Economic issues associated with the operation and evaluation of telemedicine

A thesis submitted for the degree of Doctor of Philosophy

By Hema Mistry

Health Economics Research Group

Brunel University

February 2011

Abstract

Telemedicine offers an alternative referral strategy for fetal cardiology but is currently only used for 'high-risk' pregnancies. A case-study of a cost-consequences analysis comparing telemedicine to direct referral to a perinatal cardiologist is initially presented, which highlights that for high risk women for whom telemedicine was considered no cardiac anomalies were missed using either referral method. In the light of a review of the literature on the economics of telemedicine, three of the key methodological issues (of selection bias, of patient costs and using quality-adjusted life years (QALYs)) are explored to demonstrate how the case study analysis could be improved.

Pregnant women were selected for referral based on their characteristics and risk factors; thus the cost and effects for the two groups may have been biased. Various methods identified in the literature are applied to the case study to reduce selection bias, but the analysis presented is unable to determine which method is best, given a number of limitations including the small sample size.

The analysis is extended to include estimated total patient costs. However, when patient costs are added to the total costs of pregnancy, they did not substantially increase the overall cost. The results presented provide a guideline for future researchers and pregnant women of the likely costs during pregnancy.

Given that the majority of missed cardiac anomalies were amongst low risk women, a decision analytical model is developed looking at the lifetime costs and QALYs of introducing telemedicine screening for pregnant women whose unborn babies are at a low risk of congenital heart disease. The analysis shows that offering telemedicine to all low risk women is the dominant strategy. The thesis demonstrates, within the constraints of existing data, that it would be cost-effective to provide telemedicine as part of an antenatal screening programme for all low risk women, and this would help prevent future 'missed anomalies'.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to give a special thanks to my supervisor Professor Martin Buxton for the help and encouragement throughout my thesis and for never giving up on me. I would also like to recognise the advice and support given by Professors Julia Fox-Rushby and Stephen Morris and Dr Robin Dowie, who all at some point were a second supervisor. I would also like to thank members (past and present) from the Health Economics Research Group, Brunel University and current members of the Health Economics Unit, University of Birmingham who have helped me with my work.

A large part of this work was possible because of the availability of the dataset from the TelePaed project, which was funded by the Research and Development Division of the Department of Health. However, I must also thank staff at Medway Maritime Hospital especially Daphne Brown for time and effort spent collecting the extra data; and sonographers Helen Kennet and Karen Fegan for their advice. Special thanks must also go to Dr Helena Gardiner at the Royal Brompton Hospital for always being on the end of email and answering my various clinical questions. In relation to the model in Chapter 7, I would like to thank Michael Cumper from the Grown-Up Congenital Heart Patients' Association and Dr Lorna Swann from the Royal Brompton Hospital, who helped to populate my model.

Finally, I would like to thank my family and friends, especially Kris Kay, for all their support, particularly while I have been writing this thesis.

Publications and Authorship

The work presented in this thesis has not been submitted elsewhere for a degree. Where appropriate, all sources of information are acknowledged and referenced accordingly.

As part of the TelePaed project, which was conducted in four district hospitals, the data analysis and results were presented in four reports to the Department of Health and in six peer-reviewed journal articles. My main role in the project team was to collect the unit cost information from six hospitals and analyse the data on costs (and outcomes), and questionnaires from various postal surveys. I contributed fully to the data analysis for the three groups of patients studied (pregnant women, neonates, and older infants and children), writing and the production of each of the reports and papers.

Some of the work that contributes to this thesis draws on three of the published papers and two of the Department of Health reports.

The background and the overall data analysis and results for the TelePaed project presented in Chapter 2 were summarised in the following published article (see Appendix 9):

- Dowie R, Mistry H, Young TA, Weatherburn G, Gardiner HM, Rigby M, Rowlinson G and Franklin RCG (2007) Telemedicine in pediatric and perinatal cardiology: economic evaluation of a service in English hospitals. International Journal of Technology Assessment and Health Care. 23(1): 116-125.

Detailed costing data and results specifically for pregnant women in the only hospital that used the telemedicine equipment for fetal cardiology were presented in the paper below and formed the basis for this thesis (see Chapter 2 and Appendix 10):

- Dowie R, Mistry H, Young TA, Franklin RCG and Gardiner HM (2008) Cost implications of introducing a telecardiology service to support fetal ultrasound screening. Journal of Telemedicine and Telecare. 14(8): 421-426.

The thesis also draws on three of the other publications from the project. Two were project reports submitted to the Department of Health in 2003 and 2004. The first described the background and the design of the TelePaed study (see Chapter 4):

- Dowie R, Young T, Mistry H and Weatherburn G on behalf of the TelePaed project team (December 2003) Economic evaluation of the role of telemedicine in paediatric cardiology. First report: Paediatric cardiology outpatient services. Report to the Department of Health, Brunel University, Uxbridge.

The second report focused on the configuration of the antenatal screening services in the district hospitals and the role of the telemedicine service (see Chapters 2 and 4):

- Dowie R, Mistry H, Young T and Weatherburn G on behalf of the TelePaed project team (September 2004) Economic evaluation of the role of telemedicine in paediatric cardiology. Second report: Fetal cardiology services. Report to the Department of Health, Brunel University, Uxbridge.

The third publication was a journal paper in which we analysed the pregnancy datasets from the district hospitals to compare the costs of obstetric care for women with singleton versus multiple pregnancies (see Chapters 2 and 5 and Appendix 11):

- Mistry H, Dowie R, Young T and Gardiner H (2007) The costs of maternity care for women with multiple pregnancy compared with high-risk and low-risk singleton pregnancy. British Journal of Obstetrics and Gynaecology. 114(9): 1104-1112.

Finally, the work on selection bias methodology in Chapter 5 has been presented as part of a conference paper.

- Mistry H and Morris S (2007) Methodological comparison of different approaches used to control for selection bias in non-randomised studies. HESG; University of Birmingham, January 2007.

Abbreviations

AA	Automobile Association
ATT	Average Treatment Effect on Treated
CA	Cost Analysis
CBA	Cost-Benefit Analysis
CC	Conventional Care
CCA	Cost-Consequences Analysis
CCAD	Cardiac Audit Database Registry
CEA	Cost-Effectiveness Analysis
CEAC	Cost Effectiveness Acceptability Curve
CHD	Congenital Heart Disease
CI	Confidence Interval
CMA	Cost-Minimisation Analysis
CVD	Cardiovascular Disease
CVS	Chorionic Villus Sampling
DALYs	Disability-Adjusted Life Years
DCEs	Discrete Choice Experiments
DGH	District General Hospital
DMC	Direct Medical Costs
DNMC	Direct Non-Medical Costs
DR	Direct Referral
ECMO	Extracorporeal Membrane Oxygenation
EW	Equal Weights
GDP	Gross Domestic Product
GP	General Practitioner
GUCH-PR	Grown-Up Congenital Heart Patient Representative
HADS	Hospital Anxiety and Depression Scale
HUI	Health Utilities Index
IC	Indirect Costs
ICER	Incremental Cost-Effectiveness Ratio
ICTs	Information and Communication Technologies
ICTRI	ICT Research and Development Programme
IQR	Inter-Quartile Range
ISDN	Integrated Services Digital Network
IT	Information Technology
KCH	King's College Hospital
LR	Low Risk
MR	Medium Risk
NHS	National Health Service
NICE	National Institute of Health and Clinical Excellence
NS	Not Stated
NT	Nuchal Translucency
NYHA	New York Heart Association
PedsQL	Pediatric Quality of Life Inventory
PPP	Purchasing Power Parity
PS	Propensity Score
PSav	Potential Savings
PSS	Personal Social Services
QALYs	Quality-Adjusted Life Years
QCH	Queen Charlotte's Hospital
RBH	Royal Brompton Hospital
RCT	Randomised Controlled Trial
RD	Random Draw
ROC	Receiver Operating Curve
SAB	Short-Acting B2 Agonist Inhalers
SD	Standard Deviation

SE	Standard Error
SF-12	Short-Form 12
TM	Telemedicine
VAS	Visual Analogue Scale
VAT	Value Added Tax

CHAPTER 1: GENERAL INTRODUCTION AND WHAT IS KNOWN ABOUT TELEMEDICINE

1.1 Introduction

The term telemedicine is defined as “the use of electronic information and communication technologies to provide and support health care when distance separates the participants” [Institute of Medicine, 1996, p16]. The World Health Organisation’s definition for telemedicine is: “the delivery of healthcare services, where distance is the critical factor, by all healthcare professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of healthcare providers, all in the interests of advancing health of individuals and their communities” [WHO, 1998]. In simple terms, telemedicine means ‘the delivery of medicine at a distance’.

Telemedicine may be as simple as two health professionals discussing a case over the telephone, or as complex as telesurgery, in which a surgeon receives visual and audio information to guide robotic instruments to perform surgery at a distance. In between these two types of telemedicine, lies the use of video, audio and data transmission technologies such as using videoconferencing to conduct a real-time consultation between two medical specialists with the patient present. Telemedicine can also be used for non-clinical applications, for example for medical education, meetings, research and administration.

There has always been a strain on National Health Service (NHS) resources and finances, not only due to the rapid increase in the costs of medical treatment, but also from the growing demand of the ageing population. Affordability is always an issue. Therefore, it is important that the issue of affordability is addressed within a framework that allows direct comparison of an intervention with other alternatives in terms of value for money. This has led to increasing pressures on healthcare budgets to show that healthcare technologies such as telemedicine, not only demonstrate their safety and efficacy, but also show that they are an efficient use of resources. Economic evaluations provide information on whether healthcare technologies are an efficient use of resources by comparing the costs and benefits of one healthcare technology, to the costs and benefits of another healthcare technology. So when healthcare budgets are limited, resources should be allocated towards those technologies where the ratio of incremental costs to incremental benefits is within a given cost-effectiveness threshold. That is, a cost-effectiveness ratio that is assumed to represent society’s willingness to

pay for an additional unit of health (e.g. a QALY). The economics of information and communication technologies (ICTs) are important, as financial resources are scarce and telemedicine as an ICT can have potentially high costs. Decisions ultimately have to be made by policy makers as where to best allocate these finite resources.

This chapter provides a general introduction to telemedicine, including a background to telemedicine and the status of funding for telemedicine services in the UK. The chapter then looks at what information there is on 'economic evaluations of telemedicine'; this is followed by the aims and objectives of the thesis; then there will be a short summary of the case study which will be used throughout the thesis: the use of telemedicine in providing specialist advice in fetal cardiology; and finally, the structure for the thesis will be outlined.

1.2 Telemedicine or Telecare?

Within the NHS it is vital to be able to access the right information required to deliver high quality services. ICTs are enabling technologies for the NHS which aims to provide high quality and efficient services. Telemedicine is one of a number of ICTs, which also include picture archiving and communication systems, hospital information systems, computerised medical records, computers to support decision making etc.

For clinical applications, telemedicine and telecare are terms that are used interchangeably to describe the delivery of healthcare. The Government in Information for Health: an Information Strategy for Modern NHS (1998) defined telemedicine and telecare as "Any healthcare related activity (including diagnosis, advice, treatment and monitoring) that normally involves a professional and a patient (or one professional and another) who are separated in space (and possibly also in time) and is facilitated through the use of information and communications technologies. Telemedicine is usually delivered in a hospital clinic or surgery, while telecare is delivered in the patient's home" [DoH, 1998, p123].

In other words, telemedicine supports the exchange of information between health care professionals and/or patients (such as for diagnosis or referral) usually in a primary or secondary care setting, whereas, telecare is a service bringing health and social care directly to a patient, usually in their own home.

Telecare is based on the assumption that the elderly, disabled or vulnerable people should live independently, in control and with dignity for longer. Telecare has the potential to reduce unnecessary hospital admissions and improve people's quality of

life and this means that care should be delivered where it is most appropriate; that is, in a patient's home. Examples of telecare include: sensors that monitor falls which will then trigger a warning to the response centre and systems which can monitor vital signs such as blood pressure, and this data is then transmitted to a response centre.

Even though both terms are used interchangeably, they have slightly different meanings in terms of where the care is delivered. The focus of this thesis will be on telemedicine and not telecare.

1.3 What is telemedicine?

The term telemedicine is composed of the Greek word *τελε* (tele) meaning 'far', and medicine. Telemedicine services aim to provide access to medical information, knowledge and expertise and to support the provision of care processes and the services which are delivered. The classical use of telemedicine in healthcare was to support services for distant or isolated populations. Now, telemedicine can be seen as a modern healthcare delivery process. The factor of distance is becoming less of an issue and the focus is now on the boundaries of separation that exists between different users i.e. between hospitals and general practices. Telemedicine has diversified into a wide range of applications, for clinical (i.e. patient care) and non-clinical purposes (i.e. medical education). For purposes of this thesis, the focus will be on the use of telemedicine for clinical applications.

Telemedicine can be conducted in two ways: real-time or store-and forward. Real-time telemedicine requires the presence of both medical specialists and the patient at the same time and a communications link between them allows a real-time interaction to take place. Store-and-forward telemedicine involves acquiring the medical data such as ultrasound images and then transmitting this to a medical specialist at a convenient time for assessment offline. It does not require both parties of clinicians to be present at the same time and offers more flexibility. Table 1.1 summarises some of the advantages and disadvantages of telemedicine.

Table 1.1: Advantages and disadvantages of telemedicine

	Real-time telemedicine	Store-and-forward telemedicine
Advantages	Improve a local consultant or general practitioners knowledge Reduced travelling time for specialist consultants Reduced travel and hospital time for a patient	Less organisation is required i.e. images can be viewed at a convenient time Cheaper than real-time telemedicine Possibly shorter consultations
Disadvantages	Higher costs of equipment and transmission More time consuming for health professionals There may be more complex scheduling and organisational factors	Lack of manipulation capabilities by the specialist i.e. they may wish to view the image from a different angle Lack of clinical information which may be necessary to make an informed decision because the patient is not present No direct contact with the patient

1.4 Background to telemedicine

Although, telemedicine in the last 20 to 30 years has become more widely acceptable in healthcare, the history of telemedicine goes back much further than that. According to a review by Zundel (1996), the first reference to telemedicine in the medical literature appeared in 1950. Gershon-Cohen and Cooley (1950) described the transmission, beginning in 1948, of radiological images such as x-rays over telephone lines between West Chester and Philadelphia in Pennsylvania, USA, a distance of 24 miles. In 1959, clinicians at the University of Nebraska medical centre used a two-way, interactive television to transmit neurologic examinations and other information across campus for medical education, mainly in psychiatry; and the medical centre was also linked to remote rural areas to assist with medical treatment [Wittson and Benschoter, 1972]. Shortly after, many other early telemedicine applications appeared which focused on the limited access of remote populations to a variety of health services where medical specialists and general practitioners (GP) were not easily reached. Use of telemedicine in urban areas also appeared to assist in emergencies. For example, in 1963 the Massachusetts General Hospital established a telecommunications link with a medical station staffed by nurse clinicians at Boston's Logan Airport [Bird, 1972].

1.5 Examples of telemedicine applications

Telemedicine has been used for a variety of healthcare applications and many of them appear in hospital based services such as radiology and psychiatry. Other applications of telemedicine in healthcare appear in primary care based services, such as virtual outpatient clinics for expert opinion and education networks. An overview of how telemedicine works for radiology, dermatology, psychiatry and cardiology are listed

below and Chapter 3, looks at the studies which have been conducted in these four areas in more detail.

Teleradiology can be defined as “the electronic transmission of radiologic images from one location to another, for the purpose of interpretation and/or consultation” [Ferrer-Roca, 1998]. Teleradiology involves the ability to send radiographic images such as x-rays, computed tomography scans and magnetic resonance imaging scans from one location to another. These digital images combined with demographic and other patient information, can be easily compressed to allow them to be sent more quickly and inexpensively. Whereas, for teledermatology this involves consultations between a patient with skin disease (and the primary health care provider) and a dermatologist for diagnosis and management advice. As with teleradiology, the production of digital images of skin disease, supplemented with demographic details, can be sent from one site to another quickly and cheaply. Images can be transmitted using either real-time or store-and-forward technology. Both transmission methods require two instruments: a microscope and a dermatoscope.

Telepsychiatry is the delivery of healthcare and exchanges information for purposes of psychiatric services across distances. Mental health interactions are more straightforward and demand little other than the participants seeing and talking with one another and there is no need for sophisticated devices such as an electronic stethoscope. Telecardiology in contrast, is the transmission of images such as echocardiograms from the local centre to the specialist centre. The transmission of images can be done in real-time or using store-and-forward techniques. The images are sent for purposes of interpretation for further assessment or clarification. The use of telemedicine in fetal cardiology will be the main focus of the thesis. The next section will provide an overview of the policies and funding opportunities which have been introduced for telemedicine in recent years.

1.6 Telemedicine and funding in the UK

The UK Government’s strategy to modernising the NHS involves the use of modern computer technology in the health service, in the hope of “giving the people of this country the best system of healthcare in the world” [DoH, 1998]. This included the publication of three key papers which are outlined below.

The first paper, the new NHS Modern and Dependable White Paper [DoH, 1997] set out plans to build a modern and dependable health service fit for the 21st century which included frontline patient services backed by more investment and better technology.

This meant providing: 1) at home: easier and faster advice and information through NHS Direct, a new 24 hour telephone advice line; 2) in the community: patients to benefit from quicker test results, up-to-date specialist advice in the doctor's surgery and on-line booking of out-patient appointments, when every GP surgery and hospitals would be connected to the NHSnet, the NHS's own information superhighway; and 3) in hospital: prompt access to specialist services so that treatment and care be undertaken efficiently and quickly. This paper marked a turning point for the NHS and set out the basis for a ten-year programme to renew and improve the NHS, including an extra £1.5 billion.

Secondly, the publication of the Information for Health report [DoH, 1998], included to support the drive for quality and efficiency in the NHS by the use of information technology (IT). This strategy was backed with £1 billion of modernisation funding during the lifetime of the strategy. The report stated that "a modern and dependable NHS needs accurate and instantly accessible information and the benefits of new high-speed, high-capacity information and communications networks" [DoH, 1998]. One of the key objectives from this paper included: "to eliminate unnecessary travel and delay for patients by providing remote on-line access such as telemedicine to services, specialists and care, wherever practicable" [DoH, 1998, p19]. The paper emphasised the role of telemedicine and telecare: by improving the quality of care, by making information faster and more easily available to patients, by making specialist advice and support more accessible to health professionals, and by bringing services closer to peoples' homes. The key deliverables from this information strategy included: a framework to be published to guide and support the development and application of telemedicine; the connection of all computerised GP practices to NHSnet; offering NHS Direct services to the whole population; a National Electronic Library for Health accessible through local intranets in all NHS organisations; and comprehensive electronic patient and health records would be available throughout the NHS to support the delivery of care.

The third paper: Modernising Government White Paper [DoH, 1999] set out the vision for modernising public services to make life better for people and businesses including: delivering better public services; developing an IT strategy for Government which would establish cross-government co-ordination and frameworks on issues such as digital signatures, smart cards, websites and call centres; and to benchmark progress against targets for electronic services.

Following on from the second White paper, the Department of Health launched an ICT research and development programme (ICTRI-1) to support the national information strategy and had a budget guideline of £2.5 million over three to four years. The main aim within this initiative was to commission broad based, multidisciplinary health service research to evaluate ways in which ICTs can help to provide clinical, managerial and patient benefits, and to improve the evidence-base for decisions on ICT investments. Within the ICTRI-1 programme, 14 projects were commissioned - five were in the area of telemedicine.

In early 2001, saw the publication of "Building the Information Core: Implementing the NHS Plan" [DoH, 2001a] which updated the Information for Health paper [DoH, 1998] and provided a clearer focus on what the priorities for successful delivery needed to be and also re-emphasised the future of telemedicine in the NHS. The Government announced an extra £700 million investment in IT. The investment included more consultations with hospital specialists being carried out in GP surgeries using video and tele-links removing the need for patients to visit hospital and ensuring swift diagnosis and treatment; ambulances equipped with video and monitoring equipment so that patients would get specialist care while being taken to hospital; and investment in electronic patient records, so allowing patients to connect with staff electronically for advice, booking appointments and to see their test results.

At the same time, the UK Telemedicine Information Service website funded by the Department of Health was launched on 15th January 2001 at the TeleMed 2001 conference in London [DoH, 2001b]. The aim of the website was to improve the take-up of telemedicine technology in the UK and to give access to information on all aspects of telemedicine by encouraging those working in the field of telemedicine to share their information and experience. The website contained information on the latest projects and developments, an email discussion list, and the option of receiving updates from the Medline database about current services.

Shortly after, the Department of Health published a report on "Delivering 21st Century IT Support for the NHS" [DoH, 2002b] which outlined targets for telecare to be available in all homes that require it by December 2010. Their aims included:

- The support of patients and the delivery of services designed around the patient, quickly and conveniently; and
- The support of staff, through effective electronic communications, better learning and knowledge management; cut the time to find essential information and make specialised expertise more accessible i.e. telemedicine.

The focus for the Department of Health in providing IT in the NHS since 2002 has changed from 'telemedicine' to 'telecare'. In July 2004, the Department of Health announced a Preventative Technology Grant, which aimed to increase the number of people who can benefit from telecare services; for people of all ages including those with long-term conditions, learning disabilities, mental health problems, those needing end of life care and also to help an additional 160,000 older people remain independent at home [DoH, 2005]. In 2006, ICTRI-2 was commissioned and this led to a new approach to the delivery of care, based on the premise that people in need of care should be able to remain independent for as long as possible and care will be delivered where it is most appropriate such as the patient's home. This new care delivery mechanism was referred to as 'Telecare'. Finally, in May 2008, the Whole System Programme was launched which aimed to determine whether technology can help people manage their own health and maintain their independence; specifically supporting home care of people with complex medical and social needs [DoH, 2009].

Due to the devolved Government in the UK, the funding opportunities available for telemedicine in each of the countries are different. As part of an ongoing collaboration between NHS Lothian and the Scottish Government, small scale telemedicine projects have been set up which have then been rolled out in the region [Bowes and McColgan, 2006]. For example, one project consisted of managing chronic obstructive pulmonary disease patients care remotely at home using a device designed to ask patients questions about their symptoms and depending on their response may prompt patients to use a device such as a pulse oximeter. This information is transmitted via an online interface, highlighting to clinicians any patients at potential risk. The project began in March 2008 with 30 patients and then in February 2009 was rolled out to 400 patients across the Lothian area [Glaser, 2009].

Clearly, IT had a crucial role to play in the modernisation of the NHS and this section has highlighted the importance of telemedicine (and telecare) over the last thirteen years. The section also showed that there has been a shift in the provision of healthcare from telemedicine to telecare and over time, policy priorities and ICTs have changed.

1.7 Economic evaluations of telemedicine

Economic evaluation is concerned with estimating opportunity costs (the value of the next best alternative foregone as a result of the decision made) so that the best use is made of the scarce resources and has been defined as "a comparative analysis of

alternative courses of action in terms of both their costs and consequences" [Drummond et al, 2005]. The basic tasks of any economic evaluation are to identify, measure, value, and compare the costs and consequences of the alternatives under consideration [Drummond et al, 2005]. In the case of telemedicine, the alternatives would usually be the conventional system of delivering healthcare and telemedicine. Costs and benefits of both alternatives would be compared to see which service provides the best use of resources. There are five main types of economic evaluation: cost-minimisation analysis; cost-consequences analysis; cost-effectiveness analysis; cost-utility analysis; and cost-benefit analysis. Each of these evaluations have been outlined in detail elsewhere and the choice of each evaluation depends mainly on the type of outcomes which arise [Drummond et al, 2005].

Many studies evaluating telemedicine have adopted a cost-minimisation approach, though the justification of using a cost-minimisation approach for telemedicine has never really been explained and it is unclear why one should demand cost savings from telemedicine interventions, when most evaluations in healthcare compare the additional costs with the additional benefits. Decision makers in health care (those who deliver and fund such health services) require assurance that telemedicine can fulfil its promise. Therefore, cost-effectiveness (also cost-consequences, cost-utility and cost-benefit) studies of telemedicine interventions are required to provide information to decision makers on whether these interventions in the various healthcare areas represent good value for money.

A few papers have outlined a framework for the economic evaluation of telemedicine. McIntosh and Cairns' (1997) described the economic issues associated with the introduction of telemedicine systems and the main challenges to their evaluation. The approach they suggested is based on a cost-consequences framework and their paper links the costs and consequences more formally within a set of evaluative questions which in turn forms the basis for an economic model evaluating telemedicine. The authors listed the main challenges to the economic evaluation of telemedicine which included: constantly changing technology; lack of appropriate study design to manage the frequently inadequate sample sizes; inappropriateness of the conventional techniques of economic evaluation; and the valuation of health and non-health outcomes. However, they did not provide additional detail on these points, whereas, the proposed framework for economic evaluation of telemedicine by Sisk and Sanders (1998) considers the use of a cost-effectiveness analysis to help to assess whether the expected health benefits are worth the investment. The authors felt that telemedicine raises particular challenges for evaluators: a telemedicine system may have multiple

uses and joint costs that are difficult to apportion to one service; the existence of a system may lead to expanded indications for use; and technological change may rapidly make an evaluation outdated. The authors believed that economic analysis can be helpful in addressing these issues; however, they cannot necessarily resolve them. Lobley (1997) found that even though there had been a large number of telemedicine trials by the mid 1990s, little information had been published on its economic costs and benefits. He believed that a framework was needed to enable decision makers to analyse the potential effects of telemedicine applications on the activities, functions and roles of the different parties involved, and the different costs and benefits for each of these groups. He found that telemedicine costs are strongly related to patient volumes.

There has also been a lot of discussion around the need for cost-effectiveness studies in telemedicine. For example, Hailey (2005) said that "Telemedicine has the potential substantially to improve the delivery of health. However, cost-effectiveness studies are needed to help define the appropriate scope and application of telemedicine in different settings". The author found that there were a lot of studies on telemedicine which looked at the feasibility and technical sides, which are helpful to initial decision making; however, in the longer-term, information on the cost-effectiveness of telemedicine applications are required.

A small number of literature reviews have been conducted looking at the cost-effectiveness of telemedicine interventions. Whitten and colleagues (2002) undertook a systematic review of the cost-effectiveness of telemedicine studies and concluded that there is no good evidence that telemedicine is a cost-effective means of delivering healthcare. Roine et al (2001) examined evidence on the effectiveness or cost-effectiveness of telemedicine studies and concluded that the "evidence regarding the effectiveness or cost-effectiveness of telemedicine is still limited". Hailey and colleagues (2002) conducted a systematic review of evidence for the benefits of telemedicine and concluded that "although useful clinical and economic outcomes data have been obtained for some telemedicine applications, good quality studies are still scarce and the generalisability of most assessment findings is rather limited". A more recent review of economic evaluations of telemedicine conducted by Bergmo (2009) found that "the majority of the economic evaluations were not in accordance with standard evaluation techniques"

Other systematic reviews of telemedicine also found no conclusive evidence that their findings could be generalised to other telemedicine applications. For example, Mair and Whitten (2000) conducted a systematic review of studies on patient satisfaction

with telemedicine, and found that “methodological deficiencies of the published research limit the generalisability of the findings”. Likewise, Hersh and colleagues (2002) conducted a systematic review of the literature to evaluate the efficacy of telemedicine for making diagnostic and management decisions and the authors found that “despite the widespread use of telemedicine in most major medical specialties, there is strong evidence in only a few of them that diagnostic and management decisions provided by telemedicine are comparable to face-to-face care”.

Furthermore, there have been reports and journal articles looking at screening for specific diseases. For example, the Health Technology Assessment report by Karnon and colleagues (2007) looked at different screening models for diseases such as cancer, diabetes and cardiovascular disease and they also reviewed models evaluating antenatal screening programmes. Whilst some of the methodological challenges raised in this report resonate with this thesis, there has only been one mention previously of the use of telemedicine in screening and the challenges associated with the economic evaluation of such technologies [Norum et al, 2007]. Some of the challenges raised by these papers and the challenges identified from the literature review will be discussed further in Chapter 3.

1.8 Focus of the thesis

A framework for economic evaluation is needed to enable decision makers to analyse the costs and benefits of telemedicine services. Many telemedicine initiatives have been funded as special projects, rather than from normal healthcare budgets. The continuation of such projects in routine practice means allocating resources towards telemedicine services that may not only have to be justified on the resulting health benefits which need to be sufficient, but also on the basis that the telemedicine service is an efficient use of scarce resources.

The main contribution of this thesis is to look at some of the problematic issues relating to the economic evaluations of telemedicine. Firstly, a cost-consequences analysis will be presented comparing two different referral methods: telemedicine versus direct referral to specialist hospital for obtaining specialist advice for fetal cardiology. This analysis was undertaken by myself (supervised by Dr Robin Dowie), but the methodology for this study was determined prior to my involvement. The next step is to see what the existing literature says about the costs and benefits of telemedicine and to discuss some of the economic issues associated with telemedicine which were highlighted in the review, particularly those that are relevant to the case study. All earlier reviews are dated now, so I conducted my own literature review. In light of the

literature review, a critique of the case study which is used in this thesis is provided and also the limitations are discussed.

This thesis will then focus on three issues which came to light from the literature review and to see how the analysis provided in the case study can be improved. These issues are not specific to telemedicine, but are perceived as a problem for the evaluation of telemedicine technologies. Firstly, from the literature review the majority of studies identified were non-randomised, and they did not account for selection bias. Selection bias refers to systematic differences in comparison groups. The case study used for the thesis was an observational study, as local clinicians selected the referral mode for women to obtain specialist advice according to local protocols (i.e. selection bias). This thesis will apply the methods which were identified in the literature to reduce selection bias between the two groups of women and to see what impact this has on the costs and effects.

Secondly, the majority of studies identified in the literature review only looked at costs from the perspective of the healthcare service. This neglects costs which fall on the patients and their families. There is a limited understanding of patient costs associated with pregnancy. In this thesis, using various assumptions, costs which pregnant women incur when visiting hospital for their antenatal appointments are calculated and it then estimates what the overall impact is on the total costs of pregnancy with the addition of these extra costs, when comparing a service with telemedicine to a service without telemedicine.

Thirdly, the majority of studies identified in the literature review only looked at costs of telemedicine and not at the benefits of telemedicine. Guidance for health technology assessment from the National Institute of Health and Clinical Excellence (NICE) prefer health benefits to be measured in terms of quality-adjusted life years (QALYs) [NICE, 2008]; so decisions for value for money can be made across different technologies and different disease areas. This thesis will attempt to calculate QALYs using various assumptions for the patient cohort in the case study, by looking at the cost-effectiveness of a screening strategy with telemedicine versus a screening strategy without telemedicine. The results for the cost-effectiveness analysis will be expressed in terms of cost per QALY gained.

Despite a large number of telemedicine trials, little information has been published on the costs and benefits of telemedicine services and the few systematic reviews published on telemedicine have not provided any conclusive evidence as to whether

telemedicine is cost-effective or not. Various studies have shown whether a telemedicine service used in a particular area of health is cost-effective in relation to the current service [e.g. Harley, 2006; Kumar et al, 2006; Daucourt et al, 2006; Peng et al, 2006]. Therefore, there is a need for an informative approach to consider the issues which can arise when conducting an economic evaluation of telemedicine services. For this reason, the main aim of this thesis is to determine the costs and cost-effectiveness of telemedicine compared to conventional care for fetal cardiology; and to look at the three economic issues which are important to telemedicine services in general and to see how these issues can be addressed in order to undertake a 'good' economic evaluation of telemedicine services.

1.9 Case study: The use of telemedicine in providing specialist advice in fetal cardiology

Throughout this thesis the application of telemedicine in fetal cardiology is used as a case study to examine some of the economic issues which may arise with telemedicine services. Almost all pregnant women in the UK are offered an anomaly scan at 18-22 weeks gestation to detect major congenital heart problems or other structural anomalies. For the great majority of women, no defects are found. In the case of a suspected abnormality, women would have to travel to specialist centres for investigation and management and only a few days may elapse between the local decision to refer and the specialist appointment, so travel arrangements have to be made quickly. There are various clinical and resource factors which make perinatal cardiology particularly suitable for telemedicine. Telemedicine offers an alternative referral strategy for fetal cardiology.

The population of interest is all pregnant women at Medway Maritime Hospital, in Gillingham, Kent who have (or should have) had an anomaly scan at 20-22 weeks gestation to detect major congenital heart anomalies. The dataset for this thesis were collected in two parts. The first set was collected from 1st May 2001 to 31st July 2002, and formed part of a wider study which was commissioned by the NHS Information and Communication Technology Research Initiative (ICTRI-1). The project also known as the 'TelePaed' study, examined from the hospital perspective whether the use of telemedicine was cost-effective in obtaining specialist advice in paediatric and fetal cardiology when compared to conventional methods. The second dataset has been collected since the conclusion of the TelePaed project, and the data were collected retrospectively for the time period: 1st May 2005 to 31st July 2006.

The additional data collected for this thesis along with the earlier data will enable a

more appropriate cost-effectiveness analysis of a screening service with telemedicine compared to a screening service without telemedicine to be conducted and also to address the economic issues associated with telemedicine services highlighted earlier.

1.11 Structure of the thesis

The thesis is made up of seven further chapters as set out below.

Chapter 2 introduces the original case study that forms the baseline for this thesis: an economic evaluation of the use of telemedicine in providing specialist advice in fetal cardiology. The chapter provides information about the case study and details the methods used and the results for a cost-consequences analysis comparing two groups of women: those who were assessed via telemedicine and those who were referred directly to obtain advice from a fetal cardiologist.

Chapter 3 steps back and reviews the literature on the costs (and benefits) of telemedicine. The first part of the chapter contains a literature review on the economics of telemedicine and aims to provide a review on the studies which focus on the economics of telemedicine and to see what data were collected within these studies. The second part of the chapter discusses the economic issues associated with telemedicine which were highlighted in the review, particularly those that are relevant to the case study.

In light of the findings from the literature review in Chapter 3, Chapter 4 provides a critique of the TelePaed case study which was presented in Chapter 2 and reflects on some of the problems with the design of the study. This Chapter then outlines the next steps for this thesis and provides details on the three issues (selection bias, patient costs and the measures of benefits) identified in the literature review and how they can be addressed in the context of an economic evaluation of telemedicine services. The following chapters will look at each of these issues in turn and to see how the case study analysis can be improved.

Chapter 5 examines the first of the three issues: that of selection bias and this chapter highlights the methods to minimise selection bias. Some of the methods identified in the literature to control for selection bias are applied to the fetal cardiology dataset to assess what the impact is on mean costs and effects for pregnant women in these two groups.

Chapter 6 extends the economic analysis presented in Chapter 2 by looking at all women (including the non-referred women) who were seen at Medway hospital during the same time period. The economic analysis will also look at the impact over time of telemedicine on both costs and effects (the more recent data will be examined here). The Chapter then addresses the second issue: patient costs, and examines what the overall impact is on the total costs of pregnancy when these additional costs are considered.

Chapter 7 covers the final issue: measures of benefits such as QALYs. Decision analytical modelling is used to look at lifetime costs and QALYs of the introduction of a telemedicine screening service for pregnant women whose unborn babies are at a low risk of congenital heart disease compared to a screening service without telemedicine. This large group of low-risk women are more likely to have a missed cardiac anomaly and the results from the decision model will be expressed in terms of cost per QALY gained.

And finally, Chapter 8 contains a general discussion drawing together the conclusions of the thesis, focusing upon its contribution to the literature, with particular reference to fetal cardiology. This Chapter will also consider the implications of the findings from the thesis in terms of their policy implications and future research priorities.

CHAPTER 2: DETAILED ECONOMIC EVALUATION OF THE 'TELEPAED' PROJECT

2.1 Introduction

Chapter 1 has provided a background to telemedicine applications in healthcare and set out the main focus for this thesis. The aim of this chapter is to provide information on the case study which is used throughout the thesis: an economic evaluation of the use of telemedicine in providing specialist advice in fetal cardiology, also known as the TelePaed project [Dowie et al, 2007; Dowie et al, 2008].

The wider TelePaed study outlined in Dowie et al (2007) is summarised here (see Appendix 9 for published paper). A telemedicine service was set up between the Royal Brompton hospital (RBH) in west London and four district hospitals (DGHs): Basildon, Colchester and Southend hospitals in Essex and Medway Maritime hospital, in Gillingham, Kent. Each of the hospitals' obstetric departments had annual deliveries of over 3,200 and the RBH paediatric cardiologists held outreach clinics at each of the hospitals, monthly at the hospital in Gillingham and every three to four months at the other three hospitals. Telemedicine equipment packages installed in each DGH consisted of a Tandberg videoconferencing system for use with six integrated services digital network (ISDN) lines. Training in use of the telemedicine equipment was provided and the specialists also provided advanced training in heart screening. The obstetricians and paediatricians in each of the district hospitals could choose how they utilised the telemedicine service, and their policies differed. In particular, only one hospital used it for obstetric referrals.

