papers acknowledge that the costs of caring for CF patients vary with disease severity.

Given the limited (either by time or by data) nature of the published evidence, likely candidates for further research would be for the cost-effectiveness of using rhDNase

[9] A more detailed description of the implications of skewed costs definitions occurs in chapter 3 and again in chapter 4.

on different (perhaps younger) age groups as this drug has been shown to be effective, but is still relatively expensive in relation to other CF drugs. Another area of interest is the cost-effectiveness of different methods of service delivery. This has yet to be studied in any detail.

It is also difficult to draw any solid conclusions from the cost-effectiveness studies as they are mostly stand-alone studies, have not been replicated and did not discuss their generalisability. Although both screening analyses advocated the use of newborn screening, there is still limited published evidence as to its cost-effectiveness and the most efficient methods to use in such programmes. Although screening has recently been rolled out in the UK, it is now possible to make sure that this decision has been made wisely using more up to date and relevant epidemiological, resource use, and cost information.

Methodological problems

The wide range in per person costs of CF care in the studies can be attributed to various methodological problems. The positively skewed distribution of cost data could possibly bias the result; therefore, it would have been optimal if all the papers were to have reported the median as well as the mean in order to give a fuller description of the distribution of per person costs. Only Sims et al did this. Also many of the papers did not report standard deviations or other useful summary statistics.

Studies should also be more explicit about the types of costs incorporated into the studies. For example, many studies stated that hospitalisation costs were included, but did not describe what that included. It is apparent that there is a need for a

standardised approach to cost-of-illness studies and that authors should re-examine the guidelines for reporting economic analyses (39).

Given the wide variation in costs, it is interesting to note that very few authors discussed the generalisability of their findings or the problems therein. For example, the economic evaluations of screening programmes may not be easily and directly generalisable due to differing health systems, currency fluctuations, and treatment practices[10]. As discussed above, evidence on COI and/or cost-effectiveness of treatments in the UK is very limited and therefore evidence from other countries must be used, but given the differences in the demographic make-up of populations and clinical practice should be analysed with caution.

Conclusion

The economic impact of paediatric CF was deemed a substantial cost by all the studies reviewed. Newborn screening was advocated by both of the studies found. However, discrepancies existed between studies and are most likely due to methodological differences.

[10] For instance, the US and the UK differ in terms of the preferred order (first line, second line, etc.) of antibiotic choices, as well as how often clinic visits occur. Even within the US, the way clinics and CF teams are arranged differ between institution 99. Iles R. Personal Communication. Cambridge2011.. The BMJ has gathered together various treatment guidelines from Europe and North America, however an in depth discussion of their differences is not within the scope of this thesis. 100. British Medical Journal. Best Practice: Cystic Fibrosis. BMJ; 2010 [February 10th 2011]; Available from: http://bestpractice.bmj.com/best-practice/monograph/403/treatment/guidelines.html.

Table 2.1: Cost-of-illness studies on paediatric CF								
Author	**Price Year**	**Country**	**Reported Currency**	**Patients (n)**	**Source of resource use data**	**Source of Cost data**	**Results***	**Notes**
Baumann U (54)	1996	Germany	€	138	Insurance claims	German public health insurance tariffs	$29848	
Ireys I (56)	1993	USA	US$	204	Medicaid charges	Medicaid charges	$18802	
Johnson JA (92)	1996	Canada	CAN$	171	US Epidemiologic Study of CF	Ministry of Health sources	$4969 (Inpatient $2553, Outpatient $927)	
Sims E (29)	2002	UK	US$	1134	UK CF database	Local and national sources; pharmaceutical company quotes	Screened population $8194, Non-screened population $13613, Estimated offset savings for drug costs by screening nationally $3851320	184 Screened patients, 950 clinically diagnosed
Wildhagen MF (58)	1991	Netherlands	UK £	81	Hospital records	Hospital charges	$17874 total costs, $14985 direct costs, Hospital care $10625, home IV antibiotics $9368, home care $1348	
Wildhagen MF (57)	Not Stated	Netherlands	UK £	21	Questionnaire and Hospital records	Not stated	Non-hospital costs $11429	
*All results are annualised and adjusted to 2010 US$ using the OECD PPP Method and/or consumer price index for Health.								

Table 2.2: Cost descriptions studies found by systematic review.									
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Time Horizon**	**Results***	**Notes**
Cost Description									
Farrell P (55)	1998	Ireland	US$	CF NBS	Single Study (Cohort study)	Not stated	1 year	Cost per CF diagnosis with sweat tests $4192.	
Gregg RG (59)	Not Stated	USA	US$	CF NBS (IRT vs. IRT/DNA 100ng/ml cut off vs. IRT/DNA 110ng/ml cut off)	Single Study (Cohort study)	Not stated	1 Year	Cost per CF patient detected for IRT was $12991, IRT/DNA (100 cut-off) $15539, IRT/DNA (110 cut-off) $14677. While costs per patient were similar, either IRT/DNA approach eliminates more psychosocial costs and is preferred.	1993 was used as the price year for conversion
Lee D (60)	2000	USA	US$	CF NBS	Single Study (Cohort study)	One CF Centre	1 Year	Total cost for CF neonatal screening were estimated at $11052736, or $10968 per diagnosis, or $2.72 per birth.	
Rosenberg M (61)	Not Stated	USA	US$	CF NBS (IRT/DNA (DeltaF508) vs. IRT/DNA (CFTR))	Single Study (Cohort study)	One CF Centre	1 year	Total cost for IRT/DNA (DeltaF508) was $10210 per diagnosis, or $2.81 per birth, IRT/DNA (CFTR) was $13936 per diagnosis or $4.53 per birth.	2005 was used as the price year for conversion

Table 2.3: Cost consequence studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Cost-consequences										
Ho S (63)	Not Stated	Canada	Can $?	4 types of nebulisers	Literature	Not Stated	Dead Volume, respirable fraction, drug output within RF, time to complete nebulisation	N/A	For Salbutamol, Updraft II cost Can $4.60, Acorn II Can $5.94, Misty-Neb $5.55, WJ Can $5.37. Updraft II generated the lowest costs	1998 price year was used for conversion.

Table 2.3: Cost consequence studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Iles R (64)	2001	UK	UK £	Tobramycin nebuliser solution vs. usual therapy	Single Study (retrospective cohort analysis)	NHS reference data	No explicit outcome measure	2 one year periods	Mean total cost per patient was $51257 when TNS was administered and $44266 in the year preceding TNS treatment. Clinical benefit (reduced hospital admissions and IV antibiotics) were reduced with TNS treatment.	

Table 2.3: Cost consequence studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
LeLorier J (66)	Not Stated	Canada	Can $	Tobramycin nebuliser solution vs. placebo	Canadian CF Foundation surveys, 2 US-based RCTs	Ontario and Quebec Health ministries and Statistics Canada	Hospitalisations	24 Weeks	Total weighted cost-savings for using TOBI in Quebec were from $1984-7967, and for Ontario $1361-6847 (95%CI). The authors state this would substantially offset the Canadian acquisition cost of TOBI ($8707).	1999 price year was used for conversion
Kretz S (65)	Not Stated	US	US$	Specialised vs. non-specialised home care	Case study	Hospital records	Exacerbation free periods and relative wellness	45 Months	Direct costs for non-specialised care $76782 vs. specialised home care $47656.	1996 price year was used for conversion.

Table 2.3: Cost consequence studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Menzin J** (67)	1992	Germany, France, Italy, UK	US $	rhDNase therapy for respiratory tract infections vs. placebo	Single Study (RCT) with expert opinion	Per diem costs in Germany and France CF Centres, accounting data for Italy and UK	RTI-related hospitalisation	24 Weeks	Cost saving in RTI-related care of rhDNase over normal treatment over 24 weeks was $1462 in France, $1276 in Germany, $1180 in Italy, and $941 in the UK.	

Table 2.3: Cost consequence studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Oster G** (68)	1992	US	US $	rhDNase therapy for respiratory tract infections vs. placebo	Single Study (RCT)	Hospital discharge summaries and itemised bills	RTI-related hospitalisation	24 Weeks	Mean costs of RTI inpatient care was $7436 (control), $5671 (rhDNase once daily) and $6516 (rhDNase twice daily). When cost of rhDNase was included, increase cost of therapy to $13464, offsetting therapy costs by 18-37%.	Some important cost items were left out.

Table 2.3: Cost consequence studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Suri R (71)	1999	UK	UK £	rhDNase therapy (alternate day vs. daily) vs. hypertonic saline	Single Study (RCT)	Finance depts., British National Formulary, literature	FEV % Predicted	Not Reported	$10649 for daily rhDNase and $8015 for HS, Daily vs. alternate day rhDNase $3272 vs. $1604. Although cheaper, HS was less effective than either daily or alternate day rhDNase.	
**These papers retrospectively modelled the cost of rhDNase into their measures as the drug did not exist at the time of the effectiveness data.										

Table 2.4: Full and partial economic evaluation studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Cost-effectiveness analysis										
Christopher F (44)	1998	UK	UK £	rhDNase therapy vs. normal therapy	Literature, expert opinion	Literature	LYG	Lifetime	Discounted cost per LYG for all patients was $101049 ($39576-109711). <FEV 70% predicted, discounted LYG was $30796	
Grieve R (62)	2000	UK	UK £	rhDNase therapy (alternate day vs. daily) vs. hypertonic saline	Single Study (RCT)	Finance depts., British National Formulary, literature	FEV1	12 Weeks	ICER of daily rhDNase vs. HS was $204. Alt day rhDNase compared with HS was $164. If WTP was $369 per 1% gain in FEV1, .91 (daily), and .88 (alternate) prob that rhDNase was more cost effective than HS.	

Table 2.4: Full and partial economic evaluation studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Perras C (69)	1996	Canada	Can $?	rhDNase therapy vs. normal therapy	RCT, retrospective cohort	Unknown	Hospitalisations	15.9 months	Costs associated with one less hospitalisation from the use of Dnase were $38875/6 Months, $15477/12 Months, and $7210/15.9 months. To be cost neutral, the price of Dnase would need to be $16.	Unknown methodology
Suri R (73)	2000	UK	UK £	rhDNase therapy (alternate day vs. daily) vs. hypertonic saline	Single Study (RCT)	Finance depts., British National Formulary, literature	QoL, pulmonary exacerbations	12 Weeks	Total health service costs were $10487 for rhDNase and $7893 for HS. Costs and benefits were not combined. But Daily rhDNase was more effective than HS.	

Table 2.4: Full and partial economic evaluation studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
Suri R (72)	1999	UK	UK £	rhDNase therapy (alternate day vs. daily) vs. hypertonic saline	Single Study (RCT)	Finance depts., British National Formulary, literature	FEV% Predicted	12 Weeks	Total incremental cost of Daily rhDNase vs. HS was $2635, Daily rhDNase vs. alt day rhDNase was $867 and alt day rhDNase vs. HS was $1779. ICER per 1% gain in FEV1 was $206, $399, and $166 respectively.	
Simpson N (70)	1998	UK	UK £	CF NBS vs. no screening	Literature review, authors opinion	Literature review	QALYs	Lifetime	Incremental cost per QALY gained with neonatal screening for CF over clinical screening was $13445. Authors concluded it was cost effective from the UK perspective	

Table 2.4: Full and partial economic evaluation studies found by systematic review.										
Author	**Price Year**	**Country**	**Reported Currency**	**Intervention**	**Source of effectiveness**	**Source of cost data**	**Outcome measure**	**Time Horizon**	**Results***	**Notes**
van den Akker-van Marle ME (74)	2004	Netherlands	€	No screening vs. CF NBS (IRT+IRT, IRT+DNA, IRT+DNA+IRT, IRT+DNA+DGGE)	Literature review	Literature review	Life Years	1 Year	Incremental cost per LY gained was $30501 with IRT+IRT, $47105 with IRT+DNA, $48950 with IRT+DNA+IRT and $40585 with the IRT+DNA+DGGE strategy. Authors conclude IRT+IRT appears to be the most cost effective screening strategy.	

Chapter 3: Methodology

Introduction

This thesis uses various techniques that are based on long-standing (and vigorously criticised) economic theory. This chapter serves as a critical discussion of economic theory relevant to this thesis and a description and review of the evaluation tools that are available for analysing data in order to inform decision making. First is a discussion of various cost side issues pertaining to COI and economic evaluations. Second is a discussion of issues surrounding benefit valuation in relation to economic evaluation. Next is a section discussing some of the tools used to analyse cost and benefit information.

Economic evaluation is driven by two main theoretical approaches which have been referred to as welfarism and extra-welfarism(1). Both of the approaches share the assumption that individuals are utility maximisers; that is, we all want to make our lives better and therefore make decisions based on what we think will give us the greatest utility. These decisions generally take place in the market for goods and services, with the market mechanisms ensuring that sellers and buyers achieve the highest possible utility out of their transactions. These theories are also used to drive decisions that impact society as a whole. Decision makers (often the government) try to derive the best basket of goods that maximises the utility of an entire society. However, there are often problems in making the leap from the individual to the society, given that there are bound to be differences in what is utility maximising between individuals. This is particularly difficult in areas where there are market imperfections.

Obtaining the best possible health outcome for society from health interventions means that certain conditions must prevail. One condition is that production processes must be efficient. There are generally two ways in which production can be inefficient - if large amounts of inputs are used where smaller amounts can be used, and if inefficient combinations of inputs are used. For example, why employ 40 people, when 10 can do the job? Or, if a relatively expensive doctor was doing the same job that a lower-paid nurse could do with the same result, it could be considered inefficient. It is also quite possible that production is inefficient because an intervention is being used that is not the one with the lowest costs (to produce a given level of service).

If services are being produced efficiently, then the second requirement to have total efficiency is to choose those services that will maximise welfare (i.e. health or utility). Combating these inefficiencies is the driving force behind most economic evaluations, or, comparing the benefits of different services and their relative cost.

When people trade in the market, economists assume that they will do so until they reach their goods and budget constraint, and that they will buy and sell until they have the maximum benefit. In a perfect market, everyone is assumed to maximise their welfare and there is optimal use of resources for any distribution of income. When the market reaches this equilibrium by definition, no one can be made any worse off without someone else being made better off. Or more formally, the marginal rates of substitution between all pairs of goods for all pairs of individuals are equal and equal to the marginal rate of transformation between all pairs of goods and the ratios of prices. In production, the marginal rate of technical

substitution between all pairs of inputs is equal to the ratio of prices for inputs. This state of being is called the *Pareto optimum*. (101)

If the economy was not at an optimum equilibrium point, the only method of making *Pareto improvements* is through voluntary trades that will not make anyone worse off, but at least one person better off. A Kaldor-Hicks improvement is made if an outcome is more efficient if those made worse off would, in theory, be compensated by those who gain from the exchange. This differs from Pareto improvements as the Kaldor-Hicks criterion do not require compensation to be paid, only for the possibility of compensation to exist and therefore a more efficient outcome may leave some people worse off. However, it is not often the case that Pareto improvements exist in the real world. This concept, often called the welfare approach, is the building block of cost-benefit analysis (CBA). A CBA generally seeks to use monetary measures to value the entire range of health and other consequences of a policy change and compare it with resource costs as a compensation test.

Extra-welfarism uses a form of constrained maximisation. In other words, the function is designed to get the most efficient use of limited resources. In the health context, the constraints are often defined by social objectives and a health care budget constraint. Extra-welfarist approaches are the theoretical foundation of cost-effectiveness (CEA), cost-utility (CUA), and cost-minimisation analyses (CMA). (37)

All forms of economic evaluation (CEA, CUA, CBA, CMA) of a health care intervention rely on information about costs and benefits of a particular intervention(s). Gathering information on both the cost and benefit side of an evaluation can be problematic. A cost is generally thought of by economists as the

opportunity cost forgone, and therefore one of the aims of economic evaluation is to measure opportunity costs. Most organisations try to measure costs at some level, if only to have some financial control. However, it is anything but straightforward. There are conceptual and practical problems to overcome. (101) For example, there is often limited availability of data as well as issues surrounding the treatment of overhead costs, discounting, and estimating productivity cost. Sections 3.1-3.2 illustrate the types of costs usually measured in health economics, as well as the problems associated with measuring costs.

Similarly, measuring the benefits of a technology can also be thorny. If benefits are traded in the market, then they automatically have a value given that someone considers them worth paying for. It may also be possible to measure an indirect estimate of the value people place on benefits from their behaviour. However, it is often not feasible to put any money value on the outputs of a health intervention, but it may be possible to compare standard units, such as a life year gained. Similarly to the cost side, perspective also matters when measuring benefits. (101) For example, when carrying out an economic evaluation on a screening programme, should you simply measure cases detected, the parent's benefit, or the child's benefit? How do you measure the child's benefits? Sections 3.3-3-4 discuss the issues surrounding the estimation and measurement of benefits.

Types of costs

Health service (direct) costs

Direct costs are defined by those costs in which an actual payment has been made. These typically include medical costs, such as treatment, hospitalisation and

medication costs as well as some personal costs such as travel costs to the health care provider or specialist aides.

Non-health service (indirect) costs

Indirect costs are defined as those costs in which no payment has been made, yet resources were lost. Generally, these can be separated into two groups: morbidity costs (productivity loss for the person, employer, or society due to illness) and mortality costs (present value for the productivity lost due to premature death caused by the disease).

Non-resource costs

The third type of costs, intangible costs, is defined by deterioration of quality of life in the patient. These costs are generally captured the health related quality of life measures in utility scales (e.g. SF-36, EQ-5D).

Concepts of cost

It is important to discuss the different concepts of costs in order to get any meaningful interpretation of the costs out of an analysis. Box 3.1 gives the definitions of the concepts used in costing studies and economic evaluations. All of the aforementioned types of cost can be discussed in terms of total costs, fixed costs, variable costs and therefore average or marginal costs.

Box 3.1 Definitions of cost	
Fixed Cost (FC)	= costs that do not vary with the quantity of output in the short run
Variable Cost (VC)	= costs which vary with the level of output
Total Cost (TC) (FC+VC)	= the cost of producing a particular quantity of output
Average Cost (AC)	= TC/Q, the average cost per unit of output (Q)
Marginal Cost (MC)	= (TC of x + 1 units) – (TC of x units) = the extra cost of producing one extra unit of output.

It is important to make the distinction between average and marginal cost because in economic evaluation, we are interested in finding out how much more input we need for an additional unit of output (marginal differences of cost and effect instead of average differences). For example, the extra cost of putting an additional nurse on an ICU versus the average daily cost for that nurse could produce very different results. However, in practice this issue must be explored in the context of the study as the savings from having the extra nurse may depend on other local factors such as staff availability and timing of the hours. (102) Therefore, marginal analysis is usually shown in conjunction with average results.

Study Design – Costing Issues

What costs to include?
Viewpoint/perspective

Most health economic evaluations will adopt a certain viewpoint (explicitly or implicitly). Possible viewpoints are society, health system, government, the patient, employer or agency carrying out the programme. The perspective of the study is

important, as it helps define what costs are relevant to the study. For example, in the English NHS, the cost of a nurse in a ward is relevant to the health system perspective, but not necessarily from the patient's perspective. Similarly, time off work may be a cost from the patient's and employer's perspective, but not to a health maintenance organisation.

Generally, the commissioning body of the study gives a clue about which viewpoint to adopt. However, this is not always the case and when in doubt, a societal viewpoint should be adopted as this is the broadest perspective. (1) Often times in the literature, a "societal" viewpoint is stated but actually refers to a health service viewpoint, so it is imperative to be careful when analysing literature to make sure all the relevant costs were gathered in the study.

Excluding costs

When carrying out an economic evaluation, it may also be possible to exclude costs that are common to both interventions under study. However, this is not advisable if at some point a wider analysis may be undertaken. Another possible decision facing the researcher is whether to exclude costs that may simply confirm the result of other costs. For example, if a study is looking at outpatient versus inpatient surgery, including patient costs may simply confirm the same result as hospital costs, and therefore may simply over-complicate the analysis. (1)

Estimation of costs

When the perspective has been chosen, and the relevant cost generating items of resource use have been identified, these items must then be valued. Therefore, estimation of costs relies on the quantity of items consumed, as well as applying a

relevant unit cost to those items. How information is gathered on resource use is generally dependant on the analysis in question. For example, an economic evaluation alongside a clinical trial can collect data on the report forms, whereas a retrospective analysis is likely to rely on patient notes and charts to gather resource use information.

Valuing resource items

It is important to remember that the theoretical cost of each unit of resource used is its opportunity cost (or, the value of forgone benefits because the resource is no longer available for its alternative use). However, in practical terms, market prices are used to attach a unit cost to a good unless there is a reason not to do so (for example if the resource is subsidized).

There are several possible resource items for which it is not necessarily easy to attach a unit cost. These items are generally 'non-market' items (such as leisure or volunteer time). One method of valuing these items is to use market wage rates. However, this method can be problematic for leisure time as people are generally not paid for their leisure, and therefore it can be argued that zero is a valid value or perhaps average over-time earnings. Overtime earnings are suggested because it is thought this is how much an employer would need to pay to buy out leisure time. Depending on the perspective, this dilemma may not matter. If from a health care provider perspective, the patient's time lost is not a relevant problem (unless this affects the patients utility? ref Brouwer et al various papers); however, if the analysis is from the societal perspective, these items are very important and should definitely be reported. Drummond (1) discuss an alternative approach in which the amount of

time inputs for these items (i.e. volunteer time) is reported alongside other relevant costs. This method, they argue, will enable the decision maker to see which interventions require large amounts of these difficult to value resources and therefore get an idea of the opportunity cost of changing the allocation of volunteer or family input into a programme. (This still begs the question of how you measure patient and other voluntary time inputs: adds cost to data collection and methods still are not well developed).

Adjusting market prices

Because of imperfections in the health care market, market prices do not necessarily reflect opportunity costs. For example, in the USA, hospital charges often deviate from costs as the hospital may try to cross-subsidise activities. However, it is not necessarily clear when a researcher should adjust market prices. Drummond et al (1) recommend attempting adjustments when the researcher is sure that unadjusted prices will introduce significant bias into the results and when there is a clear method of adjusting the prices.

Cost adjustments are especially important when it comes to comparisons across studies. As many studies do not adjust costs, they are not necessarily comparable with those that do, even if similar outcomes are reported. This becomes even more apparent when trying to compare across countries, as different countries often have different accounting systems and therefore costs are not necessarily comparable. Recommended practice (103) when gathering evidence for comparison or for use in economic evaluation from the literature is to adjust prices in two stage process, the first of which is to adjust for the time (i.e. from UK£ in 2000 to UK£ in 2010) and then

use exchange rate information to adjust to the target currency (UK£ to US$ for example). Shemilt et al (104) recommend using the OECD or IMF purchasing power parity method to adjust prices over time and by currency.

Time Horizon

How long to track costs often depends on the intervention follow-up time as well as what agencies are involved in the intervention. For example, if conducting a cost analysis of hip replacements from a hospital perspective, it would make sense to cost the intervention to discharge. However, if you also wanted to know about revision rates, a long-term (life-time) analysis would be optimal as it would give you information on the life-time costs associated with one patient's hip replacement. In other words, when looking at therapy specific or disease specific costs, the choice of follow-up period should not bias the results in favour of one intervention. Sometimes this could mean tracking costs over a lifetime and discounting costs to present values (see the next section for more on discounting). Figure 3.1 gives a graphical representation of these ideas.

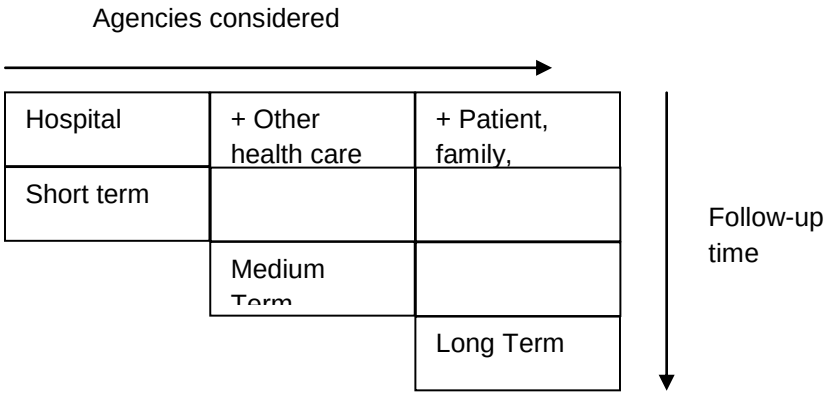


Figure 3.1 Choices for costing over time (adapted from Drummond (1))

Capital expenditures and overhead costs

Capital costs are those costs involved in purchasing the major physical assets required by a programme. Generally, this is land, buildings, and equipment. These costs are different from operating costs as they are generally fixed at a single point in time and they represent an investment that may depreciate over time (land generally does not depreciate). Capital costs can be described in terms of the opportunity cost of not using the investment money in some other way which yields positive benefits. This aspect is usually calculated by applying an interest rate equal to the discount rate of the study. Capital costs also represent the depreciation of the asset over time. This is generally calculated by various accounting practices (such as declining balance) which are usually determined by tax laws rather than a real change in the value of the asset. (33)

In practice, calculating capital costs is done by annuitizing the initial capital outlay over the lifetime of the asset. This method (105) incorporates both depreciation and

opportunity costs into the valuation. It is important to consider the annuitization and discounting of capital expenditures as generally people have a positive rate of time preference. In other words, people value money more today than they would in the future when they are likely to be better off or the resources you lay out for capital today is worth more than if you were to spend this later. Discounting future costs to present values is one method of calculating these costs. If all costs are assumed to happen at the end of each year, then the present value (P) can be calculated using the following formula:

Where n is the number of time periods, Fn is the future costs at time n, and r is the annual interest (discount) rate. The discount factor is the factor $(1+r)^{-n}$. The discount/annuity factors can generally be obtained from standardised tables for ease of use. If the assumption is that the costs occur at the beginning of each year, the first year is not discounted, but otherwise the equation is the same.

It may also be the case that most costs are easily expressed on an annual basis, but that capital costs differ from year to year. Therefore, obtaining an equivalent annual cost by amortization or annuitization would be required. This is a function of the annuity factor, the period, and the interest rate in a similar way as finding present value. Instead of finding P, you would solve for the capital outlay.

$$A = (1 / r) - (1 / (r (1 + r)^{n}))$$

Where A is the annuitization factor, r is the discount rate, and n is the length of the items useful life.

For example, if the total set-up costs for a screening programme were £200,000, and assuming a 5 year time period at a 3% interest rate, the annuity factor would be 4.5797. Therefore the equivalent annual cost would be £43,671. Or,

£200,000 = *E*(4.5797)

where E = the equivalent annual cost.

The equivalent annual cost depends on both the useful life (both physical and clinical) of the equipment/building. As clinical usefulness is highly dependent on technological change, it is generally recommended to assume short lives for clinical equipment. The annual cost also depends on the discount rate used in the calculations. Generally, there are two theories regarding how to measure the discount rate. One, the social opportunity cost method is calculated by using a weighted average of discount rates applicable to different sectors of the economy that contribute resources to the programmes under evaluation. The other, the social rate of time preference, is a measure of society's willingness to not consume today in order to consume more tomorrow. One method of calculating this is to adjust the real rate of return on long term government bonds for inflation. This would represent an aggregated individual rate of time preference. Another method is commonly called the shadow price of capital (SPC) approach, which measures the forgone private investment (or the opportunity cost of a public programme) as the present value of consumption given up.

Practically, researchers in the UK have followed the government announced discount rate for public sector projects. In areas where no advisement on discount rates has been given, the convention has been to use the prevailing discount rate in the literature (usually between 3-6 percent). Because of these differences, it is important to present costs both undiscounted and discounted. It is also important to conduct some sensitivity analysis around the discount rate and to explicitly state the base rate used.

When calculating overhead costs for a project, it is important to allocate the costs correctly in order to produce as true a picture as possible. It is first important to determine a unit of output for those departments that directly serve the programme(s) in question. The point is to try to calculate a cost per unit of output and then multiplying by the usage of each patient, therefore determining the cost per patient. A common example is an in-patient day, or a workload unit (usually hours) for staff. Then, an allocation basis must be determined for each department. For example, hotel costs for a bed day includes (among other costs) catering expenses. This means that catering costs would be in proportion to the number of patients in that bed and the number of meals a day. Similarly, hours paid to staff would be an allocation unit for administration costs.

Estimation and inclusion of productivity changes

Productivity costs in terms of cost analysis are generally seen as those costs that arise from the patient and/or family member taking time off work. These costs can be substantial, or negligible, depending on the circumstances. For example, if a patient is retired, then they are not employed and therefore do not necessarily incur

any loses. However, if a child's parent has to take time off work, then the productivity loss could be substantial.

How productivity changes are measured, and when they should be included is a matter of debate. There are two generally accepted methods of estimating productivity changes. The human capital approach is probably the most common method employed and is generally estimated by using the gross earnings of a person or estimating equivalent earnings in unemployed people. This method is frequently criticised for overestimating the value of production (106). This is because often there is negligible time off work, or perhaps in long-term illnesses, employers will hire a new person, thereby eliminating the productivity loss. In other words, productivity changes depend on the length of time in question, and the cost of replacing a worker, as well as other adjustments in the wider economy. This is similar to the argument made above that average costs are not a good estimate of losses at the margin.

In an effort to get around these issues, Koopmanschap et al (107) suggested that the friction cost method be used to estimate productivity changes. This method attempts to measure productivity loss as a function of the time it takes an organisation to return to the initial production levels. The authors note that this measure depends on the organisation, location, industry, and type of worker. (108) For example, a Tesco's cashier will theoretically take less time to train than a health economist; therefore it can be argued that the productivity losses to Tesco's due to a week off ill would be less than a pharmaceutical firm's health economist.

Of course, the relevance of productivity will change depending on the viewpoint of the study. If the study is from the healthcare provider perspective, the cost of productivity to the wider economy may not necessarily be important. For example, Gerrard and Mooney (109) argue that when the benefit of a cost-effectiveness or cost-utility study is health related, then the opportunity costs are only defined in terms of the health forgone, therefore researchers should only focus on the opportunity cost to the healthcare sector, and not other sectors of the economy. However, if a societal perspective is adopted, then the productivity change would be very relevant. Drummond (1) argues that a societal viewpoint should be taken if at all possible, as this does not enforce budgetary boundaries simply to healthcare costs. The authors encourage presenting healthcare and non-healthcare costs and benefits separately, thereby giving the decision maker a clear idea of what the opportunity cost of the healthcare budget is.

When a societal perspective is adopted, productivity changes are more difficult to untangle. For example, if a community-based programme under analysis has higher costs to the health care system compared to an institution-based programme, but the number of workdays lost is less, is it correct to deduct the production gains from the healthcare costs? It can be argued that productivity gains should be included, given there is little difference between these resource savings and labour inputs for healthcare costs. However, it can also be said that the assumption that the institution-based programme means losses for the community (through drawing people out of employment) could be false, as other (unemployed) members of the community could replace those in hospital.

Another major concern for the estimation of productivity changes is that of double-counting. This is especially an issue in relation to gains in productivity. If the value of improved health already included the value of increased productivity, then it would not be appropriate to create an additional measure of productivity. This situation typically arises in cost-utility and cost-benefit analyses where individuals are given health state information and asked to value these health states in terms of utility or money. It may be the case that respondents will factor in the impact on their income when returning to work when creating a value for the health states. However, Brouwer (110) argued that income may only have a weak connection to productivity change, especially if employees have income protection (i.e. sick pay). The author advocated an approach where individuals were asked to ignore income effects when valuing the health state, but asked separately about the changes in productivity.

Finally, the inclusion of productivity changes can raise issues surrounding equity. If you are using wage rates of individuals in the study, it is obvious that certain individuals may have a larger share of the impact than others. There are a few methods in which equity concerns can be alleviated. Firstly, expressing productivity in terms of the days of work or activity lost or gained instead of the monetary amount. Secondly, using a general wage rate to value the productivity instead of individual wages will eliminate some people having a larger impact than others.

Conclusions

It is apparent that there are many stumbling blocks in front of a researcher trying to carry out an accurate cost-description of a technology. Therefore, it is important to be clear and explicit when designing a study just what costs you are interested in and exactly what methods will be used to estimate these costs. While costs are the

main concern when carrying out a cost of illness study, for an economic evaluation, the costs only represent one part of the whole equation. The following section will discuss how health benefits are measured and valued.

For this thesis, I have adopted a rather narrow perspective of the tertiary NHS services in the east of England. Therefore, cost information will only be collected for patient's resource use at the hospital level, and the drug use. While this is rather limited in scope, it is probably the most appropriate, as the guidelines for treating CF patients clearly recommend that the majority of CF care is carried out in a specialist centre at the tertiary level of service(2). The time period of the cost of illness study was determined by data availability, and covers the period of 1998 and 2004-2007. The CEA assumes a time horizon of 50 years, as I wanted to look at the lifetime benefits of newborn screening.

Benefit measurement and valuation

How to measure and value health benefits is a contentious issue. Before looking at why, it is important to differentiate between measurement and valuation in terms of health. Many clinicians and psychometricians are seeking to measure health numerically according to one or more relevant dimensions. Economists, on the other hand, who are concerned with choices and preferences between health states, are interested in the relative ordinal value that patients place on the range of different dimensions of health or on the process of providing health. While the value of health is related to a measure of health, the concepts will not necessarily be exactly correlated. (1) For example, even if physical therapy directed at clearing the lungs results in only a small improvement in lung function, it may nevertheless be highly valued by the patient.

One of the difficulties in measuring health benefits is that they are often multidimensional. For example, improvements in life expectancy cannot necessarily be judged by looking at improvements in blood pressure management. Quality improvements in a person's health can vary in different dimensions of health, such as mobility, sensory perception, pain relief and so on. As single measures of health improvement are not necessarily adequate to describe the health consequences of health care for patients, there have been attempts to develop multi-dimensional tools to measure well-being, ability to function or social activities. (111) Many tools have been developed for use in economic evaluation to try and measure health outcomes such as the EQ-5D (112) or the Health Utilities Index (113).

Evidence about Health benefits arising from care or treatment can also be shrouded in uncertainty. Evidence is often limited (i.e. few clinical trials for anti-bacterials used as prophylaxis in CF babies), leading to even more uncertainty surrounding clinical practice. These uncertainties carry important implications for economic evaluations. (114)

Measuring Health

In order to measure health, it is necessary to define it. The commonly quoted World Health Organisation definition that health is "[a] state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity" (115) is broad and can be difficult to capture in a single measure. It can also be argued that social well-being is not necessarily health, but a single aspect of quality of life. In practice, developers of health measures often take a more narrow definition of health. Therefore, different measures of health or quality of life can differ in content and sometimes in meaning. For example, a quality of life measure that is narrowly

focused around certain symptoms would yield very different information than a quality of life measure based around the WHO definition. Brazier et al (114) argue that the specific content of the description of benefit used in the measurement or valuation instrument is of utmost importance for health policy decisions.

The features of health are often difficult to measure as the elements making up 'health' (i.e. pain, ability, symptoms) vary by individual and therefore a measurement capturing only the medical component (for example, systolic blood pressure improvement) may not necessarily be useful to capture the entirety of 'health' gained/lost by an intervention. It is commonplace within health economics and health services research that the experience of a health state should be elicited from the patients themselves, or from others on their behalf if this isn't possible (for example, an infant) in order to best reflect the experience of the disease and its treatment(114) (116).

Valuing Health

Once a measure of a health state is defined, it is essential for health economic analysis of cost utility or cost benefit to obtain a value on that health state. In other words, how does one adjust life years to reflect health levels during those years? Valuing health is generally done by using preference elicitation techniques. Both cardinal (producing responses that are on some interval scale) and ordinal techniques (discrete choice experiments and ranking and standard gamble?) are used. Different methods often result in different values, and therefore it is important to look at the different techniques in order to choose the most appropriate method.

Cardinal techniques

The three most commonly used techniques for valuing health states are the standard gamble (SG), time trade-off (TTO), and the visual analogue scale (VAS). SG and TTO are both preference-based techniques, while VAS is not. Other commonly used methods include the person trade-off (PTO).

Visual Analogue Scale

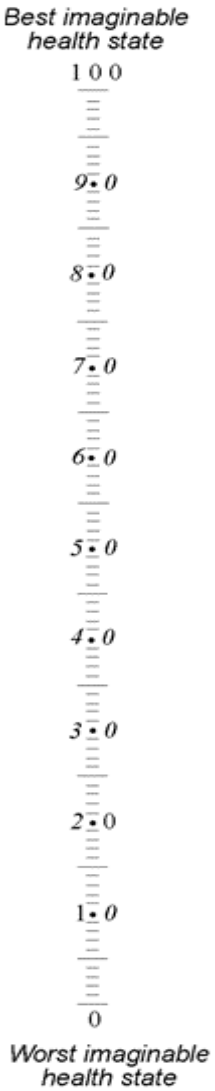
The VAS is usually represented by a line, with clearly defined end-points and clear interval properties (the difference between 10 and 30 is the same as the difference between 30 and 50 on a hundred point scale). Respondents are then asked to indicate their judgements (or values, feelings) by placing a mark or line on the scale. The distances between the marks should therefore indicate a person's relative understanding of the differences in the concepts being measured. (114, 117)

The VAS is used in many commonly used instruments, such as the Health Utility Index and the EQ-5D. To use VAS in economic evaluation, it is necessary to ensure respondents are comparable. To do this, it is imperative to define unambiguous end-points. Most commonly used are variations on 'full health' at the top end, and 'death' at the bottom end. It is useful for economic evaluation if the valuations can be placed on a zero to one scale. The scale must be clearly defined (I.e. 0-1, 0-10, 0-100), but it is unclear whether the accuracy increases with a larger scale. (1)Zero is equivalent to dead, and one to full health. However, it is possible that there are states worse than death. For these reasons, it is often asked of respondents to rate dead on the same scale along with their own health. This allows all health state valuations to be transformed using the following formula:

$$A_i = R_i - R(dead) / R(best) - R(dead)$$

where Ai= adjusted VAS rating for health state hi; R(being dead) = raw rating given to being dead; Ri = raw rating given to health state hi; R(best) = raw rating given to the best health state. (114) This allows for the value of 1 to be full health state, and zero for being dead. Any negative values would be assumed to be worse than death. Figure 3.2 illustrates what a VAS would normally look like.

Figure 3.2: Visual Analogue Scale example



The VAS can be used for valuing temporary health states (states lasting a specified period of time, assuming returning to good health or chronic health states). In this scenario, a respondent would be instructed that health states would be for a specific amount of time, and at the end of that time returning to full health. They would then be asked to place the best state at one end of the scale, and the worst at the other end. The remaining states would then be placed between the endpoints, thus indicating how the respondent felt about the different health states relative to one another.

The VAS is a useful tool for health state valuation as it has been shown to generate high response rates and relatively high completion rates. (1) (118) VAS methods are also relatively cheap and simple to administer. Mannion et al (2006) (119) and others (120) have found that the VAS is reliable in terms of inter-rater reliability and test-retest reliability.

However, there have been some concerns that the choice-less nature of VAS means that it is unlikely to reflect preferences on an interval scale. Bleichrodt and Johannesson (121) found little evidence of a stable value function. In 2001, Robinson et al (122) found that when VAS data was adjusted using the Parducci-Weddel (PW) range-frequency model, the value function was stable, and therefore allowed VAS data to be used in economic evaluations. McCabe et al (123) found that using the PW model can produce data that is not on the 0-1 scale, and therefore is not necessarily appropriate for economic evaluation.

There is some evidence that VAS methods are prone to response spreading and context effects. This is when respondents use all areas of the scale when

responding, especially where multiple health states are valued on the same scale. This can lead to similar health states being placed some distance from each other, and different health states being placed close together. If this occurs, then VAS does not produce an interval scale and therefore the numbers obtained may not be meaningful. (114) Robinson et al (122) found that some core health states were given lower values when in a set of more severe health states than when in a set of milder health states. Torrance et al (124) found some evidence of end-point bias, where health states at either end of the scale are placed further apart than when directly compared. It is unclear whether or not adjustments in the PW model would adequately remove these biases from the data.

Because the values elicited from VAS may not be directly applicable to economic evaluation research into mapping VAS values to SG or TTO utility values has been undertaken. Torrance et al (1996) and Feeny et al (2002) used VAS to elicit preferences for the various health states defined in the HUI2 and HUI3 instruments. The argument for the relationship between VAS and SG is based on the work of Dyer and Sarin (1982), which argued that utilities are made up of a combination of a value function and the relative risk attitude. This means that risk-neutral individuals will have the same values for both SG and VAS. Risk-averse respondents will have a concave relationship between VAS and SG (the respondent would prefer a certain health state with value x to an expected equivalent value). A risky individual would therefore have a convex relationship between VAS and SG. Torrance et al (125) and Feeny (113) transformed the values from the VAS into SG utilities using a power function similar to the following equation:

$$U = 1 - (1 - V)^b \text{ or } U = V^b$$

where the power term b represents a respondents constant relative risk attitude. Therefore, b>1 implies risk aversion, b=1 is risk-neutral. U is the SG utility, and V is the VAS value of the health state. However, more recently, Stevens et al (126) found that the power function was outperformed by linear, quadratic, and cubic functions when analysing the relationship between VAS and SG in the context of a UK valuation study of HUI2.