The study set out to compare the costs and outcomes of patients referred to specialists in London via the telemedicine service or by conventional means (face-to-face assessment in London or assessment in outreach clinics). Three patient groups were studied: pregnant women referred for specialist fetal heart assessment after the anomaly scan; newborn babies with a suspected heart problem; and older children and infants referred for cardiac assessment. A total of 504 patients were referred to cardiologists from the four hospitals over the 15-month period, and 117 were referred via the telemedicine service. Within the patient groups, telemedicine was used for 52 of the 248 pregnant women, 17 of the 40 newborn babies, and 48 of the 216 infants and older children.

Most pregnant women (83.1%) were referred for screening the fetus rather than to confirm a suspected anomaly. With the newborn babies, a third of the London

transfers had suspected critical congenital heart disease (CHD); the rate was lower for the telemedicine babies. For the older children, the majority of them (83.7%) were assessed for heart murmurs, most of which were normal. In terms of costs for all patient groups, the telemedicine service was generally more expensive than conventional referrals; this was part due to the technology and its operating costs. After six months, by the time the women had delivered their baby, antenatal care incorporating a teleconsultation was again more costly. This cost differential was mainly due to the different policies followed by the four hospitals over the frequency of scheduling antenatal visits for women in the later months of pregnancy. After six months for the newborns, telemedicine remained the cheaper option and for the infants and older children, there was relatively little difference in the mean costs for the two referral strategies, although telemedicine was cheaper.

The second paper [Dowie et al, 2008], which is the basis for this thesis, focused on the one hospital which used the telemedicine equipment for fetal cardiology: Medway hospital in Gillingham (the study details are outlined in more detail later in this chapter and a summary is provided below; see Appendix 10 for published paper). The paper covered all pregnant women over a 15-month period who were referred for detailed ultrasound examination of the fetal heart with a perinatal cardiologist who was based at the RBH after a routine anomaly ultrasound scan. Criteria for specialist referral included: traditional CHD risk factors such as diabetes and a family history of CHD.

Referred women formed two groups: a telemedicine group where a pre-recorded videoed anomaly scan was relayed to the specialist in the absence of the patient and a direct referral group where women were seen face-to-face for a detailed assessment in London by the specialist. Women were followed up from time of the anomaly scan until they delivered or in a few cases, after termination of pregnancy.

In total, 76 women were referred for a fetal cardiac assessment: 52 women were assessed via the telemedicine link and 24 women were sent directly to London. The ultrasound department adopted the `store-and-forward' method for using the telemedicine service. Sixty-four women were referred for detailed screening and 47 of these were assessed during videoconferencing sessions. The direct referral group were more likely to have had a cardiac abnormality than the telemedicine group. In terms of total costs until delivery, there were no significant differences in the antenatal mean costs for the two cohorts.

This chapter will provide in more detail the methods used and results for a cost-consequences analysis comparing two groups: telemedicine and direct referral groups, to obtain specialist advice for fetal cardiology.

2.2 Congenital heart disease and fetal cardiology

Congenital heart disease has been defined as "a heart condition resulting from an abnormality in heart structure or function that is present at birth" [Peterson et al, 2003]. CHD is one of the leading causes of infant mortality in the UK, representing 26% of all infant deaths; this represents a rate of 1.3 congenital heart deaths per 1,000 live births [ONS, 2005]. In the UK, it is estimated that there are approximately 4,600 babies born each year with CHD - one in every 145 births [Peterson et al, 2003]. Of these, the incidence of complex CHD with serious outcomes such as hypoplastic left heart syndrome in the UK is estimated as 1.5 per 1,000 live births; and for simple heart defects such as small ventricular septal defects the incidence in the UK is estimated as 4.5 per 1,000 live births [Peterson et al, 2003]. However, there are also small proportions of complex CHD cases diagnosed prenatally which are terminated.

Fetal cardiology is concerned with the diagnosis and the management of pregnant women with a fetal cardiac anomaly. The development of ultrasound imaging in the 1980s enabled a programme of screening for CHD in pregnancy to be developed. In England and Wales, almost all women at approximately 18-22 weeks in their pregnancy are screened by trained sonographers using an ultrasound anomaly scan to detect major congenital heart problems or other structural anomalies [NCCWCH, 2008]. No up-to-date information for England and Wales was available during the study on prenatal detection rates for fetal cardiac anomalies [Dowie et al, 2004]. Prenatal detection rates across the UK are approximately 23% [Bull, 1999]; however, these rates vary across the UK. For example, the prenatal detection rate rose from 17% in 1994 to 36% in 1996 in the northern region of England following training [Hunter et al, 2000]; at Liverpool Woman's Hospital the detection rate was 23% in 1997 [Bricker et al, 2000]; whereas, one specialist fetal centre in London reported a higher detection rate of 75% between 1997 and 1999 [Carvalho et al, 2002]. This is important because many of the so-called "high-risk" pregnancies have only a modest increased risk of CHD of around 3% and so do not yield a high number of CHD cases [Carvalho et al, 2002]. As the majority of babies (85%) are born to mothers with no identifiable risk factors [Tiny Tickers Charity, 2008], screening of the low-risk pregnancies is recommended at the anomaly scan to achieve a higher prenatal detection rate for CHD [Sharland, 2004, Simpson, 2009]. Since the conclusion of the TelePaed project there have been changes in screening policies for CHD which now includes nuchal

translucency (NT) screening much earlier in the pregnancy (11-14 weeks). This screening technique was introduced to identify fetuses at high risk of Down's syndrome; however, the NT thickness is correlated with CHD which is independent of the fetal karotype [Hyett et al, 1999] and has a stronger relation with CHD than established CHD risk factors [Simpson, 2009].

For the great majority of women, no defects are found. Detection of a heart defect at the anomaly scan allows parents to choose whether to terminate or to continue with the pregnancy. For those parents who wish to continue with the pregnancy they can be informatively counselled and following delivery, operative interventions can be carefully planned. Not all CHD defects require treatments or surgery. Some mild congenital heart defects repair themselves. However, the majority of heart conditions do require treatment, and this varies depending on the type and complexity of the congenital heart defect.

The long-term prognosis of children born with heart defects depends on the type of CHD. For example, long-term prognosis for patients with two functioning ventricles is better than those with one [Kenny and Stuart, 2009]; and over the past 20 years advances in treatment and surgical techniques, has meant that surgical mortality has decreased from an average of 15% in 1990 to an average of 5% in 2000 [Gibbs et al, 2004].

Infants with serious conditions have improved morbidity levels if their treatment has been planned prenatally [Personal communication with Specialist A, October 2007]. However, there are a few cases of CHD which are not detected prenatally at the anomaly ultrasound scan: these are known as "missed cases" and only when the baby is born is the heart defect found. For example, rare conditions associated with total anomalous pulmonary vein connections can still be missed at the anomaly scan as they lie away from the heart. Subtle, mild or late developing CHD can also be missed at 20 weeks [Tiny Tickers Charity, 2008]. A 'missed' diagnosis delays the opportunity for antenatal intervention and planned delivery.

If a routine anomaly scan reveals a complex life-threatening heart abnormality, the DGH obstetrician will normally refer the mother to a specialist in fetal cardiology (perinatal cardiologist) for further assessment. Such referrals are believed to have a sensitivity of > 80% for a true abnormality [Personal communication with Specialist A, October 2007]. Fetal cardiac referrals may also be made in other circumstances: where an anomaly is detected but is clinically less severe; where there is a suspicion

that an anomaly exists; or where confirmation is needed that the fetal heart is normal, usually because the ultrasound heart images during the scan could not be visualised properly or the parents have an elevated risk for a fetal cardiac anomaly [Dowie et al, 2004].

When women are referred for a fetal cardiac opinion, they usually have to attend a fetal medicine department. These fetal medicine departments are mostly found in tertiary units (such as an obstetric unit or maternity hospital) providing regional services. Unlike paediatric cardiology, there are no networks of fetal cardiology peripheral clinics run by specialists in DGHs. There are 12 paediatric cardiology units in England and perinatal cardiologists from these units provide the fetal cardiology service in these fetal medicine departments.

As there are so few tertiary centres in the UK, many pregnant women make lengthy journeys for specialist assessments. For example, women in North Wales are referred to maternity services in Liverpool, while women in south-east England have to travel to London. Women using public transport face considerable journeys, and for all women there are not only travel costs but other costs to consider, such as babysitting. In most instances, only a few days elapse between the local obstetrician decision to refer the women and the specialist appointment, so travel arrangements have to be made quickly.

2.3 Fetal cardiology and telemedicine

Telemedicine offers an alternative referral strategy for fetal cardiology. Telemedicine allows digitised images of heart from obstetric ultrasound machines to be transmitted electronically [Begg et al, 2001] and the structures of the heart can be visualised with clarity after 18 weeks gestation. The ultrasound images are usually transmitted through ISDN telephone lines.

Teleconsultations can be held either ‘real-time’ with a sonographer conducting an examination on the patient, whilst a remote specialist jointly views the images, or via transmission of pre-recorded videotaped ultrasound images (the ‘store-and-forward’ approach) in the absence of the patient. The few reported evaluations of telemedicine in obstetric services relied on ultrasound images being relayed in real-time [Fisk et al, 1996; Chan et al 2001; Sharma et al, 2003].

Telemedicine can lead to a significant decrease in time to diagnosis compared to sending a patient to a specialist hospital. The telemedicine equipment can be used

both to confirm, and to exclude the diagnosis of fetal heart disease (used as a screening tool for women who are considered to be at risk of having a fetal cardiac anomaly). Lengthy journeys to specialist centres may be avoided.

2.4 Telemedicine case study: TelePaed project

2.4.1 Objective and setting

The main aim of this chapter is to look at the impact of the telemedicine service on the hospital costs of antenatal and maternal care received by women and also on their personal costs. As mentioned previously in 2001, under the auspices of the NHS ICTRI-1 [DoH, 1998], the RBH in London set up a telecardiology service in four DGHs for the provision of specialist advice to clinicians in obstetric and paediatric departments. However, only the district hospital in Kent used the telecardiology service for fetal referrals alongside the existing arrangements for referring women directly to London specialist centres (Kings College Hospital (KCH) and Queen Charlotte's Hospital (QCH)). The specialist fetal cardiologist (Specialist A) was based at both the Royal Brompton and Queen Charlotte's Hospitals.

The number of maternal deliveries at Medway Maritime Hospital has been steadily increasing over the years, for example from: 3,817 (1st April 2001 to 31st March 2002) to 3,982 (1st April 2005 to 31st March 2006) [Department of Health, 2003; Department of Health, 2007]. During 2001/2002 the configuration of the antenatal screening services at Medway hospital included offering anomaly scans to all women at 20-22 weeks gestation. Nuchal translucency screening for Down's syndrome was done around 12-13 weeks gestation and only offered to selected women (women with previous trisomy; multiple pregnancy; or insulin dependent diabetic). The quadruple serum screening test was offered to all women at 15-21 weeks gestation, except those offered the nuchal scan and the risk cut-off level (taking account of serum result and maternal age) was 1 in 300.

The anomaly ultrasound scan was usually performed by the obstetric sonographers who were based in the obstetric units of the DGH. After the introduction of the telemedicine service, pregnant women with isolated cardiac anomalies were referred to QCH (majority of women sent for further assessment) and pregnant women with multi-organ anomalies plus a cardiac anomaly were referred to KCH (only a few women sent for further assessment). The ultrasound department at Medway hospital adopted a 'store-and-forward' approach for using the telemedicine service, as well as continuing to refer patients directly to London. During 10 months, 10 pre-arranged videoconferencing sessions were held and sonographers videoed the heart images of

the fetuses during the anomaly scans and the videos were transmitted during monthly teleconferencing sessions. On one occasion an urgent teleconsultation was held, and on three other occasions, a serious heart abnormality was diagnosed during an anomaly scan performed within 48 hours of a videoconference session, so the images were transmitted for verification by the specialist [Dowie et al, 2008]. Approximately five video recordings were transmitted per session (median 5, range 2 to 8) and each woman's scan, diagnosis and management plan was discussed for about five minutes [Dowie et al, 2008]. If a woman was sent to London for referral, they were given a 45 minute slot: 20 minutes for the specialist scan and a further 20 minutes for counselling.

2.4.2 Patients

A prospective audit, covering all newly presenting eligible pregnant women who underwent an anomaly scan at 20-22 weeks gestation over a 15-month period was conducted and the following women were identified for the TelePaed database:

a) women with increase risk factors for fetal CHD:

- family history of CHD (mother, father, sibling or previous pregnancy);
- pregestational diabetes mellitus;
- treatment for maternal epilepsy or lithium;
- current pregnancy with multiple fetuses; or
- previous pregnancy diagnosed with Down's or other chromosomal malformation.

b) women with an elevated risk for Down's syndrome following a serum test or nuchal translucency ultrasound scan or chromosomal test:

- serum test result was above the risk cut-off level adopted by the hospital (1 in 300);
- women whose nuchal translucency result was above 3.5mm; or
- an abnormal chromosome test result.

c) A representative sample of 'average risk' women whose screening results were within normal limits was obtained by matching. So women in groups a) and b) were matched in terms of their year of birth, parity, month of gestation and calendar month of anomaly scan with average risk women [Mistry et al, 2007].

In total, 408 women were identified in the audit; however, this chapter focuses on the 76 women who were referred to the perinatal specialist. The referred women formed two groups: a telemedicine referrals group (n = 52) and a direct referrals group as a comparator (n = 24). The telemedicine group consisted of women who, at the time of their anomaly scan, were identified for referral according to the department's protocol for high risk women, and women whose fetal heart images during the anomaly scan were abnormal or were poorly visualised. The direct referral group consisted of women who were referred for the same reasons during the three months before the

telecardiology service entered regular use and those who travelled to London, during the weeks between scheduled teleconferencing sessions. The direct referral group also included women whose fetuses were strongly suspected of having a heart defect. A consultation with a fetal cardiologist was held according to Medway hospital protocols, usually within two working days [Dowie et al, 2008].

2.5 Methods

2.5.1 Resource use and unit costs

The economic analysis was undertaken from the perspective of the hospital[1]. Non-direct medical costs (patient travel costs) and indirect costs (loss of pay) were also considered in a separate analysis. The time horizon for the analysis was from the time of the anomaly scan (20-22 weeks gestation) until just after delivery[2] or where applicable after termination of pregnancy.

A project facilitator in the district hospital entered data onto a laptop computer. The audit database recorded demographic, clinical and resource use information. Resource use items for antenatal care from the time of women's anomaly scan until delivery or in some cases after termination of pregnancy included: specialist consultations either face-to-face or through telemedicine, anomaly ultrasound scans and any subsequent antenatal and specialist ultrasound scans, counselling, hospital antenatal and outpatient clinic attendances, prenatal inpatient admissions, terminations of pregnancy, and the status of NHS personnel who were involved in each resource use activity and the time taken (in minutes) for each activity. Resource use items for maternal delivery care included the inpatient admission prior to transfer to labour ward or obstetric theatre, labour/delivery bed day, mode of delivery and the mother's stay on the postnatal ward.

Unit costs in pounds (£) sterling in 2001/2002 prices for the antenatal and obstetric resource items listed above were obtained from the hospital finance department in Medway and the two fetal medicine centres in London. The unit costs obtained from Medway were attributed to resources at the district hospital, and for specialist care in London, average unit costs from the two specialist hospitals were used. The cost for the ultrasound examinations incorporated the sonographer's time as well as consumables, administrative costs, overheads and capital charges. The unit cost for the clinic attendances included nursing supervision and clerical staff time, consumables

[1] The direct costs were from a hospital perspective and GP and community costs were excluded from the analysis

[2] 'Just after delivery' only included costs of labour delivery and any postnatal stay for the mother, no costs were included for the newborns if they required any hospital stay.

and equipment, overheads and capital charges. The cost for the termination of pregnancy was inclusive of stay in the labour ward, and finally, the unit costs for bed days were inclusive of medical and nursing staff time, administrative and clerical staff time, consumables and equipment, overheads and capital charges (see Table 2.1 – these costs are in 2005/2006 prices).

The costs for medical staff time were based on NHS salary scales [Department of Health, 2002a]. Hourly rates of pay were calculated by the number of hours expected to work each year, taking annual leave and bank holidays into account [Netten and Curtis, 2002]. These hourly rates were then applied to the time that it took the relevant medical staff to carry out the task (e.g. telemedicine consultation).

Table 2.1: Antenatal resource items with unit costs

Antenatal resource items	**Unit cost (£) 2005/2006 prices***
Resource use at district hospital	
Antenatal ultrasound scan (anomaly or other)	36.16
Antenatal or outpatient clinic attendance (review clinic)	75.83
Clinicians consulted in clinic (mean time)**	
• Consultant (15 minutes)	10.92
• Doctor (15 minutes)	6.84
• Midwife (15 minutes)	3.83
Termination of pregnancy	856.28
Prenatal maternity bed day	240.32
Labour/delivery bed day	240.32
Mode of obstetric delivery	
• Normal birth	1567.91
• Forceps birth	1256.43
• Ventouse birth	1256.43
• Caesarean birth (without complications)	2253.87
• Home birth	560.84
Postnatal maternity bed day	240.32
Resource use at specialist hospital	
Antenatal ultrasound scan plus consultation with the specialist	64.58
Counselling at specialist hospital (20 minutes)	14.58

*2001/2002 costs were inflated to 2005/2006 prices using inflation indices provided by Curtis (2007)
**Costs for staff time were based on NHS salary scales [Department of Health, 2002a] and Netten and Curtis (2002) and inflated to 2005/2006 prices

In the RBH where the perinatal cardiologist viewed the pre-recorded images, the existing telemedicine suite was already equipped with a Tandberg system that was frequently used. Medway hospital was supplied by the RBH with a Tandberg videoconferencing system (model 2500) incorporating a 384Kb/s computer (CODEC) with camera, ultrasound and video cassette recorder inputs, a pan-tilt zoom camera, loudspeaker with volume control, microphone, and peripheral items [Dowie et al, 2008].

The system was designed for use with ISDN-6 lines (i.e. three pairs of ISDN lines, each line being 64Kb) which were installed in the hospital, and was mounted on a tele-trolley for ease of transfer between the antenatal and neonatal departments. Training was provided for using the telemedicine equipment, as well as, advanced training in fetal heart scanning [Dowie et al, 2008].

Setting up costs

The total cost of the telemedicine equipment including ISDN-6 line installation, equipment maintenance contract and 17.5% value added tax (VAT) was obtained from the telemedicine co-ordinator at the RBH. It was assumed that the expected lifetime of the equipment was five years in line with assumptions in another UK study of telemedicine in outpatient clinics [Jacklin et al, 2003] and an annual discount rate of 3.5% was applied as recommended by the Treasury [HM Treasury, 2003]. In order to calculate a mean cost of the telemedicine equipment per woman, the total cost of the telemedicine equipment (over a lifetime of five years and discounted at 3.5%) was divided by the total number of women assessed by telemedicine during the five-year period (1st July 2001 to 30th June 2006). The total number of Medway women seen during this five-year period was 283 [Personal communication with RBH administrative assistant, July 2008].

Operating costs

Monthly invoices from a commercial telephone company provided details of ISDN-6 line rental, call charges and VAT. Telemedicine co-ordinator costs included time spent by the co-ordinator from the specialist hospital on briefing, installation and training in the district hospital and return travel from the specialist hospital during the initial installation and set-up of the telemedicine equipment. The cost of line rental and call charges and the cost of the telemedicine co-ordinator were divided by the number of women during the TelePaed study period to derive a telemedicine cost per woman (see Table 2.2).

In practice, the per patient costs for the telemedicine equipment will be dependent on the degree to which the equipment is adopted and used in multiple disciplines for other purposes such as for other patient groups or even for meetings, research or administrative uses. The telecardiology service in the current study was also used within the neonatal unit as part of a wider study [Dowie et al, 2007]; however, when calculating the cost for each telemedicine women, the number of newborn patients who also used the telemedicine equipment was not taken into account. As such, we may over-estimate the true cost per patient.

Table 2.2: Mean costs for the telemedicine system in 2005/2006 prices

Components of the fetal telemedicine system	N = 52 women referred*
Fixed costs Telemedicine equipment, ISDN-6 line installation and equipment maintenance contract#	£60.50
Variable costs Telemedicine training and service support	£10.35
ISDN-6 line rental and call charges*	£72.32
Specialist, obstetric sonographer and co-ordinator (mean time = 5 mins per woman)	£8.73
Total mean cost per referred woman	£151.90

The costs for the telemedicine equipment, ISDN-6 line installation and maintenance contract is based on the total cost over five years (useful lifetime of the equipment) and discounted at 3.5% and then divided by the total number of patients seen over 5 years

* In total there were 55 telemedicine consultations, 3 women in the telemedicine group had 2 telemedicine consultations each.

Finally, the unit costs[3] applied to the 2001/2002 financial year, but for purposes of this thesis and to be in line with the published paper they have been inflated to 2005/2006 prices using Pay and Prices indices [Curtis, 2007].

2.5.1.1 Missing resource use information

Labour delivery mode, length of time in labour ward and length of time in postnatal ward were missing due to censoring from the audit database for 14.7% (60/408) patients. Of these 60 women, 1 was from the direct referral group and 11 women were from the telemedicine group. This was because the women had not given birth before the fieldwork had finished. Where information on labour delivery mode and delivery place were missing, the values for these items were estimated using matching. The matching was done within each risk group and the matching took into account four factors: mother's age at anomaly scan; gestation in weeks at anomaly scan; parity and the number of fetuses. An attempt was made to match on all four variables; when this was not possible, matching was based on the next three variables. If there was more than one equal match, then the most frequent delivery mode and delivery place was used. Where information on lengths was missing (continuous variables), the values for these items were imputed using regression based imputation. Standard regression analyses provide estimates of the missing data conditional on complete variables in the analysis [Briggs et al, 2003]. So within each group, the following variables: mother's

[3] One point to note is that these unit costs differ slightly from the published paper [Dowie et al, 2008] and there are two reasons for this: 1) the telemedicine equipment cost per patient is different because of the annual number of patients is different (for the paper the annual equivalent cost was divided by 52, whereas for this thesis the total cost for five years was divided by the total number of patients seen over the 5 years); and 2) the Pay and Prices Index inflator used to inflate prices from 2001/2002 prices to 2005/2006 prices is slightly different. For the paper we used an estimate [Curtis and Netten, 2006] and for this thesis, the observed figure is used [Curtis, 2007].

age at anomaly scan; gestation in weeks at anomaly scan; parity and the number of fetuses were used as predictor variables to predict missing length times [Morris, 2005].

2.5.2 Effectiveness

The main measure for effectiveness was the detection of cardiac abnormalities before birth. The project facilitator noted the findings on the audit database from the anomaly scan: 1) a normal heart or 2) an abnormal heart. Further diagnosis of the heart was confirmed at either the telemedicine consultation or at the face-to-face meeting with the fetal cardiologist in London. If a cardiac anomaly was detected, this gave the women the opportunity to be informatively counselled and to make an informed decision on whether or not to continue with the pregnancy.

2.5.3 Comparative analyses

In the comparative analyses for this chapter, all the referred women had an elevated risk factor for CHD and/or Down's syndrome and were categorised into two groups (see Figure 2.1 below):

1. Women who were referred to a specialist via telemedicine (telemedicine group); and
2. Women who were referred directly to see a specialist face-to-face (direct referral group).

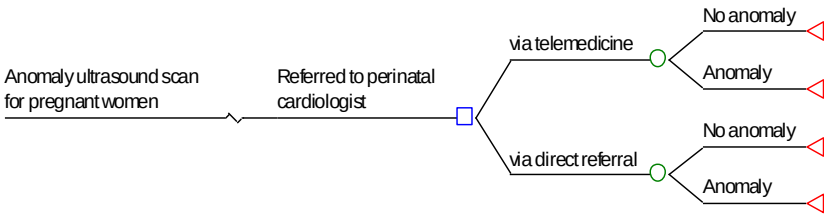


Figure 2.1: Care pathway for pregnant women referred to perinatal cardiologist

A mean cost per women in each of the two groups listed above was estimated for the following scenarios: 1) for events from the time of anomaly scan until delivery (antenatal care spanned the second and third trimesters until delivery); 2) for events during maternal delivery; and 3) for events for total obstetric care – the summation of antenatal and delivery costs (scenarios 1 and 2) [Information on scenarios 2 and 3 are not in the published paper]. Given the short time frame, costs (and effects) were not discounted.

As the cost data were skewed, bootstrapping was used whereby the distribution of costs are generated by repeated sampling of the data (to stabilise the mean), with replacement and, in the absence of any other data from the population, gives a guide

to its distribution [Manly, 1997]. Bootstrapping was performed using Stata version 10 [StataCorp, 2007] by taking 5,000 iterations of the data.

2.5.4 Sensitivity analysis

The main focus of the sensitivity analysis was to look at what the impact would be on the mean costs of the telemedicine group, if the cost of the telemedicine equipment changes. For the base case analysis, the discount rate used for the telemedicine equipment was 3.5% and a lifetime of 5 years was assumed. One-way parameter sensitivity analyses were performed on the unit cost of the telemedicine equipment, by changing the discount rate and the length of the telemedicine equipment lifetime. Depending on the lifetime of the telemedicine equipment, the number of patients seen during the 5-year period was pro-ratad. So the annual number of patients seen each year was 57 (total number of patients/equipment lifetime in years = 283/5).

2.5.5 TelePaed postal questionnaires

During the study postal questionnaires were sent to pregnant women who were eligible for inclusion in the study database between October 2001 and July 2002. The project facilitator posted the questionnaire package at least two weeks after the women had undergone the anomaly scan. The women were asked to return the completed questionnaire to the project office at Brunel University, using a pre-paid, pre-addressed envelope. Each questionnaire had a code number entered at the front. A questionnaire pack was sent to the following women:

- Women who have a higher than average risk of having a baby with a cardiac abnormality;
- Women of normal risk, but who are matched with the women at risk of fetal CHD and women with an elevated risk of Down's syndrome (see section 2.4.2); and
- Women of normal risk, but at the anomaly scan are suspected of having a baby with a cardiac abnormality.

These questionnaires included a section on patient costs and also included two health status instruments - the EQ-5D health status questions and the visual analogue scale [Brooks, 1996], and the Hospital Anxiety and Depression Scale [Zigmond and Snaith, 1983]. Respondents were sent a follow-up questionnaire after three months. For both the initial and follow-up questionnaires, one reminder questionnaire was posted three weeks later where necessary.

2.5.5.1 Patient costs

The questions for eliciting patient costs included: mode of travel to and from hospital; journey distance; travel expenditure; time spent in travelling to and from hospital; time

spent at hospital; activities foregone; loss of earnings; and out-of-pocket expenses (e.g. childcare). The women were also asked whether they were accompanied on their hospital visit, and what these companions would have been doing if they had not visited the hospital.

To calculate travel costs, for women who travelled by train, bus or taxi the actual cost supplied by them was used (inflated to 2005/2006 prices) [Curtis, 2008]. For women who travelled by car, postcode data [AA website, 2008a] were used to calculate the 'road' distance (in miles) between the home address and the district or specialist hospital. The motoring costs published by the Automobile Association (AA) [AA website, 2008b], were applied to the mileage. The cost per mile of travel (44 pence) was built on the assumption that patients had an unleaded petrol fuelled car with an engine size of 1101-1400cc and travelled 10,000 miles a year. This assumption also allowed for fixed costs, depreciation and running costs for 2005-2006. Also, if the women incurred car parking charges, this was also included in the travel cost (inflated to 2005/2006 prices) [Curtis, 2008].

The additional costs which patients may have also incurred when attending hospital were also taken into account. If women incurred costs for childcare or care of other dependents the actual cost stated by them was used (inflated to 2005/2006 prices) [Curtis, 2008]. If women were in paid work and lost pay when attending hospital, if they stated the amount of pay lost whilst they attended hospital, this figure was used. If not, then the opportunity cost of time lost from work was estimated from the mean gross weekly wage rate for women in Great Britain in April 2002, which was £383.40 [ONS, 2002]. From this figure, an hourly rate of £6.65 was estimated from the gross salary, minus 35% tax, pensions and national insurance contributions. This rate was used where women lost pay to attend hospital. Where the woman took annual leave or her appointment was outside work time, her time was valued at 40% of the mean female wage rate (£2.66 per hour) [Bricker et al, 2000]. These hourly costs were inflated to 2005/2006 prices [Curtis, 2008]. Loss of pay from companions was not taken into account, because not enough information was provided.

To calculate the time spent in the hospital by each pregnant woman, the actual time (in minutes) supplied by each woman was used. To calculate the time spent by each woman in travelling (regardless of which mode of transport they travelled to the hospital) time data supplied by AA [AA website, 2008a] were used to calculate the time (in minutes) between the home address and the district or specialist hospital. The travel times and distances published by the AA [AA website, 2008a] are calculated on

the basis of the driver driving at the maximum speed limit for each type of road (e.g. residential road at 30 miles per hour and on motorways at 70 miles per hour) [Personal communication with AA, December 2008].

2.5.5.2 Health Status instruments

The Hospital Anxiety and Depression Scale (HADS) is an easy-to-use measure of a person's present state of mind [Zigmond and Snaith, 1983]. The questionnaire provides separate measures of two dimensions: anxiety and depression. Each dimension has 7 items and the total scores for each dimension can range from 0 to 21. For each scale, a score below 8 indicates that the person is within the normal range, 8-10 indicates a possible disorder of the relevant mood and above 10 indicates a probable disorder of the relevant mood.

EuroQol EQ-5D is a standardised instrument for use as a measure of health-related utilities and is designed for self-completion by respondents. It consists of two parts: the descriptive part and the visual analogue scale. The EQ-5D descriptive system comprises 5 attributes of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) [Brooks, 1996]. Each attribute has three levels (no problems, some problems and extreme problems), generating a total of 243 possible health states, to which 'unconscious' and 'dead' have been added for a total of 245. Preferences for the scoring function were measured originally with the time trade-off technique [Dolan et al, 1996]. The scores lie on a value scale where 0 = dead and 1 = full health. The EQ-5D visual analogue scale (VAS) records the respondents' self-rated health status on a 0-100 scale where 0 is the worst state you can imagine and 100 is the best state you can imagine.

2.5.6 Statistical analysis and tests

Statistical analyses were conducted using Stata version 10 [StataCorp, 2007]. If distributions were normal, means, standard deviations (SD) and ranges are presented, and chi-squared (or Fisher's exact) and t tests were conducted (A Fisher's Exact test is conducted instead of a chi-squared test when expected cells are less than 5 for 2 by 2 comparisons). If distributions are not normal, means, SDs, medians and inter-quartile ranges (IQRs) are presented and Wilcoxon rank-sum tests conducted. All statistical tests were two-sided unless otherwise stated. A p-value ≤ 0.05 was considered to be statistically significant.

2.6 Results

2.6.1 Patient sample and demographics

Over the 15 months, 76 women were referred for a fetal cardiac assessment. Fifty-two women were assessed via the telemedicine link and 24 women were sent directly to London for face-to-face assessment. For the direct referral group, 13 women were referred before the telecardiology service entered regular use and 11 women over the following 10 months.

Table 2.3 shows the demographic characteristics for the referred women. The mean age for women in each group was not statistically different ($p = 0.560$); there were no significant differences between the groups in terms of parity ($p = 0.267$); one women in the direct referral group was expecting a multiple birth; half of the women in each group had their anomaly scan at 21 weeks; and over 50% of women gave birth between 38 to 40 weeks gestation. There were more diabetic women and women on epilepsy treatment in the telemedicine group compared to the direct referral group.

Table 2.3: Demographic characteristics and risk factors of pregnant women[±]

	Telemedicine group (n = 52)	Direct referral group (n = 24)
Maternal age: Mean (SD)	28.7 (6.1)	29.6 (6.8)
Range	17 to 44	20 to 44
Parity		
Primaparous	19 (36.5%)	12 (50.0%)
Multiparous	33 (63.5%)	12 (50.0%)
Type of pregnancy		
Singleton	52 (100.0%)	23 (95.8%)
Multiple	0 (0.0%)	1 (4.2%)
Gestation at anomaly scan*		
Mean (SD)	20.8 (0.9)	21.8 (3.0)
Range	18 to 23	19 to 33
≤ 19 weeks	4 (7.7%)	1 (4.2%)
20 weeks	12 (23.1%)	4 (16.7%)
21 weeks	28 (53.8%)	12 (50.0%)
22 weeks	7 (13.5%)	5 (20.8%)
≥ 23 weeks	1 (1.9%)	2 (8.3%)
Gestation at birth		
N[#]	50	20
Mean (SD)	38.1 (2.8)	37.6 (2.2)
Range	29 to 42	33 to 41
≤ 30 weeks	2 (4.0%)	0 (0.0%)
31-35 weeks	5 (10.0%)	3 (15.0%)
36-37 weeks	6 (12.0%)	6 (30.0%)
38-39 weeks	16 (32.0%)	7 (35.0%)
40 weeks	16 (32.0%)	2 (10.0%)
≥ 41 weeks	5 (10.0%)	2 (10.0%)
Does the mother have diabetes?		
Yes	10 (19.2%)	3 (12.5%)
No	42 (80.8%)	21 (87.5%)
Does the unborn baby have a high risk of Down's syndrome or does the mother have an elevated serum risk?		
Yes	3 (5.8%)	5 (20.8%)
No	49 (94.2%)	19 (79.2%)
*Is the mother on any anti-epilepsy drug therapy?***		
Yes	11 (21.2%)	1 (4.2%)
No	41 (78.8%)	23 (95.8%)
Does the mother have a family history of CHD or had a previous pregnancy with an anomaly?		
Yes	19 (36.5%)	8 (33.3%)
No	33 (63.5%)	16 (66.7%)

[±] Statistical tests conducted: t tests for age, gestation at anomaly scan and birth; chi-squared tests for parity, previous pregnancy with anomaly; Fisher's Exact tests for type of pregnancy, diabetes, Down's syndrome risk, and anti-epilepsy drug therapy;
[#] Not including women who had a termination of pregnancy; * $p < 0.05$; ** $p < 0.1$

2.6.2 Effectiveness and clinical results

Table 2.4 shows the clinical circumstances for using telemedicine: for screening women who met the department's protocol for identifying cardiac risk and for establishing the severity of a diagnosed abnormality. Of the 64 women referred for screening, 47 (73.4%) were assessed via telemedicine. The direct referral group were more likely to have had a cardiac abnormality (7/24 versus 5/52, Fisher's exact test, p = 0.043) due to the higher prevalence. There were no missed diagnoses amongst the two groups of women.

Table 2.4: Clinical circumstances in which the telemedicine service was used

Clinical circumstances	Telemedicine group (n = 52)	Direct referral group (n = 24)	All referrals
Screening - fetus presumed to be normal	47 (73.4%)	17 (26.6%)	64 (100.0%)
Confirmation of a cardiac abnormality			
N	5 (41.7%)	7 (58.3%)	12 (100.0%)
Mean	0.0962	0.2917	n/a
Total per referral mode	52 (68.4%)	24 (31.6%)	76 (100.0%)

2.6.3 Resource use analysis and hospital utilisation patterns for a sample population

Table 2.5 shows the resources used by each of the two referral groups. One point to note is that for the following items of resource use: prenatal maternity bed day (prior to transfer to labour ward); mode of obstetric delivery; and postnatal bed day include missing values which have been imputed (see section 2.5.1.1).

All patients had at least one ultrasound scan, usually this was the anomaly ultrasound scan, and the mean number of scans for both groups was 2.8. Almost all patients had an outpatient or antenatal clinic visit; the few patients who did not have a clinic visit during the second or third trimester were patients who had their pregnancies terminated. The telemedicine women had slightly more clinic visits than the direct referral women, a difference which was not significant (11.4 vs. 10.5, p = 0.564). About a third of patients in the telemedicine group had a prenatal admission during the antenatal period; this rate was much lower for the direct referral group. Over half of the women in each group had a normal birth; the next most common form of labour delivery was a caesarean birth. Over two-thirds of women in each group had a postnatal stay and the mean range of postnatal stay was from 3.9 days (direct referral women) to 4.5 days (telemedicine women). All women in the direct referral group and 6 of the 52 telemedicine women had a specialist ultrasound scan, and only a proportion of women who had a specialist scan received counselling. This was mainly for pregnant women whose scans indicated abnormal findings.