Mapping VAS values to TTO utilities has been fraught with difficulties. One major problem is the attribute of time preferences. A positive time preference would reduce the value of time spent in the chronic state by a larger proportion than the time spent in full health, and therefore TTO values would increase. Torrance (127) attempted to use a power function between VAS and TTO and found it explained 79 per cent of the variation. However, neither of these models used standard econometric diagnostic information on model specification and therefore it is unclear how well they actually fit the data or if they are comparable. The evidence therefore suggests that relationships between VAS and SG and TTO are not necessarily stable in terms of form of the relationship and the size of the model parameters. (114)

Standard Gamble

The SG method allows a respondent a choice between a certain health state, and the uncertainty of a gamble between two possible outcomes, one of which is better than the certain health state, and one of which is worse. For example, a respondent must choose between alternative 1: living with asthma, and alternative 2: receiving a treatment that could either a.) return them to full health, or b.) result in death. The

probability *P* is assigned to returning the patient to full health for a certain time period, and 1-*P* is the probability that death will occur. Alternative 1 will have a certain outcome of a chronic state, *h* for life. To obtain utility values, the probability *P* is varied until the individual is indifferent between the certain outcome and the gamble (perfect health or death). This is then repeated for all intermediate outcomes. SG methods can be used to consider states worse than death and for temporary states by changing state *h* for a state worse than death, and changing the gamble by assuming the choice is between perfect health and a state worse than the certain state. (111) (33) (1)

SG is often referred to as the 'gold standard' for the measurement of utility associated with health states. This is because it is based on expected utility theory (EUT). EUT is a framework for describing how individuals make decisions under conditions of risk and uncertainty whereby individuals choose between prospects in a way that maximises their expected utility. (1) This method has been used extensively throughout the literature and has also been used (via a transformation of VAS) to value the Health Utilities Index 2 and 3, and the SF-6D. (117)

There are various ways of carrying out the SG technique. Generally, they differ in terms of the procedure used to identify the point of indifference, and the method of administration (computer, interview, self-administered). One of those methods was developed by Torrance et al (128) and uses a visual aid (a probability wheel) to explain probabilities to respondents. This method works by using the wheel to iterate between values for the probability of success *P* (the ping-pong method), therefore helping the respondent choose their indifference point. This is done by using an

adjust wheel with coloured sectors. The respondent is told that the relative size of the area coloured represents the probability of achieving that alternative.

Another method, developed by Jones-Lee et al (129) does not use a visual aid, but instead uses a questionnaire with a titration method of listing the values for chances of success for treatment. The chances are presented in a 'top-down' (100, 95, 90 per cent etc.) method, or a bottom up method (0,5,10,15 per cent etc.). Respondents are asked to indicate which values of *P* they would accept and reject treatment, as well as which value of *P* they find most difficult to choose between treatment and remaining in a hypothetical health state. This method has been used in both paper and computerised form. (130)

The SG method is generally accepted to be a practical and reliable method in terms of administration and completion. Some studies have shown completion rates upwards of 95-100 per cent. (131, 132) SG methods also seem to be generally acceptable to various types of patients and clinical areas. However, its use as the 'gold standard' has been criticised, given evidence that the assumptions of the expected utility theory are violated in practice. (133) Therefore, SG methods may not adequately take into account other factors such as risk attitudes, gambling effects and loss aversion. Kahneman and Tversky (134) argued that respondents act risk-averse when choices are framed in terms of gains, and risk seeking if choices are framed in terms of losses. The authors also found that individuals overestimate small probabilities and underestimate large probabilities, which would potentially bias health state valuation tasks.

SG has shown consistency with expected rankings. Dolan et al (135) examined the performance of SG and TTO against 12 EQ-5D health states and reported that SG had high levels of consistency Other researchers (136) found similar conclusions looking at how SG compared to ranking tasks.

Time trade-off

Torrance (127) developed the TTO technique as an alternative to SG. It was designed to attempt to overcome problems surrounding explaining probabilities to respondents In TTO, respondents are asked to choose between two certain alternatives instead of a certain alternative and a gamble. In this case, respondents are given a paired comparison. For example, in a chronic health state preferred to death, alternative 1 would involve being in a specified state worse than full health for time period t. Alternative 2 would be full health for time period x where x<t. Time x is varied until indifference. Therefore, the score given to less than full health state is x/t.

For states worse than death, modifications can be made to the TTO method. To achieve this, the alternatives would be immediate death;, and spending a length of time y in state h followed by x years in full health where x+y=t. Time x is varied until the respondent is indifferent between the alternatives. The value for state h, therefore, is -x/(t-x). Or, the more time that is required in full health to compensate the time spent in h, the lower the score for state h.

Similarly, for temporary health states, the TTO method needs to be adjusted. Intermediate states are measured relative to the best state (full health) and the worst state (temporary state hj). The alternatives in this case are living in the temporary

health state hi for time period t (specified for the temporary health state) followed by a return to full health; and living in temporary health state hj for time x where x<t, followed by full health. Again, time period x is varied until indifference is reached between the alternatives. To find the required preference value for state hi, hj should be set to 0, and the calculation would be hi=1-x/t. To return the values onto a full health/death scale, the worst temporary state must be redefined as a short chronic state and valued accordingly.

Similarly to SG, there are variations on TTO methods, which use various elicitation procedures, props, and administration. For example, Dolan et al (135) use a moving sliding scale to represent life years. The TTO has been used in both self-administered format and computer-based application. (137)

Green et al (118) showed in a literature review that a variety of studies have shown the TTO technique as practical, reliable and acceptable method. However, there are two main criticisms of the method. A major source of concern is that the technique relies on two certain choices, when health care is undoubtedly uncertain (138). Cher et al (139) argued that it is possible to incorporate attitudes to risk and uncertainty into TTO. However, adjusting for risk attitude is difficult when there is not a constant attitude to risk.

Another major issue with the TTO method is the impact of time preference on the valuations. Van der Pol and Cairns (140) argue that the evidence suggests that a majority of individuals have a positive time preference for health (they give greater value to years of life in the near future than those in the distant future). If this is the case, then the assumption that people trade a constant proportion of their life

expectancy in the valuation of health states is not true, and therefore biasing the results of TTO.

Indeed, Robinson et al (141) found that there is a 'threshold of tolerability' below which health states would have to fall before respondents would be willing to sacrifice even a short period of time. This corresponds to evidence found by Dolan et al (135) that found approximately 5 per cent of respondents were unwilling to sacrifice any life expectancy in order to avoid half of the states they valued.

Person trade-off

This technique is generally used to estimate the social value of different health states. (142) PTO asks the respondent to make a choice between alternatives; however, the respondent is asked to make the decision based on groups of people instead of asking them to value their own state. In this instance, a respondent would be asked to indicate how many people in health state A are equivalent to health state B. The health states are generally described to the respondent For example, if 50 people have severe liver cirrhosis and 50 people have mild liver cirrhosis, and only one group could be help, which group would you choose? The number of people in each state is then varied until the respondent finds the groups in equivalent need of help. The undesirability of health state B is then x/y times as great as that of health state A. This process is then repeated for all other health states to be valued. This method is used by research in the context of social choices (such as the social value of a QALY) than more conventional individual perspective. (117)

The PTO method has not been widely used within the literature, and therefore the feasibility and acceptability is inconclusive. (117) However, Pinto Prades (143) and Murray and Lopez (144) found the PTO acceptable and feasible in separate studies, one pilot with 30 individuals and the other a group exercise. However, Nord (145) found that because PTO asks individuals to make decisions about others, there was some difficulty and refusal to participate (17 out of 53 respondents).

Nord also found that there was a strong random element in PTO responses, but on aggregate, they may be reliable. There is, however, little evidence to support the reliability of this method. There are also few theoretical economics links that have been made to underpin the method. However, it has been argued that PTO has at least a hypothetical advantage in economic evaluation for health care as it asks about trade-offs between people. (143)

Ordinal techniques

In the health economics literature, ordinal elicitation techniques have generally been seen as 'warm-up' exercises, instead of a method of deriving cardinal valuations. However, there is a strong methodological foundation for estimating cardinal values from ordinal information that was developed outside of health economics (in psychology, consumer marketing, environmental economics). (117)

Ordinal data collection approaches are relatively easy to comprehend to administer, and also have fewer problems with measurement errors, making them an attractive option. These exercises usually place less cognitive burden on respondents, and therefore are useful in settings or populations where education and numeracy are limited, thus have an advantage over SG and TTO. Another advantage is that preferences and judgements elicited are not contaminated by values such as risk

aversion or time preference. Ordinal methods are usually framed in a way in which respondents simply choose over health states.

As there are different types of ordinal information, different tools are required to collect it. Most commonplace are ranking exercises, contingent valuation and discrete choice experiments. In ranking tasks, respondents are asked to provide an ordering of a set of health states from best to worst (or vice versa). This can be done via an open-ended task, or via a more structured interview situation. In discrete choice experiments, respondents are asked to choose between two or more alternatives, usually described by their levels or attributes along several dimensions. (146)

Ranking

Methods of inferring cardinal values from ordinal choices began to develop in the early 1900's. Thurstone (147) proposed the *law of comparative judgement*, by which perceptions arise along some dimension of interest. In other words, given a stimulus, a respondent has a "discriminal process" that will yield a distribution which is normal if applied repeatedly. He assumed that stimuli are associated with uncorrelated but similarly varied distributions. (117) In practice, this law can be illustrated by presenting a respondent with paired comparisons repeatedly, and asking them to consider which stimuli is of greater importance on the attribute of interest.

This law provides the basic methods of transforming ordinal data into estimates of cardinal data. For health state valuation, McCabe et al (123) proposed the use of a random utility approach first described by Luce (148) to translate ordinal rankings.

This model is based on two functions: a statistical model that describes the probability of ranking a particular health state over another given the cardinal utility associated with each health state; and a valuation function that relates mean utility for a given health state to a set of explanatory variables. (117)

The random utility model is operationalised by using a conditional logit regression. The conditional logit model produces estimated valuation on an interval scale, enabling meaningful comparisons of differences. The assumptions of the model define the origin and units of the scale. This produces a general specification of

$$U_{ij} = \alpha(\mu_j + \varepsilon_{ij}) + \beta$$

where U is utility given to state j by respondent i, alpha is the normalising constant for the model coefficients, mu is a systematic component that defines i's utility, and beta is the value assigned to a state where there are the best possible levels on all of the health dimensions. It would make sense to assume that when Beta = 1, it implies that a person with no difficulties on any dimension will have an expected health state valuation of 1.

However, for alpha, there are a few more possibilities. Salomon (149) and McCabe (123) argue that alpha can have three choices. The first is normalisation using the exogenously defined value for at least one state (for example, the observed mean from TTO, VAS or SG values). The second being normalisation to produce a utility of 0 for the worst state. The final choice is normalisation to produce a utility of 0 for dead (though this only works if respondents have ranked 'dead' among health states).

Salomon used this technique to translate rank data from a study in the UK. Health states were described using the EQ-5D system, and respondents were asked to complete ranking alongside the EQ-5D descriptions of their own health, a VAS exercise, and TTO valuation. After fitting the model, the rank-based predictions were strongly correlated with the observed TTO values. McCabe et al also fitted a conditional logit model from two other valuation studies in the UK (using the HUI2 and SF-6D). In this study, rankings were compared to SG values. In the HUI2 data, the rank model was similar to the SG values; however, in the SF-6D data set, the rank based and SG values were different. The authors found that in that case, the models were sensitive to excluding the highest and lowest ranked states.

Discrete Choice

Discrete choice experiments can be applied in a variety of types of valuation problems. The method is grounded in random utility theory, and is generally used to show what individuals are willing to trade between characteristics of a treatment or service. This allows a value judgement to be made on different aspects of a service or treatment. This is of particular use in cost-benefit analysis, as a 'cost' can be included as one of the attributes, thereby allowing a willingness-to-pay aspect to be included.

There are various steps necessary to undertake a DCE. First, the characteristics of the item in question must be identified. These are often gathered from the literature or from focus group discussions. Ryan and Gerard (150) recommend that four to six attributes was acceptable. Second, levels should be assigned to the characteristics. These levels can be cardinal, ordinal, or categorical. Third, scenarios must be

created that describe all possible configurations of the characteristics and levels chosen. There is potential here to have unreasonable amounts of scenarios if a full factorial design is chosen, depending on the amount of characteristics and levels. However, a fractional factorial design that maintains orthogonality can be used. Ideally, designs should allow for interactions, minimise overlap, and be orthogonal-in-differences. (151) Fourth, as preferences for the scenarios need to be elicited using discrete choices, respondents should be presented with a number of choices and asked which they prefer for each choice. For example, respondents can choose A over B, that A is better than B on a sliding scale, that the respondent is indifferent between A and B, or that neither or no participation is preferred. Which choices given to the respondent depend on the question. However, exclusion of a 'neither' option may overestimate any obtained values. (152)

Finally, econometric techniques are used to model the data. As the choices are in binary format, the utility function is described in an additive form.

$\Delta U = \beta_1 X_1 + \beta_2 X_2 + ...\beta_n X_n$, where ΔU is the change in utility in moving from A to B, X is the differences between attribute levels A and B, and Beta are the coefficient estimates of the model. Different regression techniques are required depending on the type of questions asked (for example, nested logit for a non-participation question). (1) There are several examples where the DCE approach has yielded similar utilities as the SG, VAS, and TTO models. (150) (117)

Using the DCE approach can allow for non-health characteristics (such as process utility) to be included in the utility estimates. The technique also tells us information

about the relative valuations of aspects of a whole package. This is of particular use to health planners and clinical researchers when planning a new service.

There are theoretical concerns with all of the methods described above. Unadjusted VAS does not provide a basis for estimating preferences over health states, and methods of adjustment are still being researched. SG and TTO are both useful trade-off based valuations, but there is no real 'gold standard' choice. PTO can be a useful tool for measuring social perspective values of health states, but more theoretical underpinning is necessary, as are tests of reliability and feasibility.

Using ranking tasks and DCE to elicit utility data has also shown to be feasible, and reliable alternative to VAS, SG, TTO, and PTO. One advantage ranking and DCE techniques have over the cardinal methods are that non-health dimensions are easily added into the value judgements. Also, these techniques may be cognitively more accessible to certain patient groups than SG, TTO and PTO.

Self-reported measures of health

Eliciting self-reported levels of health dimensions using a standardised numerical scoring system is often used to describe an individual's health. Generally, these approaches contain a quantitative description of health states which contain components of health that are thought to be the most relevant to patients. These measures do not usually contain bio-medically, possibly clinically relevant measures, as these are not necessarily part of the patient's experience of their health. Self-reported measures of health can be non-preference based or preference based, as well as disease specific or generic.

Non-preference based measures

It is also the case that the instruments used to gather information on self-reported health measures differ in content. For example, the SF-36 health survey has 8 dimensions of health, uses a patient or proxy for responses, and results in a health profile. Whereas the disease-specific St. George's respiratory questionnaire has 4 dimensions, uses patient responses and results in a health profile and an index score. The content varies considerably between the two tools, covering generic concepts of functioning (physical function), to specific symptoms (wheezing) respectively. (52) Most measures of health use a simple summative scoring system which is transformed linearly onto a 0-100 scale. In the SF-36 questionnaire, the dimensions are not comparable and an equal weighting is assumed.

One major problem with non-preference based measures of health is that the instruments generally assume equal intervals between the choices and those items are of equal importance. For example, in SF-36, the difference in 'limited a lot' and 'limited a little' is assumed to be the same as 'limited a little' and 'not limited at all'. Another example would be the difference between 'mild' and 'very mild' would be equivalent to moving from 'moderate' to 'severe'. However, Brazier (114) provides evidence using visual analogue and standard gamble techniques that people are not necessarily able to perceive a difference between 'mild' and 'very mild', but there is a significant difference in perception between 'moderate' to 'severe'. This would indicate that people’s preferences are not necessarily indicated by these scales.

More sophisticated techniques have been developed to score self-reported health. For example, factor analysis weights items on their contribution to an underlying

variable with a stronger correlation indicating a stronger weight for an item or dimension. While this method may provide a method of comparing differences between populations, it does not necessarily indicate the relative importance to people in their lives. Another method used is Rasch analysis. This method is generally used when there is a group of people responding to a set of questions (items) that have categorical answers (for example, health status). The responses are then added across items to give every individual a total score. The score is then summarised across all items, therefore giving an indication that a person with a higher total score than another one is deemed to show more of the variable assessed. In other words, this method allows an investigation into the degree of severity of an item in relation to an underlying scale that the item is measuring. Rasch analysis tests the idea that the comparison of two people is independent of which items may be used within a set of items assessing the same variable and assumes "that the probability of a given patient 'passing' an item or task is a logistic function of the relative distance between the item location parameter (the difficulty of the task) and the respondent location parameter (the ability of the patient), and only a function of that difference." (153)

Similarly to factor analysis, Rasch analysis does not provide a method to value health for an economic evaluation. This is because, while it provides information on how difficult an item or dimension is for people, it does not necessarily provide information on preference-based weights which is necessary for a cost per Quality Adjusted Life Year (QALY) analysis.

Another important limitation of non-preference based measures of health is that they do not take into account mortality. This creates a statistical problem known as the

survivor effect. The effect occurs when a lower survival rate in one arm of a clinical study can increase its mean health status score in comparison with other arms in the trial. This occurs as patients who die generally have lower than average health status. So if survival is a good thing, the researcher is unable to determine which treatment is better. As healthcare interventions generally have an impact on survival and health status, this is a crucial limitation. (117)

In general, health status measures have very little scope in economic evaluation. It is for this reason that preference-based measures and multi-attribute utility scales for calculating the QALY were developed. However, health status measures are not entirely useless. In recent years, methods to map health status measures onto preference-based measures have been created.

Disease-specific measures for CF-related quality of life

A review in 2008 by Quittner et al (154) focused on measures used on children with respiratory diseases. The authors found that there were three disease specific measures: the CF Questionnaire (CFQ-R) (155), the Cystic Fibrosis Quality of Life Questionnaire (CFQoL) (156) and Questions on Life Satisfaction – Cystic Fibrosis (FLZM-CF). (157) The authors note that only the CFQ-R has a child version for ages 6-13 and a parent proxy report for the child's HRQOL. The CFQ-R and CFQoL has been validated for use in adolescents as young as 14, but the FLZM-CF has only been validated in patients over the age of 16. The authors found that all three measures had "good" internal consistency (α=0.54-0.94) and that test-retest reliability was good for all the measures (0.45-0.90 over a 10-14 day interval). The authors also state that all three measures were sensitive enough to detect different

levels of disease severity based on lung function tests. The CFQoL and CFQ-R have shown correlation (r=0.57-0.84) with the Short Form -36 item health status questionnaire (SF-36) generic utility measure (discussed in more detail below), but the FLZM-CF has not been tested. While the CFQoL and the CFQ-R have both been shown to correlate with the domains measured in the SF-36, there has been no mapping done from the domains of either measure to utility scores that would aid in the calculation of a QALY.

The authors suggest that the FLZM-CF, while short, does not allow enough flexibility for more complex interventions, but that the CFQ-R and the CFQoL are "comprehensive, reliable and valid instruments for adolescents and adults," while the CFQ-R is the only measure validated for children. Table 3.1 gives a summary of the CF HRQOL measures (adapted from (154)).

Table 3.1: HRQOL measures for cystic fibrosis				
Measure	Number of items/domains	Validated population	Pyschometrics reliability	Psychometrics validity
Cystic Fibrosis Quality of Life Questionnaire (CFQoL)	52 items 9 Domains (physical, social, treatment issues, chest symptoms, emotional responses, concerns for future, interpersonal relationships, body image, career)	14-adult	Internal consistency: α = 0.72-0.92; Test-retest r>0.80; no cross-informant reported	Discriminates between disease severity; sensitive to change following antibiotic treatment as measured by FEV%pred and BMI; Correlates with total SF-36 scores (0.64-0.74)
Cystic Fibrosis Questionnaire – Revised (CFQ-R)	Child version 35 items 8 Domains (physical, emotional state, social, body image, eating, treatment burden, respiratory and digestion) Adolescent/Adult version 50 items 12 domains (physical, emotional state, social, body image, eating, treatment burden, respiratory, digestion, role, vitality, health perceptions, and weight) Parent version 44 items 11 domains (physical, emotional state, body image, eating, treatment burden, respiratory, digestion, vitality, school, and weight)	6 – adult	Internal consistency: α = 0.34-0.74 (child); α = 0.54-0.94 (adolescent); α = 0.59-0.91 (parent) Test-retest: r = 0.45-0.90 (teen/adult) Parent-child agreement = 0.27-0.57	Teen/adult version is correlated with SF-36; child version correlated with PedsQL; discriminates between healthy children and children with CF; discriminates between disease severity; correlated with FEV%Pred, BMI, height and weight; Sensitivity to change following antibiotic treatment; used in Phase III clinical trial as primary endpoint
Questions on Life Satisfaction-Cystic Fibrosis (FLZM-CF)	9 items 9 domains (breathing difficulties, abdominal pain/digestive trouble, eating, sleep, routine therapy, adherence to daily therapy, significance for others, understanding, free from disadvantage)	16-adult	Internal consistency: α = 0.80 Test-retest r = 0.69 No cross-informant reported	Correlates with FEV%pred and amount of time spent doing treatment; discriminates between disease severity; responsive to change after inpatient stay

Preference-based measures

As has been discussed above, measures of health that give little indication of people's preferences for their health state are of limited use to analysts attempting economic evaluations. One of the major advances in health economics has been the development of the QALY for use in evaluations of healthcare technologies. The QALY attempts to measure the value of a benefit in terms of the impact on longevity with a quality of life element into a common numeraire of a year in full health. (111)

The QALY is a combination of quality of life and length of life into a single index number between 0-1, where zero is death, and one is perfect health. Certain assumptions must be made to allow the QALY to represent people's preferences over time. One is that health state values must be independent of the duration, timing, and sequence of health states. It is also common to assume that people are risk neutral. The QALY can be constructed by multiplying the value of a certain health state by the length of time spent in that health state. As outcomes are uncertain, the expected value of each possible outcome can be expressed by the value weighted by its probability. The expected outcome of the treatment can then be represented by adding the expected values of all possible individual outcomes. The net benefit, therefore, is the difference between the expected outcome with and without treatment. The benefit of an intervention at the population level is the aggregate of net benefits to individual patients. Therefore, if the expected benefit of an average patient is found, then the aggregate benefit of the population programme can be found by multiplying the individual benefit by the number of patients expected. (1) At an individual level, the QALY can be used to compare the benefits of medical interventions through the calculation of cost per QALY ratios.

One of the major criticisms of the QALY and the cost per QALY approach used in CUA is that the approach fails to capture non-health benefits. As non-health benefits cannot necessarily be converted into units of health gain, they are often ignored or converted into monetary terms and processed as negative costs. As CUAs are based on the idea that efficiency is the ratio of inputs to outputs, leaving out non-health benefits would mean it is not representative of the ratio of all inputs and all outputs. Therefore, in this case, a CUA would only be useful in determining technical efficiency (whether one programme is more efficient than another given a certain threshold) but not allocative efficiency. To be allocatively efficient CUAs using the cost per QALY measure would need to know the decision maker's (be it at the societal or local level) cost-effectiveness threshold for the good or service in question. If that threshold is not known, then basing a decision on a CUA could lead to a less than optimal decision.

Disease specific measures versus generic measures

The scope of measurement is a major aspect of the description of benefits. For example, in economic evaluation, disease specific measures are often used, as are general measures of quality of life (i.e. QALY). The argument for using disease specific measures is that the specificity of the measure enables a more sensitive account of a given dimension of health. Also, disease specific measures cover important dimensions of the given condition that may be missed by generic measures. However, disease specific measures often miss the impact of co-morbidities, side effects, and are not usually comparable across programmes. The argument for using generic preference-based measures in economic evaluations is that they allow for comparisons between interventions. Generally, these measures

are derived by using a generic descriptive system (often using a spectrum of different health states), and a single preference based algorithm for creating an index value. Generic measures may also capture co-morbidities that are missed by disease specific measures but may also be inappropriate or insensitive for many conditions (97) and therefore disease-specific measures may be more useful in certain situations.

Because both methods have their weaknesses, it has been argued that both should be used in the same study. Another approach would be to create a specific measure that takes into account co-morbidities and/or side effects or to use a vignette that incorporates disease specific treatment effects. (1)

Conclusions

Clearly measuring health is anything but straightforward. Practically, the measure a researcher chooses should be able to address the question being asked. If the researcher wishes to carry out (or have study results use in) an economic evaluation, then the measure of health needs to be a valid measure of the health dimension and status for utility estimation and preferably be able to be compared with other interventions. However, the measurement of health is only one aspect of a project. Most evaluations require some method of valuing a health outcome.

For the CEA within this study, I use a net benefit approach to evaluate the benefits of newborn screening using the QALY as the health benefit measure. I have used QALY estimates of paediatric CF from the literature. (70) However, it should be noted that using self-reported measures of health in paediatric populations is problematic, as they or a proxy may not be able to accurately reflect changes in

health. The choice of using QALY estimates was made in this case, as there are no agreed upon clinical measures of benefit, and using QALY estimates allow for comparison to other disease areas.

Analysing Health Economic Data

Once cost and/or benefit data have been collected and measured appropriately for the hypothesis in question, it is important to analyse the data to look for any patterns or other information that can help answer the question at hand. Data analysis in health economics usually involves the use of appropriate statistical and/or econometric methods. In cost studies, econometrics is often used in order to find the major cost drivers of the overall costs. For economic evaluations, decision modelling is often undertaken. It is also an important step to carry out sensitivity analysis. The next sections discuss the development of the methods to be used in this thesis.

Cost-of-illness methods

In 2006, Akobundu et al (158) carried out a systematic review of COI studies in health and analysed the various methods used within those studies. The authors found that more recent studies tended to focus on narrowly defined diseases, rather than wider burden of disease studies common in the past. The authors suggest that this is because of a possible shift in funding from government to private sources (pharmaceutical companies) and a general public interest in high-profile diseases.

COI studies have often come under criticism in terms of their ability to add to a policy making discussion. One reason is because most COI studies tend to estimate the monetary costs of illness, instead of the economic, or opportunity cost, of illness. Another was that studies typically focused on total costs, rather than marginal costs. (34) Bloom et al (159) found that published COI estimates are of

limited value when making decisions as there is wide variation in methods and within diagnosis categories. This is partially because methods for carrying out COI studies are not standardised. For example, some studies include the impact of co-morbidities, while others do not.

In general, COIs attempt to estimate either total or incremental costs. Total disease costs provide estimates on the entire healthcare expenditure of people diagnosed with the disease over a period of time. Incremental costs, on the other hand, estimate the increase in costs that is attributable only to the presence of the disease.

Total cost methods include summing all (for ease, this method will be called the SumAll) the health care expenditures on a patient with the disease, regardless of whether these are attributable to the disease, or a more specific approach, where only costs believed to be attributable to the disease are counted (SumDisease).

For incremental analysis, there are two general approaches used. One method uses a matched control situation where a group of patients with a disease, is matched via demographic or clinical characteristics to a control group. The SumAll method is then applied to both groups and the difference is the incremental cost of having the disease. Finally, regression analysis can be used to derive a COI estimate from the estimated coefficient of an indicator variable for the diagnosis of disease in the regression model. (160) This method is also useful to indicate the impact of co-morbidities.

Ideally, COI studies would include all relevant individuals and would isolate the costs specifically due to the disease of interest. However, it is rarely the case that this is

done. The SumAll method is the easiest and the most straightforward. It is also a useful method to employ for cross-country studies. This method is useful in disease areas where non-disease related costs are minimal (for example, AIDS). However, in a disease such as coronary heart disease, there may be significant costs from other co-morbidities. In this case, the SumAll method may overestimate the COI as it will attribute costs to CHD that are not directly related.

The SumDisease method, on the other hand, may underestimate the COI if it fails to include all the relevant individuals and other spillover costs. For example, a person admitted for a heart attack may be included in a study of CHD, but not for high blood pressure. Costs could be attributed to both of those diagnoses, but it is unclear in what proportion.

The incremental methods, matched control and regression, provide methods of isolating the costs specifically due to the disease. Therefore, they may provide better estimates of the true cost of illness. The matched control method assumes that there is no need to adjust for confounding factors after a matching algorithm has been applied. However, overestimates are possible using this method due to other confounding factors than medical conditions, such as demographics. Regression methods use various techniques to control for confounding factors. It has been shown that if done incorrectly, significant bias can be introduced into the parameter estimate on the disease. (161) The regression method can also introduce bias if the underlying cost distribution issues are not dealt with appropriately (usually by transforming costs into log (costs)). Model misspecification is also a problem with this method, and occurs when the statistical model imposes a structure on the underlying data that is not supported by the observed data (for instance distribution,

parameter invariance[11], etc.). (163) Given these limitations, regression methods generally work well where disease is not necessarily linked to genetics, lifestyle choice, or other unobservable factors. None of the methods for COI are perfect; however, the regression approach is the most robust in controlling for co-morbidities.

Reporting issues

Reporting of arithmetic mean COI estimates is not necessarily helpful for decision makers. Sometimes, the mean COI will also equal the COI for the average individual. However, if costs are skewed, then the average cost of a given disease will not equal the cost of disease for the average patient. Therefore, it is important also to report median costs and the skewness coefficient (or other indicators of the distribution of costs).

Econometric methods for cost of illness data

It is important to understand the associations between patient characteristics and costs following medical treatment. This helps health care administrators and clinicians make decisions when resources are limited. In order to do this, health economists frequently use multiple regression models to determine these associations. These methods can allow inter-institutional comparisons of the cost of treatment, predicting the impact of new interventions, and to evaluate the

[11] Invariance in this circumstance refers to the idea that values of item response parameters should be identical for different groups of respondents. 162. Galdin M, Laurencelle L. Assessing parameter invariance in item response theory's logistic two item parameter model: a monte carlo investigation. Tutorials in Quatitative Methods for Psychology. 2010;6(2):39-51.

relationship between certain constructs such as socio-economic status and methods of care delivery.

Types of data

The success of any econometric analysis depends on the available data. Generally three types of data can be used for empirical analysis: time series, cross-sectional, and pooled data. Time series data is a set of observations for variables that occur over time. Cross-sectional data are data on one or more variables collected at the same point in time. Pooled data generally contain elements of both time series and cross-sectional data. Panel data and longitudinal data are types of pooled data. Panel data generally consists of multiple observations over time for the same groups/individuals. Longitudinal data involves collecting repeated observations over time and can be cross-sectional, cohort, or panel data.

This data can be quantitative or qualitative. Quantitative data is generally categorized by being discrete or continuous. Discrete data can only take specific numerical values (i.e. number of siblings or shoe size) whereas continuous data can take any numerical value (i.e. mass or distance). Categorical data can be nominal (no order in categories), ordinal (order in categories) or binomial (two categories). Qualitative data (i.e. eye colour, place of birth, etc.) can sometimes be transformed into quantitative data to aid in statistical analysis by creating binomial data out of a categorical variable (i.e. a patient has blue eyes or not).

While useful tools, there are no agreed upon gold standards for regression techniques. Certain models can be ruled out due to the nature of the data collected. As the data for this thesis is longitudinal data, with both categorical and continuous

data, the rest of this section will focus on the methods used when analysing longitudinal cost and health care data (Generalised Linear Models (GLM) and General Equilibrium Equations (GEE)).

Ordinary Least Squares

OLS Linear regression using OLS estimation can be used to determine associations between patient characteristics and costs. A simple linear regression model can be described as

$$Y = X\beta + \varepsilon \text{ where } \varepsilon \sim N(0, \sigma^2) \quad (1)$$

Assuming that E[ε] = 0 and that var[ε] = σ^2, OLS results in an unbiased estimation of Beta. OLS estimates also achieve minimum variance of the linear unbiased estimators under the Gauss-Markov theorem. (163) Another advantage of using linear regression is that the model is additive, with the regression coefficients being interpretable as an increase in cost for a one unit increase in the predictor variable. However, for hypothesis testing and testing the statistical significance of the regression coefficients, the error terms must be normally distributed. As cost data is usually positively skewed and subject to outliers, the distribution of error terms tends to be positively skewed, as opposed to normal, therefore the inferences made from the significance of the regression coefficients will be biased. (164)

One method often used to overcome the positive skew of cost data is to use a log transformation. This often tends to normalise the distribution of costs, therefore satisfying the distributional assumptions of a linear regression and therefore allowing valid statistical inferences to be made of the regressors. One of the major

drawbacks of using this method is that the model is interpreted on a multiplicative scale with regression coefficients being the logarithm of the relative or proportional change in median cost with each unit increase in the predictor variable. Interpreting the relative change in median cost is difficult to interpret.

Median regression is commonly used to model data that have a skewed distribution or subject to outliers. This method contrasts with linear regression as it uses the median cost instead of the mean cost as a function of patient characteristics. As such, median regression is less sensitive to outliers than OLS regression, and therefore the data does not have to have a normal distribution. Regression coefficients can be interpreted as the change in median cost in relation to a unit change in the predictor variable. (164)

Generalised linear models

Generalised linear models are a more flexible generalisation of OLS models. The generalisation is brought about by allowing the linear model to be related to the response variables via a linking function and by allowing the variance of each measurement to be a function of its predicted value. (165)A detailed technical discussion of GLMs can be found in McCullagh and Nelder (166). Under GLM assumptions, the dependent variable, Y, is generated from a distribution function (i.e. normal, binomial, Poisson, etc.). The independent variables determine the distribution of the mean, μ, thus:

$$E(Y) = \mu = g^{-1}(X\beta) \quad (2)$$

where E(Y) is the expected value of Y; Xβ is the linear predictor, and g is the link function. The unknown parameters, β, are typically estimated using maximum likelihood methods or Bayesian techniques. (163)

The maximum likelihood method uses an iterative re-weighted least squares algorithm to come up with estimates. However this method can be prone to bias if a small sample is used. (167) Bayesian methods are used when the posterior distribution is approximated using a prior (observed) distribution. This method will usually use techniques such as Markov chain Monte Carlo (MCMC) simulation or Gibbs sampling in order to find the estimates.

GLM with a logarithmic link function has been tested for analysing non-negative continuous outcomes that are subject to skewness (cost data, for example). (168, 169) In 2003, Austin et al (170) analysed the determinants of cost in CABG therapy. The authors compared the performance of linear regression, linear regression with log-transformed cost, generalised linear models with Poisson, negative binomial and gamma distributions, median regression, and proportional hazards models. The authors found 'structural sensitivity', that is, that without a gold standard for model structure any model could be useful for identifying factors that are associated with increased costs, but the models differ in terms of assessing the magnitude of effect on cost for a given variable. The authors found the additive models were the easiest to interpret and the GLMs were best at predicting the risk-adjusted costs for the next patient admitted for surgery.

Generalised estimating equations models

In their 2006 textbook, Hedeker and Gibbons (160) argue that because GLM models are fixed effects models that assume observations are independent of each other, they are not appropriate for longitudinal data analysis. The authors suggest that extending GLM models to account for correlation that is often inherent in longitudinal data would be a more appropriate tool. Liang and Zeger (171) did this in 1986 when developing GEE models.

A basic feature of GEE models is that it is only the marginal distribution of y_{ij} (where *i* = individual subject) at each time point (*j*) that needs to be specified. In other words, GEE models assume that y_{i1} and y_{i2} are two univariate normals, rather than assuming y_{i1} and y_{i2} form a joint normal distribution. Another key feature is that GEE models treat the covariance structure of the longitudinal data as a nuisance, and instead focus directly on the regression of y on X. This means that even with a misspecification of the variance structure, GEE models yield a consistent, normal solution for the regression coefficients β.

Similarly to GLM models, a linear predictor is specified ($\eta_{st} = x'_{st}\beta$) where x'_{st} is the covariate vector for subject s at time t. Then a link function is chosen ($g(\mu_{st}) = \eta_{st}$). Common link functions include identity, logit, and log link for continuous, binary, and count data, respectively. Identity link functions assume a like for like replacement, logit links assume the probability of the response being a "yes", and log links return the natural log of the data. (160) The variance is then a function of the mean ($V(y_{st}) = \phi v(\mu_{st})$) where $v(\mu_{st})$ is a known variance function and ϕ is a scale parameter that can either be known or estimated.

The final specification in a GEE model is for the correlation structure of the repeated measures. The "working" correlation matrix is the size *n* x *n* as the assumption is that there are a fixed number of time points *n* at which subjects are measured. However, a useful feature of GEE is that a given subject does not need to be measured at every time point. Each subjects correlation matrix, R_i is the size n_i x n_i with the appropriate rows and columns removed if there are missing observations. It is assumed that R_i depends on a vector of association parameters, α, where α is assumed to be the same for all subjects, therefore representing the average dependence across all subjects.

Choosing a form for the correlation structure should be consistent with the observed correlations, however it does not have to be. While GEE allows for misspecification, it loses statistical power (efficiency) if the choice is incorrect. Typical correlation structures are detailed below where R_i (α) is the correlation matrix that is specified by the vector of parameters (α); s = subject and t= time ($s't$ = for all s and t) :

- Identity matrix: R_i (α) = I,
- Exchangeable correlation: corr (Y_{is} Y_{it}) = α; $s't$.

So that the correlation matrix for ith individual is defined as R_i (α) = corr (Y_{is} Y_{it}) = α; $s't$.

- Autoregressive correlation: corr(Y_{is} Y_{it}) = $\alpha^{|s-t|}$; $s't$. Here, α is a correlation value and thus a fraction. So, in this type of correlation, we consider that for all t > k, $\alpha^{|s-t|} > \alpha^{|s-k|}$ Then the correlation matrix can be defined as (R_i (α)) = corr(Y_{is} Y_{it}) = $\alpha^{|s-t|}$; $s't$.

- Unstructured or Pairwise Correlation: corr (Y_{is} Y_{it}) = α_{st} ; $s't$. The correlation matrix can be written as (R_i (α)) = corr (Y_{is} Y_{it}) = α_{st}; $s't$ where, $\alpha_{s,s+1} = \alpha_{s+1,s}$; s = 1, 2,…..,T

Solving GEE involves iterating between the quasi-likelihood solution for estimating β and a robust method for estimating α as a function of β. The following steps are repeated until convergence achieved:

1. With given estimates of $R_i(\alpha)$ and ϕ, calculate estimates of β using iteratively reweighted least squares where iterative estimates of α are used to yield new estimates of β.

2. With given estimates of β, obtain estimates of α and ϕ. To do this, Pearson residuals are calculated $r_{st} = (y_{st} - \mu_{st}) / \sqrt{[V(\mu)]_{tt}}$, which are then used to consistently estimate α and ϕ.

Upon convergence, hypothesis testing and confidence intervals are constructed by obtaining the standard errors associated with the estimated regression coefficient. The standard errors are obtained by taking the square root of the diagonal elements of the variance matrix. (160)

GEE models are useful for longitudinal data as various types of data can be used to estimate the impact of covariants and the structure of the variance-covariance matrix is not as critical as in other models. However, it is not a useful method if there is interest in the variance-covariance structure. GEE also makes the somewhat restrictive assumption that missing data are missing completely at random. This

may not be appropriate for all data sets. However, using Bayesian extensions to the GLM approach can get around this problem. (160)

Conclusions

Despite several limitations, cost-of-illness studies are used to inform decisions or at least to highlight the magnitude of the financial impact of diseases on the health budget. (30) The methods to gather information on costs and to analyse the cost data should be explicit to enable a reader to determine the strengths and weaknesses of the study, as well as if the study results are applicable to their circumstance.

For the COI in this study, all direct and indirect hospital-related costs to the English NHS will be collected for all registered CF patients in the east of England. A SumAll method of cost collection will be used, but with no health population control group for comparison. This method is appropriate for CF as non-CF related medical costs are likely to be minimal. Once estimates of costs for each patient are calculated, an econometric investigation into the drivers of costs will be undertaken. Descriptive statistics of cost and key demographic and clinical data will be presented for each calendar year. Given the nature of the data (time-series panel data where not all individuals are in the model the whole time), GEE models will be used to identify relationships between the various demographic and health variables and total cost per patient.

Economic Evaluation Methodologies

Cost-effectiveness Analysis

Cost-effectiveness Analysis (CEA) is a form of full economic evaluation in which both the costs and the benefits of a health intervention or programme are analysed

in comparison with an alternative programme with the same purpose. CEA is most useful in situations in which a decision maker has a limited range of options within a given area on which to spend a given budget.

One of the most important decisions to make when carrying out a CEA is what measure of effectiveness should be used. For example, if someone is looking at a newborn screening programme, it may be of interest to look at how many cases are being detected.

However, screening programmes are generally put in place to reduce morbidity and mortality, and 'cases detected' would not necessarily give an accurate measure of outcome. In this case, a life-years saved approach may be more appropriate, but it would still not account for morbidity. It is for these reasons that cost-utility analyses (CUA) are gaining in popularity, as the QALY measure is able to incorporate information on length of life and quality of life changes. However, sometimes the question needing to be answered lends itself to the use of CEA and it is important to capture a useful effectiveness measure.

Examining cost-effectiveness data sources

Finding useful and good quality effectiveness data, however, can be problematic. A major source of effectiveness data is published medical literature. This data raises issues of quality (especially freedom from bias and precision of estimates), relevance, and comprehensiveness. Guyatt et al (96) and the Cochrane Collaboration (40) both offer guidance on checking the quality and robustness of effectiveness data. Drummond (1) suggests that probably the most important aspect of study design is to choose the study design that minimises selection bias, which

can be true if a random allocation of patients into treatment groups has taken place. The 'worst' most biased evidence would come from case-series data with no controls.

However, it is often the case that even the 'best' source of evidence may not have everything necessary for an evaluation, and therefore other sources of evidence are needed. In many cases, clinical trial evidence does not exist or not in sufficient numbers to provide a precise result, and therefore other sources of evidence are the best possible sources of data (e.g. expert opinion).

A final issue to consider when judging effectiveness data is the comprehensiveness. Freemantle and Maynard (172) and Coyle and Lee (173) argue that not using all the available evidence can lead to changes in the final outcome of the analysis. One method of ensuring a comprehensive data set is to carry out a systematic review. Systematic reviews use various methodological principles (inclusion/exclusion criteria, search techniques, clinical details, statistical procedures, sensitivity analysis, etc.) to collect the most robust and relevant evidence from the published literature.

In situations where there is no good clinical evidence available, assumptions can be made on clinical evidence, and then undertaking sensitivity analysis of the economic results using different assumptions. In theory, this should yield information about whether the final result is sensitive to an estimate used in the analysis. (37)

Incremental cost-effectiveness ratios

The most common approach to presenting the results of a CEA is in terms of the incremental cost-effectiveness ratio (ICER). The ICER allows a comparison between two or more health interventions. It is calculated by the simple formula:

$$(C_1 - C_2)/(E_1 - E_2) \text{ or } \Delta C / \Delta E$$

or the change in costs divided by the change in effect. This shows how much we are paying for each additional outcome. ICERs are often illustrated on a cost effectiveness plane. Many interventions when compared to standard care fall in the northeast quadrant, as they increase cost, but also effectiveness. However, it is often the case that one intervention will dominate another. This occurs when it is both more effective and less costly than its comparator. If multiple interventions are being compared, there is the possibility of extended dominance. This occurs when a treatment's ICER is higher than that of the next, more effective, alternative. In order for a programme to be considered 'cost-effective', the ICER must fall below a certain threshold that is dependent upon a budget constraint. If the intervention lies above this threshold, then it is considered not 'cost-effective'. (1)

There are generally two schools of thought on how to set the threshold. One method assumes a fixed budget constraint, and sets a threshold that would maximise efficiency (i.e. health gains). Each health care programme would be included if it had an ICER below that threshold, and programmes would be replaced by more efficient ones as technology develops. The second method would assume that the threshold should represent some intrinsic social valuation of a QALY compared with anything else we want (willingness-to-pay). Using a societal willingness-to-pay threshold assumes that there is a flexible budget constraint, as it

implies that every intervention with an ICER below the societal WTP for a QALY (or LYG, etc.) should be financed. This approach requires decision makers to know what the societal value of a QALY, which can be problematic. Even if a context-specific societal WTP per QALY was used, the method runs a risk of reaching untenable budget requirements. Using previous decision as a benchmark for identifying the societal WTP for a QALY is an alternative method, but given decisions are generally context-specific, this method may produce misleading results. (174)

In the UK, the National Institutes of Clinical Excellence have used a relatively arbitrary £20,000-£30,000 per QALY as a threshold for many years. Other countries use similar figures (i.e. $50,000 per QALY in the US). However, until recently, there has been little empirical evidence that these figures are correct. (175) In 2011, Donaldson et al (176) published research carried out in the UK on the social value of the QALY. The authors used two methods to determine what the threshold for the ICER should be. The first method modelled the current UK value of a prevented fatality used in transport policy combined with data on fatality age, life expectancy, and age-related quality of life. The second method used surveys to combine respondents' answers to willingness to pay and health state utility questions to calculate the social value of a QALY. The modelling methods yielded results between £10-70,000 per QALY, while the survey research resulted in a value of £18-40,000 for a QALY. The authors concluded that the research did not provide "compelling" evidence that the NICE threshold should be changed in any way.

The ICER itself does cause some problems when analysing data. One reason is that the ICER does not tell anything about the size or scope of the interventions

under considerations. The second is that the ICER presents some statistical difficulties. One of those difficulties relates to the issue of negative ICERs, that is, when the comparison yields positive effects, but negative costs, or vice versa. The problem with this is that the ICER point estimate could be in the northwest or the southeast quadrant of the cost-effectiveness plane. From a deterministic analysis standpoint, this is not necessarily a problem as it is clear which quadrant the ICER is in.

However, the uncertainty around the point estimate could span more than one quadrant. Another issue with the ICER is if the change in effects is zero. This causes the ratio to be infinite. Again, this may not be a problem for deterministic analysis, as one intervention is likely to dominate in terms of cost. Again, the uncertainty surrounding the point estimate causes problems as the effects could be negative, meaning there is no way to determine the variance of the ratio. Finally, as a ratio statistic, the ICER is not easy to use as a dependent variable in regression analysis. This can hamper efforts to adjust for confounding variables. (1)

One method that is often advocated for circumventing the problems associated with exploring the variability in the ICER is non-parametric bootstrapping. This method ignores assumptions about the underlying distribution in the ICER, and instead re-samples the original data to obtain an empirical estimate of the sampling distribution of the ICER. The first step of this process is to draw a sample from the original observations from the ‘treatment group’ and a sample from the original observations of the ‘control’ group. The bootstrapped ICER is then calculated. These steps are then repeated a large number of times to empirically estimate a sampling distribution of the ICER. (177)

Another method is to avoid using the ICER altogether and adopt a net benefits approach. This approach does not use a ratio, and instead places both costs and benefits on a single scale (net monetary benefit (NMB) or net health benefit (NHB)). The net benefit is calculated by changing the difference in effects into a monetary value using a threshold willingness-to-pay for a unit of effect. The difference in costs between the interventions is then subtracted from this value. As the threshold value is rarely known in reality, it is common to present the NMB as a function of the threshold ratio. NMB is zero when the ICER is equal to the threshold ratio. (1)

Decision Analytic Modelling

As mentioned before, economic evaluation evidence comes from various sources. One of the most powerful tools in health economics are decision analytic models. This tool aims to bring all different types of evidence together for use in attempting to solve the particular problem the decision maker is facing. Decision analytic modelling has theoretical foundations in statistical decision theory, expected utility theory, and Bayesian statistics. As a method, it

- provides a structured framework which takes into account uncertainty,
- can accurately reflect patient pathways through an intervention,
- makes use of all evidence available,
- translates evidence into estimates of costs and effects (for CEAs, CUAs, and CBAs),
- enables uncertainty analysis,
- can identify priorities for future research.

Of course, to do all these things accurately, a modeller needs to ensure that all relevant options are considered, that all evidence has been gathered, and that any intermediate outcomes are linked to their final endpoints. It is also important to determine the relevant time horizon for the evaluation. Trial data rarely follows up patients long-term, and therefore, if the question is about long term effects, then certain methods of survival analysis will be necessary to extrapolate the time horizon beyond available data. Finally, it is important to make the results applicable to the decision making context. This seems obvious, but it is often the case that the evidence available does not necessarily translate into the relevant decision question. The use of decision models can translate existing evidence into data that can help inform the relevant policy decision. (37)

There are two key elements to decision analysis: probabilities and expected values. Probabilities are generally thought of as a number that indicates whether an event will or will not happen. Joint probability refers to the probability of two events occurring at the same time. Conditional probability refers to the likelihood that one event will happen given another event has already occurred. Whereas independence refers to the fact that two events are unrelated.

Expected values, on the other hand, are the expected costs and outcome of each possible pathway a patient takes in an analysis. For example, a patient can have a surgical procedure or a medicinal intervention. At the end of each of those paths lies an expected outcome and associated costs. Expected costs are those of the intervention plus the therapy weighted by the probability of the patient following each pathway. Using this approach, incremental analysis can be undertaken.

Developing a Decision Model

As with any scientific method, the first stage of developing a decision model is to determine the question. For example, it is important to identify the relevant patient groups, relevant treatments, and time horizon, to name a few considerations. As all models are simplifications of reality, it is important to explicitly state the boundaries and assumptions surrounding the model.

Determining the structure of the model is a key stage of development. Most economic evaluations of health technologies require different types of model structure given the nature of the condition. For example, do events occur once or are they recurring? Is there more than one type of adverse event under consideration? Are probabilities stable over time, or do they change?

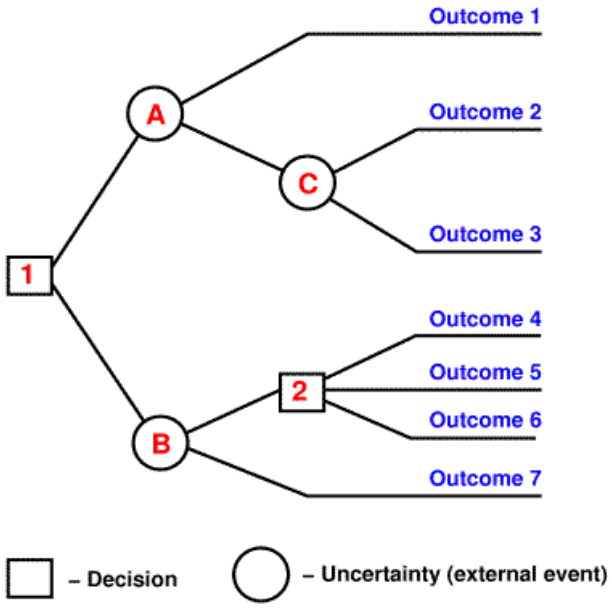
There are several different model structures available to overcome these problems. Two of the most commonly occurring methods are the decision tree and the Markov model.

Decision trees

Decision trees allow a modeller to present individuals' prognosis following an intervention along a series of pathways. Figure 3.3 illustrates a simple graphical example of a decision tree. Normal notation used for decision trees shows decision nodes as square boxes. These represent the decision being addressed by the model (i.e. to screen or not). Chance nodes, represented by circles, illustrate all the possible events that may happen to a patient after the intervention. These nodes represent areas of uncertainty. The branches all have associated probabilities.

These probabilities determine the likelihood that a patient would travel down this branch and experience the corresponding event. The pathways along which a patient travels are usually assumed to be mutually exclusive. Each pathway has a probability that is calculated by multiplying the initial branch probability by subsequent conditional probabilities (the probabilities that certain events occurred). All branch probabilities belonging to an intervention should sum to one. (37)

Figure 3.3: Sample decision tree



Each pathway in the tree also has costs (and benefits) associated with it. The costs associated with each pathway are the sum of costs for each event the patient passes through along the path. The expected cost is calculated by weighting each pathway cost by the pathway probability and then summing across all the pathways. The same method can be used to determine the benefits of the intervention. (37)

Decision trees, while common, do have limitations. Inherently, decision trees are static. In other words, they assume that the intervention takes place in a discrete period and do not include a time element. This is a significant limitation when trying to adjust survival duration for QALY analysis. However, some decision trees can include a time element through the use of Markov processes and tunnel states (for more on these techniques, see below and see Briggs et al (37)) Even when using these methods, however, it is difficult to account for patient 'history' as they move through the model over time. Another major limitation of decision trees is that they can become unwieldy when used to assess complex diseases, especially chronic diseases. For example, when modelling a cancer intervention, there are many possibilities for adverse events, and each would require its own pathway, likely with several pathways relating to the timing of events, creating a 'bushy' model structure that is time consuming to develop and to analyse.

Markov Models

Markov models were developed to tackle some of the weaknesses of the decision tree. Unlike the branch structure of decision trees, Markov models are based around health states. Time elapses in a Markov model, and the health state a patient occupies is determined probabilistically given a discrete period of time (a cycle). Cycle length is determined by the disease in question and the time period under investigation. Each state has corresponding costs and benefits and therefore expected costs and benefits can be calculated. The timing of patients moving from one health state is determined by transition probabilities. Transition probabilities can be calculated from the counts of individuals that move between health states in a given period (e.g. a year). These probabilities can be static or change over time.

The methods of calculating costs and benefits is similar to that of a decision tree, except that the weights are associated with the length of time patients spend in a certain health state. (37)

It is often the case that Markov models are embedded within a decision tree. This usually occurs when there is a discrete decision that can lead to long-term recurring events. However, it may be the case that a Markov model is not suitable for some prognoses. For example, Markov models are memory-less, that is, the nature of transition from one state to another does not depend on earlier transitions. This is difficult to justify in many chronic disease states (for example, in HIV, a previous sickness can further immunocompromise an individual, leading to a worse state in the near future). Sometimes, this can be overcome by adding states to the model. However, this can lead to a "bushy" model.

An alternative to the decision tree and the Markov models is to run a micro simulation or individual sampling model. Unlike decision trees and Markov models which use cohorts to analyse decisions, the micro simulation method uses individual patients, and tracks the patient's costs and benefits throughout the different states over time. This type of model has the advantage of retaining a 'memory' for each patient, and therefore offer greater flexibility over cohort models. (178) (179) However, these models are often computationally taxing, especially when additional probabilistic sensitivity analysis (PSA) is undertaken.

Dealing with uncertainty

There are various types of uncertainty that exist when carrying out economic evaluation. In terms of decision analysis, parameter and structural uncertainty are two major concerns.

Parameter uncertainty refers to the uncertainty surrounding the inputs into the model. This springs from the fact that measurements based on sampled date are estimated imprecisely. Generally, parameter uncertainty has been dealt with by using standard sensitivity analysis. This meant varying certain parameters and determining the effect on the model's results. This method is limited, as it will only allow for one, or perhaps a handful, of inputs to be varied at one time. The results of which are often difficult to interpret and do not provide an overall picture of parameter uncertainty. There is also no summary measure of uncertainty that can be communicated to decision makers. (37)

In light of these limitations, probabilistic sensitivity analysis (PSA) was developed. PSA allows an analyst to use probability distributions to reflect the different types of inputs. Secondly, PSA allows a study of the uncertainty in all of the input parameters simultaneously. This is often done by running a Monte Carlo simulation where the model is run a large number of times to gather a large number of sets of expected costs and effects. The distribution reflects the parameter uncertainty in the model. Finally, PSA allows confidence intervals around the outputs (ICER, or incremental net benefit) of the models. These are often presented as cost-effectiveness acceptability curves (CEACs) or by a scatter-plot of the simulations on a cost-effectiveness plane. (1, 37)

Investigation of structural uncertainty involves assessing whether the model specification is likely to be the best one, and the assumptions used to build the model are in fact sensible. One method of assessing this is by carrying out PSA which would weight each assumption relative to its plausibility. This is often called model averaging. (180) Another is by running multiple models and changing the baseline assumptions, often referred to as scenario analysis. Dealing with structural uncertainty is not necessarily straightforward. In 2006, Laura Bojke (181) conducted a review of decision analytic models used in Health Technology Appraisals to look at the methods used to address structural uncertainty. The author found that only scenario analysis had been undertaken, which was argued to leave the decision maker with a choice of multiple models without knowing what the "true" situation is likely to be. This is an area that needs further research but is beyond the scope of this thesis.

Conclusions

Despite several limitations, current methods of economic evaluation using decision analytic modelling are used for informing decisions. It is important when reporting methods and results of economic evaluations that researchers are as explicit as possible to enable future researchers as well as decision makers understand the strengths and weaknesses of the study.

For the CEA in this study, a discrete simulation model will be built, using Markov processes to allow the tracing of individuals over time as they move through the model. This model also allows for probabilistic analysis as well as an investigation of

uncertainty around the estimates. Results will be presented in terms of cost per QALY as well as the net marginal benefit of screening versus not screening.

Chapter 4: Cost of Illness

Introduction

The analysis was initiated to analyse the cost of caring for paediatric cystic fibrosis in the east of England. The analysis is a retrospective longitudinal study of the costs of tertiary cystic fibrosis care in the region. As mentioned in Chapter 3, a well-designed cost-of-illness study not only provides important information on the burden of disease, but also where costs occur, what influences costs, and how they vary.

The perspective of the study is from the NHS in the east of England. The analysis takes a bottom-up, prevalence based approach, using data gathered from the ERCFD and participating hospitals on paediatric patients with cystic fibrosis and national unit costs to estimate the cost of illness.

Methods

Study Population

Patients were recruited as part of a clinical and economics audit of the local clinical network at clinical centres in six counties in the East of England (Norfolk, Suffolk, Cambridgeshire, Essex, Hertfordshire, and Bedfordshire). Hospitals were invited to participate in the study if they cared for paediatric cystic fibrosis patients. All paediatric CF patients identified by the hospitals were invited to become part of the study and consent was taken by the consultant physician prior to inclusion. Data collection began in January 2009 and continued until December 2009.

Eastern Region Cystic Fibrosis Database (ERCFD)

Background

In 1998, the East Anglian CF Network[12] (see Figure 4.1 for a map of participating hospitals) created the ERCFD in order to achieve uniformity of care across the region and to monitor the health status and service delivery within the region (Bedfordshire, Cambridgeshire, Essex, Hertfordshire, Norfolk, Suffolk). A regional database was designed to collect routine clinical data from patient notes. The information was then disseminated to clinics across the region. The network originally hoped to collect data on an ongoing basis; however, financial constraints have only allowed data to be collected for 1998 and again in 2004, 2005, 2006 and 2007 for Cambridgeshire, Norfolk, and Suffolk. Data from this database was also routinely fed into the UK CF database and will be fed into the new national CF Registry when it comes online.

Objectives of the ERCFD

The objectives of the ERCFD are:

- To obtain longitudinal data on the health care status of children with CF
- To monitor and audit the provision of NHS services for children with CF throughout the Eastern Region
- To assess future service needs.

Data collection

[12] The East Anglian CF Network is a clinical network of health professionals in Cambridgeshire, Norfolk and Suffolk that have an interest in Cystic Fibrosis care. The network aims to standardise quality of care across the region and meet regularly for professional development and exchange clinical information and experiences.

Data was collected by James Jarrett, Julia Leigh-Smith and Fiona MacLean[13] from patient notes and annual review reports. Data on clinical and patient characteristics was collected in 1998 and then again in 2004, 2005, 2006, 2007. Appendix D shows the data collection sheet which was entered directly into the database. Any contact a patient had with the hospital was recorded as well as any drugs prescribed in Microsoft Access (182) database.

The database collects routine clinical and resource use data from all hospitals which care for paediatric CF patients on a regular basis within Cambridgeshire, Norfolk, and Suffolk. Figure 4.1 is a map of the region with the hospitals marked. The only Specialist CF centre in the region is Addenbrooke's Hospital, in Cambridge (marked by a cross in Figure 4.1). The other hospitals that care for paediatric CF patients in the region are Hinchingbrooke Hospital in Huntingdon, the Edith Cavell Hospital in Peterborough, the Queen Elizabeth Hospital in Kings Lynn, the Norfolk and Norwich University Hospital in Norwich, the James Paget Hospital in Great Yarmouth, the Ipswich Hospital in Ipswich, and the West Suffolk Hospital in Bury St. Edmunds

The clinical outcomes for assessment of severity and analysis of factors affecting cost that were used in the study were periods of hospitalisation, number of clinic visits, body mass index (BMI), weight for height, number of positive sputums for *Pseudomonas aeruginosa* (PA), *Staphylococcus aureus* (SA), and any other infections (i.e. aspergillus) and banding. Information was also collected on the

[13] JLS and FM were employed by Addenbrooke's Hospital to collect data for annual review for audit purposes for Norfolk, Suffolk and Cambridgeshire

patient's age, sex, genotype (if available), method of diagnosis (screened vs. clinical), and whether or not they were cared for in a specialist centre or had a shared care arrangement. Where necessary (in 2007, some patients already had bands assigned to them by their clinical teams) a retrospective banding exercise was undertaken by JJ based on the CF Trust suggested bands using the clinical and resource use information gathered from the ERCFD. The banding of the patients was verified by Dr. Richard Iles, consultant paediatrician at Addenbrooke's Hospital.

Figure 4.1: Map of hospitals participating in the ERCFD
(Specialist Centre marked with a cross)

Previous ERCFD findings

An in-depth analysis was undertaken on the 1998 data. (183) The authors of that study found that in general, the health status of paediatric CF patients within the region was similar to the national averages. They also suggested that patients have decent access to local CF clinical care resources, either directly or through shared care arrangements with another hospital. However, they found that those patients attending clinics serving less than 10 patients may be disadvantaged by not having a local CF team or community support. The study also found that the frequency of monitoring procedures as well as the availability of community care, social workers and psychologists varied across the region. While the effect of these factors on health status was not examined in detail, the authors suggested that any health effects should be investigated further to ensure the best pattern of care. The authors also mention that the lack of patients in the district hospitals make it difficult to justify creating specialist centres in all hospitals in the region.

Resource use

Direct costs included those for hospital care, outpatient care, and pharmaceuticals. Data on hospital care was recorded in the database in terms of the number of bed days in a hospital ward and the type of admission. Outpatient care consisted of clinic visits. A clinic visit typically consists of seeing the specialist CF team which includes a consultant physician, specialist nurse, nutritionist, physiotherapist, and psychologist. However, non-specialist centres typically only consist of a consultant physician, and a specialist nurse.

Costing

A prevalence-based costing model was applied to estimate the costs of caring for paediatric cystic fibrosis patients in the east of England. An NHS perspective was

taken, and the following cost components were included for each child with CF in 1998, 2004, 2005, 2006 and 2007: inpatient care, outpatient care (including clinics), and drug costs. Unit costs for drugs were derived from the British National Formulary for children. (184) Intangible costs were not measured. As the study is from an NHS perspective, productivity costs (such as parental time and loss of schooling) were also not considered. Cost data was adjusted to 2010 prices using the consumer price index to adjust the prices over time. As banding tariffs were not proposed until the end of 2007, previous tariffs for paediatric services were used in the initial analysis, with a subsequent analysis using the banded tariffs for each patient. A discussion of the implications of these changes is in chapter 6.

Direct costs were estimated by combining the resource use data collected for each year with a unit cost for England from national reference cost lists following a general costing formula:

$$Total\ cost = \sum_{i=1}^{n} \sum_{r=1}^{m} (frequency)_{ir} \times (unit\ cost)_{r}$$

where *i* is the *i*th individual (i = 1, ... *n*) and *r* is the *r*th service received or resource used (*r* = 1, ... *n*).

Statistical analysis

Descriptive statistics (frequency, arithmetic mean, standard deviation, and median) of cost and key demographic and clinical data are presented for each year. Histograms were made to allow for a visual analysis of the distribution of key variables (such as cost). Box plot diagrams were analysed to see if there are any significant changes in cost over time for various demographic and clinical indicators.

To find a distribution that fits the total cost data, the software @Risk (which is an extension of Excel) was used to test the best fit using the Chi-Square Statistic.

There are many different types of regression analysis, none of which are presumed to be a gold standard. Initially, a simple linear regression using backward steps (i.e. removing insignificant variables from each consecutive model) will be used on the baseline (1998) data to identify potential variables for the longitudinal analysis. For the longitudinal analysis, I have chosen to use generalized estimating equation (GEE) models. As discussed in Chapter 3, GEE models are useful for analysing data that is highly correlated and for data where respondents do not have responses for some time periods. Given that disease severity is likely to have a direct correlation on resource usage, this method seems to be appropriate.

I would hypothesize that age, shared care, and band (disease severity) would have a positive (increase) effect on cost. As patients get older, their probability of getting more ill increases, and therefore a subsequent increase in resource use and cost of care would likely follow. Although there is limited evidence, it seems likely that those patients who do not have regular access to a specialist team and receive their treatment in hospitals without a full complement of staff would be more likely to become ill, increasing the cost of care. I would fully expect that as a patient's disease becomes more severe and they move into the higher disease severity bands to be using the health service more frequently.

I would also hypothesize that being screened, and having a better BMI and weight for height measures would have a negative (decrease) impact on costs. Being identified early via screening enables health care professionals to use preventative

measures to ensure long-term health, and thus lower resource use in the long run. Similarly, a better nutritional status (BMI and weight for height) would indicate a healthier individual who would incur fewer costs.

The dependent variable for the models was total annual cost per person. Depending on the outcome of the initial regression, the likely covariates considered, in addition to year, are listed below and are all baseline measurements:

- Screened – categorical (yes or no)
- Sex – categorical (Male, female)
- Age – categorical (<4 years, 4-8 years, 9-12 years, 13-17 years)
- Shared care – categorical (yes or no)
- Band – categorical (1,2,3,4,5)
- Body mass index – continuous
- Weight for Height (Z-scores) – continuous

Model descriptions

Baseline regression analysis

The initial regression took the form of a linear backwards step regression. The model assumed the relationship between total cost and the regressors is linear. This model takes the form

$$y_i = \beta_1 x_{i1} + ... + \beta_p x_{ip} + \varepsilon_i \qquad i = 1,...,n,$$

Where ε_i is an unobserved random variable. As cost data is typically skewed the model will also be run with the natural log of costs. The backwards step regression takes out variables one or two at a time based on the F-test to create a model of "best-fit".

Generalised equilibrium equation models: identity and log link models with independent, exchangeable and unstructured correlation structures.

Using the variables identified by the initial regression, seven generalised equilibrium equation models were created. Each model carried out the same processes, but used different link, family and correlation structures that were outlined in chapter 3. For clarity, the process will be repeated again here:

First a linear predictor is specified as

$$\eta_{ij} = x_{ij}^{'}\beta$$

Where x_{ij} is the covariate vector for subject *i* at time *j*. A link function is then chosen

$$g(\mu_{ij}) = \eta_{ij}$$

with g representing the type of function (in this instance identity and log). The variance is described as the function of the mean,

$$V(y_{ij}) = \phi v(\mu_{ij})$$

Where $v(\mu_{st})$ is a known variance function and ϕ is a scale parameter. GEE models use a quasi-likelihood approach to express the GEE for β, described by Liang and Zeger (171) as

$$U(\beta) = \sum_{i=1}^{N} D_i' V[(\hat{\alpha})]_i^{-1}(y_i - \mu_i) = 0 \quad (1)$$

Where U(β) is the GEE estimator of β, $\hat{\alpha}$ is a consistent estimate of α,

$$D_i' = \frac{\delta \mu_i}{\delta \beta^T}$$

and V_i is a working or approximate covariance matrix of Y_i. The matrix can be expressed

$$V_i = A_i^{1/2} R_i(\alpha) A_i^{1/2}$$

Where, A_i = diag[var(Y_{i1}), ... , var(Y_{ij})], and is a n_i X n_i diagonal matrix and α is a vector of parameters associated with a specified model for corr (Y_i). This gives the estimating equations the form

$$U(\beta) = \sum_{i=1}^{N} X_i' A_i V_i^{-1} (y_i - \mu_i) = 0 \qquad (2)$$

The final specification for the GEE model is the structure of the correlations. A discussion of the different types of the correlation of Y is in chapter 3. Three model correlation structures will be tested, independent, exchangeable, and unstructured. Independent correlation structures assume that the correlations have no relation to one another. Exchangeable correlations assume that all of the correlations in the repeated measures are the same, or “exchangeable.” In other words, all the correlations are equal. The third correlation structure tested assumes that the correlations are unstructured. Given the regressions uses dummy, categorical, and continuous variables, it is important to take care when interpreting the results. For example, using the variable ‘band’ as an independent variable assumes that someone in band 2 costs twice as much as someone in band 1. To test the linear relationship between total cost and categorical variables like age group and band, pair-wise correlation tests will be undertaken. The models were built and analysed using STATA 11. (185)

To estimate model fit, a quasi-likelihood under the independence model criterion, or "QIC" was used. (186) (187)This test is similar to the Akaike information criterion, which is widely used to determine model fit for GLM regressions, but takes into account the quasi-likelihood theories that underpin GEE models. The Akaike information criteria is given by

$$AIC = -2LL + 2p$$

where LL is the log likelihood and p is the number of parameters in the model. The QIC is derived by making an adjustment to the above formula

$$QIC = -2Q(\hat{\mu}; I) + 2trace(\widehat{\Omega}_I^{-1}\hat{V}_R)$$

where I is the independent covariance structure used to calculate the quasi-likelihood and $\hat{\mu} = g^{-1}(x\hat{\beta})$ and g^{-1} is the inverse link function. The $\hat{\beta}$ are the coefficient estimates and $\hat{V}_R$ is a robust variance estimator obtained by a working variance structure R. $\widehat{\Omega}_I$ is obtained under an assumption of an independence correlation structure. The lower the QIC, the better the model fit. If $trace(\widehat{\Omega}_I^{-1}\hat{V}_R) \approx trace(I) = p$, then a simplified version of QIC called QIC_u can be used.

Results

Study population

Hospitals in Bedfordshire, Essex, and Hertfordshire declined to contribute to the study. Initial approaches to the hospitals in Bedfordshire, Essex, and Hertfordshire indicated that they would be willing to participate; however, after ethics approval was granted and project packs were sent to the consultants, the participating hospitals

either did not respond to enquiries from me, or declined to participate. The population of CF patients from those three counties would have allowed a “natural experiment” to take place as before 2007, they did not have routine newborn screening for CF. Also, there are no specialist centres in the three counties meaning the patients are all shared care patients with a majority of their care being carried out in London (where we did not seek ethics approval to gather information). As such, all results are from Cambridgeshire, Norfolk and Suffolk.

Table 4.1 summarises the study population results by year. The population at baseline (1998) was 170 patients. However, there were no significant demographic differences between the populations over time. The number of patients over time has been relatively stable, with new patients replacing those who transfer into adult care or die. However, there seems to be a dip in the number of patients in 2004. It is unclear why this has occurred, though the population measures are similar to the other years. Mean age for each year group was stable at approximately 8 years old and the sex ratio was similar over the time studied. The majority of patients were in some form of shared care and most of the patients in Norfolk, Suffolk, and Cambridgeshire were identified using newborn screening (132 patients compared to 38 clinically diagnosed patients). The unscreened patients in the region were either identified at birth through meconium ileus (5) or presented with symptoms and were identified as CF cases. A few patients also moved into the area from areas that did not screen for CF. There was limited information on patient’s genotypes, but the majority of those patients with information available were delta F508 homozygotes.

There were no significant differences over time in terms of clinical outcomes (lung function, weight for height, number of infections). Median lung function as measured

by FEV1 percent predicted was in the upper 80's with an interquartile range (IQR) from 71-100 depending on the year. Median weight for height measure has shown some (but not significant) improvement over time from baseline. Median infections with *Pseudomonas aeruginosa* (PA) have not changed over time. Infections of *Staphylococcus aureus* (SA) have fallen over time, while all other infections have stayed at the same median level with an IQR between 0-1 for SA, and 1-5 for other infections.

In 2007, the CF trust released the first report from the national database(10). The results from that report indicate that the population in this study are roughly representative of the national population though without having access to the data on the database, statistical checks cannot be made. However, the age structure, median BMI, lung function, and genetic structure are presented in Box 4.1.

Box 4.1: Comparison of national CF database statistics and ERCFD (2007)

Age Structure	**Nationally**		**Eastern Region**	
	N	Percent	N	Percent
0 to 3	605	23%	40	23%
4 to 7	621	23%	41	24%
8 to 11	663	25%	47	27%
12 to 15	773	29%	46	26%
Median FEV1% Predicted				
8-11	87%		85%	
12-15	82%		83%	
Delta F508 Homozygotes	54%		56%	

Table 4.1: Population Characteristics					
	1998	**2004**	**2005**	**2006**	**2007**
N	170	147	167	176	174
Male	82	68	81	88	90
Female	88	72	86	88	84
Screened	160 (94%)	132 (89%)	148 (84%)	151(86%)	147 (85%)
Not Screened	4	9	13	18	20
Meconium Ileus*	6	6	6	7	7
Delta F508 homozygotes**	21 (58%)	32 (57%)	41 (55%)	47 (58%)	45 (56%)
Mean age (SD)	8 (5.1)	7.8 (4.3)	7.93 (4.7)	8 (4.7)	8.78 (4.5)
Shared Care	144 (85%)	118 (80%)	132 (75%)	138 (78%)	135 (78%)
Median FEV% Predicted (IQR)	87.05 (72.6-100.2)	85.4 (73-100)	87.4 (74.2-100.1)	87.2 (73-100.3)	84.5 (71.2-96)
Median Weight for Height (IQR)	99.67 (57.3-163.3)	105.1 (74.7-164.1)	108.64 (73.7-221.1)	105.42 (69.9-100.3)	101.2 (66.2-160.4)
Median Infections (IQR)	3.3 (2-5)	2.8 (2-3)	3.02 (2-3)	3.12 (2-4)	3 (2-3)
Band 1	89	79	100	89	90
Band 2	60	50	48	69	61
Band 3	7	11	12	11	10
Band 4	11	7	7	7	12
Band 5	3	0	0	0	1
+indicates a significant difference from baseline year (95% level) *Meconium Ileus patients are clinically diagnosed at birth ** percentages represent percent of patients with genetic information available who are Delta F508 homozygotes					

Data analysis

Similar to the clinical and demographic data, there were no significant differences between years in terms of resource use or mean/median cost. Table 4.2 lists the total and mean resource use in terms of the number of clinic visits and days spent in hospital by patients over the time period as well as the unit costs and source of information. Table 4.3 shows mean and median costs broken down by clinic, hospital, drug and overall cost per patient per year. Confidence intervals at the 95% level and were bootstrapped.

Table 4.2: Resource use and unit costs (£2010)

Resource Use

Year	Total number of days in hospital (population mean)	Difference in total days (from previous year)	Total number of clinic visits (population mean)	Difference in total visits (from previous year)
1998	1512 (8.9)		988 (5.8)	
2004	1227 (8.3)	-285	867 (5.9)	-121
2005	1534 (9.2)	307	988 (5.9	121
2006	1346 (7.6)	-188	1081 (6.1)	93
2007	1328 (7.6)	-18	1031 (5.9)	-50

Unit Costs	**(UK£)2010**	**Source**
Clinic Visit	238.33	NHS Ref
Hospital Day	313.67	NHS Ref
Screening test	2.00	Addenbrooke's Finance
Proposed Banding tariffs for paediatric CF (provisional tariffs in place post-2007)		
Paediatric Clinic	445.63	NHS Ref
Band 1	542.02	NHS Ref
Band 2	1,893.91	NHS Ref
Band 3	2,276.79	NHS Ref
Band 4	3,368.70	NHS Ref
Band 5	1,269.77	NHS Ref

The total number of clinics increased over time, and the total number of days spent in hospital decreased, the changes are partly explained by changes in the number of children for whom data are recorded in each year. The mean number of clinic visits per child per year has remained relatively stable (between 5.8-6.1 clinic visits).The mean number of days spent in hospital on average has decreased slightly over time from 8.9 to 7.6. The increase (although not significant) in clinic visits is in line with good clinical practice, and the decrease in days spent in hospital is the preferred clinical outcome and is likely to reduce overall costs.

Table 4.3 shows the mean and median costs of the entire population (not split by any clinical indicators) by year. Figures 4.2 through 4.5 give a graphical representation of the trend in mean costs over time. Corresponding to the increase in clinic visits, mean and median clinic costs have increased over time, though not significantly. Median hospital costs have stayed at zero, meaning that over 50 percent of patients did not have a stay in hospital during any one year. Mean hospital costs have decreased from £2784 per person at baseline, to £2389 in 2007. Drug costs have also increased, from a median cost of £1450 in 1998 to £1514 in 2007. Mean drug costs have also increased from £1837 in 1998 to £2197 in 2007. This increase in drug costs is likely to be caused by more patients on expensive drugs such as rhDNase and inhaled tobramycin as well as changes in clinical care since 1998. Overall, median cost per patient has increased slightly over time (£3852.65 vs. 4249.94), though mean cost per patient has decreased slightly (£6118.16 vs. £6103.50). This is mirrored by total costs per year. At baseline, the overall cost of care in the region was £1,040,087, whereas in 2007, it was £1,062,008.