Table 2.5: Resource components and the number of women who used the resource items±

Mean per patient	Telemedicine group (n = 52)	Direct referral group (n = 24)
At district hospital		
1. Antenatal ultrasound scans	52 (100%)	24 (100.0%)
Mean number of scans (SD)	2.8 (1.5)	2.8 (1.3)
Range	1 to 6	1 to 5
2. Antenatal or outpatient clinics	50 (96.2%)	20 (83.3%)
Mean number of clinics (SD)	11.4 (6.7)	10.5 (5.6)
Range	1 to 35	1 to 24
*3. Termination of pregnancy***	2 (3.9%)	4 (16.7%)
4. Prenatal maternity bed day		
a) During antenatal period	19 (36.5%)	4 (16.7%)
Mean number of days (SD)	3.3 (2.5)	2.5 (1.3)
Median	2.0	2.5
Inter-quartile range	1.0 to 5.0	1.5 to 3.5
*b) Prior to transfer to labour ward***	17 (34.0%)	4 (16.7%)
Mean number of days (SD)	0.9 (0.8)	2.4 (1.9)
Median	0.4	2.5
Inter-quartile range	0.3 to 1.0	0.8 to 4.0
5) Mode of obstetric delivery		
Normal birth	32 (64.0%)	11 (55.0%)
Forceps birth	0 (0.0%)	0 (0.0%)
Ventouse birth	1 (2.0%)	1 (5.0%)
Caesarean birth#	16 (32.0%)	8 (40.0%)
Home birth	1 (2.0%)	0 (0.0%)
6) Postnatal maternity bed day	39 (78.0%)	18 (90.0%)
Mean number of days (SD)	4.5 (3.9)	3.9 (2.5)
Median	4.0	3.0
Inter-quartile range	2.0 to 5.4	2.0 to 6.0
At specialist hospital		
*1. Specialist ultrasound scans**	6 (11.5%)	24 (100.0%)
Mean number of scans (SD)	1.2 (0.4)	1.1 (0.3)
Median	1	1
Inter-quartile range	1 to 1	1 to 1
2. Counselling	5 (9.6%)	9 (37.5%)

± Statistical tests conducted: Wilcoxon rank sum tests for scans and bed days; t tests for clinics; Fisher's Exact tests for termination of pregnancy and mode of delivery;
Caesarean birth without complications; * $p < 0.05$; ** $p < 0.1$

2.6.4 Bootstrapped cost results in 2005/2006 prices

The bootstrapped cost results presented below in Table 2.6 are shown for the three time periods (see section 2.5.3). The mean costs per women for events from the time of anomaly scan until delivery show that the mean costs for the telemedicine group were higher than the direct referral group (£1,485 vs. £1,120), and this difference in costs between the two groups was approaching the level of statistical significance ($p = 0.060$). This was because these costs not only incorporated the cost of the telemedicine service (the marginal extra cost to the hospital of a teleconsultation over a specialist consultation was £87.32), but also each telemedicine woman had on average more antenatal attendances. This was primarily due to two reasons: firstly, because some women in the telemedicine group were scanned earlier in the second trimester of

pregnancy compared to women in the direct referral group; and secondly, due to the underlying medical conditions of the women in each group. There were significantly more diabetic women and women on treatment for epilepsy in the telemedicine group than the direct referral group: 21 (40.4%) in the telemedicine group and 4 (16.7%) in direct referral group (chi-square test: $\chi_1^2 = 4.18$, p = 0.041).

Table 2.6: Bootstrapped hospital costs per group in 2005/2006 prices±

	Telemedicine group (n = 52)	Direct referral group (n = 24)
From time of anomaly scan up until delivery (Total antenatal costs)		
N**	52	24
Mean (SD)	£1,485 (£118)	£1,120 (£109)
95% CI	£1,253 to £1,717	£906 to £1,334
During maternal delivery		
N	50	20
Mean (SD)	£2,993 (£150)	£3,246 (£185)
95% CI	£2,699 to £3,287	£2,884 to £3,607
Total costs of pregnancy		
N	52	24
Mean (SD)	£4,363 (£238)	£3,825 (£311)
95% CI	£3,896 to £4,830	£3,215 to £4,435

± Statistical tests conducted: t tests for all costs; ** p < 0.1
CI = confidence interval

Table 2.6 also shows the mean costs per group during delivery. The cost for the telemedicine group was slightly lower than the direct referral group, although not significantly so (p = 0.351). The majority of births were delivery by 'normal means', however, 40% of women in the direct referral group had caesarean sections compared to only 32% in the telemedicine group. This higher percentage maybe due to the higher number of abnormal hearts in the direct referral group and as caesarean section was the most costly mode of delivery this accounted for the slightly higher mean delivery cost for the direct referral group compared to the telemedicine group. Finally, the table also shows the total bootstrapped mean costs per group of pregnancy from the second trimester to just after delivery (or in some cases after termination of pregnancy). The overall costs for the telemedicine group were approximately £540 greater than direct referral group, a difference which was not significant (p = 0.202).

2.6.5 Sensitivity analyses results

Table 2.7 shows the impact on costs when changing the discount rate and lifetime of the telemedicine equipment. For this sensitivity analysis, it was assumed that these changes would not have an effect on the detection of cardiac anomalies. Costs are shown for the telemedicine group. Depending on the lifetime of the equipment, and assuming a 3.5% discount rate the total antenatal costs ranged from £1,531 (3 years lifetime) to £1,451 (10 years lifetime) for the telemedicine group. These one-way sensitivity analyses confirmed the robustness of the results from the main analyses.

Even if the discount rate and the lifetime of the equipment were to change and this changed the costs for the telemedicine group, this would not change the overall conclusion.

Table 2.7: Impact on mean costs when changing discount rate and lifetime of telemedicine equipment

Discount rate & lifetime of equipment	Telemedicine group (n = 52)	
	Total antenatal costs	Total costs of pregnancy
Base case		
3.5% & 5 years	£1,485	£4,363
Sensitivity analyses		
0% & 5 years	£1,489	£4,368
6% & 5 years	£1,482	£4,360
3.5% & 3 years	£1,531	£4,410
3.5% & 4 years	£1,503	£4,381
3.5% & 6 years	£1,474	£4,352
3.5% & 7 years	£1,465	£4,344
3.5% & 8 years	£1,459	£4,337
3.5% & 9 years	£1,455	£4,333
3.5% & 10 years	£1,451	£4,329

2.6.6 TelePaed postal questionnaires results

Questionnaires were sent to 235 out of the 408 women in the audit. Information on travel arrangements and expenditure when visiting the hospital for an antenatal appointment were received from 125 women, giving a response rate of 53.2%. This information was used to calculate travel costs and any other additional costs. All but 6 journeys were made to Medway district hospital: 23 questionnaires where from women whose anomaly scans were videoed for a telemedicine assessment and 96 questionnaires where from women who were managed locally in the district hospital (there were no statistical differences in demographic characteristics: age, parity, gestation at start, between the telemedicine women and the women who were managed locally).

Table 2.8 shows the representativeness of the responders was established by comparing their demographic characteristics with the characteristics of the other women who were referred for a specialist assessment. There were no significant differences between the respondents and the other eligible women in terms of demographic characteristics and risk factors as shown in Table 2.8 (apart from gestation at the anomaly scan; this difference between the two groups was approaching a level of statistical significance).

a) Demographic characteristics of responders

Table 2.8: Survey respondents compared with the rest of the observed sample[±]

Demographic characteristics & risk factors	Respondents (n = 29)	Other eligible women (n = 47)
Age		
Mean (SD)	27.1 (5.4)	28.5 (6.7)
Range	16 to 38	17 to 43
Parity		
Primaparous	10 (34.5%)	21 (44.7%)
Multiparous	19 (65.5%)	26 (55.3%)
Number of fetus		
Single	29 (100.0%)	46 (97.9%)
Multiple	0 (0.0%)	1 (2.1%)
*Gestation at anomaly scan***		
Mean (SD)	20.6 (1.1)	21.4 (2.2)
Range	18 to 22	19 to 33
Gestation at birth [#]		
Mean (SD)	37.8 (2.5)	38.0 (2.7)
Range	32 to 42	29 to 41
Does the patient have diabetes?		
Yes	7 (24.1%)	6 (12.8%)
No	22 (75.9%)	41 (87.2%)
Does the unborn baby have a high risk of Down's syndrome?		
Yes	1 (3.4%)	2 (4.3%)
No	28 (96.6%)	45 (95.7%)
Does the patient have an elevated serum result?		
Yes	0 (0.0%)	5 (10.6%)
No	29 (100.0%)	42 (89.4%)
Does the patient have a family history of CHD?		
Yes	11 (37.9%)	14 (29.8%)
No	18 (62.1%)	33 (70.2%)
Has the patient had a previous pregnancy with an anomaly?		
Yes	1 (3.4%)	1 (2.1%)
No	28 (96.6%)	46 (97.9%)

[±] Statistical tests conducted: t tests for age, gestation at anomaly scan and birth; chi-squared tests for parity, diabetes and family history of CHD; Fisher's Exact tests for number of fetuses, Down's syndrome risk, elevated serum result and previous pregnancy with anomaly.
[#] Not including women who had a termination of pregnancy; ** $p < 0.1$

For this section of the thesis, the results of the 6 women in the direct referral group, and the 23 women who were in the telemedicine group for whom a completed initial questionnaire was received will be compared. A follow-up questionnaire (sent after 3 months) was returned by 18 women in the telemedicine group and 3 women in the direct referral group.

Table 2.9 shows the demographic attributes of the responders, depending on which hospital they travelled to for that appointment (London or Medway hospital). There were no significant differences between the two groups of respondents in terms of marital status, United Kingdom as country of birth, educational experience and whether they had ever smoked. However, the direct referral group were significantly older ($p = 0.010$).

Table 2.9: Characteristics of responders±

Characteristics	Telemedicine group (n = 23)	Direct referral group (n = 6)
*Age**		
Mean (SD)	25.8 (4.6)	32.0 (6.1)
range	16 to 34	21 to 38
Marital status		
Married or living as married	19 (82.6%)	6 (100.0%)
Widowed	0 (0.0%)	0 (0.0%)
Divorced	0 (0.0%)	0 (0.0%)
Single	4 (17.4%)	0 (0.0%)
Were you born in the UK?		
Yes	21 (91.3%)	6 (100.0%)
No	2 (8.7%)	0 (0.0%)
Did you continue education after the minimum school leaving age (16)?		
Yes	15 (65.2%)	5 (83.3%)
No	8 (34.8%)	1 (16.7%)
Have you ever smoked?		
Yes	9 (39.1%)	3 (50.0%)
No	14 (60.9%)	3 (50.0%)

± Statistical tests conducted: t tests for age; chi-squared tests for smoking; Fisher's Exact tests for marital status, UK birth and continuing education; ** $p < 0.1$

b) Mode of transport

The majority of women attending an antenatal appointment at Medway hospital (87.0%) travelled by car; 4.3% travelled by bus or train; and 8.7% went by taxi or walked. For those six patients who travelled to London for an antenatal appointment, three travelled by car and the other three travelled by bus or train.

c) Journey distance

The length of each return journey was calculated using postcode data to establish the 'true' distance from the woman's home to the hospital. Table 2.10 shows the return journey length to and from hospital. Patients who travelled to London had significantly longer journeys than those women attending the local hospital ($p < 0.001$).

Table 2.10: Length of return journey (in miles) to and from and hospital for one hospital visit±

	Telemedicine group (n = 23)	Direct referral group (n = 6)
Mean (SD)*	8.2 (7.7)	88.5 (8.8)
Median	7.0	89.6
IQR	3.8 to 8.6	83.2 to 90.8

± Statistical test conducted: Wilcoxon rank-sum; * $p < 0.05$

d) Duration of time

Table 2.11 shows the duration of time spent on the different activities associated with the antenatal appointment. Women who had to travel to London naturally had a significantly longer journey time than those attending the local hospital ($p < 0.001$). The great majority of antenatal attendances lasted between 50 and 90 minutes for women who were seen in the local hospital and the median time was about an hour. For those women in London, they spent an average of 90 minutes in the hospital. This

time difference between the two groups, was not statistically significant (p = 0.291). The average time spent on antenatal visits for those who travelled to the local hospital was just over 90 minutes, whereas for women who had to travel to London for an antenatal appointment, the time spent on antenatal visits was substantially longer, at just over 4½ hours (p < 0.001).

Table 2.11: Duration of times (in minutes) for one hospital visit±

	Telemedicine group (n = 23)	Direct referral group (n = 6)
Duration of return journey (in minutes)		
Mean (SD)* Median IQR	21 (15) 18 12 to 26	184 (7) 185 182 to 190
Duration of time in hospital (in minutes)		
Mean (SD) Median IQR	75 (43) 60 50 to 90	99 (55) 90 60 to 145
Total time spent on visit (in minutes)		
Mean (SD)* Median IQR	96 (50) 87 56 to 112	283 (57) 274 232 to 335

± Statistical tests conducted: Wilcoxon rank-sum; * p < 0.05

e) Total travel costs of hospital visits

Costs associated with journeys were estimated from the mileage rates calculated for the return car journeys and car parking fees, and from actual fares reported in the questionnaires for other modes of transport (e.g. rail and bus fares). Table 2.12 shows that the cost of travel for pregnant women who had to travel to London was approximately £35 - nearly seven times that of women who travelled to the local hospital for their antenatal appointment, this difference was statistically significant (p < 0.001).

Table 2.12: Bootstrapped costs of one hospital visit (in 2005/2006 prices)±

	Telemedicine group (n = 23)	Direct referral group (n = 6)
Total travel costs		
Mean (SD)* 95% CI	£5.01 (£0.81) £3.42 to £6.59	£35.16 (£4.64) £26.08 to £44.25
Additional costs		
N Mean (SD) 95% CI	6 £38.87 (£13.52) £12.37 to £65.37	2 £34.85 (£16.73) £2.06 to £67.65
Total costs (travel plus additional costs)		
Mean (SD)* 95% CI	£15.15 (£5.12) £5.12 to £25.18	£46.78 (£10.79) £25.63 to £67.93

± Statistical tests conducted: Wilcoxon rank-sum for all costs; * p < 0.05

f) Additional costs of hospital visits

Two-thirds of respondents (65.5%) were accompanied by another adult on their hospital visit. The questionnaire asked both the patient and companion what they would have been doing if they were not attending hospital; for those who reported paid employment, what arrangements had been made to be absent from work and what

arrangements had been made if they had children or other dependents who needed looking after whilst they attended the hospital.

If they had not been attending hospital, 41.4% of the respondents would have been in paid employment; just under half the respondents (48.3%) reported they would have been undertaking household duties; and 10.3% of patients had children or adults to look after. Of those, 12 pregnant women who were in paid employment, 8 (66.7%) of them received time off with pay for their antenatal appointment (this cost being borne by the employer); 2 (16.7%) patients took annual leave; 1 (8.3%) woman had to make time up; and the other woman took time off work with loss of pay.

Of the 19 companions who accompanied women on their hospital visit, 18 of them (94.7%) would have been in paid employment; and the remaining companion would have been undertaking household duties. Of the 18 companions who were in paid employment, only 14 of them responded to the question what arrangements had been made to be absent from work. Five (35.7%) companions took annual leave, three (21.4%) companions had to make the time up, a further two (14.3%) companions had time of work without loss of pay, three (21.4%) companions had time off work with loss of pay and finally, one companion (7.1%) went outside work hours.

In total, 55.2% (16 respondents) had children or other dependents that needed looking after. Of these 16 respondents, four (25.0%) of them said that someone else had to take time off work to look after them and two of these four respondents paid someone else to look after them.

Table 2.12 also shows the total amount of additional costs for each patient, which included the loss of pay for the patient and any incidental expenses (e.g. child care), these costs have been inflated to 2005/2006 prices. Only eight (27.6%) women reported additional costs incurred during their hospital visit (two women in the direct referral group), however, there was no statistical difference in additional costs between the two groups of women ($p = 1.000$). If these costs were averaged out across all women who responded to the questionnaires in each group, for the direct referral group this additional cost was £11.62 for the 6 women and for the telemedicine group this additional cost was £10.14 for the 23 women.

g) Total costs of hospital visits

After accounting for the travel costs and any additional costs by the women, the mean total costs for women who had to travel to London were significantly greater than women who were assessed via telemedicine (£46.78 vs. £15.15, p = 0.006).

h) Health status instruments results

Table 2.13 shows the differences in the results from the HADS questionnaire between the two groups, where women were asked about their level of anxiety and depression in the past week.

Table 2.13: Anxiety and depression in the past week ±

	Telemedicine group	**Direct referral group**
Initial questionnaire	**n = 23**	**n = 6**
Anxiety scores		
Mean (SD)	8 (3)	7 (3)
Median	9	8
IQR	5 to 10	5 to 9
Anxiety categories		
No anxiety (0-7)	8 (34.8%)	2 (33.3%)
Possible anxiety (8-10)	10 (43.5%)	4 (66.7%)
Probable anxiety (11-21)	5 (21.7%)	0 (0.0%)
Depression scores		
Mean (SD)	5 (3)	2 (2)
Median	5	3
IQR	1 to 7	1 to 3
Depression categories		
No depression (0-7)	19 (82.6%)	6 (100.0%)
Possible depression (8-10)	3 (13.0%)	0 (0.0%)
Probable depression (11-21)	1 (4.4%)	0 (0.0%)
Follow-up questionnaire	**n = 18**	**n = 3**
Anxiety scores		
Mean (SD)	7 (3)	5 (3)
Median	7	7
IQR	5 to 9	2 to 7
Anxiety categories		
No anxiety (0-7)	11 (61.1%)	3 (100.0%)
Possible anxiety (8-10)	4 (22.2%)	0 (0.0%)
Probable anxiety (11-21)	3 (16.7%)	0 (0.0%)
Depression scores		
Mean (SD)	6 (4)	2 (2)
Median	5	3
IQR	2 to 8	0 to 4
Depression categories		
No depression (0-7)	11 (61.1%)	3 (100.0%)
Possible depression (8-10)	4 (22.2%)	0 (0.0%)
Probable depression (11-21)	3 (16.7%)	0 (0.0%)

± Statistical tests conducted: Wilcoxon rank-sum for all anxiety and depression scores; Fisher's Exact tests for all anxiety and depression categories

In the initial survey, the telemedicine women recorded slightly higher levels of anxiety than the direct referral women, although this finding was not statistically significant. These findings applied to both the mean anxiety scores and the anxiety categories. After three months, although the telemedicine women were still comparatively more anxious than the direct referral women, the anxiety levels for both groups had eased

(differences were not statistically significant over time). In both, the initial and follow-up questionnaires, the telemedicine women were slightly more depressed than the direct referral women, although not significantly so.

Soon after having their anomaly scan, the telemedicine women in general gave their health a lower rating on the EQ-5D VAS than the direct referral women, although this difference was not statistically significant (p = 0.302) (see Table 2.14). This may have been because there were more diabetic women in the telemedicine group compared to the direct referral group. Three months later, even though the telemedicine women still gave a lower rating for the health (EQ-5D VAS) compared to the direct referral group, this value was higher than the initial survey.

In the initial survey, when the women rated the statements on mobility, self-care, usual activities, pain or discomfort, and anxiety or depression, the telemedicine group had a higher overall mean tariff score than the direct referral group. Although, this difference between the two groups was not statistically significant (p = 0.582), this may have been because the telemedicine women were younger. Three months later, the EQ-5D tariff score for the telemedicine women had fallen over time, whereas for the direct referral women, this score had increased (see Table 2.14) and over time these differences were not statistically significant.

Table 2.14: EQ-5D visual analogue assessment and EQ-5D tariff of own health today[±]

	Telemedicine group (n = 23)	**Direct referral group (n = 6)**
EQ-5D visual analogue scale (initial survey)		
Mean (SD) Median IQR	77 (13) 80 70 to 86	84 (13) 85 70 to 98
EQ-5D tariff based on 5 dimensions (initial survey)		
Mean (SD) Median IQR	0.846 (0.146) 0.812 0.760 to 1.000	0.820 (0.159) 0.804 0.691 to 1.000
EQ-5D visual analogue scale (follow-up survey)		
Mean (SD) Median IQR	80 (12) 85 70 to 90	83 (6) 80 80 to 90
EQ-5D tariff based on 5 dimensions (follow-up survey)		
Mean (SD) Median IQR	0.831 (0.131) 0.796 0.727 to 1.000	0.932 (0.118) 1.000 0.796 to 1.000

[±] Statistical tests conducted: Wilcoxon rank-sum for all EQ-5D scores

2.7 Discussion

2.7.1 Summary of findings

This chapter has focused primarily on comparing two groups of women who were referred for a specialist fetal cardiac opinion in two different ways: by direct referral to

specialist hospital for face-to-face assessment or via telemedicine assessment where pre-recorded videoed ultrasound images were relayed to the specialist in the absence of patients'.

Over the 15-month study period, 76 women were referred for specialist opinion: 24 women were sent for a face-to-face assessment in London hospitals and 52 women's anomaly scans were pre-recorded by video and transmitted to the specialist hospital. There were no statistical significant differences in age, parity, type of pregnancy (single or multiple births) or gestation in weeks at birth between the two groups of women.

The perspective adopted for the cost analysis was that of the hospital and costs were presented in pounds sterling (£) in 2005/2006 prices. There were no significant differences in terms of resources used between the two groups, although the telemedicine women had more antenatal clinic attendances. The observed cost differences between the telemedicine group and the direct referral group may have been partly due to the additional cost of a teleconsultation (the cost difference between a specialist scan in London and a teleconsultation was £87). In addition, some of the women in the telemedicine group were scanned earlier in the second trimester and their antenatal care over the remaining months would include one or two extra antenatal visits; and also the telemedicine group contained more diabetic women. Finally, four women (16.7%) in the direct referral group had terminations compared to only two women (3.8%) in the telemedicine group, and for these women their total costs of pregnancy would have been lower. Overall, telemedicine was shown to be more costly, than direct referral. The effectiveness measure was the detection of cardiac anomalies before birth and the direct referral group had a higher prevalence of women with cardiac abnormalities in their fetus, than the telemedicine group. In terms of the quality of life data, telemedicine women gave a lower rating on the EQ-5D VAS and this may be because there were more diabetic women in this group compared to the direct referral group.

2.7.2 The use of telemedicine and why it may not have been cost-saving?

The results from this study were reported in terms of a cost-consequence analysis, as a cost-effectiveness analysis was deemed to be inappropriate. The results from a cost-effectiveness analysis are typically presented in the form of an incremental ratio e.g. cost per cardiac case detected. Whilst such an outcome may resonate with clinical decision makers, it is unlikely to be particularly meaningful to health service administrators. Unlike the cost per QALY ratio, such an outcome does not allow for comparisons across different diseases, treatments or patient groups. Furthermore, in

the absence of a net benefit calculation, it is difficult to infer whether an incremental cost effectiveness ratio, based on a clinical outcome such as a cardiac case avoided, will result in net costs or savings to the health service.

There are a number of reasons why telemedicine may appear to be not as cost-saving as direct referral. Firstly, the study was not randomised. The sonographers (or in some instances, the district obstetricians) at Medway hospital identified women 'at risk' and then chose the method of referral for these pregnant women who required further assessment and they also decided on the use of telemedicine. The two groups of women at the time of their anomaly scan were identified for echocardiographic referral according to the obstetric department's existing protocol for high risk women and they also included women whose fetal heart images during the anomaly scan were abnormal in some way or were poorly visualised. So if there was anything abnormal in the anomaly ultrasound scan, the woman was sent straight away to London for a detailed anomaly scan, unless the telemedicine clinic was scheduled to take place in the next few days. Women with apparently normal fetuses were checked by the fetal cardiologist over the telemedicine link. Hence, the higher underlying prevalence rate of cardiac anomalies in the direct referral group compared to the telemedicine group.

The use of telemedicine at Medway hospital became part of routine practice fairly quickly and this enabled monthly telemedicine clinics to be scheduled and only the occasional 'adhoc urgent' case(s) had to be sent to London. Telemedicine at Medway hospital was (and still is) being used primarily for reassurance and for detailed echocardiography. In practice, women with abnormal scans were sent to London, whereas women with 'normal' fetuses are assessed over the telemedicine link. This is important as the telemedicine link reduces the time taken to assess a woman and to screen the fetal heart (mean time: 5 vs. 45 minutes), which in most cases is usually thought to be normal by the screening sonographers (but is not always so). In most instances, fetal clinics in DGHs are overloaded with normal cases and this leaves little time and space to evaluate and counsel mothers with abnormal foetuses [Personal communication with Sonographer B, October 2007]. Practice over the years has meant that women have been selected (based on clinical factors) for further assessment who were thought to have an increased risk of CHD in their fetus.

Secondly, the antenatal costs for the telemedicine group were approximately £360 higher than the direct referral group. On average the telemedicine group, had more antenatal clinic attendances compared to the direct referral group. This was primarily due to the underlying medical conditions of the women in each group. These results

are confirmed by a separate analysis of the study's pregnancy dataset in which the NHS costs of multiple pregnancies were compared with those of low-risk and high-risk singleton pregnancy [Mistry et al, 2007]. Pre-gestational diabetes was the most influential factor driving singleton costs, where women with diabetes (£4,877) had similar total costs to women with multiple pregnancies (£4,442) and this total cost was approximately a £1,000 greater than women who had cardiac risk factors for a singleton pregnancy (£3,625) [Mistry et al, 2007].

Thirdly, given that this analysis for this section is based on a small number of referred women (n = 76) this might increase the level of uncertainty in the cost and effect calculations and limits the generalisation to other women who may need a specialist opinion. Fourthly, there was only a small number of referred women (23 (44.2%) in telemedicine group and 6 (25.0%) in the direct referral group) who completed a questionnaire with regards to their health status and these results need to be interpreted with caution. The underlying risk factors as mentioned earlier may also partly explain why the telemedicine group were slightly more anxious and depressed compared to the direct referral group (see Table 2.13). In terms of checking whether the respondents were representative of the observed patient sample (i.e. other women who were referred for specialist assessment), both groups were compared and no significant differences were found in terms of demographic characteristics or cardiac risk factors which were measured in the study. However, extrapolating to other women in the study isn't quite as straightforward as a woman's anxiety and depression levels changes throughout pregnancy [Heron et al, 2004].

2.7.3 Next steps

This chapter has provided important information on the two different referral strategies which were used to obtain specialist advice for fetal cardiology. The results highlighted that telemedicine was more costly than direct referral and also that telemedicine had a lower prevalence rate for cardiac anomalies than direct referral. In the next chapter, a literature review will be conducted in which existing evidence on the costs and benefits of telemedicine is presented and discussed. The chapter will then go onto discuss some of the economic issues associated with telemedicine which were highlighted in the review, and how important some of these issues are for the studies which are presented in the review; and finally, the issues which are relevant to the case study will be considered. Following on from the findings from the literature review, in Chapter 4, a critique of the TelePaed case study will be provided and also to reflect on some of the problems with the design of the study.

CHAPTER 3: A LITERATURE REVIEW OF STUDIES INVESTIGATING THE ECONOMICS OF TELEMEDICINE

3.1 Introduction

Chapter 2 provided an introduction to the case study which is used throughout this thesis. The objectives for this chapter are: firstly, to examine what evidence there is to support whether telemedicine interventions are cost-effective; secondly, to see what this new review adds over previous reviews on telemedicine; and thirdly, to discuss some of the economic issues that were identified in the literature review that face telemedicine technologies and how best to solve these difficulties in the context of an economic evaluation. The literature review will aim to identify methodologies that have been used and enable a comparison of these methodologies with that of the case study and explore further the economic issues which have been identified with regards to the case study.

3.2 Methods

3.2.1 Literature Search

A small number of literature reviews have been conducted looking at the cost-effectiveness of telemedicine interventions. Whitten and colleagues (2002) undertook a systematic review of the cost-effectiveness of telemedicine studies which were published until June 2000. They searched the following databases: Medline, Embase, ISI Web of Knowledge and Telemedicine Information Exchange using the following search terms: 'telemedicine or telehealth or telemonitoring or telecommunications' AND 'cost or cost-effectiveness or economic or cost analysis or budget or financial or health care costs'. Their inclusion criteria included: original research on telemedicine examining cost-effectiveness of healthcare delivery; and their exclusion criteria were: papers reporting cost benefits of telemedicine used mainly for educational or administrative purposes; papers reporting hypothetical cost analyses or modelling exercises without any associated formal clinical trial; and papers reporting economic analyses without any means of substantiating claimed resource use.

Two further reviews on telemedicine not only looked at the cost-effectiveness, but they also made conclusions with regards to the clinical outcomes of telemedicine interventions. Roine et al (2001) examined evidence on the effectiveness or cost-effectiveness of telemedicine studies and Hailey and colleagues (2002) in another review conducted a systematic review of evidence for the benefits of telemedicine, which included providing evidence on the effectiveness and economics of telemedicine applications. Further details about these three studies and other systematic reviews of

telemedicine applications (including the conclusions by the authors) are provided in more detail in the discussion section (section 3.4.2 and Appendix 5).

Instead of updating the literature search from June 2000 which was carried out by Whitten and colleagues (2002), a new search has been conducted on the costs and cost-effectiveness of telemedicine interventions, to make sure that no important papers were omitted. The search strategy aimed to identify all published studies of the costs and cost-effectiveness associated with telemedicine interventions.

As there has been an increasing amount of literature in the area of health economics and in order to widen the literature search to make sure that all articles relating to the costs and cost-effectiveness of telemedicine are retrieved, health economics databases have also been included. The following databases (with the relevant time periods) outlined below were searched and details of the search strategies including any search terms used can be found in the Appendix 1.

- MEDLINE (1950 to 12th October 2007) - articles were limited to English Language, to humans and any duplicates within Medline were deleted. The search retrieved a total of 2,252 abstracts;
- CINAHL (Cumulative Index to Nursing and Allied Health Literature) (1982 to 12th October 2007) – articles were limited to English Language, to humans and any duplicates within Cinahl were deleted. The search retrieved 838 abstracts;
- EMBASE (1974 to 12th October 2007) – articles were limited to English Language, to humans, to articles within the Embase database and any duplicates within Embase were deleted. The search retrieved 2,464 abstracts;
- ISI Science Citation Index (1970 to 12th October 2007), ISI Social Science Citation Index (1970 to 12th October 2007) and ISI Arts and Humanities Citation Index (1975 to 12th October 2007) – searches were restricted to English Language and document type was restricted to articles or abstracts of published items. Within this ISI search any duplicates which were identified were removed. Two searches were run, firstly using the keyword search term topic (TS), this identified 5,694 abstracts, and secondly, restricting the search terms to within the title (TI), this retrieved 180 abstracts. All articles regarding the costs and cost-effectiveness of telemedicine found in the topic search were also identified in the title search. Hence, the decision to use the search strategy which was restricted to title search was justified, as nothing important seemed to be omitted and the other articles which were identified in the TS search were not relevant.

In addition, the following databases outlined below were searched on 12th October 2007 using the search term "telemedicine" and limiting where possible to English Language abstracts. The number of abstracts retrieved for each database is also noted:

- Database of Abstracts of Reviews of Effectiveness (DARE) - 27 abstracts;
- NHS Economic Evaluation Database (NHS EED) -163 abstracts;
- Health Technology Assessment database (HTA) - 34 abstracts;
- European Network of Health Economic Evaluation Databases (EURONHEED) - 54 abstracts;
- Office of Health Economics: Health Economic Evaluations Database (OHE HEED) - 93 abstracts;
- Cochrane Database of Systematic Reviews - 14 abstracts;
- Cochrane Central Register of Controlled Trials - 274 abstracts; and
- Cochrane Methodology Register - 3 abstracts.

The database of Telemedicine Information Exchange which was used by Whitten and Colleagues (2002) was searched using only the economic evaluation and quality of life terms, as all the citations relate to telemedicine and the searches were restricted to English Language and journal articles only. However, this search was not included: firstly, it retrieved over 18,754 abstracts and the search truncates the results to the first two hundred citations found (even if the word 'cost' was entered as a search term on its own, this search retrieved 1,207 articles and also truncated to the first 200 articles); and secondly, the 24 articles mentioned in the paper by Whitten et al (2002) were also retrieved from the other databases in the literature search, implying that any relevant articles were not missed.

Once all abstracts were retrieved they were exported into the citation software package, Endnote [Endnote, 2002] and any duplicates identified from the different databases were deleted. All titles, bibliographic data and abstracts of the results from these searches were then read and scanned for relevance based on the inclusion and exclusion criteria. Many of the abstracts identified in the search were due to the word 'cost' or 'economics' being included in the abstract. A lot of these abstracts did not actually report any economic analyses; however, they simply stated that it was important for economic aspects to be taken into account. The inclusion criteria was articles relating to research which examined the costs and cost-effectiveness of telemedicine applications. The exclusion criterion for the abstracts included: no original economic evaluation or related type of study; studies with no cost data; effectiveness studies with no cost data; telecare or telehealth devices used specifically for use in the

home; language other than English; and articles where one project led to multiple reporting of the same results.

3.2.2 Data Extraction

Relevant full-text articles were then obtained and evaluated. A data extraction form was set up to retrieve information from the articles obtained. In addition to the author(s), title, year, volume, issue and page numbers, the fields in the data extraction form included: the aim of paper; the method of economic evaluation undertaken for the analysis; country in which the study took place; viewpoint of study; study time frame (how long the telemedicine intervention was evaluated for); what the telemedicine intervention and the comparators were; clinical application or area; study design (randomised or not–randomised); sample size of the patients in the study; the type of costs included in the study which would fall into one of the following categories: fixed, variable and other costs; costs excluded from the study; total costs for the telemedicine intervention and for the conventional comparator(s); currency in which the total costs were presented; any benefits/outcomes associated with the telemedicine intervention or the comparator; whether discounting was undertaken (not including the depreciation of the telemedicine intervention); whether any sensitivity analyses were undertaken after the base-case results were presented; number of telemedicine conferences or sessions held during study period; transmission mode for the telemedicine images; the average duration for the telemedicine consultations; strengths or limitations mentioned by the authors in their study; and the conclusion which the authors finally came to.

3.2.3 Meta-analysis

Meta-analysis is a statistical analysis of the summary of findings of quantitative studies. By pooling results together from various studies, including individual studies with small sample sizes, it is possible to increase power and precision of estimates of treatment effects and exposure risks [Mulrow, 1994]. For a fixed effect meta-analysis, it is assumed that all studies come from a common population, and that the effect size is not significantly different among the different studies. For a random effects meta-analysis, the random variation within the studies and the variation between the different studies are incorporated. Random effects meta-analysis can lead to wider confidence intervals than fixed effects models [Egger et al, 1993]. From the literature review, relevant data was abstracted in order to conduct a meta-analysis for the cost studies using a random effects procedure (the data obtained from the articles in the literature review did not have enough information to conduct a cost-effectiveness meta-analysis). This abstracted data for both the telemedicine and conventional arms included: sample

sizes; total costs; mean costs; standard deviations; 95% confidence intervals; and p values (from statistical tests) which where reported in the studies.

Once information was abstracted from the studies, using the World Bank gross domestic product (GDP) deflators [World Bank, 2006], costs in local currency units in the study year were converted to local currency units in 2007 prices. The GDP price deflators measure the change in price level of GDP relative to real output and has an advantage over consumer price indices, as it is not based on a fixed basket of goods and services. Once all local currencies were in 2007 prices, using purchasing power parity (PPP) measures [OECD, 2008] local costs were converted in UK pounds sterling in 2007 prices. "PPP are rates of currency conversion that equalise purchasing power of different currencies. That is, they attempt to eliminate the differences in price levels for the same goods between countries" [Mulligan and Fox-Rushby, 2005]. The random effects meta-analysis was computed using the comprehensive meta-analysis version 2 programme [Borenstein, 2005].

3.3 Results

3.3.1 General Findings

In total, 4,721 abstracts were retrieved after removing duplicates. After application of the exclusion criteria, 141 of these articles were obtained and were subject to a full review. After a more detailed evaluation, only 109 studies were included in the literature review results (see Figure 3.1 below). Thirty-two studies were excluded after a more detailed evaluation: one of which forms part of the case study used throughout the thesis [Dowie et al, 2007]; two studies did not provide any cost information; one study only provided time costs and not monetary costs; and the remaining studies (n = 28) used telemedicine in a home setting ('telecare'). The results from the data extraction form of these 109 studies are summarised in Appendix 2.

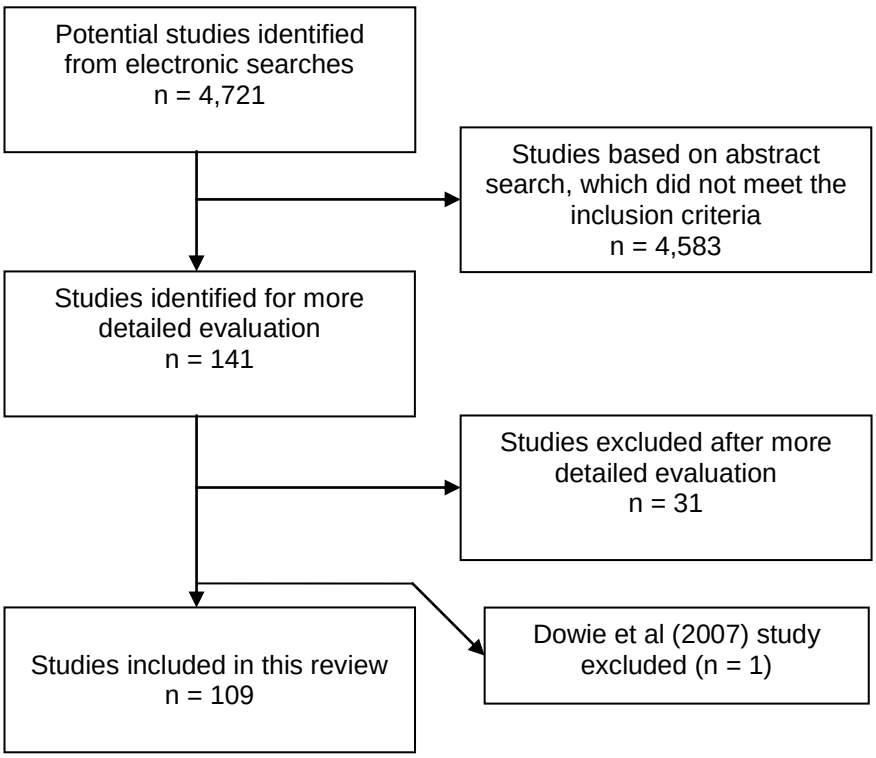


Figure 3.1: Studies eliminated from or selected for the review after applying inclusion and exclusion criteria

3.3.2 What was studied?

The greater number of studies (n = 106, 97.2%) used telemedicine for clinical applications; and three studies [46, 75, 86][4] specifically used telemedicine for clinical administration and educational purposes only. Further selection of these 109 studies highlighted the various areas of health in which telemedicine were used: ranging from specialist care for patients i.e. paediatrics, pregnant women and the elderly; through to specific health advice such as for skin disease (dermatology), through to the use of telemedicine for a range of specialties. The majority of studies which were conducted in the one health area (not including studies with multiple specialties) were in dermatology (n = 10); psychiatry (n = 9); radiology and cardiology, both of which had 8 studies each; and cancer (n = 6).