Table 4.3: Mean and median resource use and costs in 2006 UK £ *(Bootstrapped 95% confidence interval for mean; IQR for median)*

	1998 (N=170)		2004 (N=147)		2005 (N=177)		2006 (N=176)		2007 (N=174)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Costs										
Clinic Costs (95%CI)	1499.43 (1387-1612)	1290 (1290-1548)	1521.67 (1400-1643)	1548 (1290-1548)	1526.37 (1431-1621)	1548 (1290-1548)	1584.65 (1476-1693)	1548 (1290-1548)	1528.72 (1426-1631)	1548 (1290-1548)
Hosp Costs (95% CI)	2783.86 (1983-3583)	0 (0-1252)	2612.59 (1822-3402)	0 (0-0)	2875.10 (2060-3690)	0 (0-626)	2393.74 (1840-2947)	0 (0-1878)	2388.87 (1821-2957)	0 (0-0)
Drug Costs (95% CI)	1837.03 (1334-2340)	1439.76 (695-1790)	1891.47 (1336-2446)	1451.74 (695-1822)	1935.14 (1435-2435)	1398.85 (695-1556)	2093.92 (1596-2592)	1475.87 (695-1788)	2197.96 (1674-2721)	1514.02 (695-1949)
Ave total (95% CI)	6118.16 (5127-7109)	3852.65 (2986-4912)	6029.43 (5106-6952)	3694.38 (2999-4844)	6339.97 (5313-7365)	4360.02 (3497-5679)	6063.08 (5339-6788)	4355.83 (3300-5445)	6103.48 (5298-6909)	4249.94 (3558-5588)
Total cost	**1,040,087**		**886,326**		**1,058,775**		**1,067,103**		**1,062,008**	

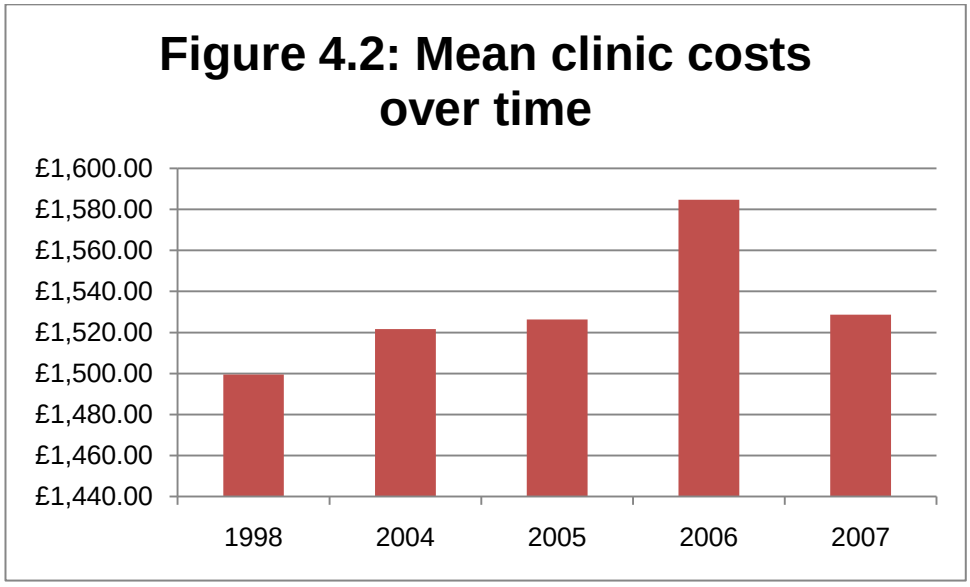


Figure 4.2: Mean clinic costs over time
£1,600.00
£1,580.00
£1,560.00
£1,540.00
£1,520.00
£1,500.00
£1,480.00
£1,460.00
£1,440.00
1998
2004
2005
2006
2007

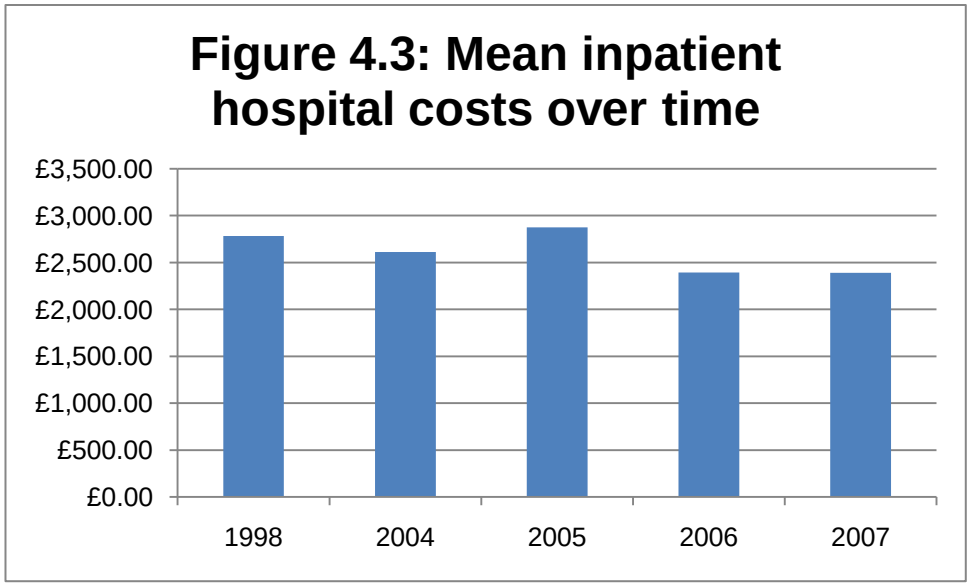
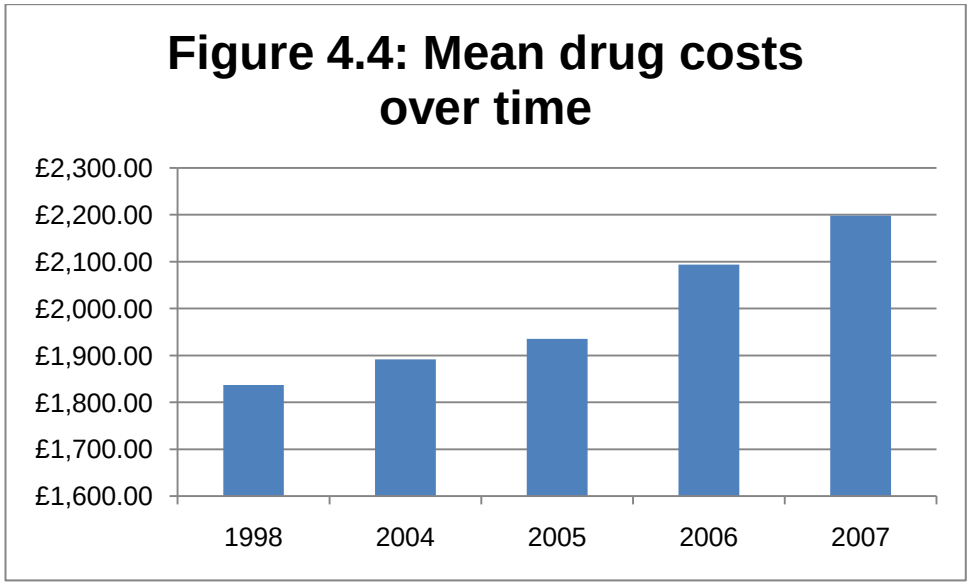


Figure 4.3: Mean inpatient hospital costs over time
£3,500.00
£3,000.00
£2,500.00
£2,000.00
£1,500.00
£1,000.00
£500.00
£0.00
1998
2004
2005
2006
2007



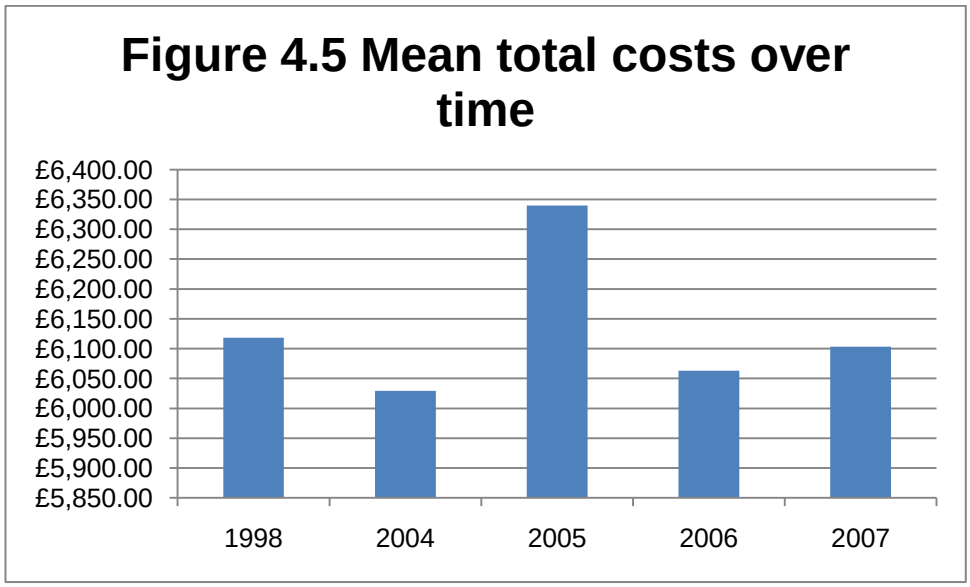
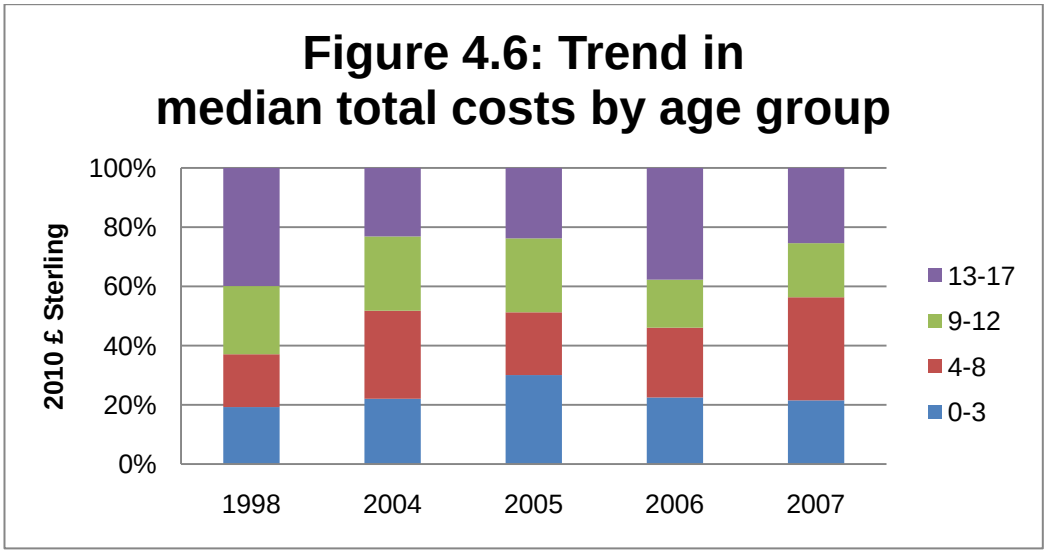


Table 4.4 gives a breakdown of median costs by age group. This table shows that the two younger age groups have an increasing median cost over time, the patients aged 9-12 seem to have a slowly decreasing median cost. It is difficult to find a trend in the oldest age group, although generally this group had a higher median cost at baseline and in 2007 (although in 2004 and 2005 it had a lower median than other age groups). While it is hard to tell an overall trend from this data, it may be

an indicator that more aggressive treatments being are being used to counteract chronic infections at a younger age and therefore causing later age groups to be more healthy. However, the spike in median costs in the 4th year of the study seems to throw doubt on that reasoning, but this may be an artefact of the data. Figure 4.6 shows a graphical representation of the trends in median total costs by age group over time. Looking at the differences in cost between groups, there may be an indication that this is the case as it looks as though when patients incur greater costs at an early age, they require fewer resources in the next age category. The regression analysis sheds more light on this aspect.

Table 4.4: Median costs over the entire time period by age group (2010 UK£)

Year	0 to 3 years (N)	4 to 8 years (N)	Difference (from 0-3)	9 to 12 years (N)	Difference (from 4-8)	13-17 years (N)	Difference (from 9-12)
1998	2,725.34 (39)	2,522.24 (55)	-203.10	3,267.50 (39)	745.26	5,646.34 (37)	2,378.84
2004	2,419.12 (31)	3,267.40 (50)	848.28	2,746.48 (51)	-520.92	2,545.80 (15)	-200.68
2005	3,326.62 (39)	2,349.89 (48)	-976.73	2,765.45 (54)	415.56	2,639.13 (26)	-126.32
2006	3,709.10 (39)	3,883.32 (50)	174.22	2,661.40 (60)	-1,221.92	6,215.78 (27)	3,554.38
2007	3,214.90 (31)	5,167.22 (50)	1,952.32	2,725.34 (57)	-2,441.88	3,786.32 (36)	1,060.98



Figures 4.7 and 4.8 look at the differences between the clinically diagnosed group and the newborn screening group in the population. It is clear that there are differences in cost between those patients that were screened and those that were not. However, it is difficult to draw firm conclusions as to the difference between groups. Indeed, running a two-sample t test showed no significant differences between groups, but this could be due to a small numbers problem. From a visual inspection of the graphs, both the mean and median costs of the unscreened patients are marginally less than the screened patients. Both groups have an increasing mean and median over time. Figure 4.9 shows the cumulative frequency plots of total cost s between clinically diagnosed and screened groups from all the data (not sorted for time). This plot reiterates the previous charts by showing that the two groups are similar in cost, with 80% of patients who were clinically diagnosed patients and patients identified by newborn screening patients cost less than £10,000 per year. Figure 4.9 shows both groups of patients, but also differences between 1998 and 2007. Again, the frequency plots show that the

groups do not have a dramatically different profile, though this figure shows that the differences between groups become even smaller between baseline and 2007. This is in contrast to the findings of Sims et al, who found clear differences between the groups. The authors found that around 75% of those who were screened had annual costs below around $4000, whereas only 50% who were clinically diagnosed had annual costs below $4000. (29)

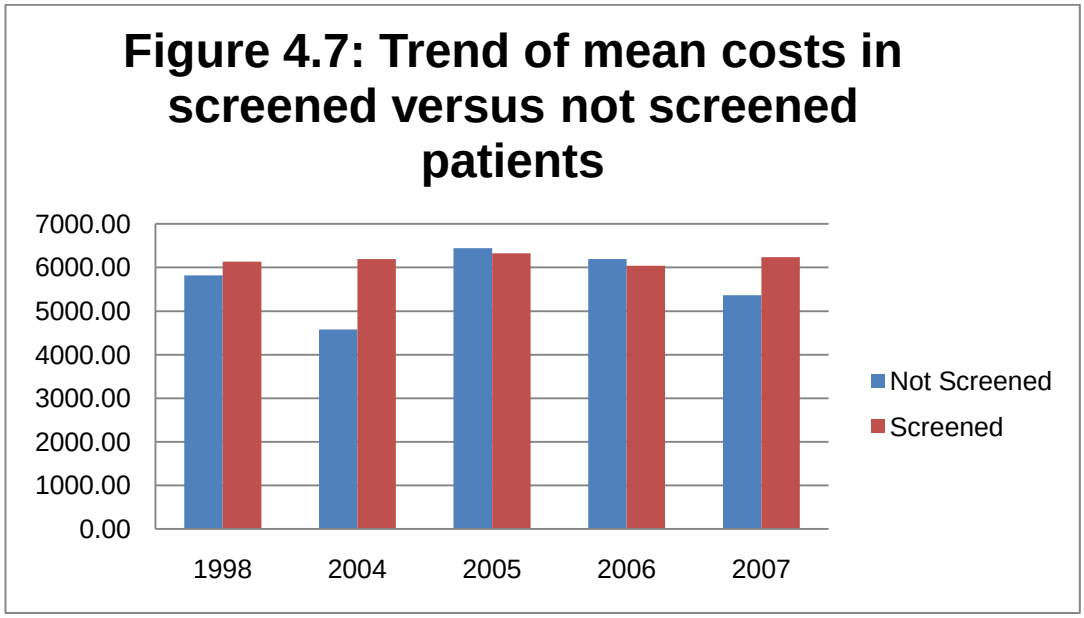


Figure 4.8: Trend of median costs in screened versus not screened patients

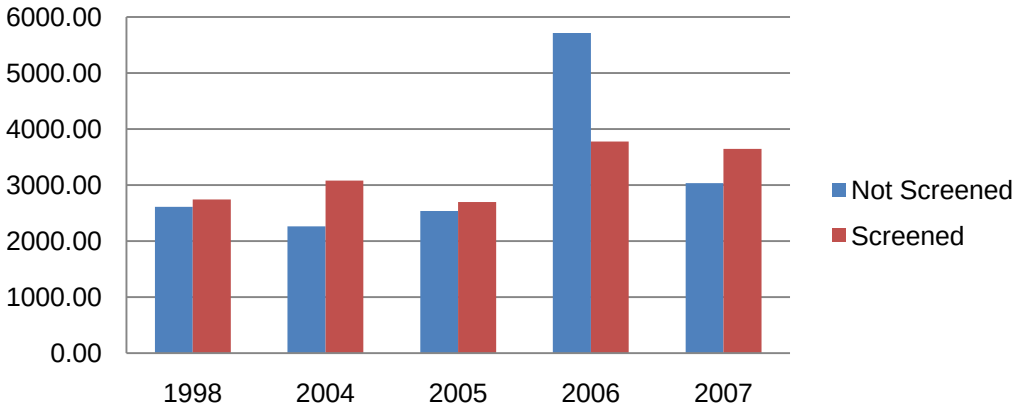
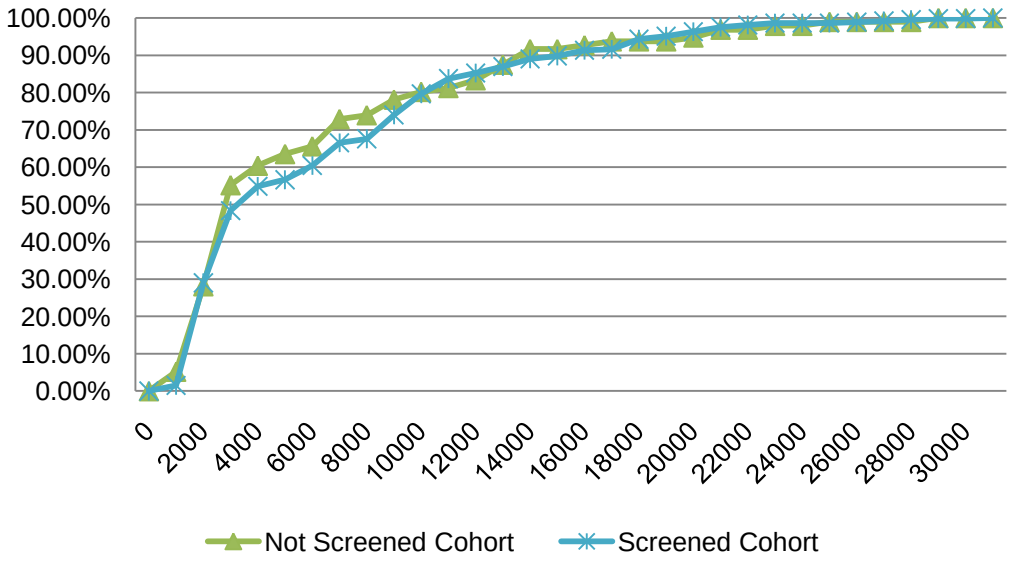


Figure 4.9: Cumulative percentage plots of total costs between screened/not screened (all observations)



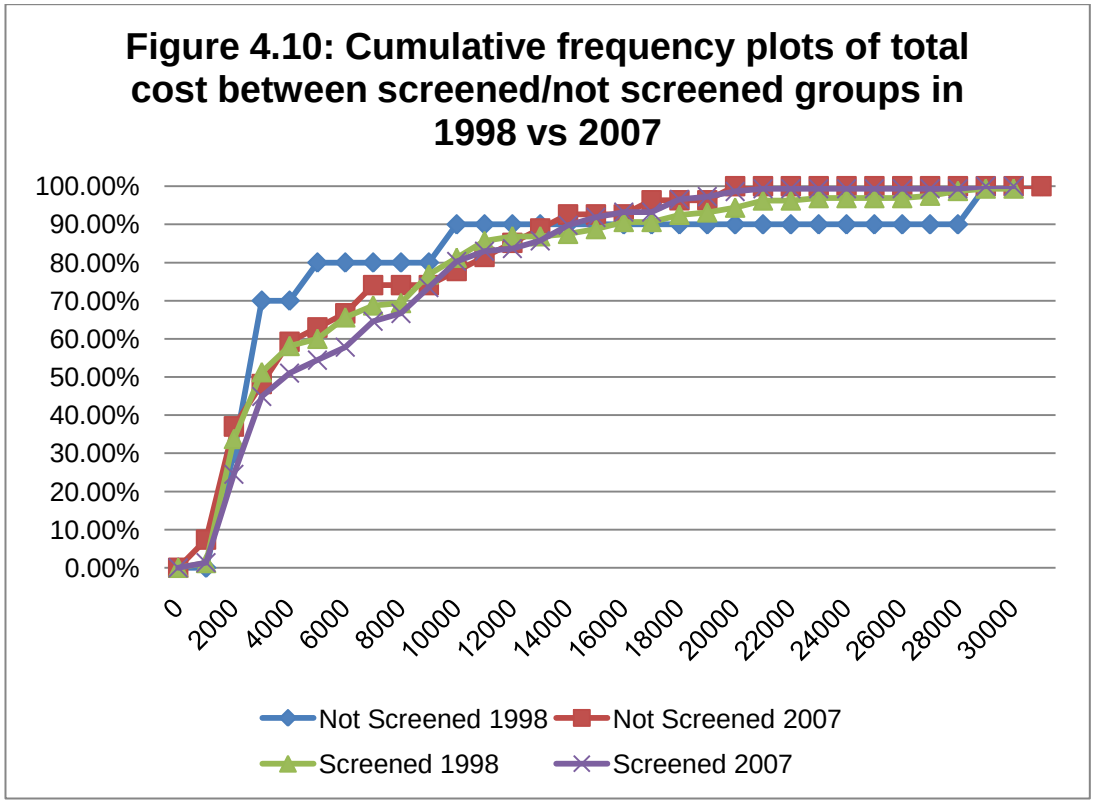


Figure 4.11 shows the box plots for total costs by diagnosis group and band. As the plot shows, there are outliers in the data. These are often patients that have been in hospital multiple times or have many different drugs, or both. These plots show that overall, cost is increasing by band, however it is not clear if there are differences between diagnosis groups.

Figure 4.11: Box plots of total costs by diagnosis type and band

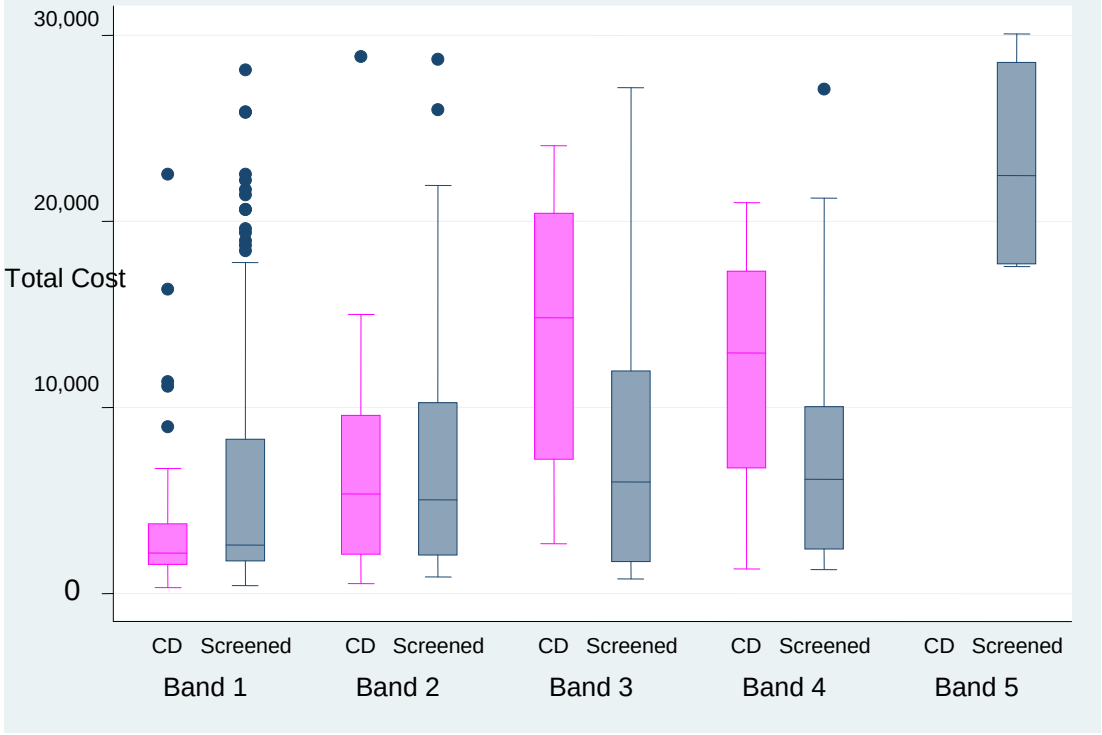
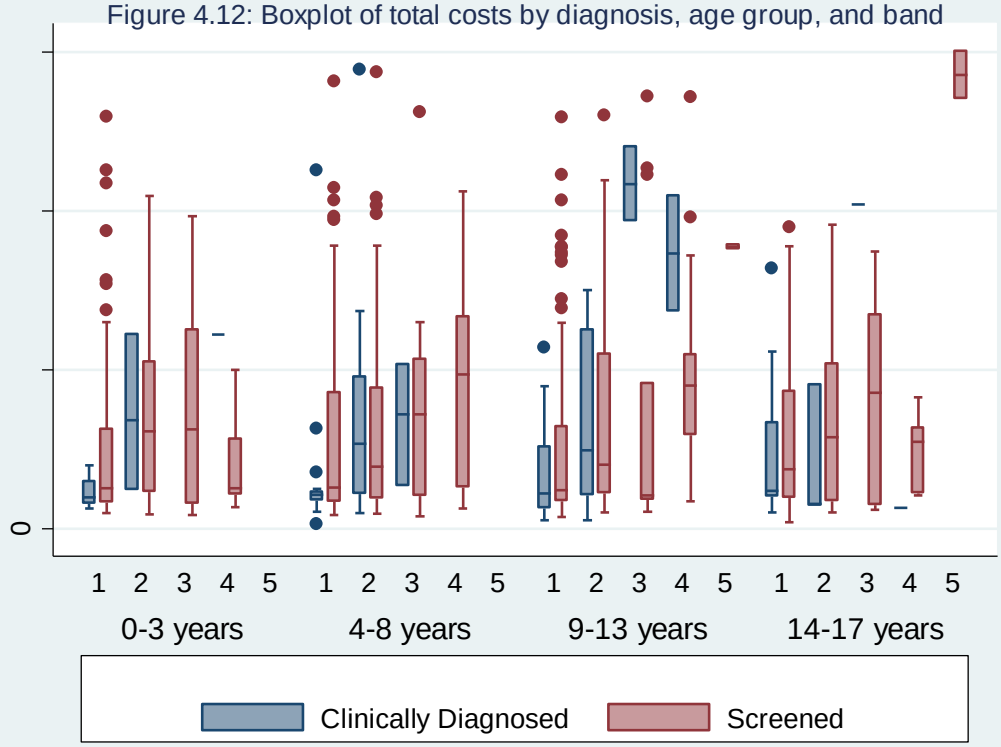


Figure 4.12 shows box plots for total costs by band, age group and diagnosis group. This plot also indicates that total costs are increasing by band and perhaps by age, and diagnosis, though it is unclear as the boxes overlap.

Figure 4.12: Boxplot of total costs by diagnosis, age group, and band

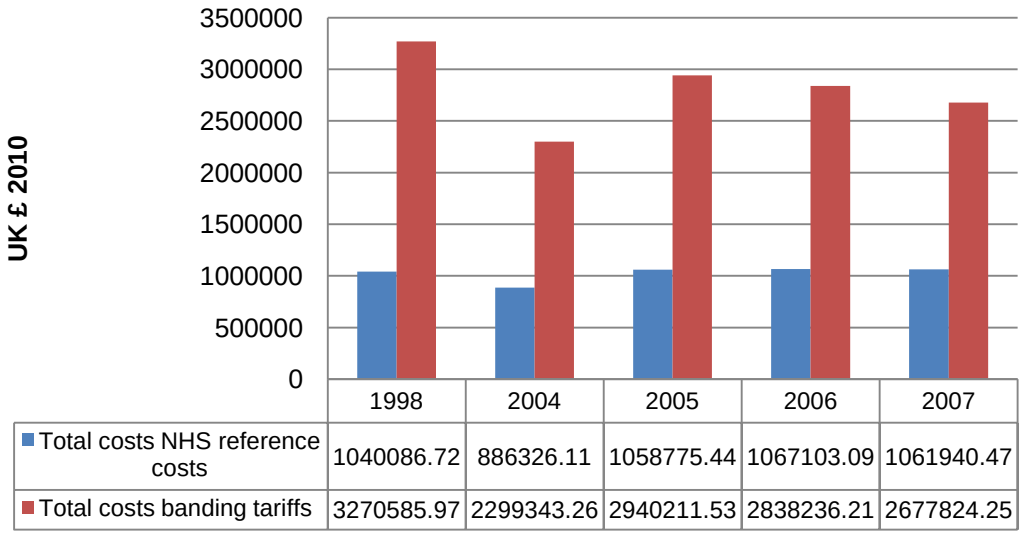


Using banding tariffs for unit costs

As discussed in the introduction, the CF Trust argued for the introduction of disease severity banding as a method of funding for CF patients. Provisional tariffs were introduced in 2007. In addition to inpatient care tariffs, outpatient clinic visits also received a tariff that was separate from general paediatric tariffs. This section gives a brief comparison of total, mean and median costs to the provider for the population based on methods using the NHS reference costs described above, which was the amount that could be claimed for CF care before the banding tariff was introduced, and the new banding tariffs which are the amount that the provider can claim for the care given within each band. Figure 4.13 shows the difference between total costs by year between the old NHS reference costs and the new proposed tariffs. There is a dramatic increase in total cost when comparing the two tariffs. Referring back to

Table 4.2, this is perhaps not surprising given the increase in clinic costs and the variability and increase of the inpatient hospital stay tariffs. Given the nature of the band-based tariffs (i.e. they're based on resource use as determined by disease severity) it may be that these are a more accurate measure of the costs actually associated with caring for CF patients. Assuming the tariffs are an accurate reflection of the cost of care, it is clear the paediatric tariffs from the NHS reference costs are a severe underestimate of the true cost and that this level of care was probably underfunded and/or was being subsidised by NHS Trusts.

Figure 4.13: Comparison of total costs using different costing sources

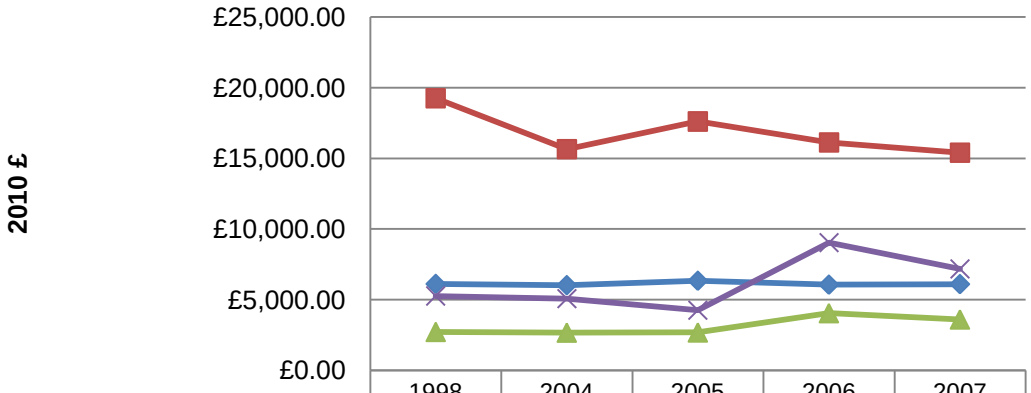


	1998	2004	2005	2006	2007
Total costs NHS reference costs	1040086.72	886326.11	1058775.44	1067103.09	1061940.47
Total costs banding tariffs	3270585.97	2299343.26	2940211.53	2838236.21	2677824.25

Figure 4.14 shows the trends between the mean and median costs using the NHS reference costs and the new banding tariffs. Given the tariffs are based on resource use associated with disease severity, it is not surprising to see that both the mean and the median cost of care using the new tariffs are higher than the reference costs. The trends generally have the same shape, however the tariff based means

have a slightly more pronounced shape. This may be due to the differences in hospitalisation costs associated with each band (i.e. one cost for hospitalisation versus five different costs). The previous analysis shows a slight increase in total costs over time, while using the new banding tariffs indicates a very slight reduction in total costs over time. Again, this is likely to be due to the change in the method of costing hospitalisations.

Figure 4.14: Comparison of mean/median total costs using different costing sources



	1998	2004	2005	2006	2007
Mean total cost using NHS reference costs	£6,118.16	£6,029.43	£6,339.97	£6,063.09	£6,103.11
Mean total cost using banding tariffs	£19,238.74	£15,641.79	£17,606.06	£16,126.34	£15,389.79
Median total cost using NHS reference costs	£2,729.85	£2,680.08	£2,692.00	£4,056.48	£3,610.83
Median total cost using banding tariffs	£5,265.73	£5,069.93	£4,267.73	£9,032.20	£7,176.21

Figure 4.15 shows the trends in terms of percentage of total costs for the cost drivers (drug costs, clinic costs, and hospital costs). It is clear that banding tariffs gives a much greater weight to hospitalisations (32.5 percent versus 72.39 percent). However, using the banding tariffs shows an increase in clinic costs (13.46 percent in 1998 rising to 17.16 percent in 2007), while the NHS reference costs shows a

relatively flat contribution of clinic costs to total costs (22.64 percent in 1998, 23.14 percent in 2007). Drug costs have an increasing contribution to total costs over time using both costing methods.

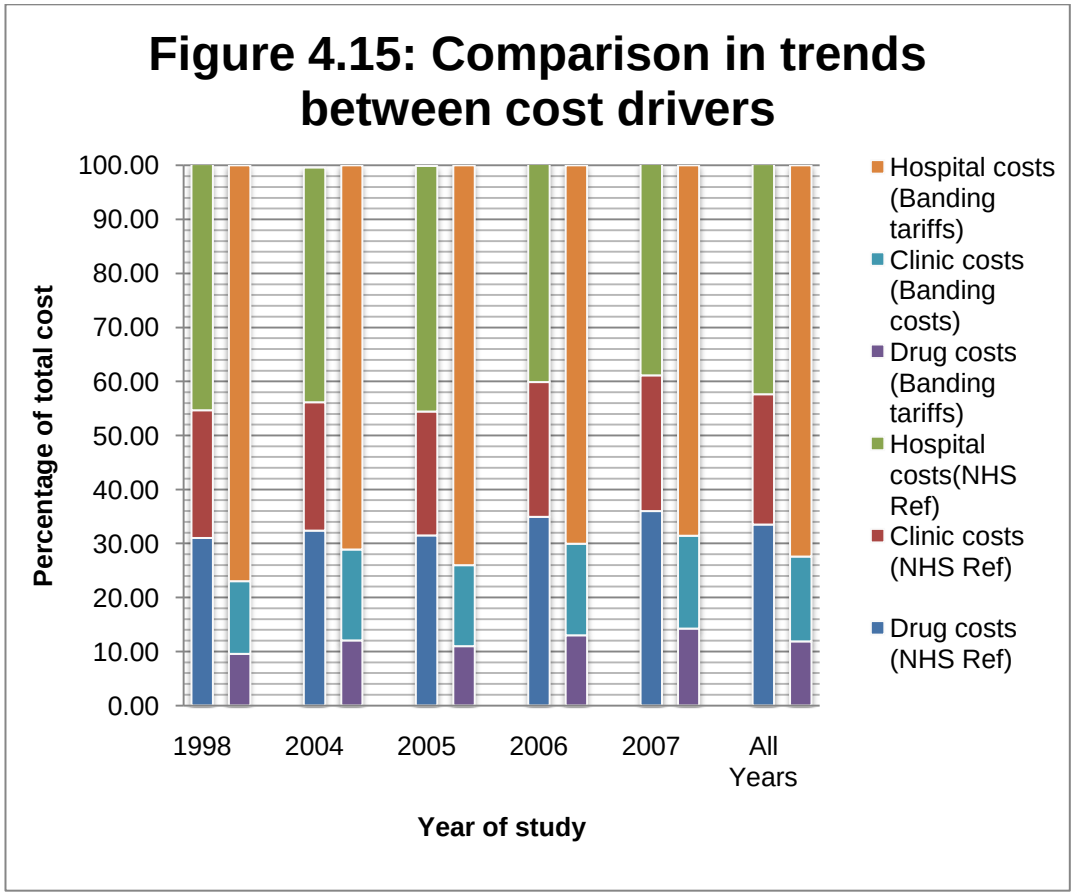
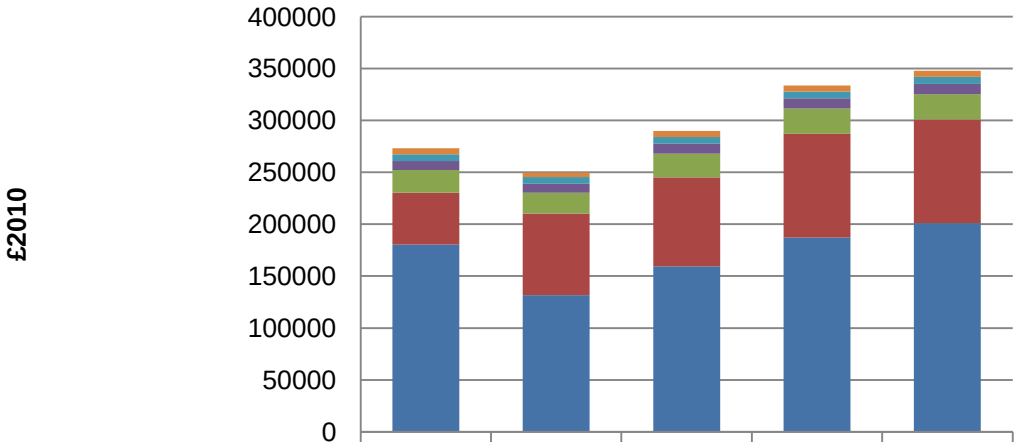


Figure 4.16 details the most expensive drugs that are driving the increase in drug costs in CF patients. Pulmozyme (DNAse) is by far the most expensive drug, and has been increasing its share of the total drug budget as more people are put on the drug. This is closely followed by inhaled Tobramycin which has nearly doubled its share of the costs over time. Creon, the pancreatic enzyme supplement is given to most CF patients and therefore its proportion of the drug budget remains relatively steady. The cost of IV antibiotics has increased slightly over the time period, while

Serevent and insulin has remained stable.

Figure 4.16: Changes the total costs of common drugs used in CF Care



	1998	2004	2005	2006	2007
Sum of Insulin	5821.2	4158	5821.2	5821.2	5821.2
Sum of Serevent	6417.84	6417.84	6417.84	6417.84	6795.36
Sum of Total IV Antibiotics	8552.91	8759.83	9760.91	10030.71	10081.61
Sum of Creon 10000	22176	20160	22848	24360	24360
Sum of Tobramycin	49862.4	78355.2	85478.4	99724.8	99724.8
Sum of DNAse	180310	131765	159505	187245	201115

Figure 4.17 shows the trends in median costs over time by age group. From a visual inspection, there does not seem to be a clear pattern for most of the age groups, with the exception of the youngest age group, which has steadily risen over time, which reflects current practice of treating younger children more aggressively to prevent the first or recurrent infections.

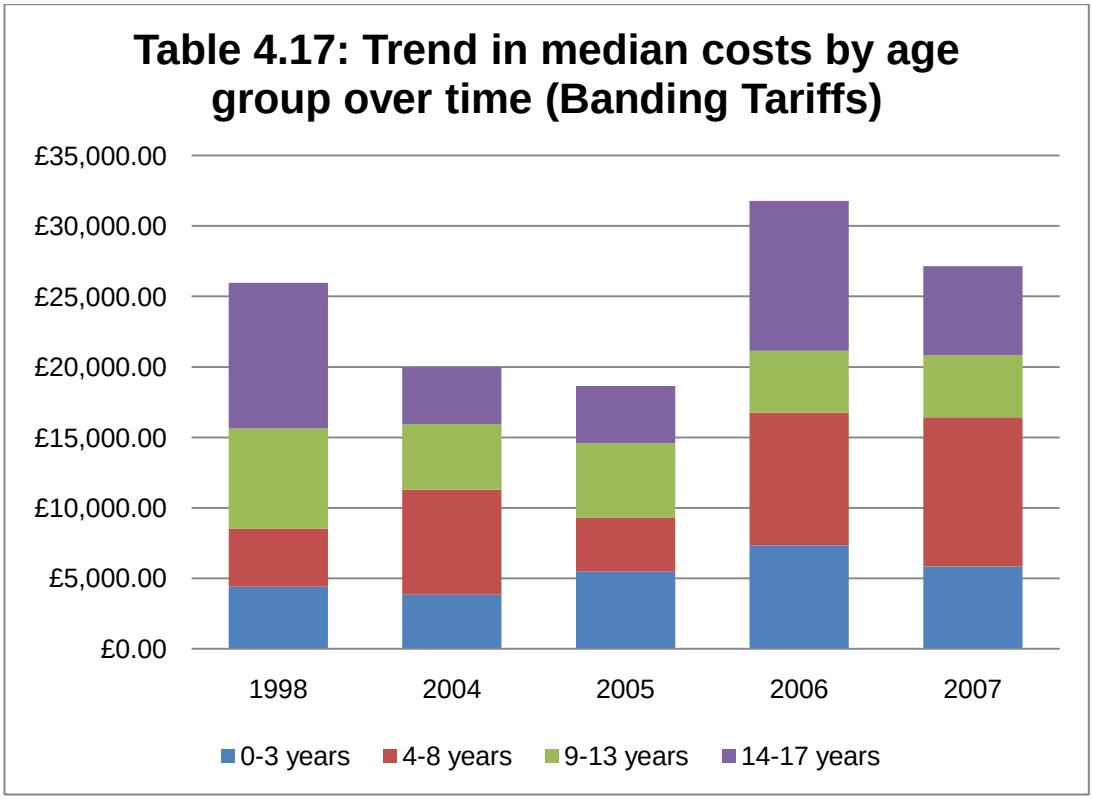


Figure 4.18 shows the cumulative percentage of total costs between screened patients and clinically diagnosed patients. This figure shows a slightly different relationship between the groups than in the previous NHS reference cost analysis. Although the differences are small, this diagram indicates that 90% of patients who are screened had costs below £45,000, whereas 90% of patients who are clinically diagnosed had costs below £55,000. Figure 4.19 shows the cumulative percentage of total costs between screened and clinically diagnosed at baseline (1998) and 2007. Again, these results are slightly different than the previous analysis using NHS reference costs. In 1998, those 90% of patients who were clinically diagnosed cost below £50,000 whereas 90% of screened patients cost below £60,000. However, in 2007, this trend is reversed, where 90% of those patients identified by newborn screening cost below £40,000, but 90% of clinically diagnosed patients

cost below £65,000. These results, however, must be interpreted carefully as there were very few clinically diagnosed patients throughout the study and therefore outliers may have an impact.

Figure 4.18: Cumulative percentage of total cost between screened/not screened (all observations)

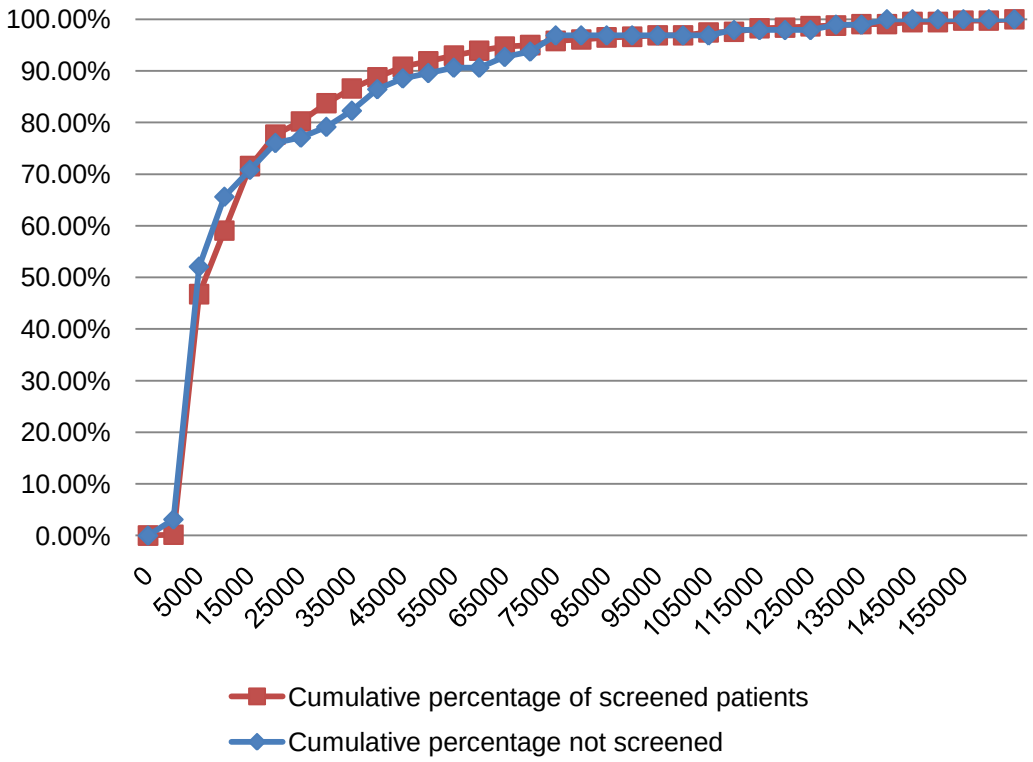
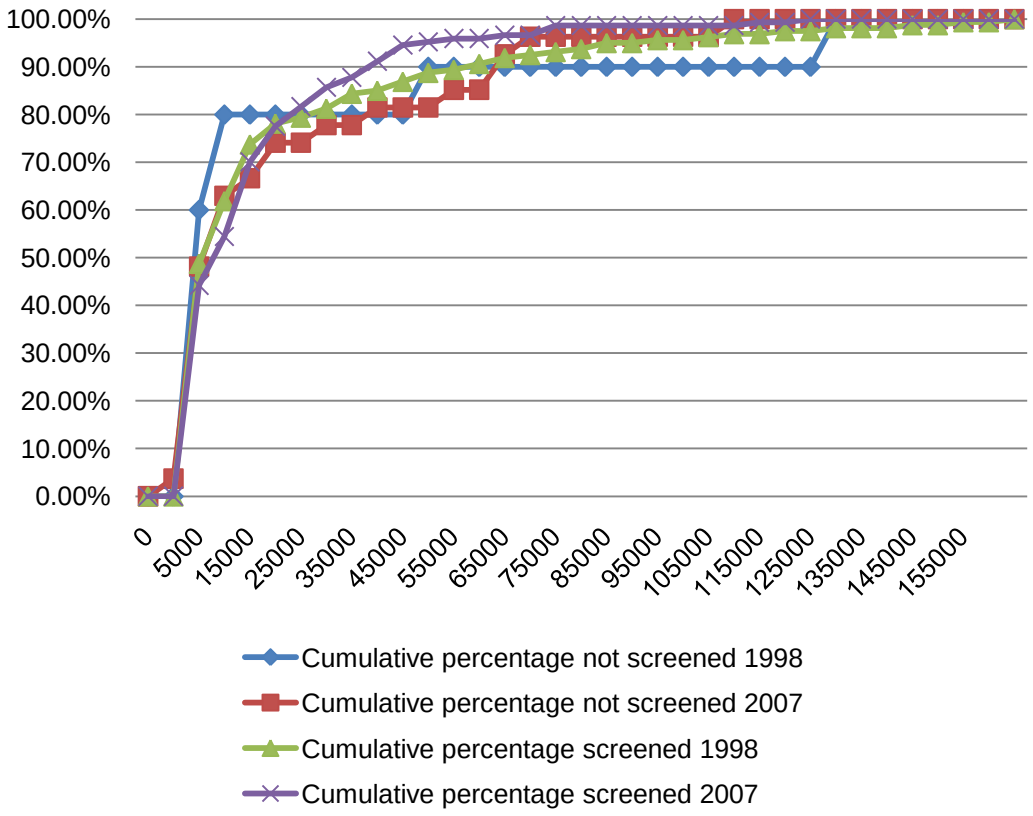


Figure 4.19: Cumulative percentage of total cost between screened/not screened at baseline and in 2007



The major differences in means and median costs between the two costing methods shown here is indicative that previous NHS payment by results tariffs may not have been adequate to cover the costs of caring for CF patients. As band-based tariffs are currently being used, the remainder of this analysis will use the banding tariffs to cost individuals.

Statistical analysis

As mentioned previously, it is important to analyse cost data in the knowledge that it is often skewed. To visually determine if the cost data for the CF cases in this study was skewed, a histogram was plotted (Figure 4.20) of total costs by year. It is

relatively clear that the distribution of cost is not normal, and that the data is skewed towards zero. To determine if taking the log of costs would be a more appropriate option for the initial regression, a histogram of the logged costs is given in figure 4.21. It is unclear from a visual inspection of the histogram of logged costs is closer to a normal distribution.

Figure 4.20: Histograms of total costs by year

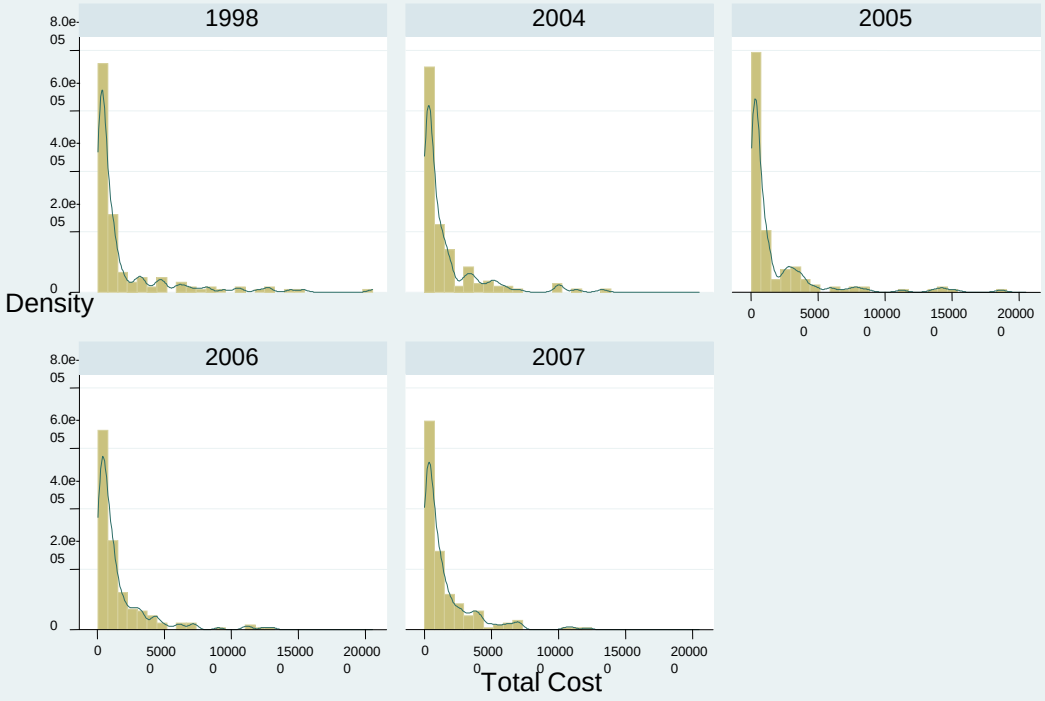
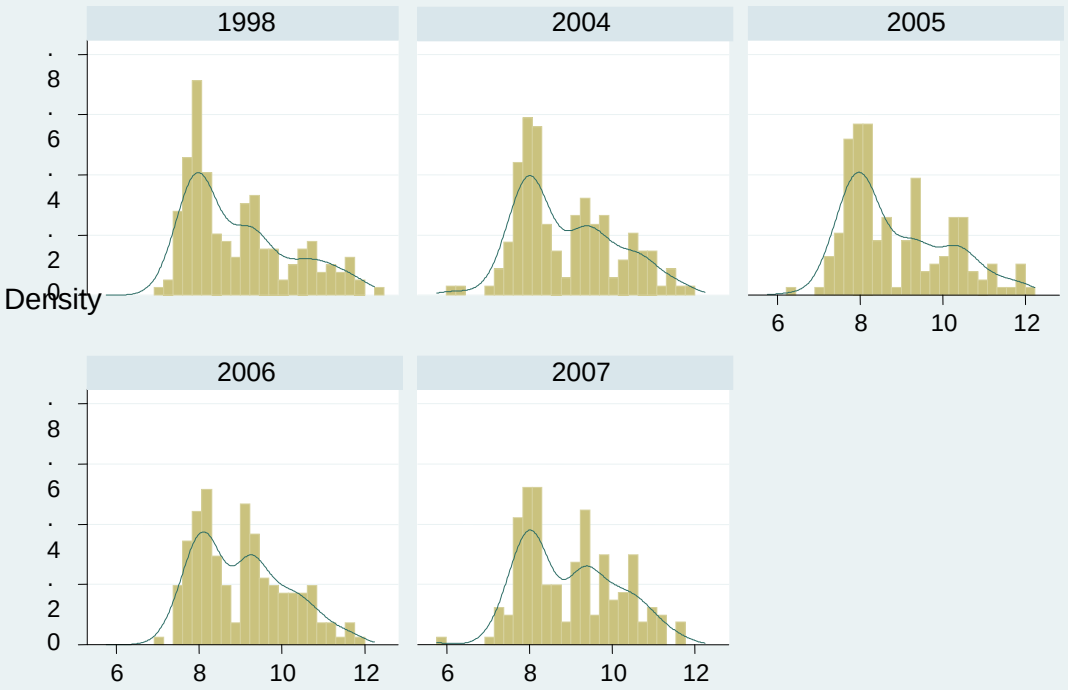


Figure 4.21: Histograms of logged total cost by year



As a visual inspection was relatively inconclusive, the data was subjected to a skewness test as well as a distributional fit test. Table 4.5 indicates the results of the variance, skewness, and kurtosis tests for each year for total cost and the log of total cost. The table also indicates the results of a distributional fit test (the lower the chi-square statistic, the better the distributional fit) carried out in Excel using the @Risk software. First, the variance in the data does not necessarily seem to be related to time, with a peak variance for both total cost and logged total cost in 2005. The data indicate that the variable of total cost is strongly skewed (2.59-3.22). The logged total cost was not as strongly skewed, with skewness close to zero, but still indicating a slight skew. Similarly, the total cost data seem to suffer more strongly from kurtosis. The initial regression analysis was run using both the skewed

distribution of cost as well as with the dependent variable being the natural log of cost. The subsequent GEE analyses were run assuming an inverse Gaussian distribution of cost as shown by the distributional fit scores for both total costs and logged total costs.

Table 4.5: Results of skewness test on total cost and natural log of total cost by year.

Year	Total Cost			Log Total Cost		
	Variance	Skewness	Kurtosis	Variance	Skewness	Kurtosis
1998	1.06	2.92	12.46	1.57	0.77	2.53
2004	5.37	2.73	11.21	1.44	0.46	2.43
2005	9.18	3.22	14.36	1.58	0.69	2.51
2006	5.1	2.86	12.2	1.2	0.49	2.37
2007	5.25	2.59	10.94	1.33	0.32	2.31

Chi-Square statistic for various distributions of total costs (for all years)

Fit	Inverse Gaussian	LogNormal	LogLogistic	Weibull	Exponen	Triang
Chi-Sq Statistic	280.90	357.56	366.17	646.34	801.35	4211.95

Chi-Square statistic for various distributions of natural logged total costs (for all years)

Fit	Inverse Gaussian	LogNormal	LogLogistic	Gamma	Weibull	Triang
Chi-Sq Statistic	302.39	304.40	329.71	330.82	479.741	1101.77

The initial regression analyses in tables 4.6 and 4.7 show that disease severity, age, and being screened may play a role in individual costs. Both models started with the independent variables of age, sex, screened, shared care, body mass index, weight for height z-scores, FEV1% predicted, and band. The initial model (Table 4.6) excluded shared care, FEV1% predicted, and sex and retained age, weight for height, bmi, screened, and band. The adjusted R-squared for the final model was 0.46. The second model (Table 4.7), regressed against logged total costs similarly

found that weight for height and band were important variables to include, but excluded the other variables. The adjusted R-squared for the logged total cost model was 0.43. From the two initial models, it appears that the predictors suggested in the methods section are the correct predictors for the more complex regression analysis.

Table 4.6: Initial regression analysis on baseline data (Total Costs)

Backwards regression beginning: total costs = age, sex, screened, shared care, bmi, weight for height, fev%predicted, band

p = 0.5935 >= 0.2000 removing sharedcare
p = 0.4960 >= 0.2000 removing fevpp
p = 0.3742 >= 0.2000 removing sex

Source	SS	df	MS	Number of obs = 101
Model	7.1958	5	1.43927922689	F(5, 95) = 18.17
Residual	7.5266	95	08	Prob > F = 0.0000
Total	1.4722	100	1.4722	R-squared = 0.4888
				Adj R-squared = 0.4619
				Root MSE = 28147

Total Costs	Coef.	Std. Err.	t	P>t	95% Conf	Interval
age	-1222.06	834.8437	-1.46	0.147	-2879.44	435.3131
weight for height	17.56619	3.553869	4.94	0.000	10.51086	24.62151
screened	-19520.7	14458.03	-1.35	0.180	-48223.5	9182.134
band	16931.98	2850.368	5.94	0.000	11273.28	22590.67
bmi	1465.243	1062.959	1.38	0.171	-644.997	3575.484
constant	-3874.39	21185.55	-0.18	0.855	-45933	38184.24

Table 4.7: Initial regression analysis on baseline data (natural logged total costs)

Backwards regression beginning: lntotalcosts = age, sex, screened, shared care, bmi, weight for height, fev%predicted, band

p = 0.6561>=0.20 removing sex
p = 0.5396>=0.20 removing screened
p = 0.3872>=0.20 removing sharedcare
p = 0.3292>=0.20 removing age
p = 0.4867>=0.20 removing bmi
p = 0.3080>=0.20 removing fevpp

Source	SS	df	MS	Number of obs = 101
Model	80.72	2	40.36	F(2, 98) = 38.95
Residual	101.56	98	1.04	Prob>F = 0.0000
Total	182.27	100	1.82	R-squared = 0.4428
				Adj R-squared = 0.4315
				Root MSE = 1.018

Ln Total Costs	Coef.	Std. Err.	t	P>t	95% Conf	Interval
weight for height	0.000186	0.000126	1.48	0.142	-0.0000637	0.000436
band	0.814989	0.100281	8.13	0.000	0.6159838	1.013993
constant	7.610276	0.198995	38.24	0.000	7.215377	8.005175

GEE results

Table 4.9 gives the results of the model fit test for GEE models. The table gives the number of parameters in the model (p), the trace, the QIC score and the QIC_u score. Several attempted models did not reach convergence, meaning any estimates from those models could be biased. The model with the lowest QIC that used total cost as a dependent variable was the inverse Gaussian family with a log link and an

independent correlation structure. The model with the lowest QIC that used logged total cost was the inverse Gaussian with the identity link and an independent correlation structure. However, all the other models that reached convergence had relatively similar QIC scores, therefore all the models will be analysed.

Table 4.10 shows the results of the 7 models identified as good fits by the QIC tests. Models 1, 2, and 3 use total costs as the dependent variable and Models 4, 5, 6, and 7 use logged total cost as the dependent variable. The independent variables were dummy variables for: age categories (0-3, 4-8, 9-12, 13-17), sex, screened or clinically diagnosed (1 and 0 respectively), shared care, and as well as variables for band (split into dummy variables), bmi, weight for height (Z-scores) and FEV1% predicted.. The model grouped the patients by patient ID (pid). Though not reported, STATA automatically reports the Wald test which indicates if we can reject the null hypothesis that all the coefficients are equal to zero (as indicated by the Prob>chi2 being significant at the 1% level). All models indicated that we can reject the hypothesis that the coefficients are equal to zero. All models also omitted the fourth dummy variable for age (13-17 year olds) and the fifth dummy variable for band (band 5) due to collinearity.

The first model assumed an inverse Gaussian family with a log link function and an independent correlation structure. The model indicated that being male is associated with a slightly higher cost. The model also indicates that at the 10% level of significance, being screened is indicative of lower costs over time. A higher FEV1% predicted is also indicative of having lower costs over time, although the effect is small. Interestingly, the age categories and disease severity bands were significant at the 1% level. The negative coefficients indicate that being in any age group or any

of the bands (except band 3, which was not significant) would reduce costs, which is counter-intuitive. As all the models show bands and age categories having a negative coefficient, this is discussed more below after describing the other models.

Model 2, which assumed an exchangeable correlation structure had similar findings to model 1, but weight for height z-scores seems to be significant at the 10% level, but the coefficient is close to zero, indicating a very slight relationship. Model 3 assumed an unstructured correlation relationship, but unlike the first two models, did not find sex and screened to be significant.

Models 4 and 5 use logged total costs with an inverse Gaussian family and an identity link, with model 4 assuming an independent correlation structure and model 5 assuming an exchangeable correlation structure. Models 6 and 7 assume an inverse Gaussian family and a log link, with model 6 assuming an independent correlation structure and model 5 assuming an exchangeable structure. All of these models also indicate that the age and band categories have a negative coefficient, and similarly to the total cost models, indicate that having a higher FEV1% predicted indicates a lower cost.

Although not significant in models 3-7, the models all indicate that being screened indicates a lower cost, though this does not necessarily agree with the indications in figures 4.7 and 4.8. Again, although not significant in most models, the signs of the coefficients indicate that a higher BMI and being male indicates a higher cost, which could reflect the difficulty in maintaining a healthy weight through pancreatic and other dietary supplements, driving up costs for those patients. Being in shared care was not significant in any of the models, but was consistently associated with a

lower cost. It is unclear from this analysis if this is due those patients receiving less clinical care or if patients in shared care are simply healthier and therefore they require less treatment.

The negative coefficients for the age categories and disease severity bands are surprising, as you would expect age and disease severity to have a positive relationship with cost. This may be due to the signs indicating the marginal effect, not the total effect (i.e. band one has a negative relationship to total costs (is cheaper) when compared with the other bands) or perhaps age/band and total cost have a linear relationship. If the variables and total cost do have a linear relationship, it is possible for the coefficients to have the wrong sign if they are split into dummy variables. (160) To determine the linear relationship between total costs and the categorical variables of band and agecat, pairwise correlations were run in STATA. Table 4.8 shows the results of these correlation tests, which indicate that band has a positive linear relationship with total costs and logged total costs (the analysis also ran correlations by year, and found a significant positive linear relationship for each year), but that age category does not have a significant linear relationship with either variable. This suggests that age category should be split into 4 dummy variables, and although this test suggests that band can be used as a variable, as it is not perfectly linear, band will also be split into 5 dummy variables, but the models will be run without the oldest age group and band 1 to test the marginal effect theory.

Table 4.8: Pairwise correlation of total costs and logged total costs (by band and age category *at 95% significance)					
Obs (834)	Total Costs	band		Total cost	Agecat
Total costs	1		Total costs	1	
Band	0.2999*	1	agecat	0.0046	1
	Logged total costs	band		Logged total costs	Agecat
Logged total cost	1		Logged total costs	1	
band	0.2808*	1	agecat	0.0186	1

Table 4.9: Model Fit Statistics					
			Total Costs		
	Inverse Gaussian with Identity link				
	Correlation	p	trace	QIC	QIC_u
	Independent	Convergence not achieved			
	Exchangeable	Convergence not achieved			
	Unstructured	Estimates Diverged			
	Inverse Gaussian with log link				
	Correlation	p	trace	QIC	QIC_u
M1 & M8	Independent	14	11.46	22.89	27.97
M2 & M9	Exchangeable	14	11.76	23.5	27.97
M3 & M10	Unstructured	14	11.93	23.82	27.97
			Logged Total Costs		
	Inverse Gaussian with Identity link				
	Correlation	p	trace	QIC	QIC_u
M4 & M11	Independent	14	14.88	-22.46	-24.22
M5 & M12	Exchangeable	14	14.97	-22.27	-24.22
	Unstructured	Convergence not achieved			
	Inverse Gaussian with log link				
	Correlation	p	trace	QIC	QIC_u
M6 & M13	Independent	14	14.99	-22.21	-24.22
M7 & M14	Exchangeable	14	15.15	-21.92	-24.22
	Unstructured	Convergence not achieved			

Table 4.10: Regression results (1st age category and last band dropped due to collinearity)

	Total Cost			Logged Total Cost			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Independent Variables	b/se	b/se	b/se	b/se	b/se	b/se	b/se
4-8 Years	-0.996**	-0.977**	-0.922**	-1.611**	-1.644**	0.162**	0.166**
	-0.19	-0.19	-0.18	-0.15	-0.15	-0.02	-0.02
9-13 Years	-0.905**	-0.881**	-0.787**	-1.532**	-1.567**	0.153**	0.157**
	-0.19	-0.19	-0.18	-0.15	-0.15	-0.02	-0.02
14-17 Years	-0.998**	-0.976**	-0.874**	-1.417**	-1.457**	0.140**	0.145**
	-0.23	-0.23	-0.21	-0.18	-0.18	-0.02	-0.02
Sex (1=Male)	0.262+	0.256+	0.215	0.053	0.04	0.006	0.005
	-0.14	-0.14	-0.14	-0.12	-0.12	-0.01	-0.01
Screened (1=Yes)	-0.369+	-0.373+	-0.28	-0.101	-0.096	-0.012	-0.011
	-0.21	-0.21	-0.2	-0.21	-0.21	-0.02	-0.02
Shared Care (1=Yes)	0.075	0.072	0.015	-0.112	-0.099	-0.013	-0.012
	-0.16	-0.16	-0.16	-0.15	-0.15	-0.02	-0.02
Body Mass Index	0.018	0.018	0.019	0.01	0.008	0.001	0.001
	-0.03	-0.03	-0.03	-0.03	-0.03	0	0
Weight for Height	0	0.000+	0.000+	0	0	0	0
	0	0	0	0	0	0	0
FEV1% Predicted	-0.006*	-0.007*	-0.007*	-0.005*	-0.005*	-0.001*	-0.001*
	0	0	0	0	0	0	0
Band 1	-1.788**	-1.805**	-1.822**	-2.167**	-2.124**	0.223**	0.218**
	-0.26	-0.25	-0.24	-0.3	-0.32	-0.03	-0.03
Band 2	-1.248**	-1.226**	-1.131**	-1.631**	-1.634**	0.163**	0.164**
	-0.27	-0.26	-0.25	-0.31	-0.33	-0.03	-0.03
Band 3	-0.426	-0.41	-0.321	-1.398**	-1.400**	0.138**	0.139**
	-0.41	-0.41	-0.37	-0.43	-0.45	-0.04	-0.05
Band 4	-0.827*	-0.788*	-0.851**	-1.087**	-1.102**	0.106**	0.108**
	-0.33	-0.32	-0.31	-0.39	-0.41	-0.04	-0.04
Constant	12.358**	12.395**	12.272**	12.843**	12.834**	2.590**	2.589**
	-0.76	-0.77	-0.76	-0.67	-0.67	-0.07	-0.07

+ p<0.10, * p<0.05, **p<0.01

Table 4.11: GEE Regression results (forced removal of oldest age group, and 1st band)

Independent Variables	Total Cost			Logged Cost			
	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
0-3 Years	0.998**	0.976**	0.874**	1.417**	1.457**	0.140**	0.145**
	-0.23	-0.23	-0.21	-0.18	-0.18	-0.02	-0.02
4-8 Years	0.002	-0.001	-0.048	-0.193	-0.187	-0.022	-0.021
	-0.23	-0.23	-0.21	-0.17	-0.17	-0.02	-0.02
9-13 Years	0.093	0.095	0.087	-0.115	-0.11	-0.013	-0.012
	-0.15	-0.15	-0.14	-0.13	-0.13	-0.01	-0.01
Sex (1=Male)	0.262+	0.256+	0.215	0.053	0.04	0.006	0.005
	-0.14	-0.14	-0.14	-0.12	-0.12	-0.01	-0.01
Screened (1=Yes)	-0.369+	-0.373+	-0.28	-0.101	-0.096	-0.012	-0.011
	-0.21	-0.21	-0.2	-0.21	-0.21	-0.02	-0.02
Shared Care (1=Yes)	0.075	0.072	0.015	-0.112	-0.099	-0.013	-0.012
	-0.16	-0.16	-0.16	-0.15	-0.15	-0.02	-0.02
Body Mass Index	0.018	0.018	0.019	0.01	0.008	0.001	0.001
	-0.03	-0.03	-0.03	-0.03	-0.03	0	0
Weight for Height	0	0.000+	0.000+	0	0	0	0
	0	0	0	0	0	0	0
FEV1% Predicted	-0.006*	-0.007*	-0.007*	-0.005*	-0.005*	-0.001*	-0.001*
	0	0	0	0	0	0	0
Band 2	0.541**	0.579**	0.691**	0.535**	0.490**	0.060**	0.055**
	-0.14	-0.13	-0.13	-0.12	-0.12	-0.01	-0.01
Band 3	1.362**	1.395**	1.500**	0.769*	0.723*	0.084**	0.079*
	-0.34	-0.33	-0.3	-0.31	-0.31	-0.03	-0.03
Band 4	0.961**	1.017**	0.971**	1.079**	1.022**	0.117**	0.111**
	-0.25	-0.23	-0.23	-0.27	-0.27	-0.03	-0.03
Band 5	1.788**	1.805**	1.822**	2.167**	2.124**	0.223**	0.218**
	-0.26	-0.25	-0.24	-0.3	-0.32	-0.03	-0.03
Constant	9.571**	9.615**	9.577**	9.259**	9.254**	2.226**	2.226**
	-0.77	-0.78	-0.77	-0.64	-0.63	-0.07	-0.07

+ p<0.10, *p<0.05, **p<0.01

Table 4.11 shows the results of the seven models that removed the last age band, and the first disease severity band. Similarly to table 4.10, the first 3 models (8-10) use total cost as a dependent variable, and the last four (11-14) use logged total cost. Models 8 and 9 indicate that being male is indicative of a higher cost, while being screened is associated with lower costs. Model 9 also indicates that weight for height is indicative of a higher cost, though the coefficient is close to zero. Model 10 also indicates that weight for height has a positive relationship with costs. All of the models (8-14) indicate that having a higher FEV1% predicted is associated with a lower cost.

Models 8-14 indicate that only the youngest age group has a significant impact on total or logged costs, and unlike the first seven models, all the age groups now have positive signs. This indicates that there is a marginal effect being shown in the model with regards to age. The results show that being in the youngest age group has a larger impact on costs than the other age groups over time. This isn't entirely consistent with the trend graphs earlier, although the youngest age group was the only group to have a continuous increase over time.

Removing the least severe band category has also changed the signs of the coefficients in all the models. This indicates that the other bands are more expensive than band one, and the negative coefficients in models 1-7 could be the result of excluding the most expensive band, therefore all the other bands were "cheaper" in comparison. However, as the pairwise correlation test found a linear relationship, the seven models were run again without splitting band into 5 dummy variables. As they were "new" regressions, the QIC test was run for each model and the results are given in Table 4.12. Similarly to the first models, the inverse

Gaussian family and a log link with an independent correlation structure (Model 15) gave the best fit for total cost as a dependent variable and in fact, had a better score than when band was split, indicating this may be a better model. For the logged total cost, the inverse Gaussian family with an identity link and an exchangeable correlation structure (Model 19) had the smallest QIC. Models 15-17 indicate that again, being male, being in the youngest age group and band is associated with increasing costs over time. Models 15 and 16 indicate that being screened is associated with a lower cost over time. All the models indicated that having a higher FEV1% predicted is associated with a small negative change in cost over time.

Figures 4.22 and 4.23 give a visual indication of the predicted versus actual total costs to see if the model outputs are comparable. Figure 4.22 plots the total cost models values of total cost versus the actual values and while they have a similar shape, the fit lines show a divergence. Figure 4.23 shows the predicted versus actual logged total cost models. While the scatter plots for the most part overlap, the linear relationships do seem to diverge slightly.

The regression analyses have shown that band (whether split into dummies or as a continuous variable) is probably the strongest driver of costs, followed by age, although the relationship with age is not exactly clear. Being screened and having a high FEV1% Predicted may result in lower costs over time, but being male may result in having higher costs.

Table 4.12: GEE Models with band as a non-categorical variable							
	Total Costs			Logged Total costs			
	Model 15 b/se	Model 16 b/se	Model 17 b/se	Model 18 b/se	Model 19 b/se	Model 20 b/se	Model 21 b/se
QIC	**20**	**21.01**	**21.26**	**-28.47**	**-28.61**	**-28.19**	**-28.25**
0-3 Years	1.025**	0.998**	0.931**	1.513**	1.537**	0.152**	0.155**
	-0.21	-0.2	-0.19	-0.15	-0.15	-0.02	-0.02
4-8 Years	0.021	0.014	-0.027	-0.173	-0.172	-0.019	-0.019
	-0.23	-0.23	-0.21	-0.17	-0.17	-0.02	-0.02
9-13 Years	0.099	0.101	0.096	-0.105	-0.102	-0.011	-0.011
	-0.15	-0.15	-0.14	-0.13	-0.13	-0.01	-0.01
Sex (1=Male)	0.262+	0.260+	0.237+	0.063	0.046	0.007	0.005
	-0.14	-0.14	-0.14	-0.12	-0.12	-0.01	-0.01
Screened (1=Screened)	-0.393+	-0.396+	-0.299	-0.105	-0.096	-0.012	-0.011
	-0.2	-0.21	-0.2	-0.21	-0.21	-0.02	-0.02
Shared Care (1=Shared Care)	0.092	0.086	0.035	-0.101	-0.088	-0.012	-0.011
	-0.15	-0.15	-0.15	-0.15	-0.15	-0.02	-0.02
Body Mass Index	0.021	0.021	0.022	0.012	0.009	0.001	0.001
	-0.03	-0.03	-0.04	-0.03	-0.03	0	0
Weight for Height	0	0	0	0	0	0	0
	0	0	0	0	0	0	0
FEV1% Predicted	-0.006+	-0.006*	-0.006*	-0.005*	-0.005*	-0.001*	-0.001+
	0	0	0	0	0	0	0
band	0.520**	0.562**	0.625**	0.424**	0.396**	0.046**	0.043**
	-0.09	-0.09	-0.09	-0.07	-0.07	-0.01	-0.01
Constant	8.979**	9.004**	8.871**	8.809**	8.826**	2.179**	2.180**
	-0.79	-0.79	-0.78	-0.64	-0.63	-0.07	-0.07
+ p<0.10, *p<0.05, **p<0.01							

Figure 4.22: Predicted versus Actual values for models using total cost

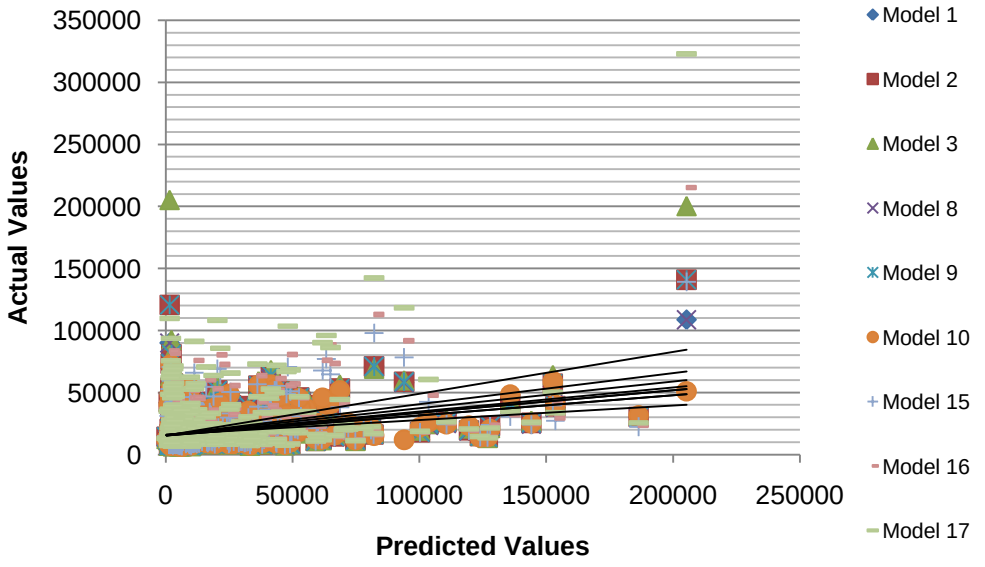
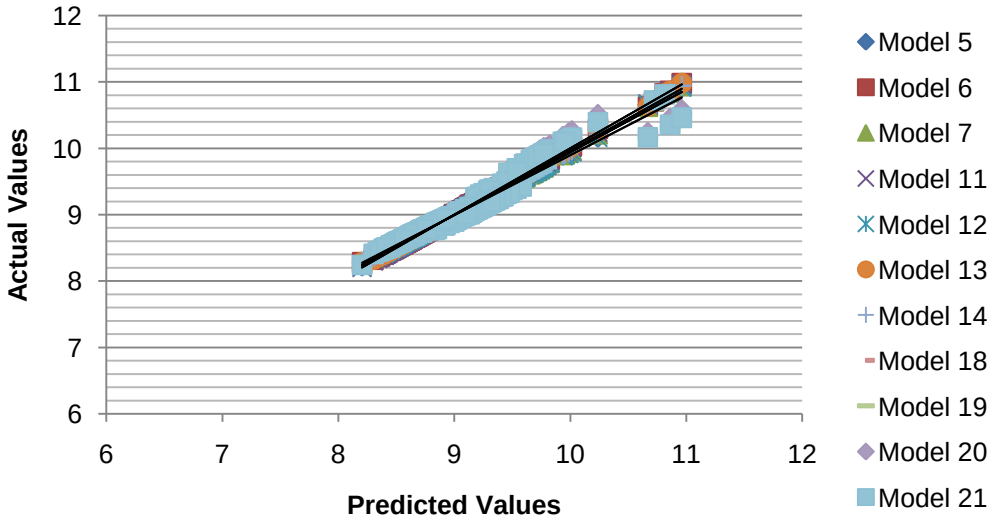


Figure 4.23: Fitted versus actual values for logged total cost models



Discussion

Given the variability of the disease course and treatment in CF, it is not surprising that costs also show variability between individuals. Aggregated cost results may be of limited use to a decision maker, but analysing the changes in mean and median costs could help plan for new patients, and help to 'insure' budgets for patients that have extreme costs. As the histograms in figure 4.20 and 4.21 indicate, a small number of patients generate high costs, which skews the distribution of cost data, possibly inflating the mean value above the median value. The median value does give an indication of what the most typical cost is for an individual patient. The number of high cost individuals may vary from year to year and this has an implication for long-term planning within departments with a tight budget.

Hospitalisation (inpatient) costs are still the greatest contributor to the hospitals cost of care, but have fallen from 77% (46% using NHS Ref costs) of costs in 1998 to 69% (39%) of costs in 2007. Drug costs have risen from 10% of the burden in 1998 to 14% in 2007. Clinic costs (outpatient) have also increased their share of the overall cost from 13% in 1998 to 17% in 2007. This is indicative of the increasing number of clinics and more expensive and/or aggressive drug treatments and the (perhaps resulting) fall in the number of days patients are spending in hospital.

Compared to the only other UK study to look at the cost of care for paediatric patients (3), median costs per patient were marginally higher. Possible reasons for

this could be that (i) the population in this study was different; (ii) the setting and method of care; (iii) data quality; (iv) availability of routine cost data for health care. The authors looked at a younger (and larger) cohort of patients than this study, and it is apparent that the older a patient is, the more likely chronic infection is going to occur, therefore increasing drug and hospital costs. While Sims et al had access to a robust national database, the current study used patient notes to record episodes of resource use and therefore may have found more cost-generating activity. However, the results of this thesis agree that patients who were screened had a lower average cost over time than those patients who were clinically diagnosed though not as pronounced (£17216/ £16766 = CD/NBS = 1.03 from this study versus $12707/7649 = 1.66 in Sims et al)

It is important to note that this study only took into account the costs to the health service at a tertiary level. It is likely that CF patients also generate a large number of primary care contacts, and therefore this study is likely to underestimate the total cost to the health service as a whole. It is also important to consider the fact that CF patients often generate costs outside of the health service. Many patients and their families have support from social services and educational services that can generate substantial costs for the public sector. This study also does not assess the out of pocket costs and productivity losses to the family, which may be substantial.

Overall, this study found no significant differences over time in terms of total cost; however the make-up of the cost has changed with changing clinical practice. As the study relied on patient notes and letters that recorded the prescribed drugs, it is possible that some resource use was missed which may influence the overall results. It is clear from the data that there seems to be a dip in the number of

patients in 2004, though the reason for this is less than clear. However, given that there are no significant differences in the population in terms of clinical outcomes or cost, this is unlikely to bias the overall results.

As mentioned in the methods section, there is no gold standard in terms of the best statistical techniques used in the analysis of cost data. However, GEE models are useful tools to analyse longitudinal data where members of the panel do not have complete data sets because they were not present when the initial or final data was collected. The models chosen here all indicate that they were a good fit and therefore the results are likely to be reliable.

Chapter 5: Cost-effectiveness analysis of newborn screening for cystic fibrosis

Introduction

One of the main aims of this study was to investigate the cost-effectiveness of newborn screening for cystic fibrosis. As the WHO (188) sets out, a screening programme should only be implemented if:

- The condition sought should be an important health problem for the individual and community.
- There should be an accepted treatment or useful intervention for patients with the disease.
- The natural history of the disease should be adequately understood.
- There should be a latent or early symptomatic stage.
- There should be a suitable and acceptable screening test or examination.
- Facilities for diagnosis and treatment should be available.
- There should be an agreed policy on whom to treat as patients.
- Treatment started an early age should be of more benefit than treatment started later.
- The cost should be balanced in relation to the possible expenditure on medical care as a whole.
- Case finding should be a continuing process and not a once and for all project.

Cystic fibrosis clearly meets the clinical aspects of these criteria. In the UK, newborn screening for cystic fibrosis was rolled out nationally in 2007. Although previous views (70) have suggested that there is a positive impact, it is still unclear if there are net economic benefits to screening in the UK, given the current costs of CF care. One method commonly employed by health economists to analyse costs and benefits of a screening programme is a cost-effectiveness analysis.

Methods

In this study, I have carried out a cost-effectiveness analysis which uses various decision modelling techniques. The perspective of the study is the NHS and the comparator is no screening.

Model design

It is the goal of this study to compare newborn screening in the UK for cystic fibrosis to not screening. The method of cost-effectiveness analysis chosen uses a decision analytic model to simulate the outcomes of the decision, based on data from the ERCFD (chapter 4) and the literature. Applying the taxonomy of the application of decision models illustrated by Brennan et al (179), the variable nature of disease progression and severity for CF, a simple decision tree is not adequate to look at both short and long-term implications of a screening programme. Similarly, a simple Markov model will not adequately take into account patient history over a long period of time and therefore is inappropriate for this study. Therefore I have adopted a discrete time Markov Chain model. This type of model is a cohort model, but assumes that fractions of the cohort can experience specific health states and that changes are deterministic. The number of individuals in each state is tracked and

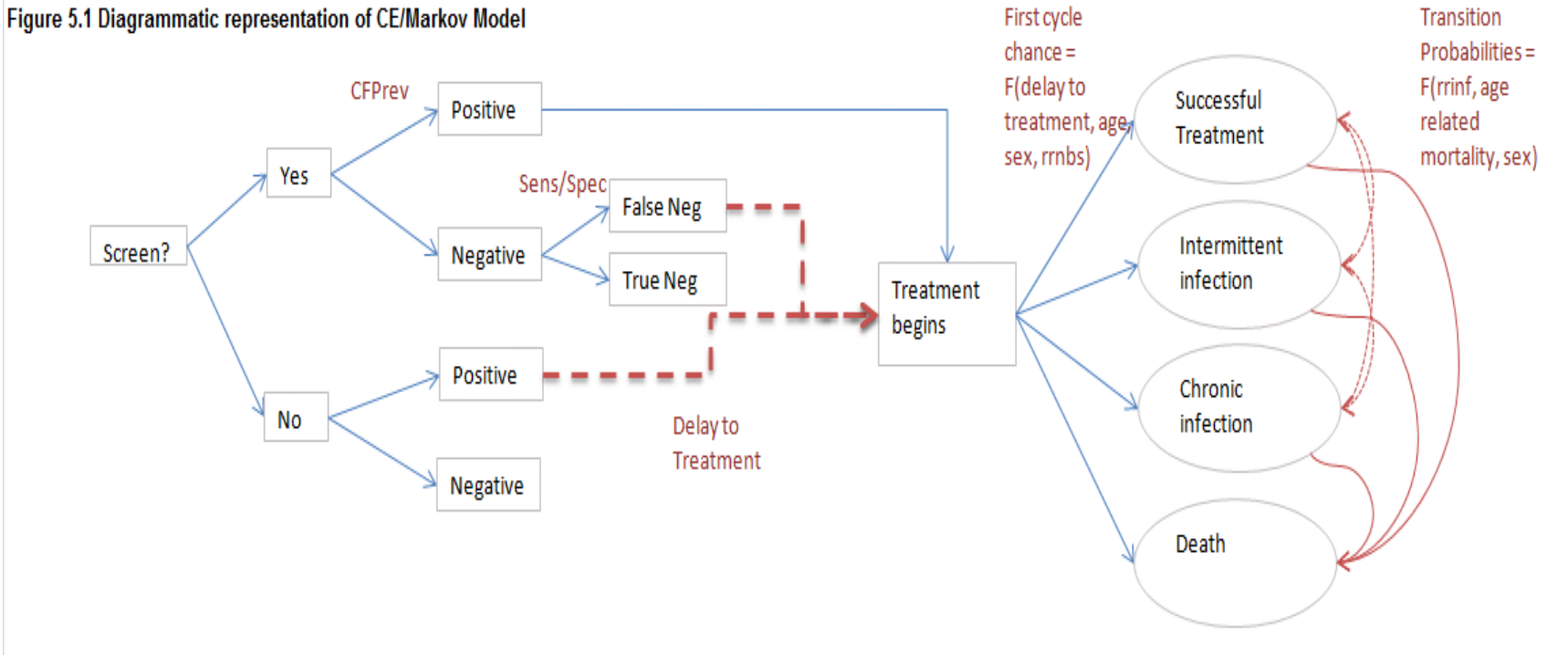
the event is processed and remembered over time. The model assumes a time period of one year, and a total of 60 years was adopted as a lifetime horizon based on epidemiological data in Dodge et al. (5)

A diagram of the model is illustrated in figure 5.1. Boxes represent a decision or clinical result. Solid lines indicate the path of the cohort, thick dashed red lines indicate delays to treatment. Ovals represent health states. The thin red dotted lines represent the probability of transition from one state to another. There were four possible health states: successful intervention (equivalent to Band 1), intermittent infections (equivalent to Band 2 and 3), chronic infection (equivalent to Band 4 and 5), and death. The bands (discussed in chapter 1) were "matched" to the chosen disease states in order to assign costs to each patient in that state. If a patient received a positive result in the screening group, it is assumed that treatment started immediately. Patients in the non-screening group also started treatment immediately upon receiving a clinical diagnosis. Patients who were clinically diagnosed at birth, either through family history, or presenting with meconium ileus were excluded from the analysis as they would receive the same treatment regardless of which arm they were in. It was assumed that any false negatives in the screening arm would be picked up clinically (and therefore has the same delay to treatment as those in the non-screened clinically diagnosed arm). The model was programmed using Microsoft Excel. (189) Appendix E contains the worksheets of the model (double click on the embedded excel worksheet to see all worksheets) as well as the Excel macros (the model is also available electronically, as the embedded model will not run macros). Table 5.6 gives a summary of the parameters

input into the model. Other tables (5.1-5.5) in this section give more detail on the methods used for estimation.