Telemedicine was compared to the conventional method of referral in 78.0% (n = 85) studies; in 18 (16.5%) studies telemedicine was compared to two methods of referral. The other methods of referral usually involved the patient travelling to specialist hospital for face-to-face assessment or a specialist visiting the local hospital to see patients in outreach clinics; in four studies [15, 18-19, 79] telemedicine was compared to what happened before telemedicine was introduced (before and after comparison);

[4] Formatting of references for this section of the thesis has changed to make it easier for the reader to refer to the table in Appendix 2.

in one study [51] telemedicine across two countries was compared; and in the final study [75], telemedicine was used for different activities.

Of those 109 studies, only 37 stated the mode of transmission used for the telemedicine service. Twenty studies [2, 9-10, 12, 14, 24, 27, 29, 39, 47, 49, 55, 59, 61-62, 77, 80, 95, 106, 108] used real-time telemedicine; 9 studies [1, 3, 20, 42, 48, 56-57, 74, 103] used store-and-forward methods; and a further 8 studies [7, 15, 38, 60, 63, 72, 100, 102] used both real-time and store-and-forward methods as a means of transmission.

3.3.3 Where, when and for how long?

Figure 3.2 below shows that the half of the studies (n = 57) were conducted in North America and 26 studies were from Europe (not including the UK or Ireland).

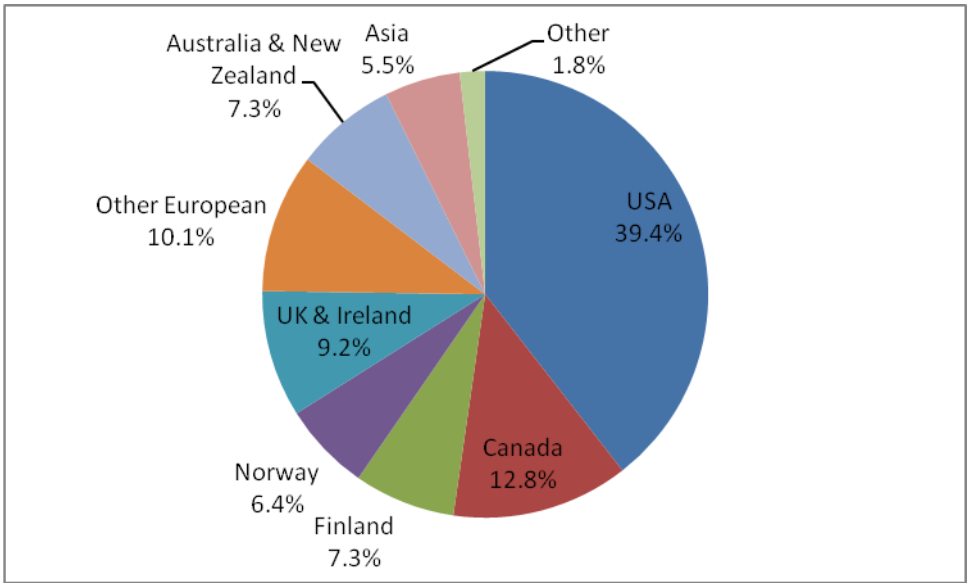


Figure 3.2: Country in which telemedicine study was conducted

The journals from which the studies were retrieved were grouped into three areas: telemedicine journals; clinical journals; or health service research or health policy journals. The majority of articles were in telemedicine journals (n = 73) (see Figure 3.3. below).

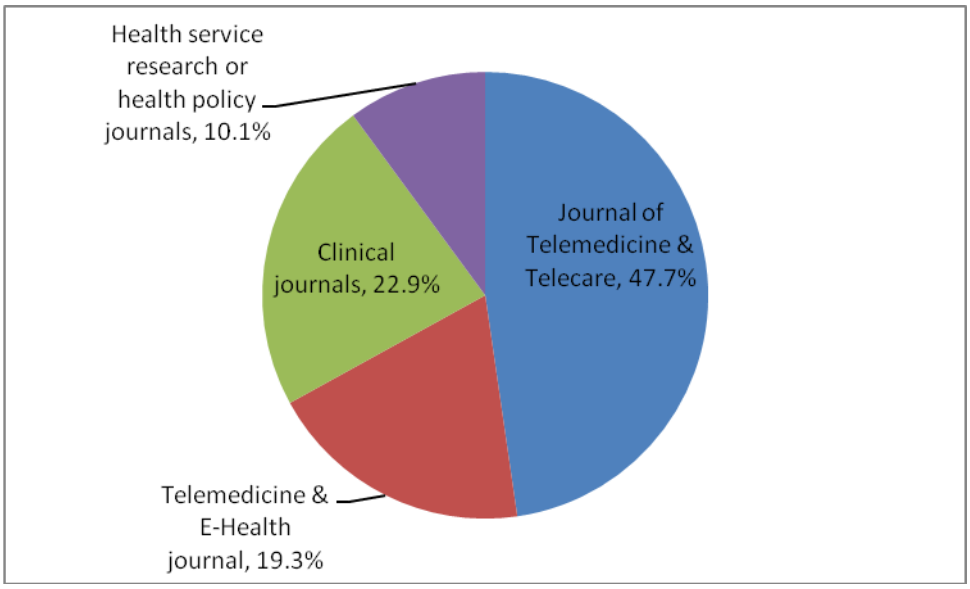


Figure 3.3: Type of journal

From 1995 to 1999, 30 (27.5%) articles were published; between 2000 and 2004, 59 (54.1%) articles were published; and between 2005 and 2007, 20 (18.3%) articles were published. The year 2004 had the most articles published on the costs and cost-effectiveness of telemedicine, 16 (14.7%).

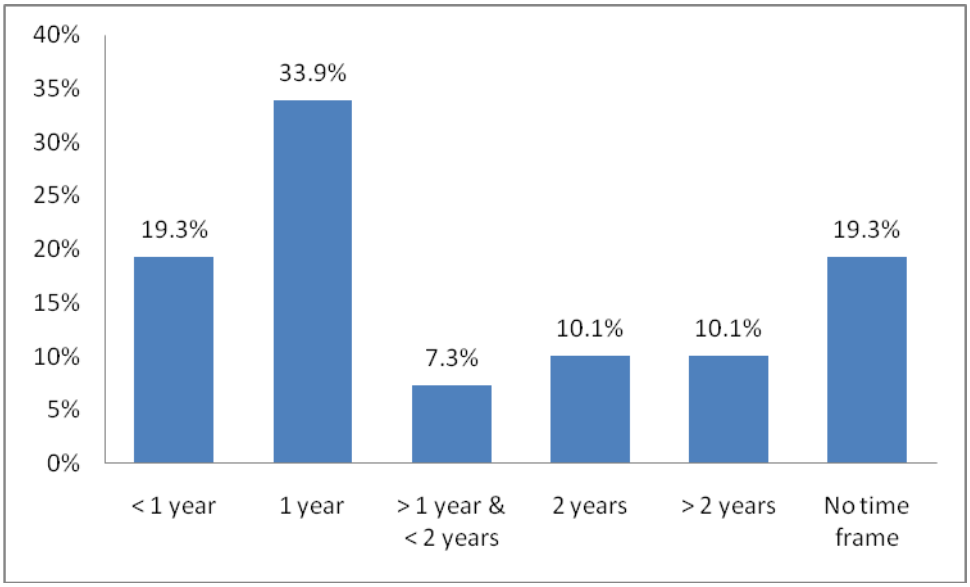


Figure 3.4: Time frame

The literature review highlighted that the majority of studies evaluating telemedicine interventions were conducted over a two-year time frame or less: 77 studies (70.6%) and 21 studies (19.3%) did not report how long the telemedicine intervention was evaluated for (see Figure 3.4).

3.3.4 Overall methods of cost-effectiveness

The articles included 71 cost analyses; 16 cost-minimisation analyses; six cost-consequence analyses; one cost-benefit analysis; 11 cost-effectiveness analyses; one joint cost analysis and cost-effectiveness analysis; and three studies used costing models. However, only 45 of the studies informed the reader of the viewpoint of the analyses: 22 studies [2, 5, 9, 12, 30, 40, 49, 60-62, 69, 73-74, 78, 83, 86-87, 89, 93, 95, 106, 108] were from a societal perspective; 21 studies [1, 3-4, 7-8, 10, 17, 21, 27, 54, 57-58, 66, 75, 81, 84, 91-92, 103-104, 109] were from the perspective of the healthcare system or prison; and only two studies [19, 90] were from the patients' viewpoint.

Eighty-five studies (78.0%) did not report about benefits/outcomes in their articles. Twenty-four studies, however, reported some outcome measure (see Appendix 2 for further details). Only two studies [3, 21] used QALYs, a further study [51] used disability-adjusted life years (DALYs), two more studies [5, 54] used life years gained and one study [49] used health state utility values as an outcome measure. Eight further studies used a different measure for effectiveness: one study [73] used safety and clinical effectiveness at 7 days after presentation; another study [103] used days to the intervention; three papers [15, 81, 105] used a range of outcomes as measures, one of them being mortality; for the other study [2] effectiveness was defined as health care access for patients who received outpatient consultations; one study [85] used differences in healing rate as an effectiveness measure and the final article [104], used three effectiveness measures including the number of cases of severe vision loss averted. Three studies [70, 80, 90] used patient satisfaction as an outcome measure; one study [84] used scales for depression, anxiety, functioning and also a generic measure short-form 12 (SF-12) for outcome measures; and the final six studies [16, 27, 31, 83, 89, 92] looked at benefits in terms of transfers or travel avoided and some of these studies also included the number of hospitalisations or consultations avoided. However, in economic terms, outcomes measures such as the number of transfers or hospitalisations avoided are not regarded as benefits, instead these should be classed as 'costs averted'.

Only seven out of 109 studies [1-3, 5, 21, 54, 75] used discounting for costs (not including the discount methods which were used to estimate the depreciation costs for the telemedicine equipment). This is reasonable as half of the studies (n = 58) reported a time frame of less than or equal to a year in which the telemedicine service was evaluated for. However, for a further 26 studies, telemedicine was evaluated for more than 12 months and costs in these studies should have been discounted. Thirty-

seven studies used some sort of sensitivity analysis for the cost components and break-even thresholds.

Table 3.1 below shows a summary of the overall conclusions provided by the authors for each study and as the table shows there was no general agreement whether telemedicine was cost-saving (or cost-effective) compared to conventional means.

Table 3.1: Overall conclusions

Conclusion	Number (%)	Study identifier
Telemedicine was cost-saving compared to the comparator(s)	39 (35.8%)	4, 6, 8, 10, 13, 16, 18, 20, 22-23, 26, 28, 31, 33, 37, 39, 41, 43-44, 47, 52, 55-57, 62, 65-66, 69, 72, 77, 80, 88, 90, 92-94, 97-98, 102
Telemedicine was cost-effective compared to the comparator(s)	8 (7.3%)	2-3, 11, 14-15, 50, 85, 101
Telemedicine was only cost-saving either above a certain number of teleconsultations or patient throughput	12 (11.0%)	9, 17, 34, 45, 59, 67, 76, 91, 95, 99, 108-109
Telemedicine had the potential to be cost-saving	14 (12.8%)	25, 27, 35-36, 42, 51, 53, 60, 63-64, 71, 79, 82, 86
Telemedicine had similar cost results to conventional methods	5 (4.6%)	29, 48, 58, 100, 105
Telemedicine had reduced patients' costs	3 (2.8%)	19, 46, 107
Inconclusive as to whether telemedicine was cost-effective or not (i.e. when telemedicine was compared to two alternatives)	4 (3.7%)	1, 32, 87, 96
Telemedicine was not cost-effective compared to the comparator(s)	12 (11.0%)	5, 7, 21, 49, 54, 73, 75, 84, 89, 103-104, 106
Telemedicine was not cost-saving compared to the comparator(s)	12 (11.0%)	12, 24, 30, 38, 40, 61, 68, 70, 74, 78, 81, 83

Finally, of the 109 studies, 100 studies (91.7%) had insufficient statistical information in order for a meta-analysis to be conducted. Results from the meta-analysis are not really informative and for illustration purposes they are presented in Appendix 3.

3.3.5 Use of telemedicine in specific health areas

The cited systematic reviews (see section 3.4.2 for more detail) deal with telemedicine literature in general, without focusing on a particular medical specialty. Because of the significantly different ways telemedicine is performed in specific medical areas, this can lead to important details being omitted in favour of general considerations. Therefore, there is a need to perform an analytical review of articles relating to the application of telemedicine in a specific medical field.

As a meta-analysis for each clinical area could not be conducted, in Appendix 4, there is a brief description of the studies within a particular clinical area (dermatology; radiology; psychiatry; cardiology), followed by an overall conclusion for the use of telemedicine in each clinical area.

3.4 Discussion

Cost-effectiveness studies are needed to help define the appropriate scope and application of telemedicine interventions or services in different settings and to provide information to the reader and the decision maker on whether the use of these interventions/services in various health areas represent good value for money. One of the aims of this literature review was to present the results of the literature search and to see whether there is evidence to support whether telemedicine interventions are cost-effective. The search strategy including search terms and the databases which have been searched are highlighted in the methods section and in Appendix 1. In total, 109 articles identified in the literature search were subject to a full review in which economic data was presented for telemedicine interventions (an overview of the results are provided in Appendix 2).

The literature search also identified all 24 studies that Whitten et al (2002) had found in their search. However, 3 of these studies were not included as they did not meet the inclusion criterion [Friedman et al (1996) and Wu et al (1995) were excluded as they were for telemedicine applications in the home (i.e. telecare applications); and Loane et al (1999) looked at time and not monetary costs]. This finding is important as it helps to understand and validate the search criteria and to highlight that the methods used for the literature search are correct.

Since the beginning of 2002 until October 2007, the search identified 55 studies that have been published on the costs and cost-effectiveness of telemedicine interventions. Ten of the 11 studies which conducted cost-effectiveness analyses; 21 of the 24 studies which reported benefit measures; six of the seven studies which used discounting; and 24 of the 37 studies that undertook some sensitivity analyses have been published during this time period. This has shown that economic tools are increasingly being used for evaluations to ensure that treatment choices reflect evidence not only on the clinical effectiveness, but also the cost-effectiveness.

The analysis of the results from the literature search of the use of telemedicine in specific health areas also gave mixed feedback (see Appendix 4) and one of the big limitations was that a meta-analysis could not be conducted for each health area separately as there was not enough statistical information. Therefore, it was not possible to summarise the results further from each health area in which telemedicine was used. This was because there were variations in the patient populations, study designs and the intervention and the comparator groups. Instead, the results for each health area were reviewed and summarised. These findings are confirmed by a cost

review conducted by Hyler and Gangure (2003) in psychiatry. They identified over 380 articles relating to telepsychiatry, but only 12 articles focused on the cost of telepsychiatry [not all cost studies identified by Hyler and Gangure (2003) were identified in my literature review: 1) they also searched the PsycINFO database for additional articles relating to costs of telepsychiatry (this database is an abstract database of psychological literature); and 2) some of the studies included in their literature review were cost feasibility studies, where costs were calculated with no actual service being delivered and two studies were based on surveys as a way to subjectively probe cost without determining it objectively]. In seven studies, telepsychiatry was worth the cost, and a further three studies telepsychiatry consultations had to be over a certain number to be financially viable. Of the remaining two studies, in one study telepsychiatry was not financially viable and in the other study a lack of clear business plan contributed to its inability to determine its cost-effectiveness. The authors concluded that "telepsychiatry can be cost-effective in selected settings and can be financially viable if used beyond the break-even point in relation to the cost of providing in-person psychiatric services."

The term telemedicine groups together all services which comprise a tele-link and can include any medical service or act which is performed from a distance. Whilst there are some benefits to grouping such interventions under a common definition the interventions included may actually differ quite significantly. For example, telemedicine may be classed as two doctors talking on the telephone about a complex issue and determining the best course of action for a patient or telemedicine could be a complex as telesurgery, in which a surgeon receives visual and audio information to guide robotic instruments to perform surgery at a distance. Even though the concept of telemedicine is simple (i.e. an exchange of information), arguably it doesn't make sense putting all tele-services together for the purposes of evaluation. Just as different approaches to evaluation may be required for different disciplines of medicine together such as paediatrics, obstetrics, and the elderly, then this too applies to telemedicine. Telemedicine is actually a broad term, and as we have seen in this Chapter when comparisons such as literature reviews (of the cost-effectiveness of telemedicine studies) or analyses need to be undertaken then the medical specialties within telemedicine need to be categorised such as real-time versus store-and-forward telemedicine, or home versus hospital based telemedicine, or telecardiology versus telepsychiatry to get meaningful conclusions. For this reason, it would be inappropriate to attempt to apply a 'one size fits all' approach to evaluation and it would be inappropriate to attempt to make broad, sweeping generalisations about the effectiveness or cost-effectiveness of telemedicine.

Finally, the results from the literature review may also be due to publication bias, that is, studies with positive results have a better chance of being published, and therefore conclusions based exclusively on published studies may be misleading. For example, the results from the meta-analysis (see Appendix 3) highlighted that there was significant heterogeneity and that there was also evidence of funnel plot asymmetry (graph has not been presented in the Appendix); that is, the smaller studies were more widely spread around the mean effect due to the large random error, highlighting publication bias amongst the studies. Furthermore, there are few published studies which suggest that telemedicine might be cost additive and/or lead to inferior outcomes to current practice.

The next section (section 3.4.1) sets out some of the economic issues which arose from the literature review. In section 3.4.2, the findings from the literature review will be compared to previous reviews; and in the final section (section 3.4.3), some of the challenges to economic evaluations of telemedicine interventions/services which will be addressed in this thesis will be discussed.

3.4.1 Economic issues arising from the literature review

The economic issues which emerged from the literature review are outlined below; these are not only relevant to most health technologies, but also to telemedicine. The Drummond et al (2005) checklist for assessing economic evaluations has been followed (but not in a strict manner), as there were some other issues which were of importance that had arisen in the literature review.

a) Inadequate details about study design and methodologies

The literature review highlighted three main issues with regards to study design: 1) whether the study was randomised or not; 2) how long the telemedicine intervention was evaluated for; and 3) reporting of methodologies.

Firstly, in primary research, the normal gold standard study design is a randomised controlled trial (RCT). In secondary research, the gold standard study is a systematic review, followed by a meta-analysis of the findings, and if appropriate modelling. The literature review indicated RCTs of telemedicine services were limited. In total, 12 of the 109 (11.0%) studies identified were RCTs [24, 40, 49, 61-62, 73, 76, 84-85, 103, 105-106] (see Appendix 2) and the rest of the studies were all of observational nature. Of those 12 studies which were randomised, for 11, the unit of randomisation was the individual patient/subject and in the remaining study [85] the unit of randomisation was

the clinical site. For example, Wootton and colleagues (2000) conducted a multicentre RCT which aimed to evaluate the health outcomes and cost-effectiveness of real-time teledermatology compared with conventional outpatient dermatological care from a societal viewpoint. Patients were randomised to either teledermatology consultation (n=102) or to conventional care (n=102). The authors found no major differences in outcomes, and real-time teledermatology was not cost-effective compared to conventional consultation.

Secondly, the majority of telemedicine studies were undertaken as pilot or short-term studies. Of these, 77 studies were undertaken for a time period of less than two years, thereby not allowing what would be the longer-term impact on patients, on costs or even outcomes (21 studies did not report a time frame). One of the main reasons why some of these telemedicine studies may only be of a short-term duration, is due to the limited funding or one-off special budgets which have been given to these applications and for these studies to continue, it may be difficult if the funding is not in place or if no further funding can be found. As some of the telemedicine studies are pilot services, the costs and benefits which have arisen in the pilot stage may not reflect the costs and benefits when the telemedicine service is in routine use. Therefore, generalising these study findings from pilot studies to other settings may not always be possible.

Thirdly, most of the studies identified in the literature search did not give enough adequate details about their study design or they had weak methodologies i.e. there was no consistency in the analyses conducted between the studies. For example, there may not be enough information on how costs (and/or outcomes) are collected, calculated and reported. If studies report the appropriate methodology, it would then be easier for the reader to understand how the author(s) came to their results and conclusions.

In the example, by Tsitlakidis et al (2005) they conducted a cost-minimisation analysis of patients assessed via telemedicine compared to the alternative option of referring patients to hospital. The telemedicine service at the Airforce health centre was linked to two remote Greek Islands. The mean duration of a telemedicine consultation was approximately 15 minutes compared with 40 minutes for a direct referral consultation. The authors did not explain whether it was the telemedicine intervention itself that allowed for this substitution of time saving. In another study by Crowe et al (1996), there was insufficient background information on the nature of the journeys or mode of transfer to enable adjustments to be made by the reader for alternative distances and conditions.

b) Study perspective

The perspective for the analysis should be clearly stated and justified, as it allows the reader to make an informed choice as to whether the information provided in the article is sufficient to answer the study question and that of the policymakers and/or funders. For example, a full economic evaluation of a telemedicine service would consider all costs and benefits to all groups of the society who would use the telemedicine service i.e. a health service and patient perspective. If a narrow perspective is adopted, one such as the health service, then the benefits to patients would be omitted. From the literature review, over half of the articles (58.7%) did not explicitly report the viewpoint of the analysis, thus it is hard to make judgements as to whether the conclusions reached by the authors are appropriate.

c) Small sample sizes

Most telemedicine studies identified in the literature review have small sample sizes. Only 78 studies (71.6%) reported patient numbers, whereas some studies did not state the sample size, but stated the number of telemedicine consultations which took place or the number of images which were transmitted. Of the 78 studies that reported a sample size, 28 studies (35.9%) had a sample size of 100 patients or less. The smallest sample size for a study was eight patients who had chronic pain and telemedicine was shown to be cost-saving when compared to conventional care [Peng et al, 2006]. These studies with small sample sizes will have low statistical power to detect differences in health outcomes and the difference between the two groups may not always be due to the telemedicine intervention.

d) Choice of alternatives

In most cases, the comparator for the telemedicine intervention is usually the most widely used current alternative method for accessing specialist care. The comparator must be able to be justified to be used in such an analysis and be reasonably accepted by all parties using the services.

Most of the time, the telemedicine intervention in literature is described quite carefully, but costs and consequences in relation to the comparator are not reported in sufficient detail to make proper comparisons and therefore assessing the costs and cost-effectiveness of the telemedicine intervention and the alternative is of limited value. For example, in the article by Bracale and colleagues (2002) they evaluated whether telemedicine services might have a role in the provision of healthcare for two remote islands. The alternative was patient transfer by boat or by helicopter ambulance. The

paper considers in length how the telemedicine services were established and how they worked, but no further details were given about the comparator.

e) Economic importance of question

The economic importance of the research question should be outlined and be phrased in a way which would consider both costs and outcomes. The literature review findings indicated that most telemedicine studies did not state a clear research question. For example, “a cost analysis of teledermatology system” would most likely look at the costs of the teledermatology service and maybe the alternative, but not the outcomes, whereas in “a cost-effectiveness analysis of a teledermatology service”, one would hope that the study would compare both the costs and outcomes of the two alternative methods. A more limited research question such as “Is a teledermatology service undertaken in a hospital in Maidstone cost-effective compared with the alternative of transporting patients to London?” places cost-effectiveness in a local context and provides a relevant comparator.

f) Type of economic evaluation adopted for the telemedicine study

The majority of studies (79.8%) identified in the literature review stated that the economic evaluation was a simple cost analysis or a cost-minimisation analysis. Thus, these studies compared the total (or mean) costs of the telemedicine intervention to the total (or mean) costs of the comparator. The studies also assumed that the outcomes are equivalent or assumed to be established without any proper scientific evidence. Studies otherwise can simply be cost studies with no claim to being an economic evaluation. Unless outcomes can be proved to be identical, then the economic evaluations should take place in the form of a cost-consequence(s), a cost-effectiveness, a cost-utility or even a cost-benefit analysis.

Depending on which of the five economic evaluations would be undertaken depends on the context of the study, the alternative and the outcome measure. None of the studies stated that they used a cost-utility analysis to compare costs and outcomes simultaneously of the telemedicine intervention and the alternative(s). In a cost-utility analysis (which is a special form of a cost-effectiveness analysis), outcomes are measured as health related preferences, which are most often expressed as QALYs and the results are expressed as cost per QALY gained. Cost-utility analysis is useful when an intervention can be expected to have an effect on the health-related quality of life and on the length of life. A cost-utility analysis uses a generic outcome measure that permits broad comparisons across different conditions and interventions. For example, in the study by Armstrong et al (2007) who conducted a cost-minimisation

analysis of interactive teledermatology with conventional care (outpatient face-to-face dermatology clinic) in the USA, the authors stated that “a cost-minimisation analysis compares the costs of multiple interventions that provide the same outcome for the purposes of identifying the lowest-cost intervention”. However, they did not provide any evidence to assume that outcomes were identical and whether this justified the use of cost-minimisation analysis for choosing between the two alternatives.

g) Costs

The information provided in the articles varied from detailed cost analyses to simply mentioning some of the costs and that a telemedicine service was cost-saving or cost-effective. For any economic evaluation the concept of opportunity cost is essential. For example, an opportunity cost to a specialist applies if they have to travel to treat patient(s) in another location, it is the missed opportunity to treat patient(s) in their own hospital. The viewpoint of the analysis is important so that the appropriate resource use and costs are identified for the telemedicine service and the comparator for measurement and valuation purposes. Most articles provided examples of the types of costs which were included in their analyses; however, some of the cost components were not given (i.e. the cost analyses were not exhaustive). Some studies did not include the telemedicine equipment costs, only the telephone charges and line rental; whereas some studies ignored patient costs and/or productivity costs. Therefore caution should be taken when using cost estimates from previous analyses as the studies are not explicitly clear in reporting their methodology and results, and allocation of costs may not be accurate. An outline of some of the telemedicine costs which should be included in an economic evaluation are listed below. The two boxes are list of costs which have been compiled from the literature review.

Box 1: Summary of the types of costs collected and reported in the articles

Direct Medical Costs

a) Equipment costs

- Videoconferencing equipment including monitors, computers, video camera, digital camera, document camera, scanner, fax machine, modem, internet connections and software.
- Equipment depreciation
- Equipment maintenance
- Installation of transmission lines – Integrated Services Digital Network or connection via Internet protocol
- Other equipment (if appropriate) such as microscope, phototherapy machine, laser equipment, endoscope, dermatoscope and electronic stethoscope.

b) Communication costs

- Line rental
- Call charges

c) Staffing costs

- Salaries and overheads
- Training costs
- Cost of travel for consultants and other staff

d) Administration and overhead costs

- Overheads i.e. cost of buildings, clinic space or telemedicine room and utilities
- Other administrative costs such as paperwork

e) Hospital and associated costs

- Inpatient stays, outpatient or GP visits, accident and emergency visits, and other hospital visits
- Tests and procedures
- Treatments and medications

f) Miscellaneous costs

- Patient transfer or transportation
- Transfer by courier
- Fuel and driver costs
- Escort costs

Direct non-medical costs

- Patient and family travel expenses
- Accommodation expenses
- Subsistence costs i.e. food and drink

Indirect costs

- Travel time costs
- Time costs
- Missed days at work
- Lost income
- Lost productivity

Box 1 above provides an overview on the different types of costs collected. Direct medical costs are the costs related to the use of resources due to either the disease or treatment and these are the costs that usually fall on the health service; direct non-medical costs are costs incurred by patients and family members, which contribute to the treatment process; and finally, indirect costs are resources lost due to treating a disease and can reflect two different costs depending on the perspective. First, if a patient perspective is adopted, this may reflect the loss of time (whether this is work or non-work time) to the individual in attending for treatment. Second, if a societal perspective is adopted, time costs incurred by individuals in receiving treatments reflect the loss of production to society, whether paid or unpaid. Some studies also explicitly noted that some costs were excluded from their cost analysis as they were common to

both arms, these are shown in Box 2. Not all costs reported in the articles appear in the two boxes.

Box 2: Summary of the types of costs excluded in some articles

Excluded costs

- Hospitalisation costs
- Treatments, medications and laboratory test costs
- Administration costs
- Electricity and overheads
- Family travel and subsistence costs
- Loss of income

Nearly all of the 109 studies reported information on direct medical costs apart from two studies [82, 102] which simply reported potential savings and 4 studies [19, 68, 77, 90] which reported direct non-medical costs and/or indirect costs. Of those 22 studies that were from a societal perspective, all studies reported direct medical costs; in addition 20 of the 22 studies also collected information on direct non-medical costs, and of those 20 studies, 13 studies also provided information on indirect costs. Of those 64 studies that did not report a perspective, four studies collected information on indirect costs, in addition to direct non-medical costs and a further 18 studies collected information only on direct medical and direct non-medical costs. The inconsistency about the costs which were collected and reported in these studies and the study perspective which has not been stated in all articles, makes it hard to make substantial conclusions about the cost-effectiveness of these telemedicine applications and to conduct a meta-analysis. These findings are confirmed by Hailey and colleagues (2002) who said “the costs included varied significantly between studies, so that the comparison of the cost estimates was not feasible in many cases”.

The telemedicine studies have had to make assumptions in the calculation of telemedicine equipment costs. For example, for the expected lifetime of the equipment in the majority of cases this was five years; the choice of discount rate (depreciation rate) varied from country to country; and maintenance charges differed depending on length of contracts. Some of the cost assumptions made in these studies may not reflect how the telemedicine service is used in practice.

Most telemedicine applications have a large initial capital outlay. For economic evaluations this capital outlay is converted into an annual cost over the expected lifetime of the equipment, to allow a comparison of the annual cost of the systems without being biased against a system which has a larger capital outlay, but a longer life [McIntosh and Cairns, 1997]. When the annual equipment costs have been calculated, the estimated annual maintenance costs and annual running costs must

also be added. Telemedicine costs are largely determined by the scale and utilisation of the service. For example, as more patients use a telemedicine service, then the costs for the service should be lower. Also, the number and level of staff at each end of the telemedicine link can increase or decrease and have an impact on costs. Thus, the level of skill mix may change and the workload may vary depending on the patient throughput.

h) Are outcomes or benefits reported?

Many telemedicine studies do not report information on patient outcomes or benefits in such a way that would enable a cost-effectiveness analysis to be performed. Indeed, telemedicine studies which were identified in the literature review were classed as cost-minimisation analyses and assumed that the clinical outcomes for both the telemedicine and the comparator arms were identical. Where cost differences are reported, rather than assume equivalent outcomes, it maybe more useful to show the degree of improvement in outcomes that would be necessary in order to justify those cost differences with a given cost-effectiveness criterion. If there are any differences between the two arms of the studies, these may be constrained by the size and duration of the study, so that these studies will have low statistical power to detect differences in health outcomes.

The benefits of telemedicine may take a while to appear and may also be difficult to quantify. Thus, the few studies that did report outcomes used surrogate measures in the short-term such as time-related measures i.e. avoidance of time lost through travel or indicators of hospital performance such as the length of stay avoided, rather than clinical improvement. However, these surrogate outcomes may be poor measures for final outcomes. Some studies looked at benefits in terms of travel or hospitalisations avoided. In economic terms, these outcome measures are not regarded as benefits, instead these should be classed as ‘costs averted’.

Also, many telemedicine studies use non-health benefits such as improvement in quality of care through improved treatment, faster and more accurate diagnosis, transfer of skills and knowledge, speed of service, improved training and education, and reassurance. The question is how do you identify, measure and most importantly value these non-health benefits and should they be included in the evaluation? McIntosh and Cairns (1997) commented “this would require a valid and reliable instrument which is sensitive enough to detect beneficial changes in the process of care of value to the patient”.

Various benefits of telemedicine services/interventions were identified in the literature and they are summarised below:

a) Improved access to healthcare

Telemedicine can increase access to healthcare, especially for those who live in remote or underserved areas. Improved or increased access to healthcare through telemedicine also means that diagnoses can be made more quickly and more accurately and this may translate not only into better overall health for the patient, but also in the reduced need for transfer to hospital and reduced length of hospital stays. In the case provided by Sicotte and colleagues (2004) they conducted a cost-effectiveness analysis of interactive paediatric telecardiology compared with conventional care (mainly patient travel to the specialist centre and the occasional outreach clinic). The main aim of the use of the telemedicine system was to ”improve accessibility in a remote area by accelerating medical decision making in terms of diagnosis and treatment, and to decrease the rate of patient transfers and visits to the tertiary centre” [Sicotte and colleagues, 2004]. In this example, access to telemedicine provided correct and earlier diagnosis of cases and the consequent confirmation of true negatives, which in turn eliminated the unnecessary patient journeys (‘averted costs’), and also provided reassurance for both the patient and doctor.

b) Patient satisfaction (including acceptability)

Aspects or indicators of patient satisfaction that are typically evaluated include: convenience, comfort during a consultation, acceptability, concerns over privacy and confidentiality, and willingness to use telemedicine in the future. Patient satisfaction with telemedicine (including patient perceptions and future needs) enables people to understand more about patients’ experience of using telemedicine, increases compliance with treatment and may also be important for the future acceptance and adoption of telemedicine. Overall, patients have documented great satisfaction with telemedicine. Television, audio and computer applications are more common now, which means that patients are more at ease and accepting of the use of telemedicine.

Simpson and colleagues (2001a) gathered information on patient perspectives on telepsychiatry through self-report questionnaires and telephone interviews and found that 89% of patients were satisfied with the telepsychiatry session. Patients felt comfortable in their interaction with the consultant and with their ability to provide the same information as they would in a face-to-face interview. Eighty-one percent of patients generally felt that telemedicine was acceptable and would recommend using telepsychiatry again. The acceptance of the telemedicine equipment was high in

relation to sound, picture, ease of use and room environment, despite the occurrence of occasional technical problems.

Acceptance of telemedicine by health care professionals is also important in any telemedicine evaluation. If clinicians are not comfortable with the technology or judge that the technology decreases their control over patient care, they may avoid using it, thereby precluding other benefits of telemedicine. Clinical acceptance of a telemedicine application may depend on the degree of confidence the clinician has in their own clinical findings (e.g. diagnosis) from using the application, as well as their satisfaction with the technology in the absence of a face-to-face interaction with the patient. Simpson and colleagues (2001b) undertook an assessment of a routine telepsychiatry service compared with providing consultations by a visiting psychiatrist. Twenty health professionals responded to the survey and overall they expressed a high satisfaction with the telepsychiatry service with respect to the referral and scheduling process, consultation report and follow-up recommendations. All psychiatric consultants agreed that telepsychiatry was an acceptable way of delivering psychiatric consultations.

c) Effectiveness (including quality of life and outcome measures)

Telemedicine has a range of possible outcome measures. They include clinical outcomes which are the results of the interventions or services used to diagnose and manage patients. For diagnosis, this includes screening, triaging and specialist consultations. For clinical management, this includes deciding on an appropriate management plan for treatment and follow-up specialist consultations for patients. For example, Whited et al (2005) conducted a cost-effectiveness analysis of a digital ophthalmology system versus traditional clinic-based ophthalmology examinations to detect diabetic retinopathy. They modelled the entire population of patients with type 2 diabetes mellitus for three federal agencies. They had three effectiveness measures: i) the number of true positive cases of diabetic retinopathy detected, ii) the number of patients identified who required treatment with panretinal laser photocoagulation, and iii) the number of cases of severe vision loss averted.

To measure quality of life, various instruments are used such as: generic measures, which use instruments to measure overall health-related quality of life such as the short form 36; utility measures, which are a special kind of generic measure but give an indication of value placed upon health-related quality of life, for example the EQ-5D; and monetary measures, which values benefits in terms of currency using ‘willingness to pay techniques’. For example, Jacklin and colleagues (2003) conducted an

economic evaluation that compared conventional outpatient consultations with teleconsultations (virtual outreach) and they used a generic measure, the SF-12 to measure overall health-related quality of life. The authors found that there were no differences in health outcomes in the two groups at six months according the physical and psychological scores of the SF-12. Castillo-Riquelme et al (2004) used a utility measure, QALYs' to measure health outcomes when comparing the cost-effectiveness of alternative methods of screening for retinopathy of prematurity in the UK, which included the existing method of indirect ophthalmoscopy by ophthalmologists and digital photographic screening by nurses. The authors used utility estimates from a study on cataracts and converted these utility values into QALY estimates over the whole expected lifetime of the babies. This latter study was one of the two studies identified in the literature review which measured health benefits in terms of QALYs for telemedicine applications and both were in the area of ophthalmology [Aoki et al, 2004; Castillo-Riquelme, 2004].

i) Was an incremental approach appropriate?