The starting assumptions, methods of disease progression, and exclusions from the model (patients who are diagnosed at or shortly after birth by meconium ileus for example) are similar to that of Simpson et al. (70) However, there are some important differences. This model uses infection status as a disease progression indicator instead of lung function, and includes sex as an added survival rate indicator. As argued in previous chapters, lung function may not give an accurate assessment of disease severity for various reasons, not least of which is that it cannot be accurately measured in children under 5. (60) (29) Using infection status is a more accurate measure as sputum samples can be gathered from all age groups and the number of infected sputum samples (over a year) gives an indication of the health of the lungs. The cost of care was based on 5 years of annual treatment costs from the ERCFD instead of one year as in the Simpson study. Also, in the analysis reported here, cost was based solely on disease severity (i.e. HRG band) and not age. Assuming the banding structure remains in place, this makes more sense as the NHS is spending money based on disease severity and not necessarily age. The model also differs from Simpson et al as I have used a probabilistic approach which allows testing for uncertainty as the model runs instead of changing one parameter at a time.

Figure 5.1 Diagrammatic representation of CE/Markov Model



Demographic characteristics

The starting population of the model is the total number of live births in England and Wales (in 2008), or 669,601. (190) The model assumed an incidence of CF of 1:2831 based on estimates from Dodge et al. (5) The model ran scenarios with the entire population in each arm, however, in the clinical diagnosis arm, it was assumed that the patients encountered a delay to treatment due to the lack of immediate diagnosis. Data from the ERCFD indicated that of those patients who were clinically diagnosed, 90% of patients were identified in the first year through the onset of symptoms. Subsequently approximately 2% were identified per year until all patients were identified (dependent on the starting population and incidence, in this case, all patients were identified by year 6). As it was a newborn screening model, all patients begin the model at age zero and progress through the model in a yearly cycle.

Table 5.1: Demographics at Baseline

Characteristic	Starting Value	Source
Population	669601	ONS (190)
Incidence CF	0.000353	Dodge et al(5)
Incidence MI	0.15	ERCFD
Clinically diagnosed in year 1	0.9	ERCFD
Subsequent Clinical diagnoses	0.02	ERCFD
Age	0	Assumption
Sex Ratio	1.05	ONS (190)

Transition probabilities and risk factors

The patient then moved through the states of the model on a yearly cycle according to a set of transition probabilities. Transition probabilities between disease states

were derived from creating a Markov matrix from the ERCFD information (See Table 5.2) in STATA, while estimates for progressing to death were based on epidemiological data from the literature (4) which used age and sex-specific survival rates. For the probabilistic analysis, the transition probabilities assumed a Dirichlet distribution (a multivariate generalisation of the Beta distribution).

Table 5.2: Transition Probabilities and Dirichlet Beta distributions estimated from the ERCFD*

	Success	Intermittent	Chronic	Death	Total
Success	0.62	0.33	0.04	0.01	1
Intermittent	0.44	0.51	0.03	0.02	1
Chronic	0	0	0.95	0.05	1
Success Beta (.62,33,.04,.01) Intermittent Beta (.44,.51,.03,.02) Chronic Beta (.95,.05)					
*probability of death calculated from life tables in Dodge et al (5)					

Patients could move from a successful state to intermittent infection, chronic infection, or death or stay in a successful state. Patients in the intermittent infection state could move to a successful state, chronic state, or death, or stay in the intermittent infection state. Patients in the chronic state could only move to the dead state, or stay in chronic infection state. Death was considered an absorbing state. Excel calculated the number of patients in each category by the following formulas:

Successful state: Patients in successful state$_{t-1}$ * (1-RRinf) + (Number of patients in intermittent infection state$_{t-1}$ * InterSuccess) + (Number of patients in chronic infection state$_{t-1}$ * ChronicSuccess) - Number of patients in successful state * (SuccessInter + SuccessChronic + SSuccessDeath)

Intermittent state: Patients in Intermittent state$_{t-1}$ + (Patients in successful state$_{t-1}$ * RRinf) + (Patients in successful state$_{t-1}$ * SuccessInter) + (Patients in chronic state$_{t-1}$ * ChronicInter) – Patients in intermittent state * (SInterDeath+InterChronic+InterSuccess)

Chronic State: Patients in Chronic state$_{t-1}$+ (Patients in successful state$_{t-1}$ * SuccessChronic) + (Patients in intermittent state$_{t-1}$ * InterChronic) – Patients in chronic state*(SChronicDeath+ChronicSuccess+ChronicInter)

Where t-1 is the population in previous cycle, RRinf is the relative risk of acquiring the first infection, SuccessInter, SuccessChronic, InterSuccess, InterChronic, ChronicSuccess, ChronicInter are all transition probabilities for the three available health states, and SSuccessDeath, SInterDeath, SChronicDeath are all mortality rates derived from life tables (shown in Appendix E).

CF patients have a higher relative mortality risk than the general population. Life table estimates (available in Appendix E) were built using data taken from Dodge et al (5) and is available for the total population and disaggregated by sex, though the model here uses the entire population for the life tables.

It has also been shown that once a patient has their first infection, they are more likely to progress more quickly to a chronically infected state. Therefore the relative risk of acquiring the first infection is important. Including parameters that allow for the estimation of the probability of this event allows the transition parameters to

change over time. To estimate these parameters, a Weibull regression[14] (table 5.3) using data from the ERCFD was used to input coefficients and standard errors in order to calculate the hazard function. This regression assumes a lognormal distribution, and therefore the hazard ratio is calculated by exponentiating the coefficient. The regression shows that the gamma, or shape parameter is greater than one, but less than 2, indicating that the hazard rate of first infection increases over time but at a decreasing rate.

As we know the covariance relationship of the parameters from the regression, a Cholesky decomposition technique was used to enable correlated random draws from a normal distribution. The Cholesky decomposition matrix is the square root of the covariance matrix. Once this is calculated, a vector of correlated variables can be generated. Z is a randomly generated lognormal vector, Tz is that vector multiplied by the Cholesky decomposition and mu+Tz is the vector of parameter means. Table 5.4 gives the co-variance matrix, correlation matrix, Cholesky decomposition matrix and the random variables used to incorporate uncertainty into the model. The calculations can be seen by clicking the “hazard tables” tab on the embedded workbook in appendix E. Using the Cholesky decomposition method enabled the transition probabilities to be calculated using a Weibull distribution,

[14] A Weibull model is a two parameter proportional hazards model, with a shape (gamma) and scale (lambda) parameter that is used to model the hazard of failure rates (in this case, infection) 37.
Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. Oxford, UK: Oxford University Press; 2006.

which is suitable for modelling data with hazard rates that increase or decrease over time.

Table 5.3: Regression to calculate hazard function for relative risk of infection

Weibull Regression – log relative-hazard form				Number of obs =	324
Group variable: pid					
No of subjects = 157					
No of failures = 121					
Time at risk = 324					
Log likelihood = -132.5275				LR chi2(4) =	17.42
				Prob > chi2 =	0.00
Time to Infection	Coefficient	Std Err	Hazard Ratio	95% Conf Int (of hazard ratio)	
Age	-0.40	.02	0.961	.9382072	1.01327
Sex	0.13	.18	1.139	.7099908	1.462772
Screened	0.16	.025	1.174	1.053336	1.155197
Gamma	0.40	.67	1.492	1.227258	1.814183
Constant	-5.49	.21	0.004	.0000987	.0015578
Lambda	.02				

Table 5.4: Covariance matrix, Cholesky decomposition, and random variable generation for probabilistic analysis (based on coefficients, standard errors and hazard ratio from regression in table 5.3)

Covariance matrix

	lngamma	cons	age	male	screened
lngamma	0.0081				
cons	-0.0057	0.0441			
age	0.0000	-0.0008	0.0004		
male	0.0000	-0.0072	0.0000	0.0324	
screened	0.0003	-0.0006	-0.0001	0.0002	0.0006

Cholesky decomposition

	lngamma	cons	age	male	screened
lngamma	0.0900				
cons	-0.0632	0.2003			
age	0.0000	-0.0039	0.0196		
male	0.0001	-0.0362	-0.0055	0.1762	
screened	0.0029	-0.0023	-0.0061	0.0004	0.0240

Random variables

	z		Tz		mu + Tz
lngamma	-0.2234		-0.0201		0.3799
cons	-1.0806		-0.2023		-5.6932
age	-0.4214		-0.0040		-0.0440
male	-0.5203		-0.0503		0.0797
screened	0.9801		0.0277		0.1877

Health economic data

The model is based upon an NHS perspective. The costs included where those for inpatient/outpatient care, and drugs, as well as the cost of screening. In the case of CF, the cost of screening is relatively cheap, as many of the costs of a screening programme are already sunk into the existing national newborn screening programme. Cost of care was assigned to different health states of the model by using data from the cost of illness study in chapter four. Details of the methods

used to gather that information is in Chapter 4. To incorporate uncertainty around the cost estimates, a lognormal distribution was assumed (this distribution was shown to be appropriate in chapter 4).This was calculated in excel by using the =LOGINV(RAND(),LN(Cost),STDEV(LN(Cost)) command. All future costs were discounted to present value at 3.5% annually and were presented for the year 2010 in UK pounds sterling (GBP).

Quality of life information was taken from a study by Simpson et al. (70) These estimates were based on previous evidence (191) that forced expiratory volume results were indicative of a change in the quality of life as measured by the Quality of Well-Being Scale (111) which can be used in economic evaluations to derive QALY estimates. While not based on infection information, these QALY estimates are currently the only available in the literature for paediatric patients. QALY estimates were modelled using a Dirichlet (Beta) distribution with alpha being calculated by (U*(U*(1-U)/(seU^2)-1)) and beta being calculated by (U*(1-U)/(seU^2)-1-alpha) where U is utility, and seU is the standard error of the utility estimates.

Table 5.5: Starting Cost and Utility parameters

Cost parameters	**Unit**	**Source**
Cost of one Success cycle	£ 10,301	ERCFD (Ch 4)
Cost of one Intermittent infection cycle	£ 22,928	ERCFD (Ch 4)
Cost of one Chronic cycle	£ 34,351	ERCFD (Ch 4)
Cost of Screening one individual	£ 2	Addenbrooke's Hospital
Cost of Treatment (pre-symptom Clinically Diagnosed)	£ 1,170	Simpson et al
Utility of Markov states per cycle		
Utility gained of one Success cycle	0.95	Simpson et al
Utility gained of one Intermittent cycle	0.75	Simpson et al
Utility gained of one Chronic cycle	0.68	Simpson et al

Table 5.6: Summary of model parameters

Name	*'live' value*	*probabilistic*	*deterministic*	*standard error*	*distribution*	*alpha*	*beta*	**Description**
Population	669601							Live births in England
	0.0003532							
Incidence CF	3							Incidence of CF
Incidence MI	0.15							MI
Clinically diagnosed in year 1	0.9							Percent clinically diagnosed year 1
Subsequent Clinical diagnoses	0.02							Percent clinically diagnosed yearly after year 1
Age	0							Average age of all patients at receipt of primary implant
Sex	2							Sex indicator (0 for female, 1 for male)
Sex Ratio	1							
cDR	3.5%							Cost discount rate
oDR	3.5%							Outcome discount rate
Transition probability variables (see diagram on page ***<Model Figure>*** for details)								
SuccessDeath	0.01	0.02	0.01	0.010	Beta	1	99	Mortality Rate for "well" patients
InterDeath	0.02	0.27	0.02	0.014	Beta	2	98	Mortality for intermittent patients
ChronicDeath	0.05	0.29	0.05	0.022	Beta	5	95	Mortality for chronic patients
SuccesInter	0.33	0.31	0.33	0.047	Beta	33	67	Chance of moving from healthy to inter
SuccesChronic	0.04	0.32	0.04	0.019	Beta	4	96	Chance of moving from health to chron state
InterSuccess	0.44	0.32	0.44	0.049	Beta	44	56	Chance of moving from inter back to health
InterChronic	0.06	0.31	0.06	0.024	Beta	6	94	Chance of moving from inter into chron state
ChronicInter	0.01	0.34	0.01	0.010	Beta	1	99	Chance of moving from chron to inter state
ChronicSucces	0.00	0.35	0.00	0.000	Beta	0	100	Chance of moving from chron to healthy state

Name	*'live' value*	*probabilistic*	*deterministic*	*standard error*	*distribution*	*alpha*	*beta*	**Description**
*Risk of acquiring first infection (see **<Hazard function>** for details)*								
cons	-5.49	-4.95	-5.49	0.21	LogNormal			Constant in survival analysis for baseline hazard
ageC	-0.04	-0.08	-0.04	0.02	LogNormal			Age coefficient in survival analysis for baseline hazard
maleC	0.13	-0.23	0.13	0.18	LogNormal			Male coefficient in survival analysis for baseline hazard
lambda	0.02							Lambda parameter survival analysis (depends on chosen mix of above coefficients)
gamma	1.49	0.36	0.40	0.09	LogNormal			Ancillary parameter in Weibull distribution
rrNBS	1.17	0.15	1.17	1.03	LogNormal			Relative risk of moving to ill state once screened
Cost parameters								
cInter	£22,928	1226	22928.45		LogNormal	9.24	1.28	Cost of one cycle in the intermittent infected state
cChronic	£34,351	17911	34351.13		LogNormal	9.70	1.33	Cost of one cycle in the chronic infection state
cSuccess	£10,301	6901	10301.12		LogNormal	8.65	1.00	Cost of one cycle in a 'success' state
cScreening	£ 2	2	2.00		LogNormal	0.59	0.02	Cost of screening
cCDTreatment	£1,170	211	1170.49		LogNormal	7.07	1.32	Cost of clinical diagnosis
Utility of Markov states per cycle								
uSuccess	0.95	0.94	0.95	0.03	Beta	49.19	2.59	Utility of being well
uInter	0.75	0.71	0.75	0.04	Beta	87.14	29.05	Utility of being intermittently infected
uChronic	0.68	0.72	0.68	0.03	Beta	163.73	77.05	Utility of being chronically infected

Data analysis

Once the parameter data was in the model, the model was programmed to run 1000 simulations for each patient. Each simulation added the costs and the QALY scores of every patient in each disease state to determine the overall cost and QALY gain/loss for the population for each simulation. This information was then used to calculate the incremental cost-effectiveness ratio:

$$\sum_{k=1\ to\ 1000}[\sum_{i=1\ to\ 669601}\{(CScreening_i\ |params = \theta_k) - (Cnonscreened_i\ |params = \theta k)$$

divided by

$$\sum_{k=1\ to\ 1000}[\sum_{i=1\ to\ 669601}\{(QScreening_i\ |params = \theta_k) - (Qnonscreened_i\ |params = \theta k)$$

Where k = the number of simulations, i = the number of individuals in the model, Cscreening and Qscreening = cost and QALYs associated with screening, Cnonscreened and Qnonscreened = cost and QALYs associated with the non screened group, and θ is the set of parameters set out in table 5.6. The ICERs from each simulation were then plotted on a cost-utility plane, which illustrates a joint distribution of costs and utility. This provides a visual indicator of the amount of uncertainty around the estimate. Knowing the joint distribution of the costs and effects allows standard statistical analysis of cost-effectiveness data. Negative ratios (those that appear in the SE and NW quadrants of the CU plane) yield negative ICERs. As mentioned in Chapter 3, the ICER fails to distinguish between two situations we would treat differently when it becomes negative. A histogram of

the ICER will be presented to look for negative ratios and intractability of the variance.

Given there is no set decision rule on whether or not to adopt screening if it is below a certain ceiling willingness-to-pay, a net-benefit approach is taken. This approach allows us to multiply the incremental effect by the ceiling ratio, which generates the maximum willingness to pay in monetary terms for the effect gain. Subtracting the incremental cost of the intervention leaves the net-benefit on a cost scale.

$$NMB = \lambda * \Delta E - \Delta C > 0$$

where NMB is net monetary benefit, λ is the threshold, and delta E and delta C are the changes in effect minus the change in cost. A positive incremental net-benefit suggests that the intervention represents good value for money, while a negative incremental net benefit suggests the intervention is not cost-effective. Using the net benefit approach allows the determination of how much of the estimated joint density falls on either side of a line representing the decision rule by plotting a cost-effectiveness acceptability curve over a range of monetary thresholds of choice (CEAC).

To understand model uncertainty, and the relative effect of different parameters in the model, the expected value of perfect information (EVPI) and the expected value of perfect parameter information (EVPPI) approaches will be used at the population level. The EVPI helps to look at the cost of making the wrong decision and the opportunity costs of error. The expected costs of uncertainty can be interpreted as the EVPI, as having perfect information eliminates the possibility of a wrong decision. A health care system's budget constraint is generally considered the

maximum that the health care system is willing to pay for additional evidence to inform the decision. For this thesis, a non-parametric approach was taken to estimate EVPI by calculating the maximum net benefit for each iteration from the simulation (1000) for a particular willingness to pay threshold (0-100,000), then take the mean over the maximum net benefits over the values of the willingness to pay threshold, or:

$$EVPI = E_\theta max_j\, NB(j,\theta) - \, max_j E_\theta NB(j,\theta)$$

Where θ is the willingness to pay threshold and j represents the alternative interventions. To ensure that the decision is applicable for all patients and across a useful time period, it is important to express the EVPI for the total patients who stand to benefit from additional information (in this case, CF patients). This can be calculated by:

$$EVPI\ for\ the\ population = EVPI. \sum_{t=1\ldots T} \frac{I_t}{(1+r)^t}$$

Where T is the period over which the information will be useful and I is the incidence over the period. EVPI associated with future patients is discounted at rate r to provide the total EVPI for the population of current and future patients.

EVPPI can give us information about what type of additional evidence would be most valuable. It is calculated in a very similar way to EVPI where:

$$EVPPI_\varphi = \, E_\varphi max_j E_{\psi|\varphi} NB(j,\varphi,\psi) - \, max_j E_\theta NB(j,\theta)$$

Where φ is the value of perfect information about a parameter, ψ is the expected net benefit of a decision by choosing the alternative with the maximum expected net

benefit when those expected net benefits are averaged over the remaining uncertain parameters. The EVPPI will use 100 inner and 100 outer loops to estimate the expected values.

Results

The model assumed that all 669,601 births in 2006 were started in the model at age 0. Once in the model, patients were either screened or not screened. Excluding a percentage of patients with MI and/or a family history meant that the number of patients with CF was 237. Table 5.7 gives a deterministic run of the numbers of patients in each health state by year. The no screening group had fewer patients spending time in the success state than the screened group (2014 versus 2312). The no screening group had more patients spending time in the intermittent state than the screened group (1521 versus 1486). The no screening group also had fewer patients spending time in the chronic state than the screened group (1356 versus1383) though the average age was slightly younger (10.72 for the not screened and 12.26 for the screened). The no screening group had more patients ending in the dead state faster, although the average age of death (calculated by averaging the ages at death) was only marginally different (21.85 in the screening group, 21.93 in the not-screened group). This may indicate that the patients in the no screening arm may progress faster through the model due to the delay to treatment.

Table 5.7: Number of people in each health state by year (total life years in each state)

Year	No Screening				Screening			
	Succ	Inter	Chron	Dead	Succ	Inter	Chron	Dead
1	132	72	9	0	237	0	0	0
2	118	81	18	1	148	78	9	0
3	113	81	27	2	127	88	20	1
4	109	79	35	4	119	86	29	3
5	106	78	43	5	112	82	38	4
6	103	76	50	7	106	78	46	6
7	96	74	57	9	101	75	53	8
8	91	71	63	12	96	71	59	10
9	86	68	65	18	91	68	65	12
10	81	64	67	24	87	64	67	19
11	77	61	68	31	82	61	68	25
12	73	58	68	37	78	58	69	32
13	69	55	69	44	74	55	70	39
14	65	52	69	50	70	52	70	45
15	62	50	68	57	66	49	69	52
16	59	47	67	63	63	47	69	58
17	56	45	67	69	60	45	68	64
18	53	43	66	75	57	42	67	71
19	49	39	57	91	54	40	66	77
20	45	36	50	104	50	37	57	92
21	42	34	44	117	46	34	50	106
22	39	31	39	127	43	32	44	118
23	36	29	35	137	40	29	39	128
24	33	27	31	146	37	27	35	138
25	30	25	28	154	34	25	31	147
26	28	23	25	161	31	23	28	154
27	26	21	23	167	29	21	25	161
28	24	19	21	173	27	20	23	168
29	21	15	10	191	25	18	20	173
30	18	13	5	200	22	14	9	191
31	15	11	3	207	19	12	5	200
32	13	9	2	212	16	10	3	207
33	11	8	2	216	14	9	2	211
34	9	7	2	219	12	7	2	215
35	8	6	1	222	10	6	2	218
36	7	5	1	224	9	5	1	221
37	6	4	1	226	7	5	1	223
38	5	3	1	227	6	4	1	225
39	2	1	0	233	5	3	1	227
40	1	0	0	235	2	2	0	233
41	0	0	0	236	1	0	0	235
42	0	0	0	236	0	0	0	236
43	0	0	0	236	0	0	0	236
44	0	0	0	237	0	0	0	236
45	0	0	0	237	0	0	0	237
Life Years	**2,014**	**1,521**	**1,356**	**5,682**	**2,312**	**1,486**	**1,383**	**5,464**

Table 5.8: Costs of health states by year

Year	No Screening				Screening			
	Cost Succ	Cost Inter	Cost Chron	Total Cost (all)	Cost Succ	Cost Inter	Cost Chron	Total Cost (all)
1	1,356,460	1,659,500	292,496	3,308,456	2,436,468	0	0	2,436,468
2	1,211,969	1,857,153	610,906	3,680,028	1,527,635	1,798,354	324,995	3,650,984
3	1,159,273	1,852,904	915,074	3,927,250	1,312,596	2,019,862	677,420	4,009,879
4	1,123,805	1,817,475	1,199,664	4,140,943	1,221,166	1,973,814	1,006,976	4,201,955
5	1,093,128	1,778,884	1,464,954	4,336,967	1,154,420	1,886,482	1,307,014	4,347,916
6	1,065,040	1,741,673	1,712,072	4,518,786	1,095,054	1,796,634	1,578,298	4,469,986
7	990,296	1,706,537	1,942,214	4,639,048	1,039,870	1,710,679	1,822,734	4,573,284
8	936,644	1,635,195	2,150,022	4,721,861	988,122	1,629,459	2,042,306	4,659,888
9	886,549	1,549,902	2,229,092	4,665,543	939,499	1,552,870	2,238,892	4,731,261
10	838,469	1,471,787	2,286,720	4,596,975	891,468	1,469,324	2,302,971	4,663,763
11	793,188	1,398,238	2,325,744	4,517,170	844,891	1,393,774	2,347,385	4,586,050
12	750,637	1,328,678	2,348,615	4,427,930	800,827	1,323,006	2,374,811	4,498,643
13	710,639	1,262,833	2,357,507	4,330,980	759,294	1,256,269	2,387,538	4,403,100
14	673,018	1,200,484	2,354,347	4,227,849	720,152	1,193,235	2,387,584	4,300,970
15	637,609	1,141,428	2,340,848	4,119,884	683,247	1,133,657	2,376,738	4,193,643
16	604,259	1,085,476	2,318,527	4,008,263	648,432	1,077,320	2,356,588	4,082,340
17	572,832	1,032,450	2,288,731	3,894,013	615,571	1,024,023	2,328,537	3,968,131
18	543,198	982,182	2,252,649	3,778,029	584,538	973,580	2,293,827	3,851,945
19	504,221	902,987	1,970,524	3,377,733	555,215	925,820	2,253,558	3,734,592
20	465,886	834,387	1,731,547	3,031,820	516,234	850,861	1,967,790	3,334,885
21	430,087	771,245	1,528,272	2,729,604	477,948	785,906	1,726,267	2,990,121
22	396,880	712,735	1,354,511	2,464,126	442,069	726,287	1,521,282	2,689,638
23	366,127	658,508	1,205,221	2,229,856	408,702	671,112	1,346,451	2,426,265
24	337,669	608,275	1,076,306	2,022,250	377,729	620,015	1,196,582	2,194,325
25	311,354	561,765	964,424	1,837,543	349,007	572,708	1,067,452	1,989,168
26	287,034	518,723	866,847	1,672,604	322,394	528,930	955,626	1,806,950
27	264,570	478,905	781,340	1,524,815	297,749	488,434	858,300	1,644,483
28	243,828	442,083	706,069	1,391,979	274,939	450,986	773,180	1,499,106

Year	No Screening Cost Succ	Cost Inter	Cost Chron	Total Cost (all)	Screening Cost Succ	Cost Inter	Cost Chron	Total Cost (all)
29	219,051	346,150	326,342	891,542	253,838	416,368	698,386	1,368,592
30	186,556	293,204	177,769	657,529	228,462	326,081	322,601	877,144
31	158,440	248,820	115,211	522,471	195,977	276,185	175,874	648,036
32	134,416	211,256	84,959	430,631	167,391	235,179	114,256	516,826
33	113,979	179,387	67,492	360,858	142,814	200,449	84,583	427,846
34	96,626	152,325	55,616	304,567	121,790	170,904	67,506	360,199
35	81,902	129,338	46,595	257,835	103,841	145,732	55,900	305,472
36	69,415	109,811	39,318	218,544	88,529	124,272	47,067	259,868
37	58,826	93,224	33,279	185,329	75,471	105,973	39,917	221,361
38	49,850	79,135	28,202	157,187	64,338	90,369	33,957	188,664
39	16,996	26,343	13,481	56,820	54,846	77,061	28,925	160,831
40	5,659	9,158	4,500	19,317	18,980	45,857	0	64,837
41	1,957	2,979	1,533	6,469	6,426	9,156	4,725	20,307
42	638	1,071	513	2,223	2,248	3,027	1,633	6,907
43	227	330	176	734	751	1,094	556	2,401
44	71	128	58	257	267	349	193	810
45	71	128	58	257	87	132	65	285
Cumulative Total	**20,749,347**	**34,875,179**	**46,570,347**	**102,194,873**	**23,811,291**	**34,061,591**	**47,497,244**	**105,370,126**

Table 5.8 shows the accrual of cost by state over time. It does not, however, show the cost of screening or the cost of caring for the pre-symptomatic clinically diagnosed patients. The total costs of care in the non-screened arm amounted to £102,194,873 and £105,370,126 in the screened arm over 45 years Reflecting the differences of numbers of patients in each state, the cost of the successful state is approximately £3 million more in the screened arm than the no screening arm. The screened arm also accrued higher costs in the chronic stage of approximately £1 million more than the non-screened arm. The non-screened arm accrued approximately £800,000 more in the intermittent state than the screened arm. Overall, the screened arm was more expensive by approximately £3 million.

Figure 5.2 gives a graphical representation of total state costs over time. Initially, the total cost group is more expensive, reflecting the number of people diagnosed from the symptoms of CF. Over time, however, newborn screening arm becomes more expensive. This reflects the findings in the above tables, as patients who are screened stay alive longer and therefore accrue more costs.

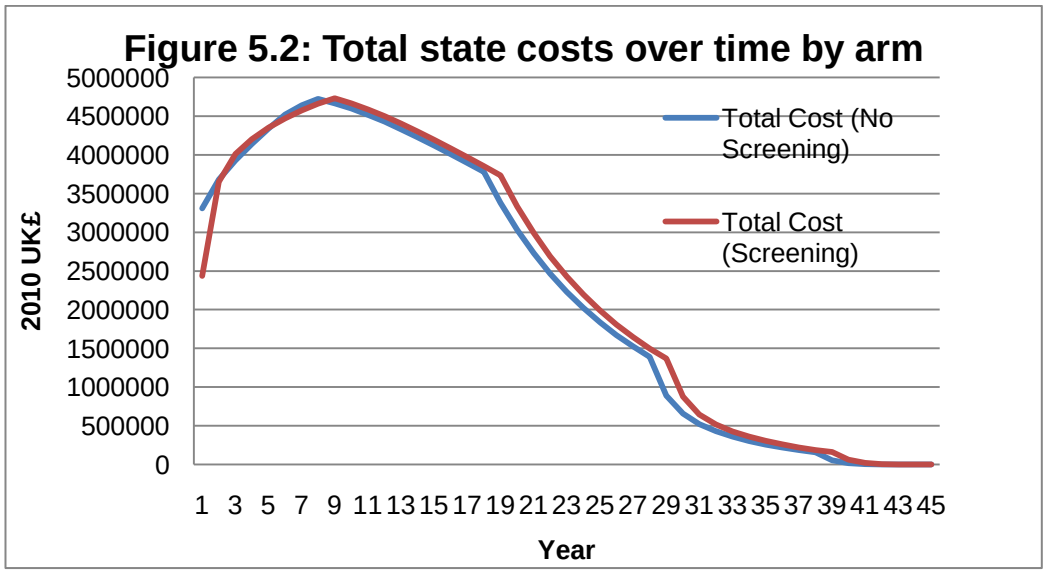


Table 5.9: Utilities gained in health states by year						
	No Screening			Screening		
Year	Utility Succ	Utility Inter	Utility Chron	Utility Succ	Utility Inter	Utility Chron
1	195	0	0	217	0	0
2	117	51	5	132	55	6
3	101	55	11	109	60	12
4	93	53	16	98	56	17
5	87	50	20	90	52	22
6	82	47	24	82	48	25
7	77	45	27	75	44	28
8	69	42	29	69	40	31
9	63	39	31	64	37	33
10	58	36	31	58	34	32
11	53	33	31	53	31	32
12	48	30	30	49	29	31
13	44	28	30	45	26	30
14	40	26	29	41	24	29
15	37	23	28	38	22	28
16	34	22	27	34	20	27
17	31	20	26	32	19	26
18	28	18	24	29	17	24
19	26	17	23	27	16	23
20	23	15	20	24	14	20
21	21	13	17	21	12	17
22	19	12	14	19	11	14
23	17	11	12	17	10	12
24	15	9	10	15	9	10
25	13	8	9	14	8	9
26	12	8	8	12	7	8
27	10	7	7	11	6	7
28	9	6	6	10	6	6
29	8	5	5	9	5	5
30	7	4	2	8	4	2
31	6	3	1	6	3	1
32	5	3	1	5	3	1
33	4	2	1	4	2	1
34	3	2	0	3	2	0
35	3	1	0	3	1	0
36	2	1	0	2	1	0
37	2	1	0	2	1	0
38	1	1	0	2	1	0
39	1	1	0	1	1	0
40	0	0	0	0	0	0
41	0	0	0	0	0	0
42	0	0	0	0	0	0
43	0	0	0	0	0	0
44	0	0	0	0	0	0
45	0	0	0	0	0	0
Cum QALY	**1,468**	**748**	**556**	**1,531**	**738**	**571**

Table 5.9 shows the number of quality adjusted life years gained by each health state over time. The screening arm accumulated more QALYs over time in the Successful and Chronic health state in comparison to the no screening arm, but accumulated fewer QALYs in the intermittent health state. This again indicates patients in the screening arm are staying in the successful state longer, and staying alive longer, and therefore reaching the chronic state.

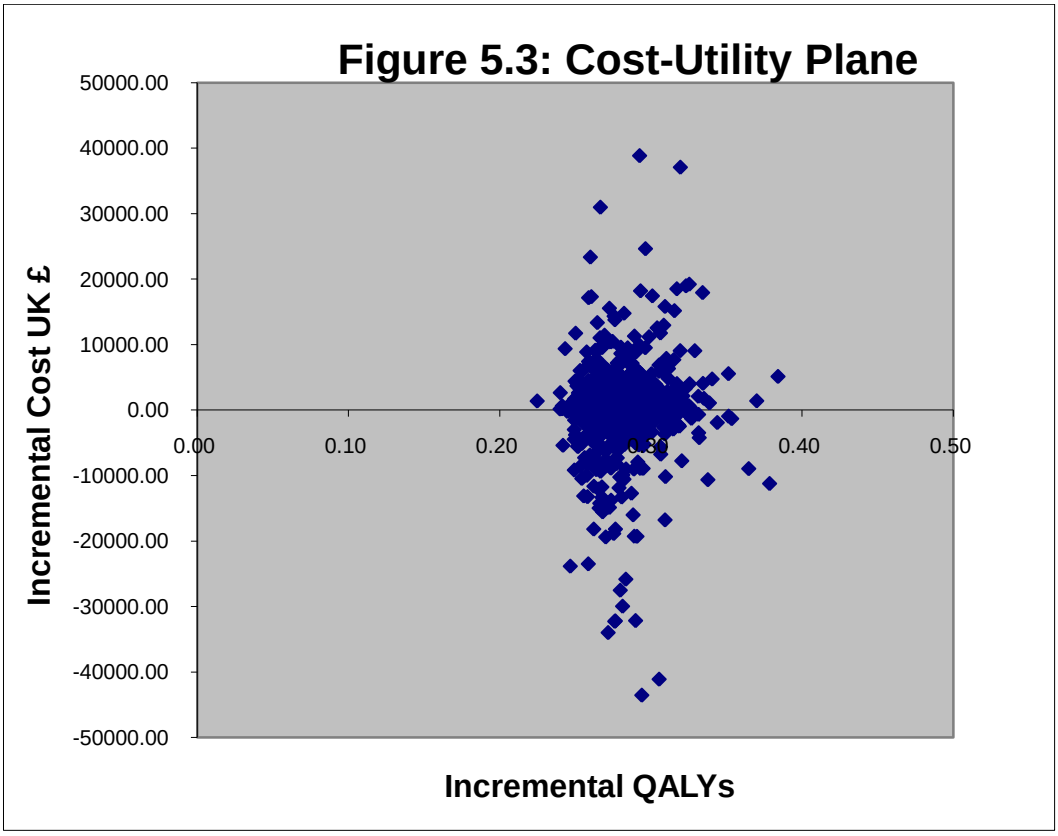
Table 5.10: Results of the probabilistic analysis

Intervention	Cost	QALYs
No Screening	£ 110,238	11.23
NBS	£ 110,417	11.51
difference	£ 179	0.28
ICERs:		
NBSvsNot	£ 641	per QALY

Table 5.10 shows the ICER calculation for the simulations run by the model. Overall, newborn screening was more expensive than not screening, though only just (£110,238 vs. £110,417). However, newborn screening seems to be more effective in generating quality adjusted life years (11.51 vs. 11.23). This meant that the ICER was estimated to be £641/QALY gained. This means that generating one additional QALY costs approximately £641. Assuming the "typical" NICE cost/QALY threshold of £30,000 per QALY, newborn screening seems to be a cost-effective option.

In order to gain a better understanding of the variance in the estimates, a cost-utility plane (Figure 5.3) was plotted with the cost and utility data from the simulations. The plot shows that simulations are spread relatively equally in the northeast and

southeast quadrants, meaning that in approximately half the scenarios, newborn screening dominates the no screening option, meaning screening was both cost saving and generated better utility gains than no screening (southeast quadrant). The diagram also shows that there were relatively few outliers within the simulations.



A visual inspection of the ICER distribution (Figure 5.4) also indicates that there are negative ICERs as well as outliers. However, the majority of ICER simulations are between 0 and £10000 per QALY gained. While a useful tool to get an idea of the variance within the replications, the distribution does not tell us about how much overall or parameter uncertainty there is within the model.

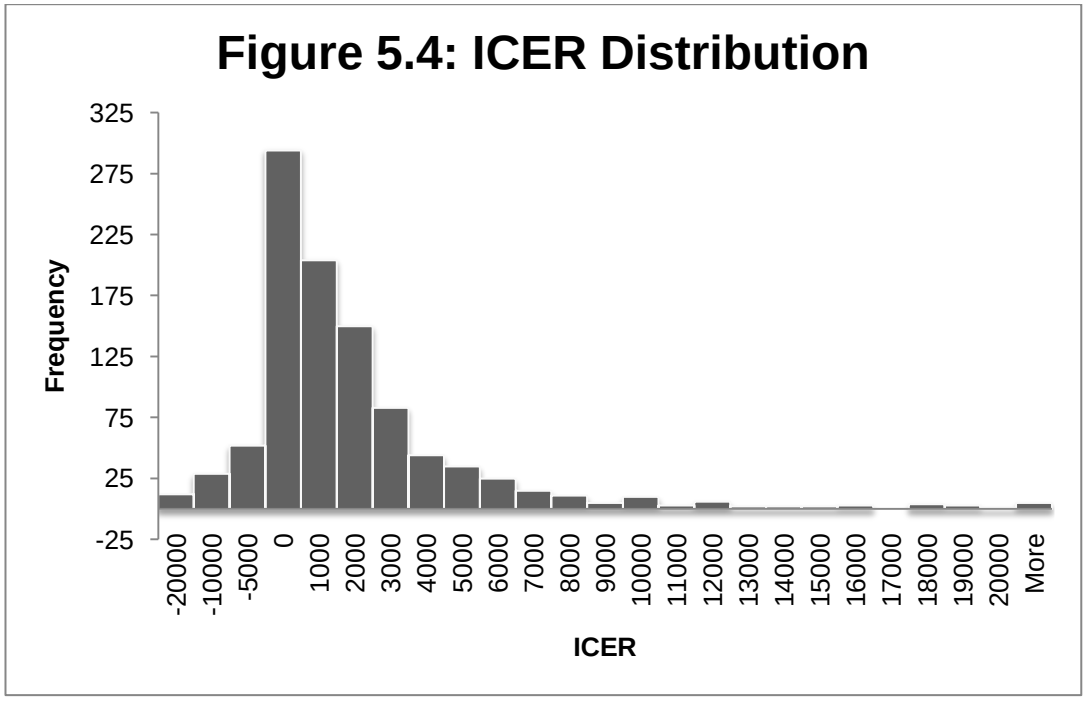


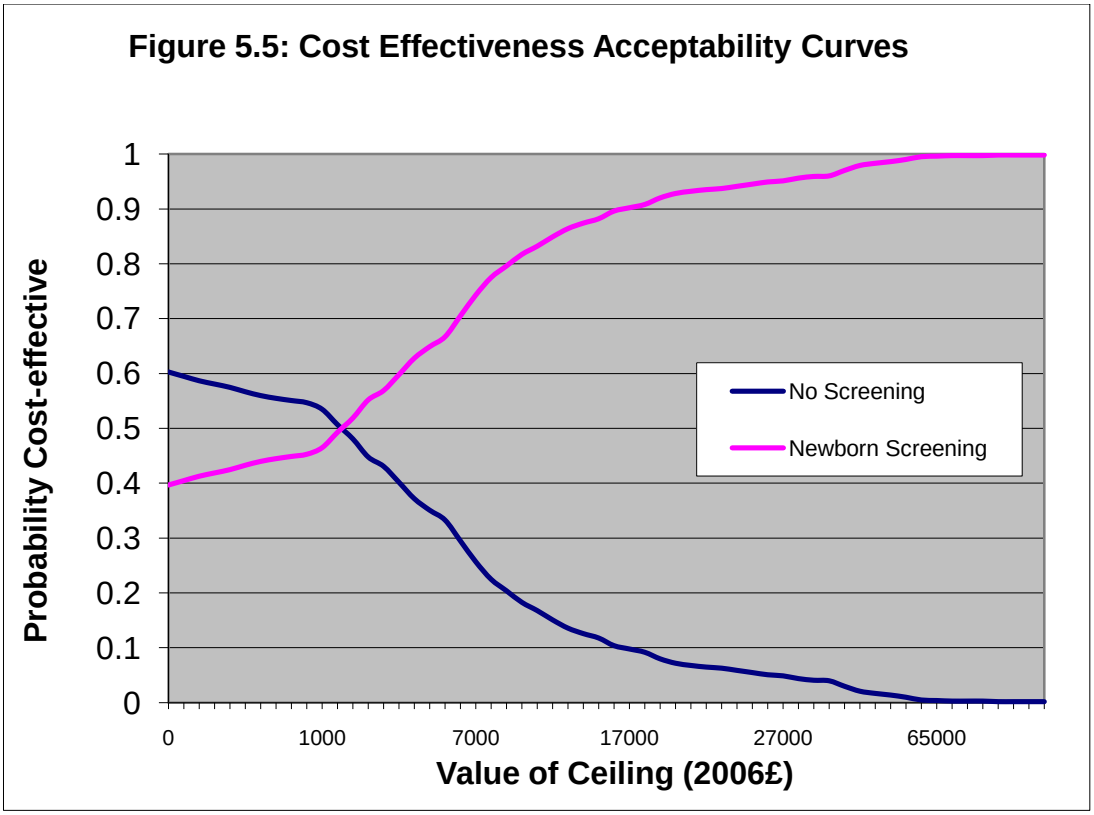
Table 5.11 gives a list of the incremental net monetary benefits for a number of ceilings. The net benefit approach indicates that if the willingness to pay threshold is above £2,000 per QALY, newborn screening is the most cost-effective option.