Results of the economic analysis can be used to inform decisions on a telemedicine application. For example, to determine whether it should continue; how it may be used in other situations; or to what extent any improvements are needed either to its costs or performance. The benefit of using an incremental approach is that it allows the decision maker to see the additional cost and/or benefit of the intervention compared to the alternative. Most studies did not provide an incremental analysis of the costs and outcomes. Only eleven (twelve) studies conducted a cost-effectiveness analysis (cost analysis and cost-effectiveness analysis) and of these, only six studies [2-3, 51, 54, 89, 103] provided an incremental ratio.

For example, in the study by Sicotte and colleagues (2004) which looked at the cost-effectiveness of interactive paediatric telecardiology in comparison with the conventional alternative (which included patient travel and outpatient clinics), they found that the total cost of telecardiology was C$272,327 and the total cost of conventional care would have been C$157,212. They stated that telemedicine, represented a supplementary cost of C$1,500 per patient. The incremental cost-effectiveness ratio of teleconsultation was estimated C$3,488 per patient journey avoided. The effectiveness measure used here (patient journey avoided) is technically not a benefit measure, but a cost which has been avoided. So in this situation, the incremental cost-effectiveness ratio was used but inappropriately, and should not be used to inform healthcare policy decisions.

Marginal costs capture the change resulting from the healthcare provision and telemedicine costs will largely be determined by the scale and utilisation of the service. In some instances, once the infrastructure of the telemedicine service is established, the cost of providing an additional telemedicine consultation may be insignificant. It is important to note that marginal costs will differ between different patient groups and different geographical locations. Most studies hypothesised about the relationship between utilisation and costs, some studies even reported patient numbers in order for it to be cost-saving, but more evidence is needed before a cost function for a telemedicine service can be properly identified.

Some telemedicine studies used break-even analysis to indicate the volume of output necessary for telemedicine applications to become cost-saving. Break-even analysis is calculated as the difference in fixed costs, divided by the variable cost-saving produced by telemedicine. So activity levels beyond break-even will lead to cost savings, as the aggregate variable cost savings will outweigh any additional fixed costs of the telemedicine service. This type of analysis can be useful for decision makers when faced with the question of whether to introduce a new telemedicine service or to continue with the existing service.

Furthermore, the majority of studies stated whether a telemedicine intervention or service was cost-effective or not cost-effective compared to an alternative. However, the authors generally did not define what they meant by cost-effective. For example, they did not provide a cost-effective threshold, that is, a level in which the cost-effective ratio should meet in order for the intervention to be regarded as cost-effective.

j) Discounting

Telemedicine may impose costs and benefits on the health service which may reach into the future. Therefore, the future stream of health benefits and costs should be discounted to their present value. Discounting reflects the fact that people place a higher value on events in the present than the future and the funds invested in the present can reap interest over time. Only seven studies undertook discounting, not including the discounting methods which were used to estimate the depreciation for the telemedicine equipment, because the majority of studies were short-term.

For the telemedicine equipment, discounting is needed as it represents an investment in the future provision of the healthcare and therefore, the equipment needs to be annuitized over the useful life of the asset, usually this being five years for most telemedicine studies identified in the literature review. However, there is no general

agreement about the expected life span of telemedicine equipment or what discount rate to use. Also, due to the rapid development of the telemedicine technologies, the likelihood of the technology becoming obsolete is high.

k) Sensitivity analyses

Sensitivity analyses are conducted to assess the robustness of the study results and are considered essential in demonstrating the validity of results. Any economic analysis of telemedicine interventions should undertake sensitivity analyses to deal with the uncertainty around key parameters (or assumptions) such as resource use, costs and outcomes. Depending on the circumstances, sensitivity analyses may be a simple one-way or multi-way sensitivity analyses, a threshold analysis or a probabilistic sensitivity analysis [Drummond et al, 2005]. The findings from sensitivity analyses will indicate how sensitive the results are to these uncertainties.

Some of the uncertainties surrounding telemedicine studies include for example: the choice of discount rate used in calculation of the equipment costs; expected lifetime of the equipment; and any anticipated future changes in equipment and transmission costs. Only 33.9% studies identified in the literature search undertook some sort of sensitivity analyses.

3.4.2 General discussion and comparison with other systematic reviews of telemedicine

The results from the literature search make it difficult to generalise these findings to other telemedicine interventions or to other ICTs. For example, an economic evaluation which has been conducted for a telemedicine intervention in a rural area is unlikely to generate the same cost-effectiveness evidence in an urban area (see study by Loane et al (2001a) as an example). It is important to note that telemedicine interventions may be cost-effective in one area, but when transferred to other areas in which local services, access, quality of care, and unit costs differ between settings and/or countries, the telemedicine intervention may not be cost-effective. This is reiterated by Håkansson and Gavelin (2000) who mentioned that “there are country-specific variations in the health systems that make it difficult to generalise the results from one country to another”.

Another limitation of these studies identified in the literature review was that most were conducted in the late 1990s or in early 2000s (63 studies were published before 2003) and these older studies may not reflect current telemedicine price structures. It is possible given the likely decrease in technology costs, that telemedicine costs for these

studies were higher than telemedicine costs in more recent studies, so that the cost calculation may be confounded by the study's year of operation or publication (this search having been undertaken in October 2007). There can also be a delay of 2 to 3 years before the publication of the study article or results. For example, the study by Bynum and colleagues (2003) was conducted during 1998 to 2002 and the results were published in 2003. The literature search shows what was happening some years ago and not what is happening today. It is also very difficult to evaluate telemedicine, as it is a constantly changing technology, and technological advances in telemedicine over the years have made the equipment less expensive and easier to use [Norris, 2001] and those studies which may not have been cost-effective or cost-saving then, may now be cost-effective or cost-saving, due to the lower prices of the telemedicine equipment or communication links (e.g. ISDN calls).

A further major limitation is that there was not enough consistency in the cost (and effectiveness) data that were collected; hence, the majority of studies could not be included in the meta-analysis. Most meta-analyses usually only summarise the clinical effects and not the costs because of the heterogeneity of the data. The thesis set out to conduct a meta-analysis of the costs of telemedicine. For example, if the same question was being asked, could the studies be pooled together? Due to the heterogeneity of the data, the majority of the studies could not be included in the meta-analysis (presented in Appendix 3 for illustration purposes only). The Cochrane Collaboration state that the decision to pool any cost-effectiveness estimates using a meta-analysis should be interpreted with caution because the metric in question may not have the same meaning across studies. That is, the resource use and costs may vary both within a country and between countries. Therefore, this limits the generalisability and transferability of resource use and cost estimates across settings [Higgins and Green, 2009]. This may be particularly pertinent to telemedicine where, as we have previously highlighted, there is a significant degree of heterogeneity in study methods and findings.

Whitten et al (2000) attempted a meta-analysis of telemedicine research studies of the costs associated with telemedicine. They searched six electronic databases and identified 551 articles for analysis, of which only 38 studies had quantitative cost data. The authors found that most of the 38 studies were inadequately designed or conducted, so they were unable to perform a traditional meta-analysis. The authors concluded "that it is premature for any statements to be made, either positive or negative, regarding the cost-effectiveness of telemedicine in general".

A few systematic reviews have assessed the cost-effectiveness of telemedicine studies and inconsistencies can be found in the studies with respect to their methods, results and conclusions about telemedicine see Table 3.2 below and Appendix 5 for more details. However, they all generally come to the same conclusion that evidence on the cost-effectiveness of telemedicine studies is still limited and that most cost-effectiveness studies of telemedicine were not generally carried out in accordance with standard economic evaluation guidelines.

Table 3.2: Summary of other systematic reviews of telemedicine

Authors	Publication year	Type of telemedicine studies considered in the review	Number of databases searched*	Number of abstracts retrieved	Number of articles that met inclusion criteria	General conclusion provided by the author(s)
Whitten et al	2002	Cost-effectiveness of telemedicine	4	612	24	"There was no good evidence that telemedicine is a cost-effective means of delivering healthcare compared to standard healthcare delivery".
Roine et al	2001	Effectiveness and cost-effectiveness of telemedicine	8	1,124	50 (16 were economic analyses)	"Evidence regarding the effectiveness or cost-effectiveness of telemedicine is still limited. Based on the current scientific evidence, only a few telemedicine applications can be recommended for broader use".
Hailey et al	2002	Benefits of telemedicine	8	1,323	66	"Although useful clinical and economic outcomes data have been obtained for some telemedicine applications, good quality studies are still scarce and the generalisability of most assessment findings is rather limited".
Hailey et al	2004	Benefits of telemedicine	6	605	48	The authors felt the results confirmed previous findings and that good quality studies are still scarce.
Mair and Whitten	2000	Patient satisfaction with telemedicine	4	?	32	"Methodological deficiencies (low sample sizes, context, and study designs) of the published research limit the generalisability of the findings".
Williams et al	2001	Patient satisfaction with telemedicine	4	125	93	"The current evidence concerning patient satisfaction with telemedicine is rather limited."
Hersh et al	2002	Efficacy of telemedicine for making diagnostic and management decisions	4	4,709	58	"Despite the widespread use of telemedicine in most major medical specialties, there is strong evidence in only a few of them that diagnostic and management decisions provided by telemedicine are comparable to face-to-face care".
Hersh et al	2006	Effect of telemedicine on diagnosis and management decisions, patient outcomes and access to care	1	4,083	106	"There are still significant gaps in the evidence base between where telemedicine is used and where its use is supported by high-quality evidence".
Whitten et al	2007	Research methodology in telemedicine studies	15	1,615	85	"Until the telemedicine field adheres to agreed standards of reporting methodological details it will be difficult to draw firm conclusions from review studies".
Bergmo et al	2009	Cost-effectiveness of telemedicine	10	779	33	"The majority of the economic evaluations were not in accordance with standard evaluation techniques".

*ISI Web of Knowledge counted as one database, but consists of ISI Science Citation Index, ISI Social Science Citation Index and ISI Arts and Humanities Citation Index

The results from the literature review do not attempt to make the case for cost-effectiveness of telemedicine and, as previously discussed, it would be inappropriate to attempt to do so. They do however, highlight some of the limitations in previous literature reviews on costs and cost-effectiveness of telemedicine interventions. Cost-effectiveness of telemedicine interventions depends not only on the service being evaluated, its comparator, the perspective of the analysis, patient group and sample size, type of economic analysis and how the costs and outcomes were measured and valued, but also on the take-up rate and the usage of the service. Decision makers and readers must be cautious as to the degree to which they can apply the results of such assessments to their own circumstances. Further research needs to be done in which telemedicine interventions and their comparators are conducted in accordance with general standards for health economic evaluations. Other issues which need to be considered include the sustainability of a telemedicine service, decisions about the equipment and telecommunications, and impact on the overall use of health resources and the measurement of outcomes.

3.4.3 Three important challenges for the TelePaed study and for economic evaluations in general

Section 3.4.1 highlighted some of the economic issues which arose from the literature review. In this section, three of the challenges to economic evaluations of telemedicine which emerged from the literature review will be discussed: non-randomised studies and the issue of selection bias; calculation of patient costs; and measures of benefits such as QALYs. These three issues will be addressed in this thesis, in terms of how the original analysis for the TelePaed study can be improved.

1) Non-randomised studies and selection bias

Due to the dominance of pilot studies or short-term studies of telemedicine interventions or services, very few evaluations have been conducted as part of an adequately powered RCT. The literature review found that only 12 studies (11.0%) were RCTs of telemedicine. In a RCT, the intervention or treatment is randomly allocated, so bias is distributed between the groups by chance. The randomisation process enables researchers to attribute the outcome to the intervention. As a result, there will be high internal validity, control of potential confounders and minimisation of random variation on costs and effects. However, practice within a RCT might be atypical (e.g. the patients, setting, protocol), with inadequate follow-up time, inadequate sample size for economic evaluations, inappropriate endpoints and there will also be limited external validity. The RCT is an essential tool in assessing the likelihood of a causal effect between an intervention and an outcome and as such, is useful in

showing a proof of concept. However, additional evidence may be required beyond this to examine the effect of an intervention in practice. An ideal clinical trial would be one which takes a pragmatic approach and the aim would be to evaluate the cost-effectiveness of the intervention under real life conditions that would prevail once the intervention was in routine use. For the pragmatic trial there would be some compromise between the goals of internal and external validity, there would still be random allocation to interventions to minimise bias, but would offer fewer restrictions in how patients are recruited and followed and thereby increasing the external validity or generalisability. In practice, further evidence from both RCTs and pragmatic studies are required to support the more widespread adoption of telemedicine.

As the majority of telemedicine studies were non-randomised, there is a potential for bias. Bias or systematic errors weaken the internal validity of studies and this can result in an incorrect estimate of the association between treatment and effect. The greatest difference between randomised and non-randomised studies is the risk of selection bias. Selection bias refers to an absence of comparability between the groups being studied. That is, the selection of patients for the telemedicine group have different characteristics from those allocated to the conventional group.

When designing the telemedicine study, if a RCT cannot be undertaken, then steps should be taken to ensure that the study is designed in such a way that the various biases which can arise are minimised. For example, in some instances it may not be ethical to randomise individual patients, so in this case, randomisation should take place at the hospital or centre level. The literature review identified only one non-randomised study which looked at whether the two comparator groups were different. Rendina and colleagues (1998) looked at whether the utilization of a telemedicine system for the interpretation of neonatal echocardiograms reduces the neonatal intensive care unit length of stay of low birthweight infants, and they estimated whether the two comparison groups were significantly different. They used a multiple regression model to distinguish the effects of telemedicine from those of other risk indicators, which might have differed in the two groups. They found a statistically non-significant reduction of 5.4 days in the length of stay of low birthweight infants ($p = 0.37$); however, they did not say whether or not selection bias was reduced after adjusting the length of stay in the regression model.

The case study introduced in Chapter 2 was not randomised (further details are provided in Chapter 4 about the study design), as it was deemed unethical to randomise individual patients. There are various methods to reduce selection bias in

non-randomised studies and Chapter 5 will focus on the issue of minimising selection bias for this case study and see what impact this has on the costs and effects.

2) Calculation of patient costs

A few of the studies identified in the literature review looked at the distance travelled by patients as a measure of access [Loane et al, 1999; Bynum et al, 2003; Doolittle, 2000]. From the telemedicine literature, access can be defined in various ways: the ease with which health services are available and/or sustained; the length of time (or travel distance) it takes to access those health services; how easily the patient can obtain information regarding health services; and equity of access. For any telemedicine study you would need to take into account: the changes to the number of patients using the service who previously wouldn't have been seen; changes in patient and/or health care professional travel plans; changes in the proportions of patients that 'did not attend'; and the changes in treatment centre's catchment area. Patients who experience reduced travel or reduced distances to services would imply they are experiencing better access to services.

For example, Bynum and colleagues (2003) in the USA evaluated patients' cost savings in a telemedicine project during 1998-2002. They assessed patients' cost savings with telemedicine regarding travel, and travel distance was measured. Their findings indicated that without telemedicine, 94% of the patients would have to travel greater than 70 miles for medical care, whereas with telemedicine, 98% of the patients travelled less than 30 miles to receive medical care; and 92% of the patients saved at least 40 miles in travel distance. This study highlighted the importance of distance when accessing a telemedicine service and the important implications on patients' cost savings.

With the introduction of telemedicine, the patient and the health care provider would want to know the distance they have to travel to receive or provide healthcare at the telemedicine site; how long this would take; the frequency of visits, and whether this would lead to an increase or decrease in their own expenditure. Patient costs are a subset of direct non-medical costs, and it is important to look at the costs that would fall on a patient, as patients can influence the uptake and usage rates of healthcare services. Fifty-four of the 109 studies (49.5%) identified in the literature review included some information on patient costs. However, the majority of these studies did not provide enough detail on how the costs were obtained and calculated. For example, Bergmo (1997) in Norway conducted a cost comparison of three different methods of providing consultations for ear, nose and throat problems. Patient travel

costs consisted of air tickets, subsistence, and the cost of a guardian who had to take a full day's absence from work to attend the consultation, because 50% of the patients were children or older persons'. However, the author did not provide any further information on the calculation of these costs.

To capture these results in an economic evaluation, a questionnaire can be given to patients and to health care professionals to find out how far they had to travel and what the time and costs implications were. The results from the questionnaire could inform future providers or other users of telemedicine of what the time and cost implications might be of setting up and using a telemedicine service. Using various data sources and assumptions, Chapter 6 focuses on the calculation of patient costs for the cohort of patients in the case study.

3) Measures of benefits such as QALYs

a) What are QALYs?

In a cost-utility analysis, outcomes are usually measured as health related preferences, which are most often expressed as QALYs gained and the results are expressed as cost per QALY gained. QALYs combine quality of life and length of life into a single index value. So for a QALY, the quality gains are from reduced morbidity and the quantity gains are from reduced mortality. The survival benefit is expressed as life years gained and the health-related quality of life benefit is described in terms of health states and a score is placed upon the health state indicating how respondents value each health state. The values are on a scale where 0 = dead and 1 = full health. The QALY represents the equivalence of being alive for a year in full health.

The quality component usually comes from individuals' preferences on health states and in turn, a utility value is obtained for each health state. Utility values can be obtained in two ways: direct or indirect measurement. Direct measurement is where individuals' value their own health directly. There are two main approaches to direct measurement these include: the standard gamble approach which is based on the axioms of expected utility theory and asks respondents to make choices that weigh improvements in health against mortality risks [Torrance, 1986]; and the time-trade off approach which is a method for valuing health states that asks respondents to make hypothetical choices that weigh improvements in health against reduced longevity [Torrance, 1986].

Indirect measurement is where individuals' value their own health using instruments such as the EQ-5D or the Health Utilities Index (HUI). For example, the EQ-5D is a

simple classification system which has five attributes and each attribute has three levels (as reported in Chapter 2, section 2.5.5.2). The five attributes and three levels result in 243 possible health states, to which unconscious and dead have been added for a total of 245 health states. This is then converted to a single-index utility score which falls on the 0 (dead) to 1 (full health) value scale, using a set or 'tariff' of values [Dolan, 1997]. Another example is the HUI and there are three versions of HUI, each includes a health status classification and a utility scoring formula. The HUI mark 1 has 4 dimensions, between 4 and 8 levels on each dimension, which results in 960 health states; HUI mark 2 has 7 dimensions (sensation, mobility, emotion, cognition, self-care, pain and fertility), between 3 and 5 levels on each dimension, which results in 24,000 health states; HUI mark 3 has 8 dimensions (vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain), between 5 and 6 levels on each dimension, which results in 972,000 health states [Horsman et al, 2003].

b) Are QALYs appropriate for telemedicine?

For cost-utility analyses, costs are measured in monetary units and the outcomes are usually units that relate to a person's health-related quality of life or proxies for measures of utility, such as QALYs. Results for cost-utility analyses are expressed in terms of cost per QALY gained and QALYs can provide a 'common currency' for comparison. QALYs are the preferred outcome measure for NICE [NICE, 2008]. NICE states that "given its widespread use, the QALY is considered to be the most appropriate generic measure of health benefit that reflects both mortality and health-related quality of life effects. It is recognised that alternative measures exist (for example, the healthy-year equivalent), but few economic evaluations have used these methods and their strengths and weaknesses are not fully understood" [NICE, 2008]. NICE believes that the EQ-5D is appropriate when calculating QALYs, because it is a generic-preference based measure which has been validated on a UK population. Also, given the comparative nature of their work at NICE and the consistency across appraisals, calculating QALYs for any technology or intervention, allows comparisons across technologies or interventions.

Thus, calculating QALYs for telemedicine services would be in line with mainstream health economics, but there is still some debate concerning whether meaningful differences in QALYS can actually be measured in the case of telemedicine. McIntosh and Cairns (1997) in their paper which looked at a framework for telemedicine applications thought that it would be difficult to estimate QALYs for telemedicine applications: "In many evaluations it will be difficult to attribute health benefits or changes in health outcomes to telemedicine itself. That is, even if changes in health

outcomes are due to telemedicine it will be difficult to quantify this" [McIntosh and Cairns, 1997]. Aoki et al (2004) believed that "telemedicine itself may not directly alter quantitative clinical outcomes such as mortality and morbidity, because it is not a direct therapeutic or diagnostic instrument". However, the authors felt that telemedicine was a useful tool for improving communication, accessibility and management and suggests that telemedicine may have an indirect impact on qualitative clinical outcomes, such as patient satisfaction.

For many economic evaluations of telemedicine interventions, QALYs may not be the preferred outcome measure as they may not capture all the appropriate health or non-health benefits of using telemedicine such as improved treatment, faster and more accurate diagnosis, transfer of skills and knowledge, speed of service, improved training and education, and reassurance. Non-health benefits of telemedicine require an instrument which is sensitive enough to distinguish between process care changes which are expected to improve health outcomes such as earlier diagnosis and process care changes that are simply about improved service delivery or reduction in costs.

One of the further difficulties with any economic evaluation of telemedicine is how to capture the changes in effectiveness. For example, due to the diversity of the use of telemedicine (telemedicine can be used for clinical purposes i.e. for checking skin diseases or for administrative purposes i.e. for meetings), it may be difficult to know which effectiveness measure to use, that will be appropriate and comparable across all telemedicine services and interventions. There may also be problems measuring and valuing this information. Such as attributing changes in effectiveness, to changes in the process of care, rather than to the nature of care itself can be problematic. Hence, QALYs may not be appropriate for telemedicine services.

QALYs may capture some of the health benefits of telemedicine such as anxiety and depression. For example, the EQ-5D questionnaire includes an 'anxiety/depression' domain; the HUI mark 3 instrument has an attribute called 'emotion'; and the SF-6D includes a 'mental health' component. However, none of the instruments which are used to estimate QALYs include non-health benefit components such as patient satisfaction, education or speed of service.

To fully capture these health (and non-health) benefits from telemedicine, for any economic evaluation of telemedicine you would need to include one of those instruments (for example, EQ-5D, SF-6D or HUI mark 3) for eliciting QALYs; but also to supplement this with questionnaires that will capture the other health and non-health

benefits of telemedicine. For example, within any economic study a questionnaire or survey could be given to patients to find out their perceptions about the telemedicine service, such as whether they were satisfied with the service they received. This survey could also incorporate questions on time and money costs. However, this may not always be practical due to time or resource constraints and in most cases decision analytical modelling is used instead. Using a decision analytical framework and various assumptions, expert opinion and literature, one can look at the longer-term benefits such as QALYs of telemedicine services.

The literature review identified two studies which used QALYs for measuring health outcomes for telemedicine – both used telemedicine for eye disease. In the first study by Castillo-Riquelme and colleagues (2004), the authors used a decision analytical model to compare the cost-effectiveness of alternative methods for screening for eye disease in newborns. The authors used utility estimates from a study on cataracts and converted these utility values into QALY estimates over the whole expected lifetime of the babies. The second study by Aoki et al (2004) looked at a cost-effectiveness analysis of telemedicine to evaluate diabetic retinopathy in a prison population. The authors used a Markov decision model and probabilities and utility values were obtained from the literature which enabled the calculation of QALYs.

QALYs are thought to be difficult to calculate for telemedicine interventions, because the benefits of reduced mortality and morbidity associated with telemedicine are not easily quantifiable. That is, QALYs may not be sensitive enough to detect small changes in health outcomes which telemedicine services are most likely to produce. Telemedicine doesn't actually kill anyone (mortality), unless they are misdiagnosed or help prolong life, but indirectly we can say that it can help to prolong life with (faster) diagnosis and treatment. Telemedicine has mainly been used as a method of screening or confirmation for diagnosis. In this thesis, using various sources, such as data from patients', data from literature, expert opinion and assumptions, QALYs will be estimated for this patient cohort (see Chapter 7 for more detail).

3.5 Summary and next steps

The results from the literature review are consistent with previous findings and there is still no further conclusive evidence that telemedicine interventions are cost-effective (or not cost-effective as maybe the case) compared with the conventional/standard/usual delivery of healthcare. Even though more studies have been added, the conclusions have remained the same.

For telemedicine interventions to be demonstrated as being cost-effective depends on the service being evaluated, the take-up and usage rate of the intervention, and its comparator, plus the research viewpoint, type of economic analysis, and the costs and outcomes which were collected. Most telemedicine studies had the potential to be cost-saving or cost-effective, but this depended on the number of patients that used the service or if the high start-up costs of the telemedicine equipment were eliminated. One of the main limitations of the independent review was that there was not enough appropriate information presented in the papers for meta-analysis of the cost-effectiveness of telemedicine studies to be conducted.

Decision makers in health care (those who fund and deliver such services) require assurance that telemedicine can fulfil its promise, not only that it is effective, but that it also represents good value for money. Telemedicine services that move on from pilot studies or short-term studies to those which are used in routine or longer-term use need to be assessed more routinely, so that the longer-term impact on costs, outcomes and the health service organisation can be quantified. As telemedicine services move into routine use other factors may need to be considered in the economic evaluation such as: sustainability of the level of telemedicine activity and the organisation of health services. Also, if the study is non-randomised then biases which arise within the study such as selection bias need to be minimised.

Finally, and most importantly, how investigators conduct and report their findings on telemedicine interventions needs to be improved. Whitten et al (2007) showed that their meta-analysis of research methodology proved difficult, because they could not ascertain all the necessary information from the published articles. The importance of missing such methodological details should not be underestimated. For example, not reporting start and end dates could potentially have an impact as we know that technology changes over time; and also not explicitly stating the research question means that readers may not know whether the study design was appropriate. Authors need to provide more transparent accounts of their study, describing in detail lengths the approaches undertaken, perspective of the analysis, sources of data, unit costs (and its components) and outcomes collected and reported, assumptions made, the reliability of their results and how generalisable their findings are to other settings.

When conducting economic evaluations of telemedicine studies, researchers should ensure that they follow guidelines such as the Drummond checklist [Drummond et al, 2005], therefore to ensure transparency in reporting of methodologies and results. This will enable readers to make comparisons to their own settings and to see whether

results can then be transferable and generalisable. Economic evaluations of telemedicine studies should contain the basics from the checklist such as:

- What is the research question?
- Who is the study for?
- The type of economic evaluation conducted
- How are costs and outcomes going to be measured and valued?
- Does discounting have to be performed?
- Are results presented in terms of an incremental analysis? and
- Was uncertainty in assumptions and results accounted for?

This applies not only to telemedicine, but for all economic evaluations of clinical services. As mentioned earlier in Chapter 1, there have been a few papers building on the Drummond checklist, looking at a framework for the economic evaluation of telemedicine (e.g. McIntosh and Cairns, 1997; Sisk and Sanders, 1998). These frameworks have looked at some of the other problems faced when conducting an economic evaluation of telemedicine such as: constantly changing technology and small sample sizes (some of these points have been discussed further in this Chapter). One important final point is that maybe telemedicine intervention and the alternative(s) are not the problem, but how the costs and outcomes are collected and reported, leads to different conclusions.

The next Chapter will look at the design of the TelePaed study, a critique of the case study will be provided and this will be followed by a discussion of the three economic issues identified in the literature review in more detail: selection bias, patient costs and the measures of benefits such as QALYs in the context of telemedicine services. The following Chapters (5 to 7) will look at how each of these three issues can be addressed in the context of the TelePaed case study in order to improve the analysis.

CHAPTER 4: DESIGN AND CRITIQUE OF TELEPAED CASE STUDY AND ECONOMIC ISSUES ASSOCIATED WITH TELEMEDICINE

4.1 Introduction

In the previous chapter, a literature review was conducted which looked at the costs and cost-effectiveness of telemedicine. The findings from the literature review identified some economic issues which may be of concern when conducting economic evaluations of telemedicine; especially in relation to the case study presented in Chapter 2. The aim of this chapter is to look at the design of the case study (including the wider TelePaed study) in more detail and to reflect on some of the problems with the design of the study. This chapter will then outline the next steps for this thesis i.e. how best to solve the three issues identified in the literature review in the context of an economic evaluation of telemedicine.

4.2 Design of TelePaed study and configuration of antenatal screening services

4.2.1 Brief overview

In 2001, under the guidance of the NHS ICTRI-1 [DoH, 1998], the RBH in west London, introduced a telemedicine service for district hospitals in southeast England that was designed for use in paediatric departments, neonatal units, and obstetric departments. The RBH already had established paediatric telecardiology links with hospitals in Greece and Portugal [Neophytou, 2000]. The four hospitals were Basildon, Colchester and Southend in Essex and Medway in Kent and each hospital was between 35 to 65 miles from central London. The paediatricians and obstetricians decided on the precise role for the telemedicine service within their hospital and they could determine how the new service could complement their existing arrangements for obtaining specialist advice. The telemedicine service could be used for face-to-face consultations with the patient present and live ultrasound images and heart sounds would be transmitted via an electronic stethoscope; or a 'store-and-forward' approach could be adopted, whereby the district clinicians would transmit pre-recorded videoed ultrasound images in the absence of patients. The main fieldwork for the TelePaed project was conducted over a 15-month period (1st May 2001 to 31st July 2002).

The project was not set out as a 'true experiment', but as an evaluation of a service delivery organisation. The project was designed to evaluate the role of telemedicine in facilitating the diagnosis and advice on the management of three patient groups:

1. Newborn babies with suspected heart abnormalities;

2. Older infants and children for whom a cardiac opinion is required. These patients would normally be seen by specialists in scheduled outreach clinics, unless the case was urgent; and
3. Pregnant women suspected of a fetal cardiac anomaly.

Ethical approval for the project was obtained from a multi-centre research ethics committee and the relevant local research ethics committees for each of the district hospitals and the specialist centre [Dowie et al, 2007].

4.2.2 Selection and randomisation of district hospitals

Two paediatric cardiologists from the RBH (who were also part of the TelePaed project team) routinely held paediatric outreach clinics in 19 DGHs and these hospitals formed the sample frame for the selection of the project hospitals [Dowie et al, 2003]. Four of these 19 hospitals were chosen for the TelePaed project and the criteria for selection focused mainly on the relatively large obstetric caseloads (over 3,500 deliveries annually) to ensure that the analyses were based on adequate numbers; their distance from the specialist hospital in London; and the presence of a district hospital consultant paediatrician with an interest in paediatric echocardiology [Dowie et al, 2003]. Of the hospitals which were chosen, Basildon and Southend only had paediatric links with the RBH; whereas, Colchester and Medway had both fetal and paediatric links with the RBH.

A pragmatic research design was adopted with the hospitals as the unit of randomisation (a cluster randomised trial) [Dowie et al, 2003]. Because of the uncertainty about how the district hospital clinicians would adopt the telemedicine service, the hospitals were randomised rather than individual patients. Also, it was deemed unethical to randomise individual patients due to the nature of the presenting conditions and for these patients any delay in the consultation was not appropriate.

The initial research protocol proposed that two of the hospitals would act as intervention sites and receive the telemedicine equipment. The other two hospitals would act as control sites and when the fieldwork finished they would also be offered the telemedicine equipment. The telemedicine equipment for each of the district hospitals would be supplied by the project, but the hospitals would incur any non-research costs such as the call charges for the teleconsultations. However, when purchasing the telemedicine equipment for the two intervention sites, it was cost-saving to purchase all four telemedicine packages in a single order. The control hospitals were also provided with their telemedicine equipment at the same time as the

intervention hospitals on the condition that it could be used for other purposes, but not for patients involved in this project [Dowie et al, 2003]. Before the introduction of the telemedicine equipment in each hospital, staff in all four hospitals underwent additional training in fetal ultrasonography and paediatric echocardiography.

Using a random number generator, a statistician in Brunel University who was independent of the project team and unaware of any preferences within the project, allocated the hospitals as:

- Intervention hospitals: Basildon and Medway
- Control hospitals: Southend and Colchester [Dowie et al, 2003].

4.2.3 Configuration of the antenatal screening services

Even though the four hospitals were similar in terms of the number of maternal deliveries each year, they were different in terms of the configuration of their antenatal screening services. For example, anomaly ultrasound scans were offered in all four hospitals, although each hospital was different at which gestational age the scan was offered. Eighteen to 20 weeks was the norm for Colchester and Southend hospitals; Medway hospital conducted anomaly scans at 20-22 weeks gestation; and for Basildon hospital anomaly scans were delayed until 23 weeks, as nuchal translucency screening for Down's syndrome was universally offered to all women (nuchal screening was selectively offered in the other 3 hospitals) [Dowie et al, 2004]. Serum screening tests were offered in all four hospitals, from about 15 weeks gestation. However, if women in Medway hospital had undergone nuchal screening they were not offered this test. In terms of serum screening programmes, the hospitals were different in terms of the choice of tests which were undertaken and the risk cut-off levels. Basildon and Southend hospitals conducted double tests (1 in 250 cut-off); Colchester hospital conducted a triple test (1 in 250 cut-off) and Medway hospital conducted a quadruple test (1 in 300 cut-off) [Dowie et al, 2004].

It was important to take into account the different screening policies across the four hospitals, as this helped determine the eligibility criteria for entering pregnant women into the TelePaed database and in turn these policies influenced the final numbers of pregnant women for the hospitals in the database.

4.2.4 Postal surveys

Postal surveys were sent out two to three weeks after the initial consultation to the pregnant women who were eligible for entry in the audit databases in the two hospitals that had links with the fetal cardiology service at the RBH (Colchester and Medway

hospitals) and to the mothers or guardians of infants and children who were seen as outreach outpatients in the four DGHs. For infants and children, the questionnaires were divided into five age groups according to the relevant health status instruments: Qualin [Manificat et al, 2000] for babies aged 1 to 3 months; 4 to 12 months; and 13 to 24 months; and Pediatric Quality of Life Inventory (PedsQL) [Varni et al, 2001]) for children aged 25 to 59 months; and 60 months (5 years) and older.

These questionnaires were used to gather information on the expenditure incurred by patients (and their families) and also information on their health-related quality of life during their first visit to the specialist. Detailed information on the health status instruments used for pregnant women and the calculation of patient costs was set out in Chapter 2 (sections 2.5.5.1 and 2.5.5.2). Follow-up questionnaires were sent three months after the initial questionnaire; however, these questionnaires only included the health status instruments.

4.2.5 Economic evaluation approach

A cost-consequences approach was adopted for the economic analysis from the viewpoint of the hospital. The aim of the economic evaluation was to identify any differences in the relevant costs and consequences for the patient groups that used telemedicine and those that did not use telemedicine.

4.3 Critique of the TelePaed study - what actually happened in practice?

4.3.1 Uptake and usage of telemedicine services

From August 2001, the consultant obstetricians and paediatricians in the two intervention hospitals were permitted to utilise the telemedicine service. Uptake of the telemedicine service in the intervention sites was slower than anticipated and not consistently used for the three patient groups. To compensate for the delay in the uptake of the technology, from February 2002 onwards the two control hospitals were invited to use the telemedicine service. The project was therefore extended by three months to allow the control hospitals to introduce the telemedicine service, notably to provide sufficient evidence on the use of new services for the research evaluation. Hence, the project was no longer a randomised trial and was classed as an observational study. However, there was always the question as to whether the TelePaed study was a randomised trial. In most instances, patients should be randomised but in some situations this may not be practical, therefore the use of cluster RCTs is recommended; however there may not be enough clusters in each arm of the study. Even by the study design for the TelePaed study, two clusters in each arm is questionable; especially when you look at it in terms of statistical power which

are required to detect statistical differences in RCTs. (It is worth noting that the original protocol proposed recruiting eight hospitals, but research finance for such a large-scale evaluation was not available). Table 4.1 below summarises the use of telemedicine for the four district hospitals.

Table 4.1: Use of telemedicine and specialist referrals from the DGHs

	Basildon	**Colchester**	**Medway**	**Southend**	**Total referrals**
Duration of TM access:	12 months	6 months	12 months	6 months	
TM service used for:					
Newborn babies	✓	X	X	✓	17
Older infants	✓	✓	✓	✓	48
Pregnant women	X	X	✓	X	52
All TM referrals	38	11	61	7	117

Key: TM = telemedicine; ✓ = TM service used; X = TM service not used

For fetal cardiac advice, this slow uptake may have been because Basildon and Southend hospitals did not have fetal cardiology links with the RBH, instead they had historical links with specialists based in another fetal cardiology unit in London; and Colchester hospital had an internal referral service provided by a consultant obstetrician with advanced fetal heart diagnostic expertise [Dowie et al, 2004]. Only Medway hospital used the telemedicine equipment for fetal cardiology assessments on a regular basis, and also the telemedicine equipment was only suitable for screening fetal hearts once the women reached 18 weeks gestation. Thus, as the study was no longer randomised, an observational design was adopted, to compare the costs and outcomes of patients referred to specialists from all four hospitals by means of the telemedicine service or by conventional methods (face-to-face consultations). The cohort of pregnant women who received specialist advice via telemedicine in Medway hospital were compared with the cohort of pregnant women in all four hospitals who received specialist advice via face-to-face consultations [Dowie et al, 2007]. However, for this thesis, methods and results for only Medway hospital are presented.

The selection of patients for the telemedicine service in Medway hospital was based on TelePaed eligibility criteria (see Chapter 2, section 2.4.2 for more detail); and as noted earlier it was unethical to randomise individual patients (however, this was never tested by an ethics committee). Thus, the pregnant women's risk factors played an important part in the district clinician's choice of referral mode. For example, women who were classed as 'high risk' were more likely to be seen in London for a face-to-face consultation, unless a telemedicine clinic was scheduled to take place within the next few days. Hence, pregnant women who were referred to see a perinatal cardiologist

for face-to-face assessments were not 'directly comparable' to the telemedicine women. As these women were not selected on a random basis for the two referral modes, they are most likely to be 'biased' i.e. that selection bias played an important part in choosing which referral mode the pregnant women were allocated to.

4.3.2 Outcome of postal surveys

With respect to fetal cardiology, postal questionnaires were sent to eligible women in Colchester and Medway hospitals. The number of questionnaires received from women in each hospital was: 125 out of 235 (53.2%) - Medway hospital and 115 out of 201 (57.2%) - Colchester hospital [Dowie et al, 2004]. The majority of questionnaire responses were from women who were screened locally in the district hospitals. Of the women respondents who were referred to a specialist fetal cardiologist, 29 women from Medway hospital (23 whose scans were videoed for further assessment and 6 women who went to London – these results were presented in Chapter 2) and 4 women from Colchester hospital.

In terms of patient costs calculation, information on the expenditure incurred by patients was only collected at one time point. This is important as over the course of their pregnancy, a woman has to travel often to the district hospital (in some instances, to the specialist hospital) for prenatal clinical attendances and ultrasound scans and these additional expenses can add up.

4.3.3 Measures of benefits

As the project set out to establish the cost-consequences of using telemedicine for specialist advice, in terms of the benefits which were recorded in the audit database for fetal cardiology, these mainly concentrated on clinical outcomes. The project aimed to collect and present the demographic characteristics, the clinical findings and specialist referral patterns according to the four DGHs [Dowie et al, 2004].

In terms of the clinical findings, the study identified the number of pregnant women with fetal cardiac anomalies (and in some instances, fetal non-cardiac diagnoses were also identified). Of those pregnant women with suspected cardiac abnormalities, they were categorised according to severity: mild (e.g. irregular heartbeats at 37 weeks); moderate (e.g. coarctation of the aorta; atrio-ventricular septal defect); and severe (e.g. hypoplastic left heart; aortic stenosis) [Dowie et al, 2004; Dowie et al, 2008]. As mentioned earlier, the incidence of fetal CHD is very low (1 in every 145 births) [Peterson et al, 2003], so only a limited number of abnormal cases would be detected in a 15-month period in a DGH with approximately 3,800 births per annum.

So in terms of measures of benefits which were recorded in the audit database, these were the number and severity of clinical fetal outcomes (i.e. normal, cardiac or non-cardiac diagnosis) and no other maternal or neonatal long-term health benefit measures were recorded such as QALYs.

4.3.4 My role in the project

In July 2002, when the fieldwork was almost complete, I took over the role as the health economist for the TelePaed project. I had no input or influence into the design of the study or how the postal surveys were conducted. My main role for the project has been outlined earlier in the section on Publications and Authorship. I contributed fully to the production of the four reports which were submitted to the Department of Health and the six papers which were published in peer-reviewed journals (two articles of which I was a lead author).

4.3.5 Applying Drummond check-list to the TelePaed study

Table 4.2 below shows the Drummond check-list for assessing economic evaluations [Drummond et al, 2005] being applied to the TelePaed case study which was presented in Chapter 2. This was to check whether the study presented in Chapter 2 was carried out in accordance with criteria required for a good 'economic evaluation'. As you can see from Table 4.2 where appropriate, the check-list was as complete as could be in accordance with the economic evaluation guidelines. On a separate note, Bergmo (2009) had stated that the Dowie et al study (2007) was one of the eight studies (out of 33 economic evaluations) which had addressed all the key issues for an economic evaluation.

Table 4.2: A check-list for assessing economic evaluations

Question	Chapter 2: TelePaed study
Was a well-defined question posed in an answerable form?	Yes – the study set out to compare the costs and outcomes of patients referred to specialists via the telemedicine service or by conventional means. The study was conducted from a hospital perspective.
Was a comprehensive description of the competing alternatives given?	Yes – the study described in detail the two referral methods for obtaining specialist advice: direct referral to London to see a specialist face-to-face for assessment or via a store-and-forward telemedicine service.
Was the effectiveness of the programmes or services established?	Yes – effectiveness data was obtained from the observational study; that is, the detection of cardiac cases prior to birth for each referral method.
Were all the important and relevant costs and consequences for each alternative identified?	Yes – the key resource use (costs) and consequences for each referral method were identified.
Were costs and consequences measured accurately in appropriate physical units?	Yes – key resource was measured in their appropriate physical units.
Were costs and consequences valued credibly?	Yes – unit costs were then applied to the key resources.
Were costs and consequences adjusted for differential timing?	Not applicable – length of study was less than a year.
Was an incremental analysis of costs and consequences of alternatives preformed?	No – as an incremental cost-effectiveness ratio would not be meaningful.
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes – one-way sensitivity analyses were applied to changing the discount rate and lifetime of the telemedicine equipment.
Did the presentation and discussion of study results include all issues of concern to users?	Yes – study results were presented in a disaggregated form; no incremental ratio was presented.

4.4 Next steps for this thesis

As identified not only from the literature review in the previous chapter, but also due to the limitations with the TelePaed study, the three economic issues highlighted earlier are of importance when conducting economic evaluations of telemedicine interventions/services and will be looked at in further detail in this thesis.

As the study was never properly randomised, the way patients were selected for referral to see a specialist may have led to selection bias and this may have resulted in biased costs and effects. In Chapter 5, selection bias will be looked at in more detail and the methods which have been identified in the literature to reduce selection bias will be applied to see what impact this has on the costs and effects. Chapter 5 will primarily look at the two groups of women who were referred for fetal cardiology (see Chapter 2 for more detail) and the costs and effects will be examined separately.

As seen in Chapter 3 in the literature review the majority of studies were pilot (short-term) or small implementation studies. Therefore, in Chapter 6, the new analysis conducted will look at all women who underwent an anomaly scan at Medway hospital during the same time period as women who were referred to a perinatal cardiologist. By looking at all women (referred and non-referred women), we are not artificially introducing selection and this is probably the best way to deal with selection bias given

the available data. Therefore, the aim will be to conduct a cost-effectiveness analysis of a service with telemedicine compared to a conventional service, that is, to look at what would happen to costs and effects if a telemedicine service was not available and also to look at the change in costs and effects over time (additional data has been collected for this purpose).

In Chapter 2, patient costs were calculated for the sample of women for whom we had data. Patient costs can have an impact on the financial situation for pregnant women, especially during their course of their pregnancy where they have to travel more than once to hospital. Usually patient costs are not included in economic evaluations and they are not in the reference case for NICE [NICE, 2008]. So in Chapter 6, the impact on patient costs for these women will be explored.

Finally, in Chapter 7, the third issue which was identified in the literature review: measures of benefits such as QALYs will be explored. Most telemedicine studies only report clinical outcomes or averted costs. Therefore, the thesis explores whether QALYs can be calculated for telemedicine services. A decision analytical model will be used to look at lifetime costs and benefits (QALYs) of the introduction of a telemedicine screening service for pregnant women whose unborn babies are at a low risk of congenital heart disease compared to a screening service without telemedicine. In this group of low-risk women, they are more likely to have a missed cardiac anomaly and the results will be expressed in terms of cost per QALY gained.

CHAPTER 5: ADJUSTING COSTS AND EFFECTS FOR SELECTION BIAS FOR THE TELEPAED DATA

5.1 Introduction

This chapter addresses the first of the three issues highlighted in the previous two chapters: selection bias. The literature review in Chapter 3 found that the majority of telemedicine studies conducted were non-randomised, and because of this there is potential for selection bias. The TelePaed case study described in Chapter 2 focused on comparing two groups of women who were referred for a specialist fetal cardiac opinion: women who were referred directly to a specialist hospital for face-to-face assessment (direct referral group) and women whose assessment was conducted via telemedicine, where pre-recorded videoed ultrasound images where relayed to the specialist (telemedicine group). As mentioned in Chapter 4, individual patient randomisation was considered unethical for this cohort due to the urgent nature of the presenting conditions. Therefore, the sonographers and district obstetricians decided whether a patient was to be assessed via the telemedicine link (used mainly for screening purposes) or to send them directly to see a specialist in a London hospital (used mainly for women who required an urgent opinion). Although, there were no statistically significant differences in demographic characteristics between the two groups, the two groups of women were different. The telemedicine women were specifically selected according to predefined eligibility criteria (see Chapter 2, section 2.4.2) and most likely the women seen directly had evidence of a cardiac anomaly. That is, patients' risk factors played an important part in the clinician's choice of referral mode. Since women were not selected on a random basis for the two referral modes, the estimation of cost and effect differences between the two groups may well have been biased. Therefore, this chapter aims to assess this bias, and to reduce its effect on the analysis.

This chapter begins with a literature review to find which methods can be applied to costs and/or effects to reduce selection bias and an overview of the methods identified. Then the various methods are applied to the TelePaed dataset to see whether it is possible to obtain more reliable estimates of costs and effects for these two groups of pregnant women.

5.2 Selection bias

The aim of this section is: 1) to describe the different types of biases that can arise in the evaluation of healthcare, focusing specifically on selection bias; and 2) to provide

an overview on the methods identified from a literature review which can be applied to reduce selection bias in healthcare.

5.2.1 Introduction to selection bias in healthcare

In primary health research, an appropriately powered RCT is often seen as the gold standard, because it has the lowest threat of bias. In a RCT, individuals are randomly allocated to two or more groups. The intervention group receives the new treatment, whereas the control group receives an alternative treatment. RCTs are considered the gold standard for evaluation of interventions and technologies because the randomisation process should balance the covariates between the comparison groups [Deeks et al, 2003]. Randomisation is the only means of controlling for unknown and unmeasured differences, as well as those that are known and measured [Kunz et al, 2005].

However, in some situations, RCTs may not be possible. In these circumstances, non-randomised studies (such as observational studies) may be the only way to assess the costs and effects of treatments. Non-randomised studies tend to have broader selection criteria for patients, and maybe more representative of the wider population; have longer follow-up periods; study costs are usually lower; and the studies are conducted under more realistic conditions. However, as participants in these studies are not randomised to either the treatment or control arm, there is potential for bias.

Bias indicates systematic error in the design or conduct of research trials that results in a distortion of the data obtained [Kielhorn and Graf von der Schulenburg, 2000] and can arise at any stage of a clinical trial. Several categorisations of bias exist in healthcare; where some have used an extensive listing of biases [Sackett, 1979; Delgado-Rodriguez and Llorca, 2004]; Feinstein (1985) and Cochrane (2005) have used four. Feinstein (1985) consolidated biases that can arise during research into four categories: a) susceptibility bias refers to differences in baseline characteristics; b) performance bias refers to different experiences of treatment; c) detection bias refers to different measurement of outcomes; and d) transfer bias refers to differential losses to follow-up. The Cochrane Collaboration Handbook (2005) highlighted four biases that can arise in non-randomised studies: 1) selection bias refers to systematic differences in comparison groups (meaning differences in baseline characteristics of individuals in different groups); 2) performance bias refers to systematic differences in care provided apart from the intervention being evaluated; 3) attrition bias refers to systematic differences in withdrawals from the trial; and 4) detection bias refers to systematic differences in outcomes assessed.

The four categories of bias for the Cochrane Collaboration are very similar to those four groups of bias by Feinstein. The term susceptibility bias is sometimes used interchangeably with the term selection bias. Feinstein defines the term selection bias as "susceptibility bias occurs if the maneuvers are received by groups whose collective baseline states have distinctly different prognostic expectations for the subsequent occurrence of the outcome event" [Feinstein, 1985]. In other words, the two groups of patients who have been selected before treatment may have baseline differences, and this can create bias even before the outcome occurs. This term is very similar to the Cochrane meaning for selection bias, where selection bias is referred to as systematic differences in comparison groups at baseline. Selection bias will be the focus of this chapter.

In a sufficiently large trial, randomisation reduces selection bias. Selection bias refers to an absence of comparability between the groups being studied and can happen during any stage of research. Selection bias arises as a result of the interaction of treatments and omitted or unobserved patient characteristics that may influence treatment choice, but independently affect health outcomes: in other words, the participants in the intervention group have different characteristics from those allocated to the control group (and these differences affect outcomes).

On average, selection bias tends to make treatment effects appear larger than they are and the size of these distortions can be as large or larger than the size of effects that are being measured [Kunz and Oxman, 1995]. Selection bias in studies may occur not only due to observed covariates, but may also be due to unobserved covariates. Both the observed and unobserved covariates can be either known or unknown to the clinicians, as well as recorded or not recorded during the study.

Despite the growing use of non-randomised studies to evaluate healthcare technologies, there is currently no 'gold standard' approach to control for selection bias in non-randomised studies. In the next section, an overview of the alternative methods used to control selection bias in non-randomised studies is provided.

5.2.2 Literature search to identify methods to reduce selection bias in healthcare evaluation

A literature search was conducted to identify methods which have been used to reduce selection bias in non-randomised studies. The search strategy including search terms

and the databases which have been searched are highlighted in Appendix 6, section A6.1.

5.2.3 Selection bias or endogeneity?

One thing that became apparent from the literature review was that the terms selection bias and endogeneity were being used interchangeably. However, the terms actually mean different things and therefore lead to two different solutions. Millimet (2001) stated "Sample selection bias and endogeneity bias refer to two distinct concepts, both entailing distinct solutions. In general, sample selection bias refers to the problems where the dependent variable is observed only for a restricted, nonrandom sample……Endogeneity refers to the fact that an independent variable included in the model is potentially a choice variable, correlated with unobservables relegated to the error term. The dependent variable, however, is observed for all observations in the data".

For example, we have a model where the dependent variable is wages and one of the independent variables is gender. There are 300 people in a dataset, where 200 are employed and 100 are unemployed.

Equation: $Wages = \alpha + \beta_1 Gender + error\ term$

With selection bias, observing wages is dependent on whether one is employed or not and the only available data is for the 200 employed people (i.e. the 100 unemployed observations are missing). In relation to endogeneity, observing wages is dependent on whether one is employed or not AND also on another variable i.e. a choice variable such as educational status (this variable is accounted for in the error term). Here, all 300 observations would be observed and pooled together. That is, the dependent variable is dependent on other things which are not included (i.e. in the error term).

In relation to the TelePaed study, this means for selection bias, one would observe the costs and effects of a woman who was seen via telemedicine, only if they were assessed via telemedicine. Likewise, one would observe the costs and effects of a woman who was seen directly, only if they had face-to-face assessment in London. If one of the groups is not observed, then this groups costs and effects are unknown, whereas, with endogeneity the problem would be identified in two ways: firstly, a choice variable as an independent variable would be included in the regression equation. The choice variable here would be one such as when the patient goes to the GP; the choice they have is on what day they see the GP or which GP they see (this would not apply in emergency situations), unfortunately we did not account for a variable like this in our data. Secondly, the entire sample would be pooled, that is, all direct referral and

telemedicine women would be treated the same; hence, as the whole sample is used here there would be no issue of selection bias. However, we know that the two referral groups are different, because selection was based on eligibility criteria. Therefore, the two groups cannot be pooled together as they are not comparable, and instead each group has to be analysed separately; hence, this is an issue of selection bias. So on this basis, the focus will be on the three methods which have explicitly said that they deal with selection bias and not endogeneity: regression analyses; propensity score methods; and sample selection models. The other methods identified in the literature review have been summarised in the Appendix 6, section A6.2.

5.2.4 Methods which can be applied to reduce selection bias in healthcare

5.2.4.1 Regression analyses

The traditional approach to control for selection bias is to use regression models. Regression models estimate how much each independent variable (such as a risk factor) relates to the dependent variable or the outcome. After estimating the regression model, the coefficients from the regression model are used to solve the regression equation for the treatment groups. When the comparison between the groups is made, the adjustments are added or subtracted from the estimated treatment effect to account for the impact of differences in each of the baseline variables according to their estimated relationship with the dependent variable [Deeks et al, 2003]. The difference in the average values of the various factors between the groups or individuals is calculated, and the mean or overall proportion is used for the entire population for each independent variable [Estrada et al, 2000]. A statistically significant coefficient of treatment (or of an interaction) involving the treatment variable indicates a treatment effect. In a regression model, the signs indicate the direction of the association between dependent and independent variables and the sizes of the coefficients indicate the strength of the magnitude of these associations.

The most commonly used regression methods are bivariate (or multivariate) linear regression analysis which attempts to model the relationship between two (more than two) variables by fitting a linear equation to the observed data and is used for continuous outcomes; and logistic regression analysis is used to model the probability of occurrence of a binary variable such as whether or not the subject has a particular symptom and the regression equation will estimate the proportion of individuals who have that symptom.

One of the main advantages of regression models is that they can include more variables than matching or stratification and can also examine the effect of each

independent variable on the dependent variable. One of the main limitations of regression analyses is that they can only adjust for observed covariates, they cannot adjust for unobserved covariates.

The pan-European SOHO (Schizophrenia Outpatient Health Outcomes) study was a three year, prospective, observational study of outcomes of antipsychotic treatment for schizophrenia in the outpatient setting. Knapp et al (2008) conducted a cost-utility analysis comparing the drug olanzapine with other antipsychotic treatments in patients with schizophrenia. The authors used multivariate regression analyses to adjust for baseline covariates (such as sex, age, weight, body mass index) to take into account the potential selection bias inherent in the study and after adjusting for selection bias the authors then estimated the incremental cost and utility gains for patients treated with olanzapine compared with the other treatments.

5.2.4.2 Propensity scores

Propensity score analysis can be used to address selection bias in the estimation of costs and effects. Rosenbaum and Rubin (1984) defined a propensity score as, "the conditional probability of assignment to a particular treatment given a vector of observed covariates". In other words, a probability model is estimated to predict the likelihood (i.e. the probability) that individuals are assigned to the treatment group, compared to the control group, based on the values of the observed variables. In observational studies we know for certain which treatment a patient receives; by estimating the propensity score, we can create a 'quasi-randomised' experiment. We may have two patients with the same propensity score, one in each group, and we can imagine these two subjects were 'randomly assigned' to each group in the sense of being equally likely to be either a treated or control subject.

The propensity score summarises all the background covariates (which are included in the calculation) such as age for each patient into a single-index variable (the propensity score). The generation of one variable allows the assessment of whether the treated and control groups overlap enough on background characteristics. When such overlap is present, the propensity score approach allows calculation of the estimated treatment versus control effects that reflect adjustment for differences in all observed background characteristics [Rubin, 1997]. So the propensity score is a balancing score i.e. the treatment and control groups have the same distribution on all observed covariates, as in a randomised experiment, and so selection bias is removed when comparisons are made between groups with the same propensity scores. If there is insufficient overlap

between the groups in a relevant variable, then the two groups are not really comparable, and propensity scoring cannot deal with the problem.

The most common method to obtain propensity scores for each patient is to conduct either a probit or a logit regression model. For this type of model, the dependent variable is treatment assignment (0 = control group; 1 = treatment group) and the independent variables are background characteristics such as age. By running a regression model for each patient in the study, a propensity score can be obtained.

Another method used to estimate propensity scores is the 'classification tree' approach. Stone and colleagues (1995) used a classification tree approach to obtain propensity scores for hospitalized and ambulatory patients with community acquired pneumonia. So essentially, a 'classification tree' partitions data into subsets (i.e. it identifies a subset of patients who have similar propensities of being selected into one of the two groups). For every partition, all observed variables remain available even if they have been used earlier in the construction of the tree. So it is possible for a single variable to reappear at several points in the tree. These subgroups are by definition the propensity score strata and a subset that is not further partitioned is called a final subset or strata. The proportion of the treated patients in each subset is used as the estimate of the common propensity for that subset.

Luellen at el (2005) said that classification trees have a number of advantages over logistic regression. Firstly, the classification algorithm automatically selects variables for the model; secondly, the algorithm also automatically detects interactions in the data; and finally, the trees' terminal nodes supply the strata, therefore eliminate the need to identify a set of stratification cut-points.

Propensity scores have the advantage of reducing selection bias in observed differences, but remain subject to bias from unobserved differences [Rubin, 1997; Laudrum et al, 2001]. Crown (2001) commented on propensity scores and said "it does not provide a direct test of the presence of selection bias, nor does it provide an estimate of the magnitude of selection bias if it is present". In other words, propensity scores only control for observed variables and by using this method you cannot tell the extent of selection bias that may still exist. Propensity scores work better in larger samples, because achieving an overlap between the treatment and control group in terms of observed characteristics increases with the sample size [Rubin, 1997], unless allocation is systematic.

One of the main limitations of the propensity scores technique is the way it handles weak covariates which are included in the propensity score estimation. A covariate which is related to treatment assignment, but not to outcomes, is treated the same as a covariate which is related to both treatment assignment and outcomes. Propensity score analysis rests on the assumption of "strong ignorability", which assumes that all variables related to both outcomes and treatment assignment are observed and are included, and that treatment assignment is unconfounded with potential outcomes conditional on observed covariates. That is, the model assumes that these individuals are sorted into different treatments as if they are randomly assigned. Foster et al (2003) said that "this assumption is strong; its plausibility depends on the particular treatment involved and on the range of covariates included in the analysis".

Once calculated, propensity scores have been applied in several ways to help reduce bias: matching, stratification and regression adjustment.

1) Propensity score matching

The propensity score matching method orders the treatment and control groups by propensity scores and then matches each patient who receives the treatment to a control patient with a similar propensity score. Matching aims to create groups that are similar in terms of observable background characteristics and this reduces the observed selection bias.

Matching on propensity scores is a way of matching on many variables indirectly, instead of matching directly on many variables, which becomes increasingly difficult with more variables. The most common method of matching is nearest neighbour matching. That is, patients in the treatment group are matched to patients in the control group who have the closest propensity score. This can be done on a one-to-one basis or a many-to-one or one-to-many basis [Dowie et al, 2003]. "Nearest neighbour" matching can be done with or without replacement. Dehejia and Wahba (2002) commented that matching with replacement allows a control group patient to be selected more than once, and this could be beneficial in terms of bias reduction and the average quality of the matching will increase. Matching without replacement is where a control group patient can only be selected once, and this could also improve the precision of the estimates as you ensure the smallest propensity-score distance between the treated and control units. Another method of matching is known as "calliper" or radius matching, which would select all control units within a pre-defined radius [Dehejia and Wahba, 2002]. After treated patients are matched to control patients, the next step is then to calculate costs and outcomes for each group.

For example, Johnson and colleagues (2005) evaluated an asthma care support program in the USA. The program aimed to reduce unnecessary hospital and emergency department visits through patient education and the use of clinical protocols. The authors used a logistic regression model to predict each person's propensity for program enrolment. Variables used in the logistic regression included demographic factors, comorbid conditions, medical service utilisation, prescription drug use, and clinical procedures. For each program participant, a non-participant was chosen with the closest (nearest-neighbour) propensity score.

2) Propensity score stratification

For the propensity score stratification method, once propensity scores are obtained for each patient, this method consists of grouping subjects into strata, so each stratum contains patients from both groups, determined by observed background covariates. This method aims to reduce selection bias, so that any differences between the groups or strata are not due to background characteristics. So patients with similar propensity scores form one group, and another group is formed from patients with similar propensity scores and so forth, with approximately the same total number of patients in each stratum (although this may not always be the case). The investigator must decide on the cut-off points for the different strata (i.e. the strata are usually divided into equal propensity score ranges). Once the strata are defined, treated and control patients who are in the same group or stratum are compared directly [D'Agostino, 1998]; for each stratum, the mean propensity scores for each group are compared.

Coyte and colleagues (2000) estimated the impact of alternative discharge strategies following joint replacement surgery on the total cost of care. There were four discharge destinations: a rehabilitation hospital with subsequent discharge to home without home care; a rehabilitation hospital with subsequent discharge to home with home care; discharge to home with home care; and discharge to home with self-care. The authors used a logistic regression to calculate the patient's propensity to be discharged to a given destination. The covariates used in the logistic regression included: age, gender, comorbid conditions, case mix group, diagnosis, length of stay, urban/rural characteristic of patient residence, whether a joint replacement revision was performed, teaching status of hospital and year. They subclassified patients on the basis of their propensity score to produce five strata in which both groups were balanced with respect to all of the observed covariates. Then treatment costs and patient outcomes due to each discharge strategy were calculated within each stratum.

3) Propensity score regression adjustment

For the propensity score regression adjustment method, once propensity scores are obtained for each patient, these are then added as an independent variable into the regression model. Inclusion of propensity scores as a covariate in the regression model takes into account the likelihood for treatments, thereby reducing selection bias. That is, by adding the propensity scores into the regression equation, the component of correlation which is due to the assignment process can be eliminated. The regression adjustment refers to a statistical procedure that adjusts estimates of the treatment effects by estimating the relationship between the dependent variable and the independent variables in each treatment group [Rubin, 1979].

Mitra and Indurkhya (2005) used propensity score analysis to estimate the cost-effectiveness of cystectomy versus no cystectomy in elderly patients with muscle invasive bladder cancer. From the observational data the authors used a logistic regression model to estimate the propensity of patients to receive surgical treatment of their bladder cancer based on their background covariates which included age, sex, race, cancer stage and grade, income, marital status and comorbidity index and score. After adjustment, none of the background covariates were statistically significant ($p < 0.05$). The propensity score values were then used in a generalised linear model to estimate net monetary benefit adjusting for propensity score.

Application of propensity scores

Propensity scores can be calculated in Stata version 10 [StataCorp, 2007] using both the 'pscore' and the 'psmatch2' programmes.

'Pscore' programme

The 'pscore' programme stratifies women into blocks according to their propensity score and the model identifies the optimal number of blocks to ensure that the mean propensity score is not different for treated and controls in each block. The programme then checks that each covariate is balanced. Each variable is tested separately for each block using a two-sample t-test with equal variances [Becker and Ichino, 2002]. The estimated propensity scores can be used to obtain estimates of the average treatment effect on the treated (ATT) using various matching methods (see below). The ATT computes the difference in average outcomes between the two groups. The pscore programme creates two new variables: 1) mypscore – individual propensity scores for each patient; and 2) myblock – block number of the estimated propensity score. The various matching methods include:

a) The *nearest neighbour matching*, first sorts all records by the estimated propensity scores and then takes each telemedicine patient, searches forward and backward for the closest direct referral patient(s). This method is usually applied with replacement, in a sense that the direct referral patient can be a match for more than one telemedicine patient and this method selects the smallest possible control group. Once each telemedicine patient is matched with a direct referral patient, the difference between the outcome of the telemedicine patients and the outcome of the matched direct referral patients is computed i.e. the ATT. If for a telemedicine patient forward and backward matches happen to be equally good, the *nearest neighbour random draw version* program randomly draws either the forward or backward matches [Becker and Ichino, 2002]; whereas, the *nearest neighbour equal weights version* program gives equal weight to the groups of forward and backward matches [Becker and Ichino, 2002].

b) With *radius matching*, in order to compute the ATT, each telemedicine patient is matched only with the direct referral patients whose propensity score falls in a predefined radius [Becker and Ichino, 2002]. Because, the radius method can generate more controls than the nearest neighbour method, this can worsen the quality of the match on the propensity score and there is also greater variability in the sample. If the radius is set quite small, it is possible that some telemedicine patients are not matched because the radius does not contain direct referral patients. On the other hand, the smaller the size of the radius, the better the quality of the matches.

c) With *kernel matching* (most commonly used), in order to compute the ATT, all telemedicine patients are matched with a weighted average of all direct referral patients with weights that are inversely proportional to the distance between the propensity scores of telemedicine and direct referrals (i.e. downweighting "distant" observations) [Becker and Ichino, 2002]. Each match can be thought of as being weighted in proportion to the number of standard deviations away from the treated individual. Closer the match in terms of propensity scores, the greater the weight placed on the match.

d) With *stratification matching*, this method consists of dividing the range of variation of the propensity score in intervals, so that within each interval telemedicine and direct referral patients have on average the same propensity score [Becker and Ichino, 2002]. The same blocks identified by the model that estimates propensity score can be used. Difference in outcome measures between the telemedicine and direct referrals in each interval is computed i.e. the ATT.

'Psmatch2' programme

The 'psmatch2' programme [Leuven and Sianesi, 2003] uses both the nearest neighbour and kernel matching approaches. However, for the nearest neighbour matching, this programme goes one step further and directly applies matching on a one-to-one basis and creates two new variables: 1) _id – a new identifier is created for all observations; and 2) _n1 – for every treatment observation it will store the observation number (i.e. the patient id) of the matched direct observation.

5.2.4.3 Sample selection models

Sample selection models controls for the bias due to both observed and unobserved factors associated with outcomes and are commonly referred to as the Heckman method [Heckman, 1979; Crown et al, 1998]. In the first stage, a probit model (for the Heckman model [Heckman, 1979; Crown et al, 1998]) or a logit model (for the Lee model [Lee, 1983]) of treatment selection is estimated. Probit models are very similar to logit models, but assume that the error term is normally distributed.

In the first stage, for example, a probit model of treatment selection is estimated (the probability of 'selecting' into one of the treatment groups of interest). Then the estimated probabilities from this probit model, are used to calculate a new variable, known as an adjustment factor (or λ or inverse Mills ratio - this serves as an adjustment for potential selection bias) for each patient, which is the probability of not receiving the treatment given that the individual was 'at risk' of receiving the treatment. As the probability of receiving the treatment approaches 1, the adjustment factor approaches 0. In other words, for patients who actually receive the treatment, if the predicted probability of receiving the treatment is high based upon observable factors, the influence of unobservable variables is small and consequently the bias is small. If the probability of receiving the treatment approaches 0, the adjustment factor approaches infinity. As a result the potential for bias is large for patients who receive the treatment, but who are predicted not to receive the treatment based upon observable factors [Crown et al, 1998].

In the second stage, we want to predict the outcome of interest such as length of stay. So in the second stage regression equation (such as an ordinary least squares), this adjustment factor is included as one of the independent variables in the outcome model. The significance and sign associated with the adjustment factor coefficient in the second model indicates the statistical significance and the magnitude of selection bias that would have been present in the estimates, if the adjustment factor had not

been included in the equation. The adjustment factor permits a direct test of whether selection bias is present and if so, what the direction of its impact is. If the coefficient on the adjustment factor, in the outcome equation is statistically significant, this indicates that selection bias is present and that the results of the treatment effect would have been biased had the adjustment not been made [Crown, 2001]. The sign on the adjustment factor also indicates the direction in which the results would have been biased. If no selection bias is present, the sample selection model will produce the same regression coefficients as ordinary least squares estimation [Vogel et al, 2002].

The way sample selection models accounts for unobserved variables (bias) is similar to the instrumental variables method, where an instrumental variable is required (see Appendix 6, section A6.2.4 for more details). So for example, in the regression model there is a variable for whether the mother has a family history of CHD but not for whether the mother has a previous pregnancy with an anomaly. So in this case, the family history of CHD variable (identifier) is used as a proxy for previous pregnancy. That is, the family history of CHD variable would be the instrumental variable in this case. This variable would have to be significant in the first part of the model, but not in the second part of the model. Most articles claim that sample selection models deal with unobserved variables; however, they are not explicit in saying how they do this.

The main problem with sample selection models is multicollinearity. If the observed variables used to model the probability of receiving treatment are the same as those used to model the outcome variable in the second equation, the adjustment factor will be highly correlated with the observed variables in the second equation, leading to severe multicollinearity [Crown et al, 1998]. On the other hand, if it is possible to identify different sets of observable variables in the two stages (even if there is overlap between them), sample selection models tend to be more effective in controlling for selection bias [Crown et al, 1998]. Another drawback of sample selection models is that they require strong assumptions. Firstly, the two regressions have to be simultaneously estimated i.e. jointly model the selection into treatment and outcome; and secondly, it also assumes that the error terms are jointly normally distributed.

Sample selection models by themselves, may not be very useful in making conclusions about treatment effects, because they focus only on the treatment group. A variant of the sample selection model is a treatment effects model which pools the sample for the treated and non-treated outcomes, and therefore the treatment effect variable in this equation can be included. The adjustment factor used in this type of analysis includes values for both the treatment and non-treatment groups [Crown et al, 1998].

Crown et al (1998) used sample selection models to control for selection bias and to estimate the effects of alternative antidepressant therapies on a variety of cost measures such as total charges, physician charges, antidepressant and non-antidepressant charges. In the first stage, a probit regression model was conducted which showed the probability of being prescribed fluoxetine versus another antidepressant and allowed for the creation of the adjustment factor. In the second stage regression, the authors entered the adjustment factor as an additional variable in the equation which predicted expenditures for physician visits using the observations for fluoxetine users only. The results showed that if the authors did not control for selection bias, then there were unobserved variables that differed between the fluoxetine and non-fluoxetine users that were correlated with expenditures and as a result estimates in the expenditure equation would have been biased.

5.2.4.4 Summary of the methods identified to control for selection bias

In summarising the three methods which were identified to reduce selection bias, the regression analysis method and the propensity scoring methods are very similar; however the sample selection method is not; see the few examples below (these examples have not been selected systematically from the literature). Shah and colleagues (2005) conducted a systematic review to determine whether propensity scores gave different results from traditional regression modelling when adjusting for bias in observational studies. They found that both methods produced similar results in terms of the strength or statistical significance of association between exposures and outcomes, although propensity scores gave slightly weaker associations. However, they found that many of the reviewed studies did not implement propensity scores well. Vogel and colleagues (2002) looked at stroke patients' gain in overall, motor and cognitive functional status during rehabilitation which was measured using the functional independent measure. They used both sample selection models and regression analyses to correct for selection bias. They found statistically significant evidence of selection bias, and there was considerable differences in the results obtained from both the standard multiple regressions and the sample selection models. Dusheiko et al (2004) examined the effect of practice's fundholding status on the waiting times of its patients using methods to correct for selection bias including regression analysis, propensity score matching and the Heckman selection model. The authors found that the result from the sample selection method was significantly different from regression and propensity score results.

Table 5.1 highlights the key requirements of each method; strengths and limitations of each of the methods; and whether any of the methods are similar or different to other methods identified to control for selection bias.