Table 5.11: Net Monetary Benefit

Ceiling	Net Monetary Benefit
-£10,000	-£2,973
-£8,000	-£2,414
-£6,000	-£1,856
-£4,000	-£1,297
-£2,000	-£738
£0	-£179
£2,000	£380
£4,000	£938
£6,000	£1,497
£8,000	£2,056
£10,000	£2,615
£12,000	£3,173
£14,000	£3,732
£16,000	£4,291
£18,000	£4,850
£20,000	£5,409

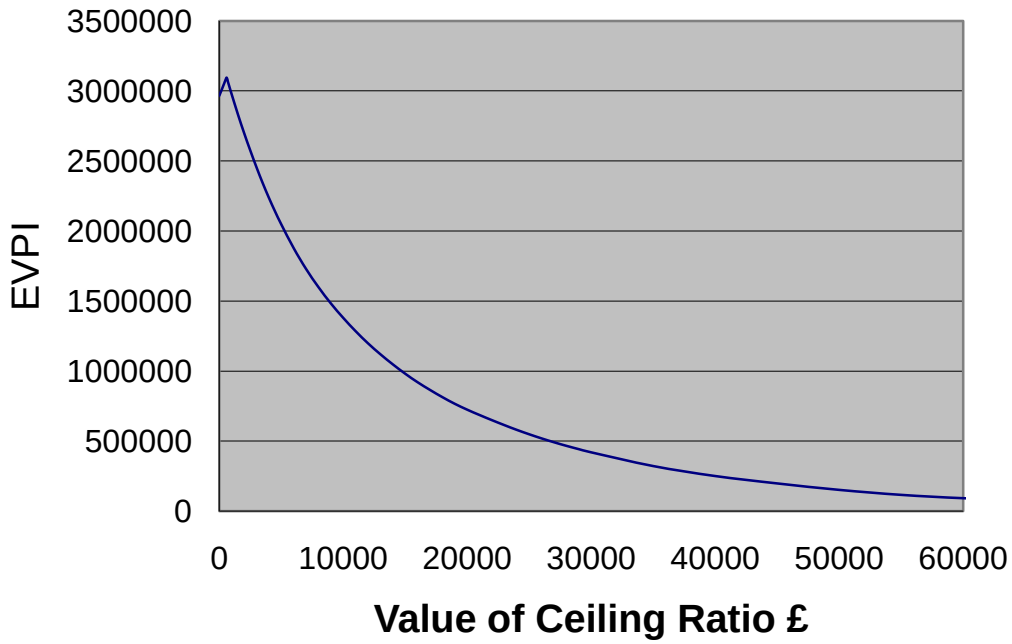
While this is consistent with the ICER, a CEAC was created to estimate the probability of the interventions being cost-effective given a certain ceiling. Figure 5.5 shows the CEACs for this analysis. This figure illustrates that until a ceiling of approximately £2,000, no screening has a higher probability of being the more cost-effective option. Thus, if a decision maker is not willing to pay anything for a gain in QALYs, then not screening has a 60% chance of being the cost effective option but there is a 40% chance that will be the wrong decision. However, after £2,000 per QALY gained threshold, newborn screening becomes more likely to be cost effective. However, there is still uncertainty surrounding the decision (until a willingness to pay threshold of approximately £65,000 per QALY gained).

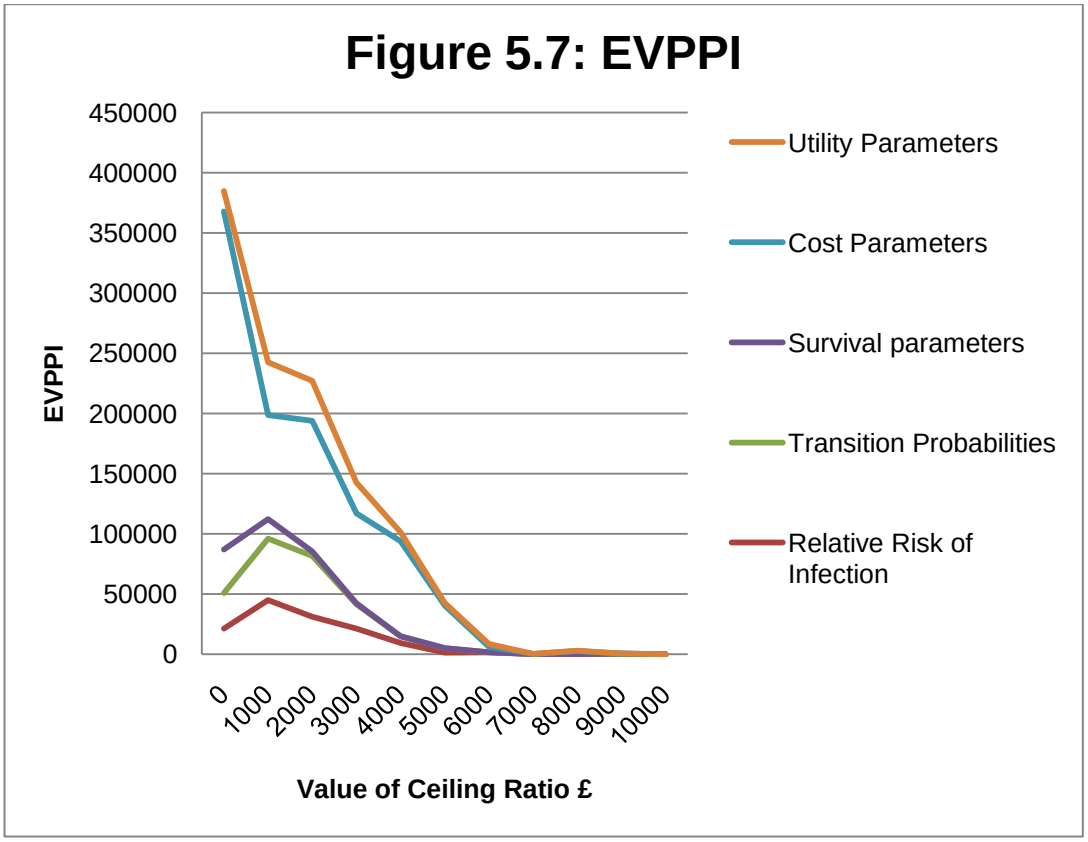
Figure 5.5: Cost Effectiveness Acceptability Curves



It is obvious from the cost-utility plane and the CEACs that there is a substantial amount of uncertainty in the decision rule for ceiling values under £18,000. To aid decision makers in determining if further research should be done to eliminate this uncertainty, the population EVPI was calculated (Figure 5.6). The population EVPI was calculated by multiplying the maximum net benefit by the effective population of CF patients in the England and Wales for one year's birth cohort over 15 years (cumulative population 1957). The EVPI has its highest point (~£31,000,000) at a ceiling ratio of £2000, which is where the two CEACs meet and uncertainty is highest. If the cost of research is assumed to be £30,000 per QALY, there is very little value in conducting further research. However, if the cost of research is closer to £2000, then it may be worthwhile to carry out the research to reduce the amount of uncertainty.

Figure 5.6: Population EVPI





Although it is unlikely that new research would be funded at a £30,000 per QALY threshold, it is worth checking what parameters may be worthwhile to investigate. The EVPPI (figure 5.6) shows that the utility and cost parameters contribute a large amount of uncertainty to the model. Further information on survival parameters and disease progression (transition probabilities) may also benefit the model. However, all the parameters or groups of parameters have an EVPPI of £0 after a ceiling ratio of £10,000, indicating that if research costs on any of these groups of parameters are more than £10,000, it is unlikely that it would be worthwhile.

Discussion

Compared to no screening, our model estimated that the incremental cost effectiveness of £641 per QALY gained. This is significantly cheaper in comparison to the incremental cost-effectiveness ratio of £6864 per QALY gained as estimated by Simpson et al (70) in 1998. This may be due to differences in cost estimates, updated epidemiological evidence, and model design. However, both models suffer from uncertainty of evidence on the impact of screening on various health indicators (such as height and weight and lung function) and how they affect quality of life. The quality of life estimates used in this model were relatively conservative with a relatively high value placed on the chronic state. However, a more accurate reflection of quality of life for this particular study would probably make newborn screening even more cost-effective.

The survival estimates are based on the most recent and reliable age-specific survival data, but there is still very limited information on the interactions between CF and the normal aging process as well as how and when patients move from one health state to another. There is also limited information on the use of the Quality of Well-Being scores to estimate QALYs, and therefore these may not be indicative of the true state of health. Indeed, the EVPPI shows that the utility estimates contribute the most amount of uncertainty to the model. The recent emergence of the disease and age specific CFQoL measure is likely to improve estimates of utility by disease state for children with CF.

The model and data inputs would satisfy most of the criteria for good practice in decision modelling. (192) While the model has not been externally validated, running a probabilistic model with uncertainty analyses should control for biases in the estimates.

Due to the narrow costing approach, it is likely the cost estimates in the model are underestimates of the total cost of CF care to the economy, though they are representative from the tertiary NHS perspective. This may, however, impact on the generalisability of the study outside the NHS and the UK. Care would need to be taken as newborn screening approaches differ from country to country and area to area.

In conclusion, this model presents the most up to date evidence on the cost-effectiveness of newborn screening in the UK. As the screening programme was rolled out in 2007 nationwide based on the evidence available at the time (70), it seems that the Department of Health's decision was the correct one based on a £30,000 per QALY threshold. Indeed, both the Simpson study from 1998 and this one indicate that it is highly likely that it would be a cost-effective decision at even lower thresholds.

Chapter 6: Discussion

In this final chapter I will outline the main results for each of the objectives of this thesis. I will then discuss the methodological limitations of the approaches in this thesis and how I have tried to address them. This will be followed by a discussion of the policy implications of the findings of the thesis as well as areas of further research.

Overview of thesis

Rationale and objectives

Cystic Fibrosis is the most common genetic disorder in the UK. It is a chronic, life-shortening, multisystem disease. CF disease severity differs between individuals, though it is not entirely clear why this is the case. Due to the variability in morbidity and range of types of interventions that can be helpful, caring for CF patients requires a team of medical professionals: including consultants, specialist nurses, physiotherapists, nutritionists and psychologists.

Given the complex and intensive nature of CF treatment, it is important for NHS decision makers to have an idea of where costs fall and how they vary when treating CF patients so that the proper funding is in the right place to provide the right incentives to ensure care is of the highest standard. As CF disease severity (and therefore resource use) differs between patients, the best way of gathering information on patient costs is by looking at individuals resource use over time.

One of the major clinical developments in the last 10 years was the introduction of nationwide newborn screening for CF in 2007 (although some areas such as the

counties in this study have been screening for 20+ years). Newborn screening allows for earlier intervention, which in theory could lead to a better long-term prognosis for the patient. Newborn screening and the subsequent treatment has resource and cost implications for the NHS.

The first aim of this thesis was to collect and analyse data on patient resource use and subsequent costs to the NHS over time. The second aim of the thesis was to analyse the cost-effectiveness of a newborn screening programme for CF versus clinical diagnosis.

Main findings

Review of the current evidence on the economics of paediatric CF:

The systematic review of the economics of paediatric CF (Chapter 2) highlighted a number of limitations of existing published research. The age of some of the cost-of-illness and cost-effectiveness studies meant that they are likely to be of limited use now, given changes in clinical such as the introduction of rhDNAse. Most of the cost studies also only reported the arithmetic mean, which could possibly lead to an overestimate of costs given the skewed nature of cost data. There were also wide variations in the type of health system, and treatment regimes which limits the applicability of the studies in the UK context.

The cost of paediatric CF in the East of England

Chapter four describes the findings of the cost-of-illness analysis. The analysis showed that in 2007, the total cost of caring for 174 paediatric patients was £1,062,008. Median costs per year varied from £3853 to £4249 over time. Despite a

slight fall in the number of days spent in hospital, inpatient stays contributed the most to the cost of care, followed by drug costs and clinic costs (outpatient visits). These findings were marginally higher than those of Sims et al (3). Using the banding tariffs as unit costs instead of NHS reference costs shows a total cost of just under £2.7 million in 2007, with median costs ranging from £5266 to £7176.

Analysis using generalised estimating equation regressions indicated that costs were higher for the youngest (0-3) age group over time, that being screened at birth generated lower costs over time. Regression analysis also showed that cost was higher for those patients who were between 0 and 3 years old than those between 4 and 8 years old. This may indicate that more intensive treatment earlier is indicative of healthier patients later on. The models also suggested that, although not the conventional significance level of .90, that having better nutrition (as measured by BMI and weight for height) also contributed to lower costs over time.

Cost effectiveness of newborn screening for CF

Chapter 5 discusses the results of a model built to assess the cost-effectiveness of newborn screening versus clinically diagnosing patients. Compared to no screening, the model estimates that the incremental cost effectiveness of newborn screening is approximately £641 per QALY gained. From the CEAC analysis, the probability that accepting newborn screening at this level has a 50% probability of being the correct decision, rising to 100% at a threshold around £65,000. This indicates that the government was willing to pay at least that amount per QALY gained, as they have implemented the programme nationwide. Subsequent analysis shows that there is significant uncertainty surrounding some of the estimates, particularly around the

utility and cost estimates. However, as the cost of research would likely exceed £10,000, where the added value of finding more information becomes zero, it is unlikely that more research is good value for money, at least to answer the question of whether or not newborn screening for CF is cost-effective.

Methodological considerations

This thesis raises a number of methodological issues for consideration, most notably in the area of the statistical analysis and interpretation of longitudinal panel data, and in the methods used to model the cost-effectiveness of newborn screening.

Methodological limitations

As Chapter 3 points out, there are many considerations and many pitfalls when conducting cost analysis and economic evaluations. This section does not intend to go into as much detail as Chapter 3, but discusses the impact this thesis' limitation may have had on the conclusions drawn in this thesis and how I have attempted to minimize the impact.

Point-of-view considerations

The analyses in this thesis adopted the viewpoint of tertiary NHS services in the East of England. This is, admittedly, a narrow perspective. Paediatric cystic fibrosis patients and their families use the primary care setting for some of their care, as well as incurring out-of-pocket costs and probable productivity losses as a result of the illness. In the Netherlands in 1991, Wildhagen et al (57) indicated that this could be up to 16% of the total cost of care. Wildhagen carried out another questionnaire based study that found the mean non-hospital costs of care for paediatric patients to

be $11,429 annually. This means that the total cost of paediatric CF to the economy could be much larger than the estimates presented here (in both the cost-of-illness analysis and the cost-effectiveness analysis). This narrow perspective was purposefully chosen; the funding for CF is distributed to tertiary NHS facilities, therefore knowing the costs directly related to resource use in these facilities is most relevant to the intended audience of this thesis.

Cost-of-illness considerations

Initially, I had hoped to gather data from across the six counties in the eastern region to enable a "natural experiment" situation as three counties (Cambridgeshire, Norfolk and Suffolk) had been carrying out newborn screening for decades, and the other three (Essex, Hertfordshire and Bedfordshire) had not. However, participation in Essex, Hertfordshire, and Bedfordshire, (approximately 7 hospitals) was non-existent. When contacted after initial ethics approval was granted to get local R&D approval, none of the hospitals responded to email or phone inquiries. I am not sure if they decided not to participate due to having to take patient consent, or due to other administrative issues (e.g. sponsoring me for honorary contracts so I could collect data). If this was the case, a "research passport" situation would have eased the administrative burden, as gaining ethics approval and honorary contract in one hospital would have enabled me to work throughout the region. Unfortunately this system was only put in place in 2009, after data collection had ended. This led to the study having nearly half the numbers of patients to analyse, which can lead to biased estimates, especially in the mean. To minimise the impact of small numbers in the cost-of-illness study, bootstrapped confidence intervals for the mean and interquartile ranges for the median were estimated.

A related possible limitation is the representativeness of the data gathered. Information gathered from a small sample in a relatively rural region may not be nationally representative. However, the paediatric CF population in this region seems to be nationally representative in terms of age and sex as well as in terms of disease severity when compared with the information gathered from Sims et al(3), Dodge et al(4), and the CF Trust (10).

As mentioned in Chapter 3 and 4, there are many methods of analysis available to investigate the relationship between a dependent variable and other covariates. I chose to use a generalised estimating equation models. This method does not allow for investigation of the variance-covariance structure, but instead focuses on the relationship between the covariates and the dependent variable. Given that I was only interested in the impact the variables had on total cost over time, this seemed the most appropriate model choice. The method also allows for individuals to enter and leave the sample at different times, meaning that it did not discard or treat as incomplete, those individuals who did not have all five years of data available. GEE models have been used in healthcare before to analyse longitudinal data sets (Marshall 2011, Asche CV 2010) and the models have had a model fit test designed for them (186) (187). The models in chapter 4 had a very low QIC, indicating that they were a good fit and therefore the results were reliable. It would be ideal to gain access to the information in the national database to increase the number of observations in the data set and compare the results to those found in chapter 4.

Cost-effectiveness analysis considerations

Given the complex nature of cystic fibrosis, any model created to estimate the costs and effects of a technology will probably deviate from reality. However, as mentioned in Chapter 3, there is often a trade-off between model complexity and reality. As with econometric analysis, models rely on robust data. Therefore, it is important to address the uncertainty that surrounds any parameter in the model. By using a probabilistic cohort simulation model, I was able to address both concerns. This type of model uses Monte Carlo simulations to repeatedly estimate cohort costs and outcomes over time, with each variable having a distribution that each simulation pulls from.

In terms of model validation, the cost-effectiveness model was not externally validated by any primary screening versus not screening research studies. However, the model did have face validity, as the estimates produced did make sense. It could also be argued that the model was predictively validated, as some of the model's inputs were taken from the Simpson et al (33) modelling study. The model also had between model validity, as the results were not remarkably different from those of Simpson et al.

Conclusions

The analyses in this thesis have led to three main conclusions: the costs of caring for paediatric CF patients have increased over time; disease severity (as measured by band) and age are significant drivers of those costs; and that newborn screening for CF is a cost-effective intervention when compared with clinical diagnosis on suspicion or appearance of signs.

While there are implications of the small numbers of patients in the sample for the results of the cost-of-illness analysis, the results seem robust when compared with other estimations on the cost of caring for the paediatric CF patients in the UK. (3) The results also are within the range of those found in the literature review, which included patients in the US and Europe. However, as mentioned in Chapter 2 and Chapter 3, comparing results across different health care settings must be done with caution. Although the method of capturing disease severity in this thesis may be unique to the UK (at least in terms of the banding system) the move away from using lung function alone as a disease severity indicator is becoming more common (99). Also, the outcomes of the regression analysis also seem to chime with the literature.

Finally, the results of the cost-effectiveness analysis indicate that the government has made the correct decision in implementing a national newborn screening programme for CF in the UK. While the model indicated that there was uncertainty in most of the parameters, the cost of research to eliminate the uncertainty was prohibitive.

The methods used for the cost-of-illness study are not necessarily new, however the application of GEE models for longitudinal CF analysis does seem to be the first of its type. GEE models are particularly useful for investigating longitudinal data where patients do not have information available at every time point. There is an opportunity to explore the data more fully as the database becomes more populated and the bounded nature of paediatric CF care (0-16) makes the GEE method an attractive option for looking at longitudinal relationships.

Information from the ERCFD in terms of cost, transition parameters, linear relationships, and risk factors were used to populate the model in chapter 5. The cohort model built to look at the cost-effectiveness of newborn screening was robust in terms of allowing for disease variability and the probabilistic analysis of the various parameters being input into the model. This type of model is useful for investigations of interventions that have lifetime cohort cost and utility implications.

Policy implications

The findings of this thesis have implications for existing NHS policies on the care of CF patients. The cost-of-illness analysis has implications for various funding policies, and the cost-effectiveness findings confirm current NHS policy as good value for money.

Cost-of-illness

As mentioned in Chapter 1, funding for the care of paediatric CF patients is still allocated on a case-by-case basis. The cost-of-illness study in this thesis has illustrated that the costs of caring for patients with CF can vary from year to year (and more likely month to month). These results also indicate that young patients accrue more costs due to preventative treatment. These findings echo those of the CF Trust, who have argued that the Payment-by-Results system is not appropriate for CF, given the high costs of sustaining patients' health. I would agree that funding a package of care surrounding a patient's banding is an appropriate method of funding hospital trusts and specialist centres as a patient who is relatively healthy costs considerably less to care for than a patient who is in band 3 or 4. The banding system recognises this distinction, whereas the old system would have charged the

same tariff for both patients, which could have resulted in a shortfall in funding for the more expensive patient that may have had to have been subsidized by the NHS in other ways.

The results of the regression analysis indicate that disease severity is the major driver of costs. While not surprising, the related finding that those patients who are the youngest are accruing greater costs than the next two age groups (4-13 year olds) indicates that the preventative measures taken for younger children may actually reduce costs and (by association) health care use for a period of time. However, it is not clear from this thesis' analysis whether and what clinical interventions are resulting in the reduction in costs in childhood and adolescence. A more in depth and longer term analysis of longitudinal data from the ERCFD and the national database may be able to answer these questions.

It was also unclear from the regression analysis what the relationship is between total cost and being in a shared care setting. While currently the CF Trust advocates for patients to be seen in a specialist centre so patients can access the best quality care, it is not clear from the data in this study that users of district hospitals are more likely to require higher intensity care.

Developments since 2007 (5) have indicated that funding will be allocated via banding tariffs, and that patient information is key to acquiring funding. Hospitals will be allocated funding based on the information about their patients they enter onto PORT CF (a national database). The implications of this policy are interesting in light of the inconclusive data on the impact of shared care on individual total cost. On the face of the policy, hospitals would have an incentive to try to enter more

patients on the higher bands, therefore getting more funding. However, as the PORT CF is also a clinical audit tool, poorly performing hospitals may be asked to refer more of their patients to the specialist centre, and therefore lose a significant amount of funding. If that is the case, specialist centres could become increasingly burdened by large patient numbers and will either need to expand or perhaps set up new specialist centres to deal with the influx of patients. Hopefully, this policy will encourage hospitals and specialist centres to work more closely together (perhaps in more formal clinical networks) to ensure quality and continuity of care across all hospitals.

Newborn screening

The results from Chapter 5 go some way towards confirming that newborn screening for CF is good value for money. The government rolled out nationwide screening in 2007 with the only economic evidence assuming a UK setting available being the paper by Simpson et al (33). I attempted to take their analysis further with a more robust model and used updated epidemiological data and the results of this thesis suggest an even better cost per QALY estimate. The results of the regression analyses in Chapter 4 also support that screening does indicate lower costs over time.

Recommendations for future research

Given the current state of public finances in England, it is likely that there will be a push from the government to streamline services. While this thesis attempted to give an overview of the hospital and drug costs associated with paediatric CF care, there are still many unanswered economic questions. For example: what is the most cost-

effective method of caring for paediatric CF patients? Are the most efficient treatments being used? Are there areas of wastage that could be contained? What is the economic impact of CF on the family? Answering these questions, particularly questions surrounding the cost-effectiveness of treatments and the most cost-effective method of care could help produce a more efficient and effective service for paediatric CF.

As shown in Chapter 2, there are still relatively few economic evaluations or cost studies surrounding many of the therapies used in CF care. The evidence that is available is now relatively old, with new therapies and techniques for delivering therapies (such as nebulised Tobramycin) having emerged in the last decade have not been adequately scrutinised. As the cost of some therapies can be quite high, it would be useful to investigate the cost-effectiveness of therapies such as rhDNAse and Tobramycin. Cost effectiveness methods such as those used in chapter 5 could be appropriate to estimate the effectiveness of the drugs over time in terms of both clinical outcomes and quality of life.

Another priority area where there is a gap in published research highlighted by Chapter 2 is on information relating to the type of care paediatric patients receive. In the UK, care for CF patients in both paediatric and adult populations generally revolves around specialist centres. The CF Trust advocates as much Specialist Centre care as possible, but the evidence suggesting that this is better is based on one paper (126). However, within the Eastern Region and many other areas in the UK, due to the large geographical area that the specialist centres need to cover, paediatric CF care is shared between various NHS centres (in accordance with UK

CF Trust guidelines). In the east, shared care can be generally separated into five different models:

1) All care at the specialist centre (SC);

2) most care at district general hospital (DGH) with annual review done at SC;

3) most care at DGH with SC team travelling to DGH to do annual review;

4) an even split of care between SC and DGH;

5) most care at DGH with SC being used when required (25).

In various places (the East and Southwest of England, for example) there are coordinated clinical networks that care for CF patients where there is only one specialist centre and several district hospitals without specialist teams. Research into the clinical and cost-effectiveness of these networks would benefit Strategic Health Authorities who will no doubt be searching for ways to streamline services. The ERCFD is currently being expanded with patient data back to 1992. This could be a valuable resource for a natural experiment, using regression methods to look at the clinical outcomes and resource use of patients before and after the clinical network began in 1998.

There is little agreement or knowledge on the clinical effects or benefits of different models of care, and no economic evaluation of shared CF care has been published. van Koolwijk et al (127) found that there were little differences in clinical outcome amongst the patients in their shared care study. Assuming our clinical outcomes are the same for the various regimes within the region, it makes sense to compare the

value of different regimes in terms of patient and/or patient's family's values. No study has analysed patient preferences for shared care, or tried to value a service as a whole. The lack of clinical and economic evidence for shared care arrangements indicates a hole in the evidence base for decision makers not only for CF but also for other paediatric chronic diseases. In order to conduct a cost-effectiveness analysis on the different methods of care, a model could be developed using information from the ERCFD, however the small numbers in the database may make it problematic to see the effects. However, the PORT CF database also has information on the type of care patients receive and is probably a better source of information for a study of this type as it would allow for wider geographical analysis (i.e. is shared care more cost-effective in remote areas like the Scottish Highlands than in rural areas such as East Anglia).

While this thesis has examined the cost to the tertiary NHS services, caring for paediatric CF patients also has an impact in the primary care setting, as well as potentially significant out-of-pocket and productivity costs to the family. The impact of paediatric CF from these points of view has not been studied in the UK. Investigating both of these could have major policy implications for how services are arranged. For instance, if it is found that a significant portion of care for CF patients is done in the primary care setting, then perhaps a proportion of the specialist service money available to the tertiary care centres could be transferred to the primary care setting.

Related to both the evaluation of therapies for CF care and to the most effective service delivery method is the lack of information on CF quality of life in paediatric patients. It is only recently (2005) that a validated, disease specific measure been

developed(128) for paediatric CF. To date, this measure has only been used in the USA and on children above 6. Although more research on Quality of Life did not seem like a good investment for the cost-effectiveness of newborn screening in this study, an analysis of paediatric patients (and parental proxies) quality of life with CF could drastically improve economic evaluations of both treatments and service delivery.

Conclusion

This thesis has shown that the cost to the NHS of caring for paediatric patients with CF has increased slightly over time, but that the makeup of these costs has changed alongside clinical practice. Results from the econometric analysis seem to indicate that aggressive treatment approaches to keep disease severity low is an appropriate clinical tactic. The results also indicate that patients who were screened under the newborn screening programme used fewer NHS resources. This finding was supported by the cost-effectiveness analysis of newborn screening for CF versus clinical diagnosis, which indicated that newborn screening was a cost-effective intervention.

As with any research, there are always more questions than answers. Further research on methods of service delivery, the costs to families and/or primary care, and information on the cost-effectiveness of new therapeutics is advocated. Also welcome would be more in depth information on quality of life in CF patients, as well as information on the cost of caring for adult CF patients.

Appendix A: Medical Glossary *(193)*

Antibiotic: a group of drugs used to treat infections caused by bacteria.

Anti-inflammatory agent: these drugs act by inhibiting the formation of prostaglandins, which are mediators of inflammation. They act as analgesics to relieve pain and as inhibitors of inflammation

Autosomal recessive: genetic trait that is linked to one of the 44 non-sex linked chromosomes. For a trait to be autosomal recessive, those affected must have the faulty gene on both copies of the chromosome pair.

Bacteria: simple, single celled organisms which can be both beneficial and harmful.

Biofilm: a large body of microorganisms that adheres to tissues

Body Mass Index (BMI): A method used to define degrees of over and underweight. Calculated by dividing a person's weight by the square of their height.

Bolus: a rapid intravenous injection of fluid or a drug.

Bronchoscopy: the use of a bronchoscope to see the interior of the bronchial tubes within the lungs.

Burkholderia Cepacia Complex: common bacteria that infects animals

Carrier: a person who does not necessarily express a disease or genetic trait, but can pass on infection or genetic disorders.

Chromosome: rod shaped bodies found in the nucleus of cells that contain genes which establish the traits of an organism

Chronic disorder: a persistent or recurring condition or group of symptoms.

Cirrhosis: a condition where normal tissue is replaced with fibrous tissue

Cystic Fibrosis Transmembrane Regulator (CFTR): chloride channel that aids in the regulation of salt and water movement through the cell membranes. Cystic fibrosis patients have gene mutations which encodes for this channel, resulting in thickened secretions in mucus membranes

Enzymes: a protein that acts as a catalyst for the body's metabolic processes.

Exacerbation: a clinical event that indicates a worsening of a patient's condition

Failure to thrive: a patient does not gain weight or height normally

Gastrostomy: an operation on the stomach by which, when the oesophagus is blocked (i.e. by a tumour) an opening is made from the front of the abdomen into the stomach, so that fluid food can be passed into the organ.

Heterozygous: an individual having dissimilar members of the pair of genes coding for a given characteristic

Homozygous: an individual having similar members of a pair of genes coding for a given characteristic

Inflammation: the reaction of the tissues to any injury, which may be the result of trauma, infection, or chemicals.

Intravenous (IV): an intravenous injection is one that is given into a vein

Jejunostomy: an artificial opening in to the intestine to allow for a feeding tube.

Malabsorption: condition characterised by faulty absorption from the intestine of foodstuffs such as fat, vitamins, and mineral salts.

Mucolytic: the term used to describe the property of destroying or lessening the strength of mucus.

Mucus membrane: The general name given to the membrane which lines many of the hollow organs of the body and are generally lubricated by mucus.

Nasal gastric feeding: a plastic or rubber tube is passed into the stomach through the nose, pharynx and then the oesophagus used to pass food and/or drugs into the stomach.

Oral Glucose Tolerance Test (OGTT): medical test in which glucose is given and blood tests are run afterwards

Pancreatic insufficiency: condition in which the pancreas does not produce sufficient pancreatic enzymes which aid in the absorption of nutrients

Pseudomonas aeruginosa: common bacteria that causes disease and damage to tissues in animals

Pulmonary: relating to the lungs

Pulmonary function tests: tests to assess how the lungs are functioning. They can range from simple spirometry (measuring breathing capacity) to sophisticated physiological assessments including Vital Capacity (VC) which tests the maximum volume of air that can be expelled slowly and completely after a maximum deep breath; forced vital capacity (FVC) is similar technique using forced maximum

exhalation; and functional residual capacity (FRC) is the volume of air in the lungs at the end of a normal expiration. A dynamic technique, the forced expiratory volume (FEV1) test is the amount of air forcefully exhaled during the first second after a full breath.

Screening: testing a population of apparently healthy people to identify those who may have a treatable disease.

Spirometer: a device used to test how the lung is working (see pulmonary function tests).

Sputum: secretions expectorated by coughing

Sputum culture: testing the sputum for organisms such as bacteria and fungi

Staphylococcus aureus (Staph)

Surfactant: a surface active agent lining the alveoli of the lungs which does not let alveoli collapse at the end of respiration

Sweat test: measures the concentration of chloride that is excreted in sweat. It is commonly used to test for Cystic Fibrosis.

Tolerability: the response to a particular amount of a drug or physiological messenger

Toxicity: how toxic a drug is to a patient

Virus: a group of infective agents which are more difficult to treat than bacteria.

Appendix B: Literature searches

Pubmed:

PubMed search 1: Cost$ AND cystic fibrosis

("economics"[Subheading] OR "economics"[All Fields] OR "cost"[All Fields] OR "costs and cost analysis"[MeSH Terms] OR ("costs"[All Fields] AND "cost"[All Fields] AND "analysis"[All Fields]) OR "costs and cost analysis"[All Fields]) AND ("cystic fibrosis"[MeSH Terms] OR ("cystic"[All Fields] AND "fibrosis"[All Fields]) OR "cystic fibrosis"[All Fields])

March 2010 result: 583 references

PubMed search 2: Search (#1) AND (Economic evaluation OR cost benefit OR cost effectiveness OR cost utility)

AND ((("economics"[MeSH Terms] OR "economics"[All Fields] OR "economic"[All Fields]) AND ("evaluation studies"[Publication Type] OR "evaluation studies as topic"[MeSH Terms] OR "evaluation"[All Fields])) OR ("cost-benefit analysis"[MeSH Terms] OR ("cost-benefit"[All Fields] AND "analysis"[All Fields]) OR "cost-benefit analysis"[All Fields] OR ("cost"[All Fields] AND "benefit"[All Fields]) OR "cost benefit"[All Fields]) OR ("cost-benefit analysis"[MeSH Terms] OR ("cost-benefit"[All Fields] AND "analysis"[All Fields]) OR "cost-benefit analysis"[All Fields] OR ("cost"[All Fields] AND "effectiveness"[All Fields]) OR "cost effectiveness"[All Fields]) OR (("economics"[Subheading] OR "economics"[All Fields] OR "cost"[All Fields] OR "costs and cost analysis"[MeSH Terms] OR ("costs"[All Fields] AND "cost"[All Fields] AND "analysis"[All Fields]) OR "costs and cost analysis"[All Fields]) AND utility[All Fields]))

March 2010 result: 226 references

Cochrane Library/NHS Economic Evaluation Database:
Search: Cystic Fibrosis AND Pediatrics
March 2010 result: 29 results

Embase:
Search 1: cystic fibrosis AND (pediatric OR paediatric) {Including Related Terms}

Search 2: limit 1 to (abstracts and human and english language and yr="1990 -Current")
Search 3: 2 and cost.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
March 2010 result: 258

Appendix C: Literature Review Pro-forma (adapted from (38, 39))

First Author

Type of Evaluation

Price Year

Intervention

Population

Location (setting)

Currency

Outcome measure

Types of costs gathered

Source of effectiveness data

Source of cost data

Modelling

Time horizon

Results

Notes

Appendix D: Data collection sheet for the ERCFD

THE EASTERN REGION CYSTIC FIBROSIS SURVEY PATIENT DATA

Demographics

Patient Name
Patient Address
Sex
Date of Birth
Shared Care
- Shared care hospital
- Consultant

Diagnosis

Date of diagnosis
Age at diagnosis
Presenting features
- IRT Screening
- Genetic Screening
- Affected Sibling (Name)
- Meconium Ileus
- Diarrhoea/Steatorrhoea
- Failure to thrive
- Respiratory
- Prolonged Jaundice
- Prolapsed Rectum
- Other

Annual review
- Date
- Does not have one

Care discontinued (date)
- Died
- Not CF
- Moved
- Transfer to Adult Clinic

Clinics/Admissions

Clinics attended
- Number
- Location

Hospital Admissions and length of stay
- Admission
  - Bed days
  - Reason
  - Location

Health Outcomes

Complications
- Diabetes

- Liver disease
- Oesophageal Varices
- Other

Naso Gastro Feeding
Gastrostomy
Porta Cath or Equivalent
Home Oxygen
Height
Weight
Lung Function Tests
- FEV1 (%Pred)
- FVC (%Pred)

Bacteriology
- *Staphylococcus Aureus*
- *Haemophilus influenzae*
- *Pseudomonas aeruginosa*
- *Burkholderia cepacia*
- *Aspergillus fumigates*

Drugs taken
- Drug (all drugs listed on next page)

Dosage
- Length of time

Acetylcysteine
Acyclovir
Acyclovir IV
AeroBec
Aerolin
Amikacin
Amikacin IV
Amikin
Amoxil
Amoxycillin
Amphotericin
Amphotericin IV
Ampicillin
Ampicillin IV
Aspirin
Atrovent
Augmentin
Augmentin IV
Augmentin Duo
Azetreonam
Azetreonam IV
Aziathioprine
Azithromycin
Azithromycin IV
Azlocillin
Azlocillin IV
Beclofort
Beclomethasone
Becodiscs
Beconase
Becotide
Benzylpenicillin
Betnasol
Bricanyl
Budesonide
Calcium
Carbamazepine
Cefaclor
Cefadroxil
Cefixime
Ceflucoximine
Cefotaxime
Cefotaxime IV
Ceftazidime
Ceftazidime IV
Ceftriaxone
Ceftriaxone IV
Cefuroxime
Cephradine
Cerumol
Chloramphenicol
Chlorpeniramine
Cidomycin
Cimetidine
Ciprofloxacin
Ciprofloxacin IV
Ciproxin
Cisapride
Clarithomycin
Clarithomycin IV
Clarityn
Clindamycin
Co-amoxiclav
Codydramol
Co-fluampicil
Colistin
Colistin IV
Co-trimoxazole
Colomycin
Colomycin IV
Creon 10000
Creon 25000
Cyclosporin
Daktarin
Desmopression
Destolit
Dioctyl
Distaclor
DNAse
Domperidone
Dopamine
Efexor
Ephedrine
Erythromycin
Fabrol
Fersamal
Flagyl
Flixonase
Flixotide
Flucloxacillin
Flucloxacillin IV
Fluconazole
Fluoride
Fluticasone
Folic Acid
Frusemide
Frusemide IV
Fucidin
Ganciclovir
Gastrografin
Gaviscon
Gentomycin
Gentomycin IV
Humulin
Hydrocortisone
Hydrocortisone IV
Hypostop
Iboprofen
Imipenem IV
Immodium
Insulin
Intal
Ipratroprium Bromide
Iron
Itraconazole
Ketotifen
Klaricid
Lactulose
Lamotrigine
Lansoprazole
Loperamide
Loratadine
Losec
Magnesium Carbonate
Maxolan
Meropenem
Meropenem IV
Metanium
Methylphenidate
Metronidazole
Metronidazole IV
Movicol
Naprosyn
Naseptin
Neoral
Nifedipine
Nizatidine
NSAID
Nystatin
Oestrogel
Ofloxacin
Omeprazole
Ondansetron
Otrivine
Oxis
Oxygen (at home)
Paracetemol
Parvolex
Penicillin
Penicillin IV
Periactin
Phenytoin
Phyllocontin
Piperacillin
Piperacillin IV
Piptazobactum IV
Piriton
Pizotifen
Plesmet
Potassium
Prednisolone
Prepulsid
Prozac
Pulmicort
Ranitidine
Rifamipicin
Salbutamol
Salmeterol
Senokot
Septrin
Serevent
Singulair
Sodium
Slow Sodium
Sodium
Valporate
Steroids
Sudacrem
Sytron
Tagamet
Taurine
Teicoplanin
Teicoplanin IV
Terbutaline
Theophyline
Thyroxine
Tobramycin
Tobramycin IV
Tocoplurol
Trimethoprim
Triple Saline
Unichem
Ursodeoxychloric acid
Vancomycin
Ventolin

Appendix E: Excel Workbook and Macros for cost-effectiveness model

Double Click Icon to open workbook:

Macros for probabilistic analysis, CEAC, EVPI, EVPPI

```
Sub Probabilistic()
'
' Probabilistic Macro
' Macro recorded 12/08/2008 by James Jarrett
'

Application.ScreenUpdating = False
Sheets("Parameters").Select
Range("D3").Select
ActiveCell.FormulaR1C1 = "1"
Application.DisplayStatusBar = True
Sheets("Simulation").Select
Dim Index
Dim Trials
Index = 0
Trials = 1000

Do

    Range("C4:AA4").Select
    Selection.Copy
    Range("C6:AA6").Select
    ActiveCell.Offset(Index, 0).Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Index = Index + 1
    Application.StatusBar = "Simulation " & Index & " of 1000 trials"

Loop While Index < Trials

Application.DisplayStatusBar = False
```

```
Sheets("Parameters").Select
Range("D3").Select
ActiveCell.FormulaR1C1 = "0"
Sheets("Simulation").Select
Range("A1").Select
Application.ScreenUpdating = True


End Sub

Sub CEACurve()
'
' CEACurve Macro
' Macro recorded 17/11/2008 by James
'

Application.DisplayStatusBar = True
Sheets("Simulation").Select
Dim Index
Dim Trials
Index = 0
Trials = 58

Do

    Range("AK6").Select
    ActiveCell.Offset(Index, 0).Range("A1").Select
    Selection.Copy
    Range("AE1").Select
    ActiveSheet.Paste
    Range("AG4:AI4").Select
    Application.CutCopyMode = False
    Selection.Copy
    Range("AL6").Select
    ActiveCell.Offset(Index, 0).Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
        Index = Index + 1
    Application.StatusBar = "Calculation " & Index & " of " & Trials

Loop While Index < Trials

Application.DisplayStatusBar = False
Sheets("Simulation").Select
Range("AK1").Select


End Sub
```

```
Sub EVPI()
'
' EVPI Macro
' Macro programmed September 2008 by JJ
'

'
Application.DisplayStatusBar = True
Sheets("Simulation").Select
Index = 0
Trials = 60

Do

    Range("AU6").Select
    ActiveCell.Offset(Index, 0).Range("A1").Select
    Selection.Copy
    Range("AE1").Select
    ActiveSheet.Paste
    Range("AU1").Select
    Application.CutCopyMode = False
    Selection.Copy
    Range("AV6").Select
    ActiveCell.Offset(Index, 0).Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
        Index = Index + 1
    Application.StatusBar = "Calculation " & Index & " of " & Trials

Loop While Index < Trials

Application.DisplayStatusBar = False
Sheets("Simulation").Select
Range("AK1").Select

End Sub


Sub Survivalpartial()
'
' Survivalpartial Macro
'

' Set number of outer loops

    oloops = 100
```

```
' Backup existing formula

    Sheets("Parameters").Select
    Range("B29:B33").Select
    Selection.Copy
    Range("N29:N33").Select
    ActiveSheet.Paste

' Set up outer loop and then record one realisation from distribution

    Index2 = 0

    Do

    Application.ScreenUpdating = False
    Application.DisplayStatusBar = True

    Sheets("Parameters").Select
    Range("C29:C33").Select
    Application.CutCopyMode = False
    Selection.Copy
    Range("B29:B33").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False

' Now run standard simulation (same as MCsimulation macro - but total simulations
reduced to 100)

    Sheets("Parameters").Select
    Range("D3").Select
    ActiveCell.FormulaR1C1 = "1"
    Sheets("Simulation").Select
    Index = 0
    Trials = 100

    Do

        Range("C4:AA4").Select
        Selection.Copy
        Range("C6:AA6").Select
        ActiveCell.Offset(Index, 0).Range("A1").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False

        Index = Index + 1
        Application.StatusBar = "Realisation " & Index2 & " of 100 for parameter
Revision Cost and simulation " & Index & " of 100"
```

```
    Loop While Index < Trials

        Application.DisplayStatusBar = False
        Sheets("Parameters").Select
        Range("D3").Select
        ActiveCell.FormulaR1C1 = "0"
        Sheets("Simulation").Select
        Range("A1").Select

' And record the resulting net benefit given that realisation of the parameter of
interest

        Sheets("Simulation").Select
        Range("AC4:AE4").Select
        Selection.Copy
        Range("Ax6:az6").Select
        ActiveCell.Offset(Index2, 0).Range("A1").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False

' Close outer loop

        Index2 = Index2 + 1

    Loop While Index2 < oloops

' Record final partial EVPI (average over the outer loop realisations)

    Sheets("Simulation").Select
    Range("BC7").Select
    Selection.Copy
    Range("BC13").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False


' Now restore the original rrNP1 formula

    Sheets("Parameters").Select
    Range("N29:N33").Select
    Selection.Copy
    Range("B29:B33").Select
    ActiveSheet.Paste
    Range("N29:N33").Select
    Selection.ClearContents
    Range("D1").Select
    Selection.Copy
```

```
    Range("C1").Select
    Selection.PasteSpecial Paste:=xlFormats, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Application.CutCopyMode = False
    Range("C3").Select

Application.ScreenUpdating = True
Application.DisplayStatusBar = False

End Sub

Sub infpartial()
'
' infpartial Macro
'
Dim Index As Integer
Dim Trials As Integer
Dim oloops As Integer
Dim Index2 As Integer


' Set number of outer loops

    oloops = 100


' Backup existing formula in Utility cells

    Sheets("Parameters").Select
    Range("B34").Select
    Selection.Copy
    Range("D8").Select
    ActiveSheet.Paste

' Set up outer loop and then record one realisation from distribution

    Index2 = 0

    Do

    Application.ScreenUpdating = False
    Application.DisplayStatusBar = True

    Sheets("Parameters").Select
    Range("C34").Select
    Application.CutCopyMode = False
    Selection.Copy
    Range("B34").Select
```

```
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False

' Now run standard simulation (same as MCsimulation macro - but total simulations
reduced to 100)

    Sheets("Parameters").Select
    Range("D3").Select
    ActiveCell.FormulaR1C1 = "1"
    Sheets("Simulation").Select
    Index = 0
    Trials = 100

    Do

        Range("C4:AA4").Select
        Selection.Copy
        Range("C6:AA6").Select
        ActiveCell.Offset(Index, 0).Range("A1").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False

        Index = Index + 1
        Application.StatusBar = "Realisation " & Index2 & " of 100 for Utility parameters
and simulation " & Index & " of 100"

    Loop While Index < Trials

        Application.DisplayStatusBar = False
        Sheets("Parameters").Select
        Range("D3").Select
        ActiveCell.FormulaR1C1 = "0"
        Sheets("Simulation").Select
        Range("A1").Select

' And record the resulting net benefit given that realisation of the parameter of
interest

        Sheets("Simulation").Select
        Range("AC4:AE4").Select
        Selection.Copy
        Range("AX6:aZ6").Select
        ActiveCell.Offset(Index2, 0).Range("A1").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False

' Close outer loop
```

```
        Index2 = Index2 + 1

    Loop While Index2 < oloops

' Record final partial EVPI (average over the outer loop realisations)

    Sheets("Simulation").Select
    Range("BC7").Select
    Selection.Copy
    Range("BC11").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False


' Now restore the original Utility formulae

    Sheets("Parameters").Select
    Range("D8").Select
    Selection.Copy
    Range("B34").Select
    ActiveSheet.Paste
    Range("D8").Select
    Selection.ClearContents
    Range("D1").Select
    Selection.Copy
    Range("D8").Select
    Selection.PasteSpecial Paste:=xlFormats, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Application.CutCopyMode = False
    Range("C3").Select

Application.ScreenUpdating = True
Application.DisplayStatusBar = False


End Sub

Sub Costpartial()
'
' Costpartial Macro
'
' Set number of outer loops

    oloops = 100

' Backup existing formula

    Sheets("Parameters").Select
```

```
    Range("B39:B42").Select
    Selection.Copy
    Range("N39:N42").Select
    ActiveSheet.Paste

' Set up outer loop and then record one realisation from distribution

    Index2 = 0

    Do

    Application.ScreenUpdating = False
    Application.DisplayStatusBar = True

    Sheets("Parameters").Select
    Range("C39:C42").Select
    Application.CutCopyMode = False
    Selection.Copy
    Range("B39:B42").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False

' Now run standard simulation (same as MCsimulation macro - but total simulations
reduced to 100)

    Sheets("Parameters").Select
    Range("D3").Select
    ActiveCell.FormulaR1C1 = "1"
    Sheets("Simulation").Select
    Index = 0
    Trials = 100

    Do

        Range("C4:AA4").Select
        Selection.Copy
        Range("C6:AA6").Select
        ActiveCell.Offset(Index, 0).Range("A1").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False

        Index = Index + 1
        Application.StatusBar = "Realisation " & Index2 & " of 100 for parameter
Revision Cost and simulation " & Index & " of 100"

    Loop While Index < Trials

        Application.DisplayStatusBar = False
```

```
        Sheets("Parameters").Select
        Range("D3").Select
        ActiveCell.FormulaR1C1 = "0"
        Sheets("Simulation").Select
        Range("A1").Select

' And record the resulting net benefit given that realisation of the parameter of
interest

        Sheets("Simulation").Select
        Range("AC4:AE4").Select
        Selection.Copy
        Range("Ax6:az6").Select
        ActiveCell.Offset(Index2, 0).Range("A1").Select
        Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
            False, Transpose:=False

' Close outer loop

        Index2 = Index2 + 1

    Loop While Index2 < oloops

' Record final partial EVPI (average over the outer loop realisations)

    Sheets("Simulation").Select
    Range("BC7").Select
    Selection.Copy
    Range("BC14").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False


' Now restore the original rrNP1 formula

    Sheets("Parameters").Select
    Range("N39:N42").Select
    Selection.Copy
    Range("B39:B42").Select
    ActiveSheet.Paste
    Range("N39:N342").Select
    Selection.ClearContents
    Range("D1").Select
    Selection.Copy
    Range("C1").Select
    Selection.PasteSpecial Paste:=xlFormats, Operation:=xlNone, SkipBlanks:= _
        False, Transpose:=False
    Application.CutCopyMode = False
```

```
    Range("C3").Select

Application.ScreenUpdating = True
Application.DisplayStatusBar = False
'
End Sub
Sub Partials()
'
' Partials Macro
'
' Select simulation sheet and enter 4000 as ceiling ratio

    Application.ScreenUpdating = False
    Application.DisplayStatusBar = True

    Sheets("Simulation").Select
    Range("AE1").Select
    Application.CutCopyMode = False
    ActiveCell.FormulaR1C1 = "4000"

    Range("C6:AA1005").Select
    Selection.ClearContents
    Range("Ax6:AY1005").Select
    Selection.ClearContents

' Run the separate partial macros

    infpartial
    Survivalpartial
    Costpartial
    Transitionpartial
    Utilpartial
```

Reference List

1. Drummond M, Sculpher M, G T, O'Brien B, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes. Oxford, UK: Oxford University Press; 2005.
2. Cystic Fibrosis Trust. The Clinical Care of Children and Adults with Cystic Fibrosis Bandings and Associated Costings. 2007.
3. UK Newborn Screening Committee. A Laboratory Guide to Newborn Screening in the UK for Cystic Fibrosis. London2009.
4. Cystic Fibrosis Trust. Standards of Clinical Care for Children and Adults with Cystic Fibrosis in the UK. London2001.
5. Dodge J A, Lewis P A, Stanton M, Wilsher J. Cystic fibrosis mortality and survival in the UK: 1947-2003. Eur Respir J. 2007;29(3):522-6.
6. Cystic Fibrosis Trust. The UK CF Database http://www.cystic-fibrosis.org.uk/background.htm

[webpage] April 2010 [26/01/2011].
7. Cystic Fibrosis Trust. CF Registry http://www.cftrust.org.uk/aboutcf/publications/cfregistryreports/. [26/01/2011].
8. Rosenstein B J, Cutting G R, Cystic Fibrosis Foundation Consensus Panel. The diagnosis of cystic fibrosis: a consensus statement. Journal of Pediatrics. 1998;132:589-95.
9. McCormick J, Green M W, Mehta G, Culross F, Mehta A. Demographics of the UK cystic fibrosis population: implcations for neonatal screening. European Journal of Human Genetics. 2002;10:583-90.
10. Cystic Fibrosis Trust. UK CF Registry Annual Data Report 2007. Bromley, Kent2007.
11. Department of H. Specialised Services National Definitions Set. London: Department of Health2009.
12. Chalkley M MD. Choice of contracts in the British National Health Service: An empirical study. Journal of Health Economics. 2008;27(5):1155-67.
13. Barnes R. Commissioning of cystic fibrosis services in England. J R Soc Med. 2006;99 Suppl 46:36-45.
14. Robson M, Abbott J, Webb K, Dodd M, Walsworth-Bell J. A cost description of an adult cystic fibrosis unit and cost analysis of different categories of patients. Thorax. 1992;47(9):684-89.
15. Corey M ELLHKM. Longitudinal analysis of pulmonary function decline in patients with cystic fibrosis. J Pediatr. 1997;131(6):809-14.
16. Corey M MFJWMLH. A comparison of survival, growth, and pulmonary function in patients with cystic fibrosis in Boston and Toronto. J Clin Epidemiol. 1988;41(6):583-91.
17. Cystic Fibrosis Trust. Nutritional Management for Cystic Fibrosis. Consensus Document. London2002.
18. Kerem E, Corey M, Gold R, Levison H. Pulmonary function and clinical course in patients with cystic fibrosis after pulmonary colonisation with Pseudomonas aeruginosa. J Pediatr. 1990;116:714-9.

19. Pamukcu A, Bush A, Buchdal R. Effects of Pseudomonas aeruginosa colonisation on lung function and anthropometric variables in children with cystic fibrosis. Pediatric Pulmonology. 1995;19:10-5.
20. Littlewood JM, Miller M G, Ghoneim A G, H RC. Nebulised colomycin for early pseudomonas colonisation in cystic fibrosis. Lancet. 1985(i):865.
21. Valerius NH, Koch C, Hoiby N. Prevention of chronic Pseudomonas aeruginosa colonisation in cystic fibrosis by early treatment. Lancet. 1991;338:725-6.
22. Frederiksen B, Koch C, Hoiby N. Antibiotic treatment of initial colonisation with Pseudomonas aeruginosa postpones chronic infection and prevents deterioration of pulmonar function in cystic fibrosis. Pediatr Pulmonol. 1997;23:330-5.
23. Langton Hewer SC, Smyth AR. Antibiotic strategies for eradicating Pseudomonas aeruginosa in people with cystic fibrosis. Langton Hewer Simon C, Smyth Alan R Antibiotic strategies for eradicating Pseudomonas aeruginosa in people with cystic fibrosis Cochrane Database of Systematic Reviews: Reviews 2009 Issue 4 John Wiley & Sons, Ltd Chichester, UK DOI: 101002/14651858CD004197pub3. 2009(4).
24. Ratjen F, Doring G, Nikolaizik W H. Effect of inhaled tobramycin on early Pseudomonas aeruginosa colonisation in patients with cystic fibrosis. Lancet. 2001;358:983-4.
25. Abbott J, Dodd M, Bilton D, Webb A K. Treatment compliance in adults with cystic fibrosis. Thorax. 1994;49:115-20.
26. van der Schans C, Prasad A, Main E. Chest physiotherapy compared to no chest physiotherapy for cystic fibrosis (Cochrane Review). Oxford: Update Software: The Cochrane Library; 2001.
27. Rogers DF. Mucoactive agents for airway mucus hypersecretory diseases. Respir Care. 2007;52(9):1176-93.
28. Cystic Fibrosis Trust Diabetes Working Group. Management of cystic fibrosis related diabetes mellitus. London2004.
29. Sims EJ, Mugford M, Clark A, Aitken D, McCormick J, Mehta G, et al. Economic implications of newborn screening for cystic fibrosis: a cost of illness retrospective cohort study. Lancet. 2007;369(9568):1187-95.
30. Segel J. Cost of Illness - A Primer: RTI International RTI-UNC Center of Excellence in Health Promotion Economics2006.
31. Cooper NJ, Sutton A, Ades A, Paisley S, Jones D, Working Group on the Use of Evidence in Decision Models. Use of evidence in economic decision models: practical issues and methodological challenges. Health Economics. 2007;16(12):1277-86.
32. Jefferson T, Demicheli V, Mugford M. Elementary Economic Evaluation in Health Care. London: BMJ Books; 2000.
33. Bowling A. Research Methods in Health. UK: Open University Press; 2003.
34. Shiell A, Gerard K, Donaldson C. Cost of illness studies an aid to decision-making. Health Policy. 1987;8:317-23.
35. Maetzel A. Cost of illness and the burden of disease. Journal of Rheumatology. 1997;24:3-5.
36. Byford S, Torgerson D J, Raftery J. Cost of illness studies. BMJ. 2000 13 May;320:1335.

37. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. Oxford, UK: Oxford University Press; 2006.
38. Centre for Reviews and Dissemination. Systematic Reviews: CRD's guidence for undertaking reviews in health care. University of York; 2009 [cited 2007 6/01/2007]; Available from: http://www.york.ac.uk/inst/crd/index_guidance.htm.
39. Drummond M F, Buxton M, et al. Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ. 1996;313:275-83.
40. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions. 2006.
41. OECD. OECD Purchasing Power Parities Department. Available from: www.oecd.org/std/ppp.
42. Balinsky W, Zhu CW. Pediatric cystic fibrosis: evaluating costs and genetic testing. J Pediatr Health Care. 2004;18(1):30-4.
43. Brice P, Jarrett J, Mugford M. Genetic screening for cystic fibrosis: an overview of the science and the economics. J Cyst Fibros. 2007;6(4):255-61.
44. Christopher F, Chase D, Stein K, Milne R. rhDNase therapy for the treatment of cystic fibrosis patients with mild to moderate lung disease. J Clin Pharm Ther. 1999;24(6):415-26.
45. Conway SP. Recombinant human DNase (rhDNase) in cystic fibrosis: is it cost effective? Arch Dis Child. 1997;77(1):1-3.
46. Goa KL, Lamb H. Dornase alfa. A review of pharmacoeconomic and quality-of-life aspects of its use in cystic fibrosis. Pharmacoeconomics. 1997;12(3):409-22.
47. Krauth C, Jalilvand N, Welte T, Busse R. Cystic fibrosis: cost of illness and considerations for the economic evaluation of potential therapies. Pharmacoeconomics. 2003;21(14):1001-24.
48. Murray J, Cuckle H, Taylor G, Littlewood J, Hewison J. Screening for cystic fibrosis. Health Technol Assess. 1999;3(8):i-iv, 1-104.
49. Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV, et al. Neonatal screening for inborn errors of metabolism: cost, yield and outcome. Health Technol Assess. 1997;1(7):i-iv, 1-202.
50. Radhakrishnan M, van Gool K, Hall J, Delatycki M, Massie J. Economic evaluation of cystic fibrosis screening: A review of the literature. Health Policy. 2008;85(2):133-47.
51. Rogowski W. Genetic screening by DNA technology: a systematic review of health economic evidence. Int J Technol Assess Health Care. 2006;22(3):327-37.
52. Tullis DE, Guyatt GH. Quality of life in cystic fibrosis. Pharmacoeconomics. 1995;8(1):23-33.
53. Young SS, Kharrazi M, Pearl M, Cunningham G. Cystic fibrosis screening in newborns: results from existing programs. Curr Opin Pulm Med. 2001;7(6):427-33.
54. Baumann U, Stocklossa C, Greiner W, von der Schulenburg JM, von der Hardt H. Cost of care and clinical condition in paediatric cystic fibrosis patients. J Cyst Fibros. 2003;2(2):84-90.
55. Farrell P, Joffe S, Foley L, Canny GJ, Mayne P, Rosenberg M. Diagnosis of cystic fibrosis in the Republic of Ireland: epidemiology and costs. Ir Med J. 2007;100(8):557-60.
56. Ireys HT, Anderson GF, Shaffer TJ, Neff JM. Expenditures for care of children with chronic illnesses enrolled in the Washington State Medicaid program, fiscal year 1993. Pediatrics. 1997;100(2 Pt 1):197-204.

57. Wildhagen MF, Verheij JB, Verzijl JG, Gerritsen J, Bakker W, Hilderink HB, et al. The nonhospital costs of care of patients with CF in The Netherlands: results of a questionnaire. Eur Respir J. 1996;9(11):2215-9.
58. Wildhagen MF, Verheij JB, Verzijl JG, Hilderink HB, Kooij L, Tijmstra T, et al. Cost of care of patients with cystic fibrosis in The Netherlands in 1990-1. Thorax. 1996;51(3):298-301.
59. Gregg RG, Wilfond BS, Farrell PM, Laxova A, Hassemer D, Mischler EH. Application of DNA analysis in a population-screening program for neonatal diagnosis of cystic fibrosis (CF): comparison of screening protocols. Am J Hum Genet. 1993;52(3):616-26.
60. Lee DS, Rosenberg MA, Peterson A, Makholm L, Hoffman G, Laessig RH, et al. Analysis of the costs of diagnosing cystic fibrosis with a newborn screening program. J Pediatr. 2003;142(6):617-23.
61. Rosenberg MA, Farrell PM. Assessing the cost of cystic fibrosis diagnosis and treatment. J Pediatr. 2005;147(3 Suppl):S101-5.
62. Grieve R, Thompson S, Normand C, Suri R, Bush A, Wallis C. A cost-effectiveness analysis of rhDNase in children with cystic fibrosis. Int J Technol Assess Health Care. 2003;19(1):71-9.
63. Ho SL, Coates AL. Effect of dead volume on the efficiency and the cost to deliver medications in cystic fibrosis with four disposable nebulizers. Can Respir J. 1999;6(3):253-60.
64. Iles R, Legh-Smith J, Drummond M, Prevost A, Vowler S. Economic evaluation of Tobramycin nebuliser solution in cystic fibrosis. J Cyst Fibros. 2003;2(3):120-8.
65. Kretz SE, Pantos BS. Cost savings and clinical improvement through disease management. J Case Manag. 1996;5(4):173-81.
66. LeLorier J, Perreault S, Birnbaum H, Greenberg P, Sheehy O. Savings in direct medical costs from the use of tobramycin solution for inhalation in patients with cystic fibrosis. Clin Ther. 2000;22(1):140-51.
67. Menzin J, Oster G, Davies L, Drummond MF, Greiner W, Lucioni C, et al. A multinational economic evaluation of rhDNase in the treatment of cystic fibrosis. Int J Technol Assess Health Care. 1996;12(1):52-61.
68. Oster G, Huse DM, Lacey MJ, Regan MM, Fuchs HJ. Effects of recombinant human DNase therapy on healthcare use and costs in patients with cystic fibrosis. Ann Pharmacother. 1995;29(5):459-64.
69. Perras C, Otten N. Pulmozyme: Clinical and Economic Impacts. Canada: Canadian Coordinating Office for Health Technology Assessment (CCOHTA)1996.
70. Simpson N, Anderson R, Sassi F, Pitman A, Lewis P, Tu K, et al. The cost-effectiveness of neonatal screening for cystic fibrosis: an analysis of alternative scenarios using a decision model. Cost Eff Resour Alloc. 2005;3:8.
71. Suri R, Grieve R, Normand C, Metcalfe C, Thompson S, Wallis C, et al. Effects of hypertonic saline, alternate day and daily rhDNase on healthcare use, costs and outcomes in children with cystic fibrosis. Thorax. 2002;57(10):841-6.
72. Suri R, Metcalfe C, Lees B, Grieve R, Flather M, Normand C, et al. Comparison of hypertonic saline and alternate-day or daily recombinant human deoxyribonuclease in children with cystic fibrosis: a randomised trial. Lancet. 2001;358(9290):1316-21.

73. Suri R, Wallis C, Bush A, Thompson S, Normand C, Flather M, et al. A comparative study of hypertonic saline, daily and alternate-day rhDNase in children with cystic fibrosis. Health Technol Assess. 2002;6(34):iii, 1-60.
74. van den Akker-van Marle ME, Dankert HM, Verkerk PH, Dankert-Roelse JE. Cost-effectiveness of 4 neonatal screening strategies for cystic fibrosis. Pediatrics. 2006;118(3):896-905.
75. Antioch K, Walsh M. Funding issues for Victorian hospitals: the risk-adjusted vision beyond casemix funding. Aust Health Rev. 2000;23(3):145-53.
76. Antioch KM, Walsh MK. Risk-adjusted capitation funding models for chronic disease in Australia: alternatives to casemix funding. Eur J Health Econ. 2002;3(2):83-93.
77. Beech R, Bekker H. Planning the development of cystic fibrosis gene carrier screening. J Health Serv Res Policy. 1996;1(2):81-92.
78. Beech R, Rona RJ, Mandalia S. The resource implications and service outcomes of genetic services in the context of DNA technology. Health Policy. 1994;26(3):171-90.
79. Cairns JA, Shackley P. Assessing value for money in medical screening. J Med Screen. 1994;1(1):39-44.
80. Goodman DM, Mendez E, Throop C, Ogata ES. Adult survivors of pediatric illness: the impact on pediatric hospitals. Pediatrics. 2002;110(3):583-9.
81. Hunter MF, Heaf DP. Allowances for care for children with cystic fibrosis. Arch Dis Child. 1993;68(1):144-6.
82. Wilfond BS, Parad RB, Fost N. Balancing benefits and risks for cystic fibrosis newborn screening: implications for policy decisions. J Pediatr. 2005;147(3 Suppl):S109-13.
83. Donaldson C. Using economics to assess the place of screening. J Med Screen. 1994;1(2):124-8; discussion 8-9.
84. Donaldson C. Eliciting patients' values by use of 'willingness to pay': letting the theory drive the method. Health Expect. 2001;4(3):180-8.
85. Donaldson C, Shackley P, Abdalla M. Using willingness to pay to value close substitutes: carrier screening for cystic fibrosis revisited. Health Econ. 1997;6(2):145-59.
86. Donaldson C, Shackley P, Abdalla M, Miedzybrodzka Z. Willingness to pay for antenatal carrier screening for cystic fibrosis. Health Econ. 1995;4(6):439-52.
87. Miedzybrodzka Z, Semper J, Shackley P, Abdalla M, Donaldson C. Stepwise or couple antenatal carrier screening for cystic fibrosis?: women's preferences and willingness to pay. J Med Genet. 1995;32(4):282-3.
88. Miedzybrodzka Z, Shackley P, Donaldson C, Abdalla M. Counting the benefits of screening: a pilot study of willingness to pay for cystic fibrosis carrier screening. J Med Screen. 1994;1(2):82-3.
89. Shackley P, Cairns JA. Evaluating the benefits of antenatal screening: an alternative approach. Health Policy. 1996;36(2):103-15.
90. Kaplan RM, Anderson JP, Wu AW, Mathews WMC, Kozin F, Orenstein D. The Quality of Well-Being Scale: Applications in AIDS, Cystic Fibrosis, and Arthritis. Med Care. 1989 March;27(3):Supplement.
91. Quittner AL, Schechter MS, Rasouliyan L, Haselkorn T, Pasta DJ, Wagener JS. Impact of socioeconomic status, race, and ethnicity on quality of life in patients with cystic fibrosis in the United States. Chest. 2010;137(3):642-50.

92. Johnson JA, Connolly MA, Jacobs P, M M, Brown NE, Zuberbuhler P. Cost of care for individuals with cystic fibrosis: a regression approach to determining the impact of recombinant human DNase. Pharmacotherapy. 1999;19(10):1159-66.
93. Dodge JA, Morison S, Lewis PA, Coles EC, Geddes D, Russell G, et al. Incidence, population, and survival of cystic fibrosis in the UK, 1968-95. Arch Dis Child. 1997;77(6):493-6.
94. Congleton J, Hodson M, Duncan-Skingle F. Quality of life in adults with cystic fibrosis. Thorax. 1996;51:936-40.
95. Doull IJM, Ryley HC, Weller P, Goodchild MC. Cystic fibrosis-related deaths in infancy and the effect of newborn screening. . Pediatr Pulmonol. 2001;31:363-66.
96. Guyatt GH, Sackett DL, Cook DJ. Users guides to the medical literature II. How to use an article about therapy or prevention. Journal of the American Medical Association. 1993;270:2598-601.
97. Fuchs HJ, Borowitz DS, Christiansen DH, Morris EM, Nash ML, Ramsey BW, et al. Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. The Pulmozyme Study Group. N Engl J Med. 1994;331(10):637-42.
98. The Pharma Letter. Pulmozyme launched in UK. 1994 [cited 2011 January 15th, 2011]; Available from: http://www.thepharmaletter.com/file/2982/pulmozyme-launched-in-uk.html.
99. Iles R. Personal Communication. Cambridge2011.
100. British Medical Journal. Best Practice: Cystic Fibrosis. BMJ; 2010 [February 10th 2011]; Available from: http://bestpractice.bmj.com/best-practice/monograph/403/treatment/guidelines.html.
101. McPake B, Normand C. Health Economics: An International Perspective. New York: Routledge; 2008.
102. Sherry KM, McNamara J, Brown JS, Drummond M. An economic evaluation of propofol/fentanyl compared with midazolam/fentanyl on recovery in the ICU following cardiac surgery. Anaesthesia. 1996;51(4):312-7.
103. Shemilt I, Mugford M, Byford S, Drummond M, Eisenstein E, Knapp M, et al. Incorporating economics evidence. In: Higgins J, Green S, editors. Chochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley; 2008.
104. Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. Evidence and Policy. 2010;6(1):51-9.
105. Richardson AW, Gafni A. Treatment of capital costs in evaluating health care programmes. 1983. p. 26-30.
106. Sculpher M, Gafni A. Recognizing diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. Health Econ. 2001;10(4):317-24.
107. Koopmanschap M, Burdorf A, Jacob K, Meerding WJ, Brouwer W, Severens H. Measuring productivity changes in economic evaluation: setting the research agenda. Pharmacoeconomics. 2005;23(1):47-54.
108. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J Health Econ. 1995;14(2):171-89.
109. Gerard K MG. QALY league tables: handle with care. Health Economics. 1993;2:59-64.

110. Brouwer WB, Koopmanschap MA, Rutten FF. Productivity costs measurement through quality of life? A response to the recommendation of the Washington Panel. Health Econ. 1997;6(3):253-9.
111. Bowling A. Measuring Health – a review of Quality of Life measurement scales. Milton Keynes: Open Clinical Press; 1991.
112. Brooks R. EuroQol: the current state of play. Health Policy. 1996;37(1):53-72.
113. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002;40(2):113-28.
114. Brazier J, McCabe C. 'Is there a case for using visual analogue scale valuations in CUA' by Parkin and Devlin. A response: 'yes there is a case, but what does it add to ordinal data?'. Health Econ. 2007;16(6):645-7; discussion 9-51.
115. World Health Organization. The molecular genetic epidemiology of cystic fibrosis: Report of a joint meeting of WHO/ECFTN/ICF(M)A/ECFS. Geneva: World Health Organization2004.
116. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2(14):i-iv, 1-74.
117. Brazier J. Valuing health States for use in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):769-79.
118. Green C, Brazier J, Deverill M. Valuing health-related quality of life: a review of health state valuation techniques. PharmacoEconomics. 2000;17(2):151-65.
119. Mannion AF, Junge A, Fairbank JCT, Dvorak J, Grob D. Development of a German version of the Oswestry Disability Index. Part 1: cross-cultural adaptation, reliability and validity. European Spine Journal. 2006;15(1):55-65.
120. O'Brien B, Viramontes JL. Willingness to pay: a valid and reliable measure of health state preference? Medical Decision making. 1994;14:289-97.
121. Bleichrodt H, Johannesson M. An experimental test of a theoretical foundtation for rating scale valuations. Medical Decision Making. 1997;17:208-16.
122. Robinson A, Loomes G, Jones-Lee M. Visual analogue scales, standard gambles and relative risk aversion. Medical Decision Making 2001;21:17-27.
123. McCabe C, Brazier J. Using rank data to estimate health state utility models. Journal of Health Economics. 2006;25(3):418-31.
124. Torrance GW, Feeny D, Furlong W. Visual analog scales: do they have a role in the measurement of preferences for health states? Medical Decision Making. 2001;21:329-234.
125. Torrance GW, Feeny D, Furlong W, Barr R, Zhang Y, Wang Q. Multi-attribute utility function for a comprehensive health status classification system: health utilities index mark 2. Med Care. 1996;34:702-22.
126. Stevens KJ, McCabe C, Brazier J. Mapping between visual analogue and standard gamble data: results from the UK Health Utilities Index 2 valuation. Health Econ. 2006;15(5):527-34.
127. Torrance GW. Social preferences for health states: An empirical evaluation of three measurement techniques. Socio-Economic Planning Sciences. 1976;10:129-36.
128. Torrance GW. Measurement of health state utilities for economic appraisal: a review. Journal of Health Economics. 1986;5:1-30.

129. Jones-Lee M. Personal willingness to pay for prevention: evaluating the consequences of accidents as a basis for preventive measures. Addiction. 1993;88(7):913-21.
130. Stein K, Dyer M, Crabb T, Milne R, Round A, Ratcliffe J, et al. A pilot internet 'value of health' panel: recruitment, participation and compliance. Health and Quality of Life Outcomes. 2006;4(90).
131. Stewart ST, Lenert L, Bhatnagar V, Kaplan RM. Utilities for prostate cancer health states in men aged 60 and older. Med Care. 2005;43(4):347-55.
132. Gerson LB, Ullah N, Hastie T, Triadafilopoulos G, Goldstein M. Patient-derived health state utilities for gastroesophageal reflux disease. Am J Gastroenterol. 2005;100(3):524-33.
133. Schoemaker PJ. The expected utility model: its variants, purposes, evidence and limitations. J Econ Lit. 1982;20:529-63.
134. Kahneman D, Tversky A. Prospect theory: an analysis of decisions under risk. Econometrica. 1979;47(2):263-91.
135. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. J Health Econ. 1996;15:209-31.
136. Rutten-van Molken MPMH, Bakker CH, van Doorslaer EKA, van der Linden S. Methodological issues of patient utility measurement: experience from two clinical trials. Med Care. 1995;33(922-37).
137. Lenert L. The reliability and internal consistency of an internet-capable computer program for measuring utilities. Quality of Life Research. 2001;9:811-7.
138. Mehrez A, Gafni A. The healthy-years equivalents: how to measure them using the standard gamble approach. Medical Decision Making 1991;11:140-6.
139. Cher D, Miyamoto J, Lenert L. Incorporating risk attitude into Markov-process decision models: importance for individual decision making. Medical Decision Making. 1997;17:340-50.
140. van der Pol M, Cairns J. Comparison of two methods of eliciting time preference for future health states. Soc Sci Med. 2008;67(5):883-9.
141. Robinson A, Dolan P, Williams A. Valuing health status using VAS and TTO: What lies behind the numbers. Social Science and Medicine. 1997;45(8):1289-97.
142. Nord E. Cost-value analysis in health care: making sense out of QALYs. Cambridge: Cambridge University Press; 1999.
143. Pinto Prades JL. Is the person trade-off a valid method for allocating health care resources. Health Econ. 1997;6:71-81.
144. Murray CJL, Lopez AD. The Global Burden of Disease. Boston, MA: Harvard University Press; 1996.
145. Nord E. The person-trade-off approach to valuing health care programs. Medical Decision Making 1995;15:201-8.
146. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. BMJ. 2000;320(1530).
147. Thurstone LL. The method of paired comparisons for social values. Journal of Abnormal and Social Psychology. 1927;21:384-400.
148. Luce R, Raiffa H. Games and decisions. New York: Wiley; 1957.
149. Salomon J. Reconsidering the use of ranknings in the valuation of health states: a model for estimating cardinal values from ordinal data. Poupulation Health Metrics. 2003;1(1).

150. Ryan M, Gerard K. Using discrete choice experiment to value health care programmes: current practice and future research reflections. Health Economics. 2003;7:373-8.
151. Louviere JJ, Henscher DA, Swait JD. Stated choice methods. Analysis and application. Cambridge: Cambridge University Press; 2000.
152. Morey ER, Rowe RD, Watson M. A repeated nested logit model of atlantic salmon fishing. American Journal of Agricultural Economics. 1993;75:578-92.
153. Tennant A, McKenna SP, Hagell P. Application of Rasch Analysis in the Development and Application of Quality of Life Instruments. Value in Health. 2004;7(Supplement 1):S22-S6.
154. Quittner AL, Modi A, Cruz I. Systematic review of health-related quality of life measures for children with respiratory conditions. . Paediatr Respir Rev. 2008;9(3):220-32.
155. Quittner AL, Buu A, Messer MA, Modi AC, Watrous M. Development and Validation of the Cystic Fibrosis Questionnaire in the United States: A Health-Related Quality-of-Life Measure for Cystic Fibrosis. Chest. 2005;128:2347-54.
156. Gee L, Abbott J, Conway SP, Etherington C, Webb AK. Development of a disease specific health related quality of life measure for adults and adolescents with cystic fibrosis. Thorax. 2000;55:946-54.
157. Goldbeck L, Schmitz TG, Henrich G, Herschbach P. Questions on life satisfaction for adolescents and adults with cystic fibrosis: development of a disease-specific questionnaire. Chest. 2003;123:42-8.
158. Akobundu E, Ju J, Blatt L, Mullins C D. Cost-of-Illness Studies: A Review of Current Methods. PharmacoEconomics. 2006;24(9):869-90.
159. Bloom B, Bruno DJ, Maman DY, Jayadevappa R. Usefullness of US Cost-of-Illness Studies in Healthcare Decision Making. PharmacoEconomics. 2001;19(2):207-13.
160. Hedeker D, Gibbons R. Longitudinal Data Analysis. New Jersey: John Wiley & Sons, Inc.; 2006.
161. Lee DW, Meyer JW, Clouse J. Implications of controlling for comorbid conditions in cost-of-illness estimates: a case study for osteoarthritis from a managed care system prespective. Value in Health. 2001;4(4):329-34.
162. Galdin M, Laurencelle L. Assessing parameter invariance in item response theory's logistic two item parameter model: a monte carlo investigation. Tutorials in Quatitative Methods for Psychology. 2010;6(2):39-51.
163. Gujarati D. Basic Econometrics. London: McGraw-Hill International Editions; 1995.
164. Verbeek M. A Guide to Modern Econometrics. Sussex, England: John Wiley & Sons Ltd; 2008.
165. Blough DK, Ramsey SD. Using generalized linear models to assess medical care costs. . Health Services and Outcomes Research Methodology. 2000;1(2):185-202.
166. McCullagh P, J N. Generalized Linearl Models. London: Chapman and Hall; 1989.
167. Cooper NJ, Sutton A J, Mugford M, Abrams K R. Use of Bayesian Markov Chain Monte Carlo methods to model cost-of-illness data. Med Decis Making. 2003;23(1):38-53.

168. Manning WG. The logged dependent variable, heteroscedatsticity, and the retransformation prolem. Journal of Health Economics. 1998;17:283-95.
169. Blough DK, Madden C, Hornbrook M. Risk Modeling using generalised linear models. Health Economics. 1999;18:153-71.
170. Austin PC, Ghali WA, Tu JV. A comparison of several regression models for analysing cost of CABG surgery. Statistics in Medicine. 2003;22(17):2799-815.
171. Liang KY, Zeger S. Longitudinal data analysis using generalized linear models. Biometrika. 1986;73(1):12-22.
172. Freemantle N, Maynard A. Something rotten in the state of clinical and economic evaluations? Health Economics. 1994;3(2):63-7.
173. Coyle D, Lee KM. Evidence-based economic evaluation: how the use of different data sources can impact results. London: BMJ Books; 2002.
174. Smith P. Measuring value for money in healthcare: concepts and tools. York: Centre for Health Economics University of York 2009.
175. Cleemput I, Neyt M, Thiry N, De Laet C, Leys M. Threshold values for cost-effectiveness in health care: Belgian Health Care Knowledge Centre2008.
176. Donaldson C, Baker R, Mason H, Jones-Lee M, Lancsar E, Wildman J, et al. The social value of a QALY: raising the bar or barring the raise? BMC Health Services Research. 2011;11(8).
177. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. London: Chapman & Hall; 1993.
178. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. Journal of Health Services Research and Policy. 2004;9:110-8.
179. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Economics. 2006;15(12):1295-310.
180. Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian Model Averaging: A Tutorial. Statistical Science. 1999;14(4):382-417.
181. Bojke L. Defining and Characterising Structural Uncertainty in Decision Analytic Models. CHE Research Paper 9: The University of York 2006.
182. Microsoft Access. Microsoft; 2003.
183. Legh-Smith J, Iles R. Short and long-term outcomes of neonatal screening for cystic fibrosis within the Eastern Region of the UK. Cambridge: Addenbrooke's Hospital1999.
184. British National F. British National Formulary for Children. London: Pharmaceutical Press; 2010.
185. Stata Statistical Software. Release 11 ed. College Station, TX: StataCorp; 2009.
186. Pan W. Akaike's information criterion in generalized estimating equations. Biometrics. 2001;57:120-5.
187. Cui J. QIC program and model selection in GEE analyses. The Stata Journal. 2007;7(2):209-20.
188. Wilson JMG, Jungner G. Principles and practice of screening for disease. Geneva, WHO1968.
189. Microsoft Excel. 2007 ed: Microsoft.
190. Office of National Statistics. Population Statistics. 2009; Available from: www.statistics.gov.uk.

191. Orenstein DM, Nixon PA, Ros EA, Kaplan RM. The Quality of well-being in cystic fibrosis. Chest. 1989;95(2):344-7.
192. Weinstein M, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices - Modeling Studies. Value in Health. 2003;6(1):9-17.
193. Black's Medical Dictionary. Dr. Harvey M, editor. London: A & C Black; 2010.