Table 5.1: Summary of methods identified to control for selection bias

Method	Main idea behind method – 'concept'	Key requirements	Why is this method different from the others?	Is this method similar to other methods?
1) Regression analysis	Estimates how each independent variable relates to the dependent variable and controls for the observable differences between treated and non-treated subjects.	- Independent variables included in regression should not be highly correlated - Assumes linearity	- Regression model uses all observations/variables - The regression predictions for patients who receive the treatment shift either up or down by the amount of the regression coefficient for the treatment variable and the method assumes that the slope coefficients are identical for those who do and do not receive treatment.	Similar to propensity score regression adjustment if all the same variables are used; doesn't include the propensity score in regression equation. However, results obtained should be similar
2) Propensity score	- Easy to estimate using a probit or logit model of treatment selection to obtain propensity scores - Summarises all background characteristics into a single-index variable (the propensity score). - Balances the covariates in the two groups i.e. enough overlap in background variables means that the groups are comparable and so selection bias is removed when comparisons are made between groups with the same propensity scores.	- Groups must overlap enough before matching, stratification or regression adjustment takes place - Rests on the assumption of "strong ignorability" (this means that no systematic, unobserved, pre-treatment differences exist between the two subjects) - Used with binary outcomes - Probability of treatment is correlated with outcome term	- A single index variable is created - To adjust for selection bias, covariates that independently affect the chance of being treated are identified and either matching, stratification or regression adjustment using this score will, on average, remove all of the bias from the background covariates and in turn will obtain less biased treatment effect estimates.	Similar to sample selection models first stage when calculating treatment selection probabilities
2a) Propensity score matching	Once propensity scores are obtained, then subjects can be matched based on their propensity score	Tries to match each treated patient to a control patient with a similar propensity score	There are various matching schemes i.e. nearest neighbour can be done with or without replacement	Not the same as the regression method, because propensity score matching may not use all patients
2b) Propensity score stratification	Once propensity scores are obtained, strata or groups are formed, so subjects in each stratum can be compared directly	Rosenbaum and Rubin (1984) suggest that 5 strata (or groups) are sufficient to remove 90% of the bias.	- A single-index variable makes it easier to compare and create strata - Method can be used also for missing data imputation	
2c) Propensity score regression adjustment	Once propensity scores are obtained they are added as an independent variable in the regression equation		It's only the inclusion on the propensity score variable as an extra independent variable, which makes it different to the multivariate regression method	Similar to regression analysis if all the same variables are used, and results should be similar
3) Sample selection models	In the first stage, a treatment selection model is estimated and the estimated probabilities are used to construct an adjustment factor. In the second stage, this factor is included as one of the variables in the outcome model.	- If variables are different in the two stages, then these models will be more effective in controlling for selection bias - Need some valid exclusion restrictions (i.e. variables that predict treatment but not outcome)	- The inclusion of adjustment factor in the second stage permits a direct test of whether selection bias is present and if so, what the direction of its impact is - The adjustment factor captures the effects of unobserved variables on the outcome variable, which enables estimation of unbiased estimates for the parameters associated with the observed variables.	Similar to propensity score methods in calculating the treatment selection probabilities in the first stage

Method	What are the main weaknesses?	What are the main strengths?	Tests to check whether bias is reduced
1) Regression analysis	- Can't control for unobserved variables, only controls for observed variables - Can't be sure that the covariates among the two groups are balanced (i.e. variables may not overlap enough) - The regression may not be as robust if variables are omitted (omitted variable bias) - If there is multicollinearity between independent variables, this may increase the standard error of the estimates	- Easy to understand and implement - Can use more variables than matching or stratification methods - Can be used to predict values - Estimates a 'mean effect' for each independent variable - Doesn't require any strong assumptions - Has more power than sample selection models	- RESET test can be conducted to check that the model is not mis-specified (or suffer from omitted variables bias) - Hosmer-Lemeshow test can be done to check how well the model fits the data for logistic regressions
2) Propensity score	- Can't control for unobserved variables, only controls for observed variables. - Handling of weak covariates - As the covariates are combined it may possibly obscure important interactions, as it only estimates the overall treatment effect (can't see the effect of each independent variable on the treatment - such as in a regression analysis). - May not use all the observations for the analysis	- Easy to implement, are efficient and more robust - Doesn't assume any functional form e.g. a linear relationship between each outcome and covariate within each group - Can control for observed variables better than some methods - Requires few assumptions - Has lower standard errors than regression models - Easy to understand and explain using histograms than sample selection models	No specific tests
2a) Propensity score matching	- Matching on a one-to-many or many-to-one basis may reduce sample size - Depending on the criteria for matching some subjects may not be matched (matching range may be too narrow) and can lead to a smaller sample size - A smaller sample size, therefore there is less power, and the model may not be as robust	- Makes it easier to match subjects on one variable (a score), than matching across various variables - Attempts to use all subjects in the matching process - Can calculate separate propensity scores for each pair of treatments	No specific tests
2b) Propensity score stratification	Needs a large sample size to create five strata	Can calculate separate propensity scores for each pair of treatments	No specific tests
2c) Propensity score regression adjustment	If there is multicollinearity between independent variables, this may increase the standard error of the estimates.	Estimates a 'mean effect' for each independent variable	No specific tests
3) Sample selection models	- Only controls for the unobserved variables in the treatment process - Problems with multicollinearity if observed variables in two stages are the same - Requires strong assumptions to estimate the model - May not be as robust as propensity scores	- Controls for both observed and unobserved variables in treatment process - Doesn't assume any functional form e.g. a linear relationship between each outcome and covariate within each group - Sample selection models explicitly address bias caused by the regressor with omitted variables, by adding a term into the regression	Adjustment factor which is created, can check whether selection bias is present

5.3 Methods for data analysis

The previous section concluded that regression analysis, propensity score analysis and the Heckman method (sample selection model) were all appropriate methods to reduce selection bias. This section will attempt to apply each of the three methods to the dataset which was used in Chapter 2 and to see what impact these methods have on the estimation of costs and effects. The cost variable used in this analysis is the total costs of pregnancy and the effect variable used in this analysis is the detection of cardiac cases before birth (i.e. true positives), which has been coded 1 for a cardiac anomaly and 0 for no anomaly. The first step was to identify variables such as demographic characteristics and risk factors from the dataset to use to 'adjust' costs and effects for selection bias and to hypothesise which of these variables will have an influence on outcomes. In Chapter 2, section 2.4.2 the TelePaed' eligibility criteria was set out.

1) Binary variables

- Method of referral to specialist (telemedicine coded 1 and direct referral coded 0):
- Parity (first or subsequent pregnancy) – primaparous mothers may receive extra antenatal care, leading to higher antenatal costs. They may also have higher obstetric costs because they are more likely to have a hospital birth than a home birth to minimise the risk of complications arising during delivery and they may also spend longer periods on postnatal wards after delivery than multiparous mothers [Mistry et al, 2007];
- Type of pregnancy – mothers expecting multiple births have higher antenatal and obstetric costs than mothers of singletons overall, because of the increased risk of both maternal complications and fetal difficulties (such as twin-to-twin transfer syndrome) [Mistry et al, 2007];
- Pre-gestational diabetes - as the disease may precipitate maternal complications, [CEMACH, 2005] women with diabetes require extra antenatal care resulting in higher antenatal and obstetric costs;
- Down's syndrome risk and/or an elevated serum risk – Down's syndrome may also bring about further maternal complications [Wald et al, 2003] and if the woman has an elevated serum level above the cut-off level they may require extra care resulting in higher antenatal and obstetric costs;
- Epilepsy - this disease may also precipitate maternal complications [Morrow et al, 2006], so women on anti-epileptic therapy may require additional care resulting in higher antenatal and obstetric costs;

- Family history of CHD - maternal CHD may also lead to complications, so these women may require extra antenatal and obstetric care, leading to higher antenatal and obstetric costs [Uebing et al, 2006]; and
- Previous pregnancy with an anomaly – women who have had previous pregnancy with an anomaly may be given extra monitoring (i.e. antenatal care), leading to higher antenatal costs.

2) Continuous variables

- Mother's age – as mother's age increases, obstetric costs increase since age is a confounder with respect to preterm delivery, induction of labour and caesarean section [Mistry et al, 2007];
- Gestation (in weeks) at time of anomaly scan - if women have an anomaly scan before 19 weeks, they may have additional antenatal check-ups before delivery, leading to higher antenatal costs;
- Gestation (in weeks) at birth – if a mother is delivered prematurely because of complications, she may require additional obstetric care, which results in higher obstetric costs; and
- Differential distance - defined as the distance in miles from patients' home to specialist hospital minus the distance in miles from patients' home to nearest hospital. This variable shouldn't have an impact on health service costs; thus this variable is linked to treatment in the sense where does a patient receive specialist care?

The demographic characteristics and risk factors for the two groups of women have previously been presented in Chapter 2 (see Table 2.3). All variables were checked for normality (i.e. whether they were normally distributed) before being used in the analysis.

All variables listed above were included in the analyses conducted in this chapter apart from type of pregnancy. Type of pregnancy was excluded as only one mother was expecting a multiple birth and inclusion of this variable in a logistic regression model would have removed this patient from the analysis. This is because there is no relationship with the dependent variable as only one mother was expecting a multiple birth. Technically this woman should have been removed from the analysis, however as the sample size was small for the direct referral group (n = 24), this woman has been included in the analyses.

Gestation in weeks at birth as a continuous variable included six patients' who had a termination of pregnancy (coded 0 for these patients) and including this variable in the models will distort the results. Therefore, to take into account termination of pregnancy, this variable was recoded into a categorical variable: 0 = termination, 1 = pre-term birth (babies born before 37 weeks of pregnancy) and 2 = full-term birth (babies born between 37 and 42 weeks of pregnancy) [Steer, 2005] and has been renamed as pregnancy duration.

Four variables were combined, as only a few pregnant women had certain risk factors. Including these risk factors in as separate variables in each of the models would mean that there is a high chance of these patients who have the disease in question being removed from the analysis due to the low variation (i.e. there is no or very little relationship with the dependent variable and Stata would remove them from the model). Firstly, the variable whether the mother has an elevated serum risk was combined with the variable does the unborn baby have a high risk of Down's syndrome (the variable has been renamed as Down's). An elevated serum test is a blood test which assesses the women's risk (low or high) of having a baby born with Down's syndrome. Even if a high risk is detected, it is not always certain that the baby will have Down's syndrome, therefore further tests are needed to confirm the diagnosis [Wald et al, 2003]. Secondly, the variable representing a family history of CHD was combined with whether the patient had a previous pregnancy with an anomaly (the variable has been renamed as family history). For their previous pregnancy, the pregnant woman's fetus or baby would have had a chromosomal abnormality. This abnormality could either have been a fetus with Down's syndrome or a CHD defect which was terminated or they had a child which was born with Down's syndrome or a CHD defect [Carvalho et al, 2002].

The comparative analyses between the two groups are presented. For the regression analyses, firstly, as total costs of pregnancy were skewed, a generalised linear regression model was conducted; and secondly, as cardiac abnormalities was a binary variable a logistic regression was conducted. All statistical analyses were conducted in Stata version 10 [StataCorp, 2007] and a p-value ≤ 0.05 was considered to be statistically significant for the comparative analyses.

5.4 Results from methods applied to reduce selection bias in this case study

5.4.1 Observed results

During the period May 2001 to July 2002, a total of 76 pregnant women were referred for specialist opinion following a routine anomaly scan: 52 (68.4%) were assessed via

the telemedicine link, and 24 women saw a specialist in London. As seen in Chapter 2, the overall total costs of pregnancy for the telemedicine group were higher than direct referral group, a difference of £539 which was not significant (p = 0.202) and the direct referral group were more likely to have had a cardiac abnormality (p = 0.043).

5.4.2 Regression analyses

Table 5.2 below shows the generalised linear regression model results to examine the relationship between costs and referral mode and the logistic regression model results to examine the relationship between effects and referral mode, controlling for all other variables: age, gestation at time of anomaly scan, pregnancy duration (termination was set as the base case, pre-term or full-term birth), parity, diabetes, Down's, epilepsy and family history, in order to see whether referral mode is a significant predictor of costs and/or effects. Pregnancy duration and diabetes were significant predictors of the total costs of pregnancy. Women assessed by telemedicine had higher costs than direct referral women (an extra £84), this is in the same direction as the observed cost results, but of a much smaller magnitude and is not significant. For the cardiac anomalies which were detected before birth, the fitted model explained 39% of the variation and only pregnancy duration was a significant predictor. Women who were assessed by telemedicine had a lower underlying prevalence rate of cardiac anomalies than direct referral women (0.6000), this is in the same direction as the observed effect results. This means that the women assessed by telemedicine were 60.0% less likely to have a cardiac anomaly (detected) compared to the direct referral group.

Table 5.2: Results from regression analyses taking into account demographic characteristics and risk factors

	Coefficient	Standard error	z statistic	p-value
Generalised linear model: Total costs of pregnancy				
Referral mode	83.55	373.10	0.22	0.823
Mother's age	15.95	26.31	0.61	0.544
Gestation	-33.22	89.87	-0.37	0.712
Pre-term	3708.41	706.46	5.25	< 0.001
Full-term	3328.01	656.54	5.07	< 0.001
Parity	-364.76	357.67	-1.02	0.308
Diabetes	1220.61	512.50	2.38	0.017
Downs	-483.98	609.60	-0.79	0.427
Epilepsy	-382.14	535.25	-0.71	0.475
Family history	-567.14	441.24	-1.29	0.199
Constant	1559.98	2205.68	0.71	0.479
	Odds ratio	**Standard error**	**z statistic**	**p-value**
Logistic regression: Cardiac anomalies Pseudo R^2 = 0.393, Likelihood ratio χ^2_{10} = 26.02, p = 0.004				
Referral mode	0.6000	0.6156	-0.50	0.619
Mother's age	1.0533	0.0789	0.69	0.489
Gestation	1.2668	0.2324	1.29	0.197
Pre-term	0.0232	0.0392	-2.23	0.026
Full-term	0.0320	0.0487	-2.26	0.024
Parity	0.2084	0.2413	-1.35	0.176
Diabetes	0.2714	0.3781	-0.94	0.349
Downs	0.0996	0.1652	-1.39	0.164
Epilepsy	0.2621	0.3685	-0.95	0.341
Family history	0.1172	0.1538	-1.63	0.102

Both regression models were checked to see whether they were well specified. For the generalised linear model, the p value from the link test was not significant (p = 0.743), indicating that the model was well specified. After the logistic regression model, the Hosmer-Lemeshow test was conducted to consider the classification power of the logistic regression, by regrouping data according to predicted probabilities and then creating equal size groups [Hosmer and Lemeshow, 1989]. The Hosmer-Lemeshow test was not significant (χ^2_8 = 6.68, p = 0.571), indicating a good fit for the logistic model.

5.4.3 Propensity score analyses

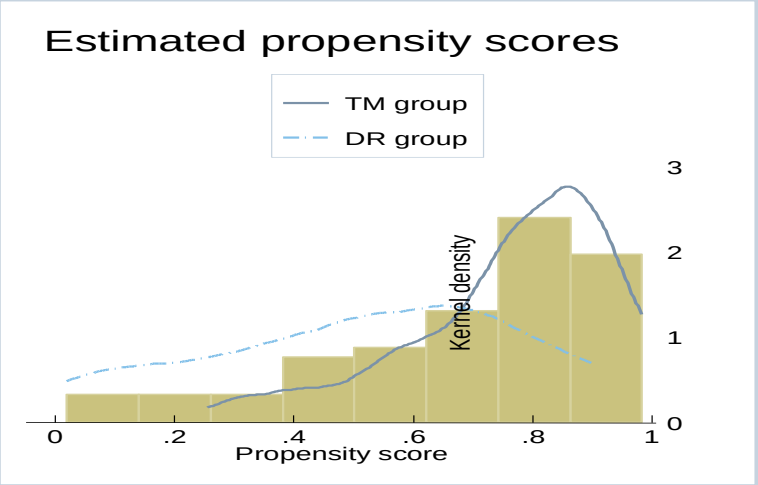
Propensity scores were estimated using a logit model, where the dependent variable was the method of referral to specialist (telemedicine = 1 and direct referral = 0). The independent variables used to estimate the propensity scores were: age, gestation at time of anomaly scan, pregnancy duration, parity, diabetes, Down's, epilepsy and family history. The results from the model (see Table 5.3) showed that only gestation in weeks at time of the anomaly scan was a significant predictor in the calculation of propensity scores ($p \leq 0.05$).

Table 5.3: Propensity score logistic regression model

	Odds ratio	Standard error	z statistic	p-value
Pseudo R^2 = 0.223, Likelihood ratio χ^2_9 = 21.13, p = 0.012				
Mother's age	0.9505	0.0484	-1.00	0.318
Gestation	0.5899	0.1555	-2.00	0.045
Pre-term birth	4.2297	5.4104	1.13	0.260
Full-term birth	11.5210	14.6854	1.92	0.055
Parity	2.2990	1.5599	1.23	0.220
Diabetes	1.0032	1.0331	0.00	0.998
Downs	0.1317	0.1601	-1.67	0.095
Epilepsy	3.6454	4.8615	0.97	0.332
Family history	0.5296	0.4803	-0.70	0.483

The Hosmer-Lemeshow test was not significant (χ^2_8 = 12.92, p = 0.115), indicating a good fit for the logistic model. Figure 5.1 below shows that there was some overlap between the estimated propensity scores for both groups.

Figure 5.1: Checking whether the estimated propensity scores overlap



Once the propensity scores were obtained, then matching was applied in three different ways. The first two methods of matching were available in Stata version 10: pscore and psmatch2, which used various matching methods to obtain estimates of the ATT. The ATT estimates the difference in average costs or effects between the two groups. For the final method, the estimated propensity scores were matched by 'hand'.

5.4.3.1 'pscore' matching

The 'pscore' program identified that five blocks were needed to ensure that the mean propensity score is not different for telemedicine and direct referrals and also the balancing property was satisfied. The estimated propensity scores were used to obtain estimates of the ATT using various methods.

Table 5.4: Average Treatment effect on Treated (ATT) results for the different matching methods using 'pscore'

Method of matching	Number of matched cases[5]		ATT	95% confidence interval using bias corrected method
	Telemedicine group	Direct referral group		
Total costs of pregnancy				
Nearest neighbour (random draw)	52	14	156.13	-1237.74 to 1097.08
Nearest neighbour (equal weights)	52	14	156.13	-1604.31 to 1488.18
Kernel	52	19	295.18	-550.31 to 1227.41
Stratification	52	19	182.83	-927.44 to 1058.23
Radius (0.10)	52	19	264.98	-469.00 to 1864.78
Cardiac anomalies				
Nearest neighbour (random draw)	52	14	-0.212	-0.8250 to 0.1087
Nearest neighbour (equal weights)	52	14	-0.212	-0.7447 to 0.0536
Kernel	52	19	-0.202	-0.6303 to 0.0488
Stratification	52	19	-0.162	-0.5918 to 0.1111
Radius (0.10)	52	19	-0.173	-0.5272 to 0.0449

In Table 5.4 above, using nearest neighbour matching (either random draw or equal weights version) 52 telemedicine patients have been matched to 14 direct referral patients. Both the nearest neighbour matching gave the same ATT results, but the 95% confidence intervals are different. For the other three methods, 52 telemedicine patients have been matched to 19 direct referral patients. The ATT suggests that by using either of the nearest neighbour matching methods, those patients in the telemedicine group had higher costs (£156 higher) compared to patients in the direct referral group. With the other three methods the costs were again higher for the telemedicine group, ranging from £183 (stratification matching) to £295 (kernel matching). For all five matching methods, the telemedicine group was less likely to have cardiac anomalies than the direct referral group. For example, with the stratification method, telemedicine group was 16.2% less likely to have a cardiac anomaly compared to the direct referral group.

5.4.3.2 'psmatch2' matching

The 'psmatch2' programme using the nearest neighbour approach directly applies matching on a one-to-one basis. In total, 24 telemedicine patients were matched to all 24 direct referral patients. Twenty-eight (36.8%) patients were not matched using this method. The 'pmatch2' programme also calculates the ATT and results are shown below in Table 5.5.

[5] The pscore programme identifies the number of matched cases in each group, however, it does not provide a list of which telemedicine unit(s) is matched to which direct referral unit(s).

Table 5.5: Average Treatment effect on Treated (ATT) results for the different matching methods using 'psmatch2'

Method of matching	Coefficients		ATT	95% confidence interval using bias corrected method
	Telemedicine group	Direct referral group		
Total costs of pregnancy				
Nearest neighbour	3961.76	3824.78	136.99	-717.85 to 917.25
Kernel	4363.31	4048.41	314.90	-555.69 to 1245.95
Cardiac anomalies				
Nearest neighbour	0.1250	0.2917	-0.1667	-0.6429 to -0.0500
Kernel	0.0962	0.3055	-0.2094	-0.6512 to 0.0606

The results from the psmatch2 method found that using the nearest neighbour matching, the telemedicine group had higher costs than the direct referral group (£137 higher), however, these results are based only on the 24 matched cases. However, taking all cases in account, for the kernel matching the telemedicine group still had higher costs than the direct referral group (£315 higher), which is very similar to the costs from the pscore kernel matching programme (£295). With both matching methods, the telemedicine group was less likely to have cardiac anomalies compared to the direct referral group. Likewise, the psmatch2 ATT value for kernel matching for cardiac anomalies was similar to the pscore ATT value for kernel matching.

The psmatch2 programme also provides an indication of whether the propensity score method is a good estimator and provides a comparison between the unmatched and matched samples. The propensity score is a good estimator if after matching, the mean average standardised bias and Pseudo R^2[26] are lower than before matching. For the nearest neighbour matching the mean average standardised bias has fallen from 40.94 to 25.73 and the Pseudo R^2 has fallen from 0.224 to 0.085, and likewise for the kernel matching the mean average standardised bias has fallen from 40.94 to 15.11 and the Pseudo R^2 has fallen from 0.224 to 0.067, implying that propensity score matching is a good estimator.

5.4.3.3 Propensity score matching by 'hand'

Using both the 'pscore' and 'psmatch2' matching methods, not all patients were matched. So the task of matching telemedicine cases to direct referral cases was undertaken (by hand), using the nearest neighbour approach without replacement [Dehejia and Wahba, 2002] for the estimated propensity scores. For this approach, all patients are first sorted by their estimated propensity score, and then matching for telemedicine patients by searching forward and backward for the direct patient(s) with

[6] R^2 values range from 0 to 1, which makes sense because it is a proportion and a squared correlation. The higher this value this indicates a better fitting model. However, Pseudo R^2 cannot be interpreted the same as an R^2 because most Pseudo R^2 do not range from 0 to 1. Furthermore, Pseudo R^2 cannot be interpreted independently, it only has meaning when compared with another Pseudo R^2 in the same dataset.

the closest score. Matching patients using this approach was based on a one-to-many, or a many-to-one basis [Dowie et al, 2003]. This approach meant that a telemedicine patient could be matched to more than one direct patient with a similar propensity score or a direct patient could be matched to more than one telemedicine patient. If no exact match was found for a patient (i.e. a propensity score to 5 decimal places), then matching was based within 4 decimal places, then within 3 decimal places, then within 2 decimal places and finally within 1 decimal place, similar to an approach used by Gum and colleagues [Gum et al, 2001].

All 52 telemedicine patients were matched to 24 patients who were seen by direct referral. After weightings were applied, the number of cases in each group was 13. In order to make a direct comparison between the telemedicine case(s) and the direct case(s), for example, for the total costs of pregnancy for the patients forming the cases, they were adjusted in accordance with their 'weights'. So if two telemedicine patients formed a case, then the cost for each patient was multiplied by the weight (0.5) and summed together to obtain a final cost per case. In terms of effects, the same method was applied, but if the number was below 0.5 this was coded as 0 (no cardiac anomaly) and if it was equal to or above 0.5 this was coded as 1 (cardiac anomaly). Results from the propensity score matching are shown below in Table 5.6.

Table 5.6: Results from propensity score matching[±]

	Telemedicine group	Direct referral group
Total costs of pregnancy (Bootstrapped)		
Number of cases	13	13
Mean (SD)	£4,016 (£371)	£4,072 (£258)
95% CI	£3,290 to £4,743	£3,567 to £4,577
Confirmation of abnormality		
Number of cases	2	4
Mean	0.1538	0.3077

[±] Statistical tests conducted: paired t tests.

After propensity score matching, the telemedicine group had lower mean costs than the direct referral group and the difference in mean costs for the two groups was reduced to £56 (observed incremental cost was £539), and again this cost difference was not statistically significant (p = 0.907). The telemedicine group was less likely to have cardiac anomalies compared to direct referral group. The mean difference in effectiveness between the two groups was also smaller after propensity score matching (-0.1955 to -0.1538) and this difference was not statistically significant (p = 0.370).

Table 5.7 below presents the demographic characteristics and risk factors for the pregnant women after applying propensity score matching. Thus, after applying propensity score matching, the variable gestation in weeks at time of anomaly scan was no longer statistically significant, indicating that the two groups were 'balanced'.

Table 5.7: Demographic characteristics and risk factors after propensity score matching[±]

	Telemedicine group (n = 13)	Direct referral group (n = 13)
Parity		
Primaparous	6 (46.2%)	6 (46.2%)
Multiparous	7 (53.8%)	7 (53.8%)
Does the mother have diabetes?		
Yes	2 (15.4%)	3 (23.1%)
No	11 (84.6%)	10 (76.9%)
Does the unborn baby have a high risk of Down's syndrome or does the mother have an elevated risk serum?		
Yes	1 (7.7%)	2 (15.4%)
No	12 (92.3%)	11 (84.6%)
Is the mother on any anti-epilepsy drug therapy?		
Yes	1 (7.7%)	1 (7.7%)
No	12 (92.3%)	12 (92.3%)
Does the mother have a family history of CHD or had a previous pregnancy with an anomaly?		
Yes	7 (53.8%)	6 (46.2%)
No	6 (46.2%)	7 (53.8%)
Maternal age		
Mean (SD)	28.2 (6.2)	27.6 (5.5)
Range	17 to 43	20 to 36
Gestation (in weeks) at anomaly scan		
Mean (SD)	21 (0.6)	21 (1.2)
Range	20 to 22	20 to 25
Gestation (in weeks) at birth[#]		
Mean (SD)	35.8 (5.7)	35.3 (7.1)
Median	38	38
IQR	34 to 39	36 to 39

[#]Does not include terminations

[±] Statistical tests conducted: t tests for age and gestation at anomaly scan; chi-squared tests for parity and previous pregnancy with anomaly; Kruskal Wallis tests for gestation in weeks at birth; Fisher's Exact tests for diabetes, Down's syndrome risk, and anti-epilepsy drug therapy.

5.4.3.4 Propensity score stratification

For this part of the analysis five blocks were used for the stratification and the overall mean cost and effect results of the five blocks are shown below.

Table 5.8: Results from propensity score stratification[±]

	Telemedicine group	Direct referral group
Total costs of pregnancy (Bootstrapped)		
Number of blocks	5	5
Mean (SD)	£3,814 (£455)	£3,696 (£314)
95% CI	£2,922 to £4,705	£3,081 to £4,312
Confirmation of abnormality		
Mean	0.2000	0.2000

[±] Statistical tests conducted: paired t tests.

After propensity score stratification (see Table 5.8), the telemedicine group had higher mean costs than the direct referral group and the difference in mean costs for the two groups after propensity score stratification was reduced from £539 to £118, and this

cost difference was not statistically significant (p = 0.855). For the telemedicine group the mean cost was much lower than observed mean cost (a difference of £667). The telemedicine group was as likely to have the same detection rate of cardiac anomalies as the direct referral group.

5.4.3.5 Propensity score regression adjustment

Once the propensity scores were obtained, they were then entered as a covariate into a regression equation (see Table 5.9 below). Pre-term birth and diabetes were significant predictors of the total costs of pregnancy. Women assessed by telemedicine had higher costs than direct referral women (an extra £57). This is similar to the results from the generalised linear model (see Table 5.2). For the cardiac anomalies which were detected before birth, the fitted model explained 43% of the variation and only pre-term birth was a significant predictor. For women who were assessed by telemedicine there was a decreased likelihood of having a cardiac anomaly compared to direct referral (0.6123), this is in line with results from the logistic regression model (see Table 5.2). Furthermore, the pscore value in each model is for each individual patient, which was positive and not significant.

Table 5.9: Results from the propensity score regression analyses

	Coefficient	Standard error	z statistic	p-value
Generalised linear model: Total costs of pregnancy				
Referral mode	56.69	376.30	0.15	0.880
Mother's age	31.76	34.25	0.93	0.354
Gestation	130.78	243.59	0.54	0.591
Pre-term birth	3135.70	1061.65	2.95	0.003
Full-term birth	2452.50	1375.97	1.78	0.075
Parity	-647.54	530.17	-1.22	0.222
Diabetes	1168.89	519.31	2.25	0.024
Downs	234.58	1164.98	0.20	0.840
Epilepsy	-720.53	711.73	-1.01	0.311
Family history	-412.46	491.60	-0.84	0.401
Pscore	2080.20	2870.02	0.72	0.469
Constant	-2906.35	6547.71	-0.44	0.657
	Odds ratio	**Standard error**	**z statistic**	**p-value**
Logistic regression: Cardiac anomalies Pseudo R^2 = 0.427, Likelihood ratio $\chi^2_{11} = 28.32$, p = 0.003				
Referral mode	0.6123	0.6338	-0.47	0.636
Mother's age	1.2043	0.1672	1.34	0.181
Gestation	4.7532	5.2460	1.41	0.158
Pre-term birth	0.0003	0.0013	-2.08	0.038
Full-term birth	0.0001	0.0002	-1.78	0.075
Parity	0.0164	0.0412	-1.64	0.102
Diabetes	0.1755	0.2618	-1.17	0.243
Downs	18.1081	78.4508	0.67	0.504
Epilepsy	0.0071	0.0241	-1.45	0.146
Family history	0.5607	1.0055	-0.32	0.747
Pscore	4808925	0.0001	1.29	0.197

Including a variable which is a function of all variables may lead to double counting or the over-identifying restriction in the model. If this is the case, then a variable needs to be removed. For both models, the least significant variable in both cases was referral

mode (this was not an appropriate variable to remove from the equations), instead the next least significant variable was removed from each model. The incremental costs and effects were similar to the model in Table 5.9 i.e. the referral mode coefficient for costs was £55 and for the odds ratio for effects was 0.5737. This means that the results in Table 5.9 where all variables were used is okay to use as the model was not over-identified.

5.4.4 Results from the Heckman sample selection model

The variables used in the sample selection model included: referral mode, age, gestation at time of anomaly scan, pregnancy duration, parity, diabetes, Down's, epilepsy, family history and differential distance. This latter variable was included in the analysis as it may have an influence on the referral mode which is chosen (that is how quickly a patient would be assessed), but has no impact on costs and effects. The results from the Heckman sample selection model are shown in Table 5.10. Thus, when looking at the correlation between differential distance and method of referral, differential distance was very weakly correlated with method of referral (r = 0.066). So for this part of the analysis, differential distance is not likely to be a good identifier (variable) for the model; however, there were no other variables in the dataset which were potential identifier variables. The model only computes values for those 52 patients who were assessed by telemedicine (treated patients). The model assumes that for the 24 direct referral patients, their outcomes are censored.

The first part shows the regression equation i.e. the model predicting total costs of pregnancy. Pregnancy duration was the only significant predictor of costs. The constant value was big (£8,527) and not statistically significant. The second part shows the results of the probit analysis of the selection process i.e. the model predicting referral method. Gestation in weeks at time of the anomaly scan and full-term birth were the only significant predictors of referral mode. The final part of the output provides the selection bias statistics. Rho gives an estimate of the correlation between the error terms of the first and second equations (rho = -0.2775). Sigma is the standard error of the residuals for the total costs of pregnancy equation and lamba is rho*sigma. Finally, for the model a likelihood ratio test of independent equations is conducted and the p value for rho = 0.591. As the p value is not significant, the model indicates there is no evidence of sample selection bias (if the p value was significant, there is evidence of sample selection bias). As there is no evidence of sample selection bias, a multivariate regression model should be conducted.

Table 5.10: Results from Heckman sample selection model

	Coefficient	Standard error	z statistic	p-value
Total costs of pregnancy: Wald $\chi^2_9 = 32.77$, p < 0.001				
Mother's age	30.86	37.00	0.83	0.404
Gestation	-347.32	240.43	-1.44	0.149
Pre-term birth	3407.00	1215.05	2.80	0.005
Full-term birth	2921.97	1195.59	2.44	0.015
Parity	-809.24	499.52	-1.62	0.105
Diabetes	1139.08	708.63	1.61	0.108
Downs	-1020.61	1058.83	-0.96	0.335
Epilepsy	-299.56	693.34	-0.43	0.666
Family history	-427.06	619.57	-0.69	0.491
Constant	8527.03	5009.23	1.70	0.089
Referral mode				
Mother's age	-0.0285	0.0286	-1.00	0.319
Gestation	-0.3290	0.1574	-2.09	0.037
Pre-term birth	0.9343	0.7677	1.22	0.224
Full-term birth	1.5266	0.7462	2.05	0.041
Parity	0.5220	0.4133	1.26	0.207
Diabetes	-0.0624	0.5604	-0.11	0.911
Downs	-1.2745	0.7057	-1.81	0.071
Epilepsy	0.7207	0.7497	0.96	0.336
Family history	-0.4361	0.5328	-0.82	0.413
Distance	0.0012	0.0225	0.05	0.958
Constant	6.8326	4.0612	1.68	0.092
rho	-0.2775	0.3970		
sigma	1362.91	150.92		
lambda	-378.26	561.87		
Likelihood ratio test: $\chi^2_1 = 0.29$, p = 0.591				

A Heckman probit model was also conducted for the effect variable: cardiac anomaly or no cardiac anomaly. Firstly, the model dropped three variables: Down's, epilepsy and family history - as these three variables predicted the dependent variable perfectly (i.e. there was perfect correlation between the variables, that is, the Heckman model was identifying which group the variables go into) and leaving them in the model caused numerical instability in the estimation. So the Heckman probit model was rerun without these three variables and results from the model were not able to be estimated. Thus, the Heckman probit sample selection model was really unstable for this cohort and therefore results could not be obtained and have not been presented here.

The Heckman model for costs has been presented here for illustration purposes. Essentially to conduct this model you need some good variable(s) like an instrumental variable (i.e. variables that predict treatment, but not outcome). However, the dataset did not have any other variable(s) which were appropriate and differential distance is a rather a poor variable for this type of model.

5.4.5 Comparison of results from the various methods

Table 5.11 below shows a comparison of the results obtained from the different methods and the aim of this Chapter was to find a method which would obtain a minimum estimate of difference between the two groups and to reduce the observed

selection bias. The results from the generalised linear model analysis showed that the cost difference between the two groups was reduced compared to observed estimates (costs: a reduction to £84). The results also highlighted that only diabetes and pregnancy duration were significant predictors of costs and again pregnancy duration was a significant predictor of cardiac anomalies which were likely before birth. These results also highlight that after adjusting for regression analyses the differences between the two groups was smaller than the observed results.

Using the various propensity score matching methods in order to estimate the ATT, the cost differences between the two groups ranged from £156 to £315; whereas the effect differences between the two groups ranged from -0.1620 to -0.2120. These differences may be due to the way the matching method works and the number of cases they select for matching. After propensity score matching (by hand), the cost differences between the two groups were reduced to £55 and likewise the effect difference was reduced to -0.1538, thereby increasing the homogeneity and reducing the variance. Even though both the propensity score matching method (by hand) and the regression method used all patients, the cost and effect differences between the two methods may be due to the fact that the propensity score model was based on a weighting method (and the matching that was used) and the resulting small sample meant that there was a lot more uncertainty in the cost and effect results compared to the regression method. The results from the propensity score stratification method showed that the cost differences between the two groups was lower than the difference between the observed costs and telemedicine was actually cheaper (difference: -£117); however, there was no difference in effects between the two groups. The propensity score regression adjustment results highlighted that the cost differences between the two groups were slightly smaller than the observed difference (£57).

The Heckman selection model results showed that the mean difference in costs between direct referral and telemedicine was £747 (an increase of £208 compared to the observed difference) and the effect differences between the two groups could not be computed. All methods except the Heckman selection model applied to this dataset reduced the cost and effect differences between the two groups. This indicated these methods may have increased the homogeneity and reduced the variance in the adjusted costs and effects; that is, these methods may have reduced the observed selection bias between the two groups for this dataset.

Table 5.11: Comparison of results from the different methods

Results	Difference in total costs of pregnancy	Difference in the detection of cardiac anomalies
Observed	£539	-0.1955**
Regression	£84	-0.5108**
PS matching ('pscore') – ATT (difference)		
• Nearest neighbour (RD)	£156	-0.2120
• Nearest neighbour (EW)	£156	-0.2120
• Kernel	£295	-0.2020
• Stratification	£183	-0.1620
• Radius	£265	-0.1730
PS matching ('psmatch2') – ATT (difference)		
• Nearest neighbour	£137	-0.1667
• Kernel	£315	-0.2094
PS matching (by hand)	-£55	-0.1538
PS stratification	-£117	0.0000
PS regression	£57	-0.4905**
Heckman model	£747	n/a*

* A value hasn't been included here as it could not be computed; ** These values have been reported as coefficients instead of odds ratio
PS = propensity score; RD = random draw; EW = equal weights

5.5 Discussion

To obtain unbiased estimates of cost and effect differences, we need large, adequately powered RCTs. However, this is not always possible and non-randomised studies are often used. The women in this dataset were selected for referral to see a specialist according to specific criteria, and this may create selection bias. So in this instance, the observed costs and effects which were presented in Chapter 2 are said to be biased. This chapter has provided a quick overview on what selection bias is and the methods which can be applied to reduce selection bias. These methods were then applied to the dataset to reduce selection bias in the estimation of costs and effects. One point to note is that these methods cannot eliminate selection bias, we can only reduce it.

First, a generalised linear model and a logistic regression model was used to see whether referral mode is a significant predictor of costs and effects, respectively; however, this method in reality only deals with observed covariates. Second, the various propensity score methods (matching, stratification and regression adjustment) were used to balance the sizes and compositions of the two referral groups in order to reduce the element of bias in the estimation of costs and effects for telemedicine and direct referral patients. This method also only looks at observed variables and not at unobserved variables which may also influence costs and/or effects. Third, the Heckman sample selection method was applied to the dataset which attempts to deal with both observed and unobserved variables, to see what impact this may have on

both costs and/or effects and to see whether selection bias between the two referral groups was reduced.

However, the analysis cannot prove which of these methods is more accurate for this dataset, only some direction on which method may be the best way forward can be provided. Propensity score matching may be a more reliable way of obtaining cost and effect estimates, because after matching the groups were similar in terms of background characteristics (i.e. 'balanced') and the psmatch2 method indicated that the mean average standardised bias and Pseudo R^2 values had fallen after matching, indicating that the matching method was a good estimator. With regards to the regression or the Heckman models, they may not have explicitly balanced the covariates among the groups; therefore the two groups may not be similar. Regression models can indicate differences in costs and effects between a dependent variable (e.g. referral method) and other covariates and the results from the regression analysis indicated that referral method is not a significant predictor of costs and/or effects. The propensity score technique cannot indicate differences between the dependent variable and individual covariates, because all covariates are collapsed into a single index variable, possibly obscuring important interactions. An advantage of using propensity score matching is that matching does not have to assume linearity (i.e. assume a constant relationship between an outcome and the covariate within each treatment group), whereas with regression analyses a linear relationship is assumed. However, both regression analyses and propensity score methods only controlled for observed variables and both of these methods cannot control for unobserved variables. The Heckman method was not a good model for this dataset, as the dataset did not have any good identifier variables; that is, variables that could be used as proxy variables for unobserved variables which may have caused selection bias in this dataset.

Deeks et al (2003) considered different methods for evaluating selection bias in non-randomised studies and none of the methods which were applied (regression, stratification and propensity scoring) successfully removed bias in cohort studies. They found that most methods applied to reduce selection bias were not standardised and also some covariates are sometimes missing (not at random) which in itself can also lead to bias. They also highlighted that some methods were not relevant or meaningful in some contexts and that adequate adjustment for selection bias can only be made in an unrealistic situation when selection depends on a single factor that is measured and included in the model.

In a recent simulation study comparing propensity scores with multivariable regression models, the authors concluded that propensity scores performed better in situations with less than 8 cases per covariate [Cepeda et al, 2003]. Peduzzi et al (1996) stated that usually 10 cases per covariate are considered to be a minimum requirement for stable estimates in multivariate regression models. For the regression models, nine variables were included (including the dependent variable), so in total the number of cases required for stable estimation should be 90, however, the number of cases in this dataset was just under at 76. Apart from this specific condition there is little, if any practical guidance for researchers regarding when the use of propensity scores will produce different, and in particular, better estimates compared with conventional regression models. The small sample size may have also created an additional problem for the propensity score matching. Small sample sizes can increase the variance of estimated effects, making identification of significant effects difficult. Also, fewer matches may be available, therefore by picking distant matches increases the variance.

The fact that several types of propensity score matching techniques exist, then raises the question of which is the most appropriate one? The literature does not offer guidelines for making this choice. In principle, for matching it shouldn't matter which matching method is used, the answer should be similar. If the control group is small, then matching with replacement is the preferred option, as the average quality of matching increases and the bias decreases. However, if the control group is large, then kernel and radius matching work better [Baser, 2006].

Propensity score methods and Heckman model can be used in the same situations, as the ultimate goal for both approaches is the same, to turn an observational comparison of two groups into a quasi-randomised experiment. However, the Heckman model claims to allow for unobservable differences (but doesn't show how) and was really unstable for the effect variable, indicating that this type of model was not suited to this type of data. This may be partly due to the small number of observations (women) and also partly due to the identifier variable which had a low correlation with referral mode.

The analysis was confined to patient-related observed variables that were recorded routinely in hospital records. There may be other characteristics which were not recorded in the dataset such as social class, education, income, ethnicity or smoking which may affect the cost and effect results. Other unobserved variables which may have a possible impact on costs and effects, but have not been included in the data collection include: patient's preference for referral method, clinician's preference for

referral method and the quality of care which the patient receives at the specialist or local hospital.

One of the main limitations of this analysis is that all the models were conducted on a small sample size; however, in practice most of these models are usually conducted on bigger sample sizes. The small sample size of the dataset and the exclusion of variables such as health status, ethnicity and clinician's choice of referral mode may have affected the precision of the costs and effects estimates. Nevertheless, there is some confidence in the adjusted costs and effect estimates, as they reduced the difference (incremental) in costs and effects between the two groups.

After adjusting for selection bias, most of the adjusted cost differences were smaller than the observed differences between the two groups (except for the Heckman method); whereas for effects, these differences were similar to the observed effect differences. This means that reviewing the literature from non-randomised telemedicine studies and also from small non-randomised telemedicine studies, where both types of studies have not been adjusted for selection bias, the results from these studies should be interpreted for caution. This is because we don't know how much selection bias was in the dataset and what the direction of the bias is. As the results from this study have shown, the direction (and/or magnitude of association) of differences between the two groups can go either way.

There is no set criterion to judge which method is the most appropriate in controlling for selection bias in observational studies. Below are some pointers which may help:

a) Three types of variables(s) are required: 1) variables that only affect costs and/or effects but not the dependent variable; 2) variables that only affect the dependent variable but not costs and/or effects; and 3) variables that affect costs and/or effects and the dependent variable.
b) Does the method work best for observed or unobserved variables or for both?
c) Does the method work best for recorded or unrecorded variables or for both?
d) Are the coefficients obtained from each model behaving in the way we would expect them to? That is, the signs indicate the direction of association between the variables and the sizes indicate the strength of magnitude for these associations.
e) Conduct goodness of fit tests and provide summary statistics such as R^2. For example, for the logistic regression the Hosmer-Lemeshow test can be conducted.
f) How does each method reduce selection bias and what tests can be applied to see whether selection bias has been controlled for? The sample selection model, for

example, provides an adjustment factor and the p-value for the test to shows how much selection bias is still present.

g) What is the most appropriate sample size for each of the methods? Cut-off points are required to define small, medium and large sample sizes. This well also help determine which is the best sample size for each method to see either an economically important or policy relevant difference.
h) Finally, we need to compare the results obtained from these models with results from existing literature and see whether the results seem plausible.

As these results have varied greatly between each of these different methods, this means that a researcher cannot simply use the results from one method of correcting for selection bias to represent all methods especially for this dataset; rather, an argument has to be made concerning which model is best to accept. Also, due to the size and nature of this dataset, the bias adjustment still does not provide a clearer picture of how much selection bias remains in the dataset and what the direction is. However, I believe that the propensity scoring methods worked better for this dataset, because after propensity score matching, the two groups were similar in terms of background characteristics and the adjusted cost differences were smaller. After all, there is no method to check after adjustment for the other two methods whether the groups were balanced.

This chapter has provided a review of one of the three economic challenges associated with telemedicine which was identified from the literature review: selection bias. The analysis in this chapter still cannot conclude which of these methods is the most accurate; however, for this dataset the propensity scoring methods seemed the most reliable methods in reducing selection bias, although the research would suggest that using all these approaches appropriately in order to reduce bias. The next chapter (Chapter 6) will look at the second economic challenge associated with telemedicine: patient costs and the following chapter (Chapter 7) will look at the final issue: benefit measures such as QALYs.

CHAPTER 6: EXTENSION OF THE ECONOMIC EVALUATION OF THE TELEPAED PROJECT AND THE CALCULATION OF PATIENT COSTS

6.1 Introduction

This Chapter will be presented in two parts. In the first part of this Chapter (section 6.2), the analysis presented in Chapter 2 will be extended; that is, looking at the costs and effects not only of the referred women but also the non-referred women, who were seen during the same time period at Medway hospital. In the second part of this Chapter (section 6.3), the second economic issue associated with telemedicine which was highlighted in Chapters 3 and 4: patient costs, will be discussed.

One of the main limitations of the dataset presented in Chapter 2, was that it focused on a small number of women who were referred to a perinatal cardiologist after a routine anomaly scan (n = 76). As the allocation of women to these two groups was based on eligibility criteria (i.e. not random), the issue of selection bias and how to deal with selection bias was the main focus of Chapter 5. Therefore, the aim of this Chapter is to look at the whole screening programme within the hospital and to conduct a cost-effectiveness analysis by comparing a service with telemedicine to a service without telemedicine. This Chapter extends the analysis conducted in Chapter 2 and includes all women screened: 1) those women referred to a perinatal cardiologist and 2) those 'non-referred women' who were managed in the DGH who had also undergone an anomaly scan at 20-22 weeks gestation at Medway hospital during May 2001 to July 2002. Thus, by including all women who had undergone an anomaly scan during this period; is the best way to deal with selection bias given the available data, as it does not artificially introduce 'selection'. However, looking at the total population during this period, also raises a number of issues which this Chapter will aim to address.

Firstly, only a sample of non-referred women's costs and outcomes were observed, so costs and outcomes will be extrapolated to all non-referred women who had undergone an anomaly scan during the same time period. Secondly, to conduct a cost-effectiveness analysis of a service with telemedicine compared to a service without telemedicine, what assumptions would have to be made if there was no telemedicine service? If there was no telemedicine service, the women who are in the telemedicine group would be reclassified to be either managed in the DGH or be sent directly to London for face-to-face assessment; so the key assumptions that need to be made for this comparative analysis of both costs and outcomes need to be outlined. Thirdly, what happens to costs and outcomes over time i.e. when the telemedicine service is in

a steady state? Additional data have been collected for this purpose. Finally, costs and outcomes will be combined into a cost-effectiveness analysis.

6.2 Extension of the TelePaed project and additional data for thesis

As mentioned in Chapters 2 and 4, the RBH in London set up a telecardiology service in four DGHs for the provision of specialist advice to clinicians in obstetric and paediatric departments [Dowie et al, 2007]. The main fieldwork for the TelePaed project was conducted over a 15-month period. However, only one of the district hospitals (Medway Hospital) used the telecardiology service for fetal referrals alongside the existing arrangement for referring women directly to London specialist centres. The eligible pregnant women identified for the project (see Chapter 2, section 2.4.2) and the resource use and unit costs (see Chapter 2, section 2.5.1) have been discussed earlier.

The aim of this section is to see whether a service with telemedicine in addition to standard care (direct referral to specialists in London or for care to be provided for in DGH) is a cost-effective alternative compared to a service without telemedicine (consisting of only standard care); and also to see the change over time. So for this cost-effectiveness analysis, additional data on resources used and effectiveness data for pregnant women a few years after implementation of the telemedicine service (when the service was in a steady state) were also collected. This phase covered women who had an anomaly scan in 2005/2006. Permission was obtained from the Deputy General Manager for Children's and Women's Care at Medway hospital, on 10th October 2007, to obtain this additional dataset.

This additional dataset in an anonymised form included the following information on patient demographics: date of birth; parity; number of fetuses; and risk factors such as diabetes, whether on lithium or anti-epileptic therapy, Down's syndrome risk, elevated serum risk, family history of CHD or previous pregnancy with abnormality, or no risk factor. Resources used during antenatal care from time of anomaly scan for each pregnant women included: type of antenatal scans and dates, types of antenatal and outpatient clinic attendances and dates, dates of specialist consultations either face-to-face (direct referral to specialist) or by telemedicine, type of termination procedure and date (if applicable), counselling for those women who were seen in London, and any prenatal inpatient admissions including dates of admission and discharge. Resources used during maternal delivery included: length of inpatient stay prior to transfer to labour ward, mode of delivery, delivery place and length of stay in labour/delivery ward and length of stay on the postnatal ward (all lengths of stay included admission and

discharge dates). This information was collected and entered onto data extraction forms (see Appendix 7: Data Extraction Form) by the project facilitator at Medway hospital and then the data were entered onto Microsoft Excel spreadsheets to carry out subsequent analyses.

6.2.1 A few years after implementation of telemedicine at Medway hospital

The time period for this additional data collection and analysis was from 1st May 2005 to 31st July 2006 (this time period was then comparable to that of the earlier data). This later time period was chosen to see what the effect on costs and outcomes would be when the telemedicine service was in a steady state of use. The average referral number during 2005/2006 for women being sent directly to KCH (multi-organ plus cardiac anomaly referrals) or to QCH (cardiac anomaly referrals) was approximately 30 patients, the rest of the women who required a specialist opinion were assessed via telemedicine (approx. n = 70) [Personal communication with Sonographer A, October 2007]. The sonographers also received additional training from staff who were based at the Tiny Tickers Charity and also from Specialist A.

In terms of utilising the telemedicine service during this time period, there was a more established pattern. Telemedicine clinics were scheduled monthly, on average they lasted 45 minutes to an hour and the average number of patients seen in each clinic was between 12 and 15. So each patient was reviewed for approximately 4 to 5 minutes. If a woman was sent to London for referral, they were given an hour slot for the specialist scan and counselling.

Specialist A at RBH provided the project facilitator with dates of each patient's specialist contact (either via telemedicine or direct referral), plus the patient's id and diagnosis. This enabled the project facilitator to identify all women who were seen by the specialists in London. A random sample of non-referred women at 'medium risk' and 'low risk' whose care was provided for in the DGH were identified for this time period (see section 6.2.3.1 for categorisation of women into risk groups) by the project facilitator using hospital appointment calendars (approximately ten women each month).

6.2.2 Effectiveness

For the economic evaluation there were two measures of effectiveness: 1) detection of cardiac abnormalities before birth, and 2) number of cardiac abnormalities which were 'missed' at the anomaly scan (or at other subsequent antenatal scans) and only detected at birth.

1) Detection of cardiac abnormalities before birth

The project facilitator noted the findings from the anomaly scan: 1) a normal heart or 2) an abnormal heart. Further diagnosis of the heart was confirmed at either the telemedicine consultation or at the face-to-face meeting with the fetal cardiologist in London (referred women only). If a cardiac anomaly was found at one of these time points, this gave the women the opportunity to be informatively counselled and to decide whether or not to continue with the pregnancy.

2) Number of cardiac abnormalities which were 'missed' at the anomaly scan and only detected at birth

The project facilitator noted the findings of the newborn's birth outcome. This included whether the birth was a live birth or a still birth; and also the newborn's cardiac status: 1) a normal heart or 2) an abnormal heart. If the baby was born with a cardiac defect and this was not detected prenatally, then this was recorded as a 'missed' anomaly.

To make sure that the number of all cardiac anomalies detected before birth (even those that were terminated) and the number of all missed cardiac anomalies were correct, with the help of Specialist A, the Central Cardiac Audit Database Registry (CCAD) for the specified time periods in the analysis was checked. The CCAD is a UK clinical audit project involving cardiac surgeons, cardiologists and cardiac anaesthetists. The registry holds patient-specific data on major heart diseases that need coronary intervention (surgery or catheterisation) and their outcomes.

6.2.3 Statistical analysis

6.2.3.1 Comparative analyses

For the comparative analyses, the women who formed the dataset were categorised into three main groups:

1. Women who had an elevated risk factor for CHD or were suspected of having a fetal anomaly and were referred to a specialist via telemedicine (telemedicine group);
2. Women who had an elevated risk factor for CHD or were suspected of having a fetal anomaly and were referred directly to see a specialist face-to-face (direct referral group); and
3. Women whose care was managed in the DGH:
   a. Medium risk women who had an elevated risk factor for CHD or Down's syndrome, but were not considered 'urgent' to be referred to a perinatal

cardiologist and therefore their care was managed in the DGH (medium risk women);

b. Low risk women ('normal women') included women with a low risk of having a baby born with CHD and whose care was managed in the DGH (low risk women).

6.2.3.2 Without a telemedicine service

To determine what would happen to the costs and effects if the telemedicine service was not available, all women assessed via telemedicine were reclassified as either women who are 'seen by direct referral' in London or women whose care is managed in the DGH as 'medium risk' women. Specialist A confirmed that no women who were assessed by telemedicine would come under the category of 'low risk'. Specialist A also confirmed that the following women would be seen directly in London: diabetic women; women on epileptic or lithium therapies; women with a previous child requiring surgery for CHD or a family history of CHD; women with an increased nuchal translucency risk (> 3.5mm); and women with a suspicion of cardiac anomaly. Conversely, women with an increased risk of Down's syndrome; women with multiple pregnancies; and all other women who were seen by telemedicine would be managed in the DGH as 'medium risk' women. So on the cost side, the cost of a telemedicine consultation would be replaced by a specialist consultation cost, if the patient was assumed to have been seen directly. For those patients whose care would now be managed in a DGH, the cost of a telemedicine consultation was removed. On the effect side, Specialist A also asserted that all telemedicine women whose fetus had a cardiac anomaly would have all been assessed by direct referral, so there would not have been any missed cardiac cases if the telemedicine service was not available.

6.2.3.3 Regression analyses for total costs of pregnancy for the observed sample population

The two datasets also obtained demographic, clinical and resource use data for a sample of 'medium' and 'low' risk women. This information was used to calculate the total costs of pregnancy for these women (see Chapter 2, section 2.5.1 for more information). Using this information and to determine what risk factors were important predictors of the total costs of pregnancy, multiple regression models were fitted to observed caseloads of women. Firstly, unadjusted total costs of pregnancy are presented (these are the same as the observed costs); and secondly, a multiple regression model for the total costs of pregnancy is presented, adjusting for all risk factors which included: parity; mother's age at anomaly scan; number of fetuses; gestation in weeks at time of the anomaly scan; and also whether the woman had one

of the following risk factors: diabetes; an elevated serum risk; a high risk of Down's syndrome; family history of CHD; or a previous pregnancy with an anomaly.

6.2.3.4 Extrapolating sample costs and effects to population costs and effects

As mentioned in the previous section, for both time periods (2001/2002 and 2005/2006), due to the time and resource constraints only a sample of medium and low risk patients' resource use and effectiveness data have been collected (all referred women in each time period were identified). For each of the time periods, the total number of births (including still births), maternities, multiple pregnancies and terminations were obtained from Medway hospital [Personal communication with Medway Hospital administrative assistant, September 2008]. During the 2001/2002 period, all possible women who were of 'medium risk' of a fetal CHD anomaly (excluding seven women) had been identified in the TelePaed audit. Sonographer B and the project facilitator confirmed that on average, approximately the same number of women during each time period would be classed as 'medium risk' (this included women who would have had a multiple pregnancy[7] or a CHD or Down's syndrome risk factor which was not of concern, therefore they could be managed in the DGH). So the proportion of women during this period who were classed as 'medium risk', this ratio was used for the latter time period for the medium risk women. The remainder of women in each time period were then allocated to the low risk group.

The next step was to extrapolate the sample costs to the population costs (i.e. to the rest of the women who were managed in the DGH – medium and low risk women for whom we had no resource use data for), taking into account the different numbers of women in each group. All women with missed cardiac anomalies in their fetuses and all women with cardiac anomalies which were detected before birth had been identified in the two samples. Multiple regression models were fitted to observed caseloads of women in each sample (in 2001/2002 and 2005/2006) to predict costs for women with missing resource use and cost data. Total costs of pregnancy were predicted for each risk group, in each time period, for patients with 'no cardiac anomaly'. The costs were adjusted for the following risk factors: parity; mother's age at anomaly scan; number of fetuses; gestation in weeks at time of anomaly scan; diabetes; an elevated serum risk; a high risk of Down's syndrome; family history of CHD; and a previous pregnancy with an anomaly. The mean costs were obtained for each group (in each time period) and were then allocated to the rest of the women in each group (in each time period) who did not have a cost.

[7] Multiple pregnancy was a factor for the TelePaed eligibility criteria, but not for the Medway eligibility protocol.

6.2.3.5 Comparison of cost scenarios

The time horizon for each study period was from the time of the anomaly scan (i.e. 20-22 weeks gestation) until just after delivery or where applicable after termination of pregnancy. To keep costs for all different time periods in the same financial year, cost results are presented in 2005/2006 prices. Given the one-year short time frame, costs (and effects) were not discounted.

So to summarise, costs and effects were calculated for four main time periods:
1) 2001/2002 period – observed costs and effects with a telemedicine service;
2) 2001/2002 period – costs and effects adjusted for a service without telemedicine;
3) 2005/2006 period – observed costs and effects with a telemedicine service; and
4) 2005/2006 period – costs and effects adjusted for a service without telemedicine.

The cost analysis compares a service with telemedicine to a service without telemedicine:

a) 1 vs. 2
b) 3 vs. 4

As there was no difference in the number of missed cases for each time period[8], a cost-effectiveness analysis was not appropriate and results for a cost-consequences analysis study are presented.

6.2.3.6 Statistical tests

Statistical analyses were conducted using Stata version 10 [StataCorp, 2007]. If distributions were normal: means, SD and ranges are presented and chi-squared, t tests and F tests conducted. If distributions are not normal: means, SDs, medians and IQRs are presented and Kruskal Wallis tests (if more than 2 groups) or Wilcoxon sign-rank tests (if two groups) are conducted. All statistical tests were two-sided unless otherwise stated. A p-value ≤ 0.05 was considered to be statistically significant. As cost data were skewed, bootstrapping was used (see Chapter 2, section 2.5.3 for further information).

[8] There was no difference in effectiveness when comparing 2001/2002 with and without a telemedicine service and when comparing 2005/2006 with and without a telemedicine service; however, we must note that there is a major difference in effectiveness 'over time' and this will be discussed later in the Chapter.

6.2.4 Results

6.2.4.1 Patient numbers for the sample and total populations

The total number of births recorded for the two 15-month time periods: 2001/2002 and 2005/2006 were 5,211 and 5,492 respectively (see Table 6.1). These figures also take into account women who gave birth to more than one baby.

Table 6.1: Total number of births, maternities and type of pregnancy

	1/5/2001 to 31/7/2002	1/5/2005 to 31/7/2006
Total number of births	5,211	5,492
Total number of maternities	5,114	5,407
Type of pregnancy		
Single	5,018	5,323
Twin	95	83
Triplet	1	1

During the two time periods, the total number of women (total population) seen by a specialist face-to-face in London, or via telemedicine, or managed in the DGH - medium and low risk are shown in Table 6.2 below.

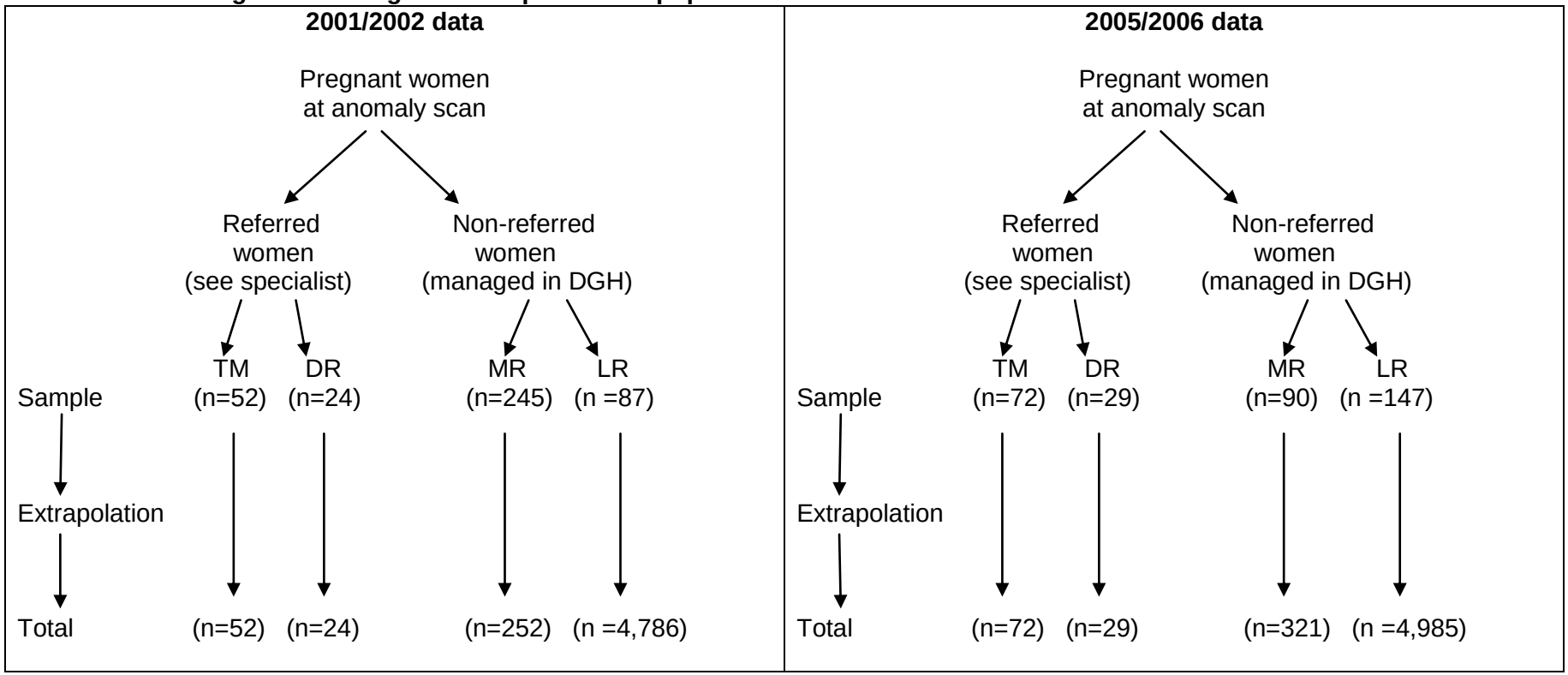
Table 6.2: Total number of maternities, cardiac cases detected and number of missed cardiac anomalies for each group in each time period

	1/5/2001 to 31/7/2002				1/5/2005 to 31/7/2006			
	Total N	Cardiac cases detected*	Missed cardiac anomalies	Sample N	Total N	Cardiac cases detected	Missed cardiac anomalies	Sample N
TM	52	7	0	52	72	9	0	72
DR	24	8	0	24	29	15	0	29
MR	252	7	2	245	321	0	0	90
LR	4,786	3	5	87	4,985	0	1	147
Total	5,114	25	7	408	5,407	24	1	338

Referred women: TM = telemedicine; DR = Direct referral;
Non-referred women: MR = medium risk; LR = Low risk;
* 3 Cases (1 in DR and 2 in TM) were really insignificant cardiac anomalies which had corrected themselves before the baby was born
Sources: 2001/2002 data from TelePaed dataset; 2005/2006 data from Sonographers A and B.

The number of women in the two samples for whom data were available are also shown in Table 6.2. The number for the sample of medium risk women in each of the two time periods was different, because for the earlier time period the project facilitator collected this data prospectively, so that (nearly) all the medium risk women during this time period were identified; whereas for the latter time period, the project facilitator picked women randomly from hospital calendars and they could either belong to the medium or low risk group, hence not all medium risk women were identified during this extra data collection and also Medway eligibility criteria did not include women with multiple pregnancies. Figure 6.1 below shows the process and numbers of pregnant women in both the sample and total populations for the two time periods.

Figure 6.1: Process diagram showing from sample to total population



Key: TM = telemedicine; DR = Direct referral; MR = medium risk; LR = Low risk;

Table 6.3: Referred women according to risk status of pregnancy

	2001/2002		2005/2006	
	Telemedicine n = 52	Direct referral n = 24	Telemedicine n = 72	Direct referral n = 29
Risk status of pregnancy				
Down's/cardiac risk	46 (88.5%)	18 (75.0%)	63 (87.5%)	21 (72.4%)
Low risk	6 (11.5%)	6 (25.0%)	9 (12.5%)	8 (27.6%)
Cardiac anomalies detected by risk status				
Total number of cardiac anomalies detected	7 (13.5%)	8 (33.3%)	15 (20.8%)	9 (31.0%)
Down's/cardiac risk	3 (42.8%)	4 (50.0%)	7 (46.7%)	5 (55.6%)
Low risk	4 (57.1%)	4 (50.0%)	8 (53.3%)	4 (44.4%)

Table 6.3 shows that according to risk status of pregnancy the overall number of detected cases for the referred women in the two 15-month periods. The table clearly highlights the relatively large detection of abnormalities amongst the women who were classed as 'low risk' (i.e. these women had no risk factors for a fetal anomaly, and only when the anomaly scan was conducted an anomaly was found; hence these women were then referred to a perinatal cardiologist). When looking at the numbers detected and missed for the women who are of 'low risk' (see Tables 6.2 and 6.3), these numbers are: 2001/2002 – 8 detected and 5 missed (13 in total) and 2005/2006 – 12 detected and 1 missed (13 in total). This emphasises the importance of screening for CHD in low risk women (Chapter 7 will discuss this further).

6.2.4.2 Patient demographics for a sample population

Tables 6.4a to 6.4b show the demographic characteristics for the observed samples of pregnant women. For 2001/2002 period, the mean age for women in each group was similar and not statistically different ($p = 0.142$); there were no significant differences between the groups in terms of parity ($p = 0.470$); a third of patients in the medium risk group had a multiple pregnancy; for all groups, over half of the women had their anomaly scan at 21 weeks; and 50% of women gave birth between 38 to 40 weeks gestation. For the 2005/2006 time period, there were significant differences in the mean age of women ($p < 0.001$), the women of medium risk who were cared for in the DGH were slightly older than the other three groups of women; there were no significant differences in terms of parity in each group ($p = 0.139$); nearly all women in this sample were expecting one baby; over half of the women (except the direct referral group) had their anomaly scan at 21 weeks; and similar to the earlier time period, 50% of women gave birth between 38 to 40 weeks gestation.

When looking at demographic characteristics across the two time periods, the mean age for all women was 29 years and there were no statistically significant differences in maternal age ($p = 0.358$); there were significant differences in parity ($p < 0.001$), in the 2001/2002 period there where more multiparous women compared to the other time

period where the ratio between primaparous and multiparous was approximately 50-50; there were also significant differences in terms of the type of pregnancy (single versus multiple births) as there where more multiple pregnancies in 2001/2002 data collection ($p < 0.001$ – Medway criteria did not include multiple pregnancies); there were no significant differences in the mean number of weeks gestation when the anomaly scan was conducted ($p = 0.124$); and there were no significant differences in the mean number of weeks gestation at birth ($p = 0.132$).

Table 6.4a: Demographic characteristics of pregnant women for 2001/2002 period

2001/2002 period	TM group (n = 52)	DR group (n = 24)	MR women (n = 245)	LR women (n = 87)	All women (n = 408)
Maternal age					
N	52	24	245	87	408
Mean (SD)	27.7 (6.1)	28.6 (6.8)	29.6 (5.6)	29.4 (5.3)	29.3 (5.7)
Range	16 to 43	19 to 43	16 to 43	16 to 42	16 to 43
Parity					
N	52	24	245	87	408
Primaparous	19 (36.5%)	12 (50.0%)	93 (38.0%)	39 (44.8%)	163 (40.0%)
Multiparous	33 (63.5%)	12 (50.0%)	152 (62.0%)	48 (55.2%)	245 (60.0%)
Type of pregnancy*					
N	52	24	245	87	408
Singleton	52 (100.0%)	23 (95.8%)	157 (64.1%)	87 (100.0%)	319 (78.2%)
Multiple	0 (0.0%)	1 (4.2%)	88 (35.9%)	0 (0.0%)	89 (21.8%)
Gestation at anomaly scan*					
N	52	24	245	87	408
Mean (SD)	20.8 (0.9)	21.8 (3.0)	21.0 (1.2)	20.9 (0.7)	21.0 (1.3)
Range	18 to 23	19 to 33	17 to 32	19 to 24	17 to 33
≤ 19 weeks	4 (7.7%)	1 (4.2%)	11 (4.5%)	1 (1.1%)	17 (4.2%)
20 weeks	12 (23.1%)	4 (16.7%)	49 (20.0%)	18 (20.7%)	83 (20.3%)
21 weeks	28 (53.8%)	12 (50.0%)	141 (57.6%)	58 (66.7%)	239 (58.6%)
22 weeks	7 (13.5%)	5 (20.8%)	37 (15.1%)	7 (8.0%)	56 (13.7%)
≥ 23 weeks	1 (1.9%)	2 (8.3%)	7 (2.9%)	3 (3.4%)	13 (3.2%)
Gestation at birth*					
N	50	20	240	87	397
Mean (SD)	38.1 (2.8)	37.6 (2.2)	37.7 (3.1)	39.2 (1.8)	38.1 (2.9)
Median	39	38	38	39	39
IQR	37 to 40	36 to 39	36 to 40	39 to 40	37 to 40
≤ 30 weeks	2 (4.0%)	0 (0.0%)	5 (2.1%)	1 (1.1%)	8 (2.0%)
31-35 weeks	5 (10.0%)	3 (15.0%)	38 (15.8%)	2 (2.3%)	48 (12.1%)
36-37 weeks	6 (12.0%)	6 (30.0%)	53 (22.1%)	7 (8.0%)	72 (18.1%)
38-39 weeks	16 (32.0%)	7 (35.0%)	61 (25.4%)	35 (40.2%)	119 (30.0%)
40 weeks	16 (32.0%)	2 (10.0%)	42 (17.5%)	24 (27.6%)	84 (21.2%)
≥ 41 weeks	5 (10.0%)	2 (10.0%)	41 (17.1%)	18 (20.7%)	66 (16.6%)

[±] Statistical tests conducted: t tests or F tests for age, gestation at anomaly scan and birth; chi-squared tests for parity and type of pregnancy.

* $p < 0.05$

Table 6.4b: Demographic characteristics of pregnant women for 2005/2006 period

2005/2006 period	TM group (n = 72)	DR group (n = 29)	MR women (n = 90)	LR women (n = 147)	All women (n = 338)
Maternal age*					
N	72	29	90	153	338
Mean (SD)	28.6 (6.1)	27.8 (6.4)	33.8 (6.4)	28.1 (6.0)	29.7 (6.6)
Range	17 to 43	17 to 42	18 to 44	17 to 42	17 to 44
Parity					
N	72	29	90	147	338
Primaparous	33 (45.8%)	19 (65.5%)	54 (60.0%)	74 (50.3%)	180 (53.3%)
Multiparous	39 (54.2%)	10 (35.5%)	36 (40.0%)	73 (49.7%)	158 (46.7%)
Type of pregnancy*					
N	72	29	90	147	338
Singleton	70 (97.2%)	29 (100.0%)	81 (90.0%)	147 (100.0%)	327 (96.7%)
Multiple	2 (2.8%)	0 (0.0%)	9 (10.0%)	0 (0.0%)	11 (3.3%)
Gestation at anomaly scan*					
N	72	29	90	147	338
Mean (SD)	21.3 (1.7)	20.1 (2.0)	21.2 (0.9)	21.1 (0.8)	21.1 (1.2)
Range	20 to 33	17 to 25	19 to 24	19 to 26	17 to 33
≤ 19 weeks	0 (0.0%)	12 (41.4%)	1 (1.1%)	2 (1.4%)	15 (4.4%)
20 weeks	13 (18.1%)	2 (6.9%)	13 (14.4%)	16 (10.9%)	44 (13.0%)
21 weeks	41 (56.9%)	10 (34.4%)	50 (55.6%)	97 (66.0%)	198 (58.6%)
22 weeks	12 (16.7%)	3 (10.3%)	19 (21.1%)	27 (18.4%)	61 (18.0%)
≥ 23 weeks	6 (8.3%)	2 (6.9%)	7 (7.8%)	5 (3.4%)	20 (5.9%)
Gestation at birth*					
N	72	27	86	147	332
Mean (SD)	37.8 (2.8)	38 (3.1)	38.1 (3.8)	38.9 (2.2)	38.4 (2.9)
Median	38	38	40	39	39
IQR	37 to 40	37 to 40	37 to 40	38 to 40	38 to 40
≤ 30 weeks	1 (1.4%)	1 (3.7%)	6 (7.0%)	1 (0.7%)	9 (2.7%)
31-35 weeks	10 (13.9%)	1 (3.7%)	7 (8.1%)	6 (4.1%)	24 (7.2%)
36-37 weeks	10 (13.9%)	5 (18.5%)	10 (11.6%)	14 (9.5%)	39 (11.7%)
38-39 weeks	30 (41.7%)	12 (44.4%)	20 (23.3%)	58 (39.5%)	120 (36.1%)
40 weeks	12 (16.7%)	3 (11.1%)	24 (27.9%)	42 (28.6%)	81 (24.4%)
≥ 41 weeks	9 (12.5%)	5 (18.5%)	19 (22.1%)	26 (17.7%)	59 (17.8%)

[±] Statistical tests conducted: t tests or F tests for age, gestation at anomaly scan and birth; chi-squared tests for parity and type of pregnancy.
* $p < 0.05$

6.2.4.3 Effectiveness results for the total population

Table 6.2 also showed the total number of cardiac anomaly cases detected before birth for each of the two time periods. These numbers include pregnant women who went on to have a termination. In total, the number of cardiac anomaly cases detected before birth for the two time periods (2001/2002 and 2005/2006) were 25 (0.5%) and 24 (0.4%) cases, respectively.

The table also showed the total number of 'missed' cardiac anomalies not detected prenatally. None of the 'missed' cases were from women in either the direct referral or telemedicine groups. The 2001/2002 time period indicated that there were a total of seven missed cardiac anomalies, whereas during the 2005/2006 period there was only one missed cardiac anomaly. During the earlier time periods, temporary staff were called in to scan pregnant women when there was a staff shortage or staff were on sick leave. These temporary staff members who had varied training and experience, may have led to an increase in the number of missed cases [Personal communication with Sonographer A and Sonographer B, October 2007]. The lower number of missed anomalies in the 2005/2006 period could also reflect the fact that staff were more confident and were better trained in detecting fetal heart anomalies. This is better reflected in Table 6.5, which shows that over time the detection rate at the anomaly ultrasound scan has improved and the number of 'missed' cardiac cases has fallen.

Table 6.5: Number of 'missed' cardiac anomalies based on year when anomaly scan was undertaken (data obtained from Medway hospital and Central Cardiac Audit Database)

Year	Total number of births	Total missed cases	Rates of missed cases per 1,000 births
2000	4,035	7	1.73
2001	4,092	7	1.71
2002	4,184	3	0.72
2003	4,216	3	0.71
2004	4,265	1	0.23
2005	4,260	2	0.47
2006	4,449	0	0.00
2007	4,623	1	0.22

6.2.4.4 Resource use analysis and hospital utilisation patterns for a sample population

Tables 6.6a to 6.6b show the frequency of resources used by a sample of women in each group in each time period. One point to note for Table 6.6a (2001/2002 period) for the following items of resource use: prenatal maternity bed day (prior to transfer to labour); mode of obstetric delivery; and postnatal bed day includes missing values which have been imputed. Of the 60 women who had missing values imputed, 1 patient was in the direct referral group; 11 women were in the telemedicine group; 27

women were in the medium risk group; and 21 women were in the low risk group (see Chapter 2, section 2.5.1.1 for more details).

For the sample data on average, for the women in the 2001/2002 period (Table 6.6a) compared to the women in the 2005/2006 period (Table 6.6b), they had slightly more antenatal scans (2.9 vs. 2.4 scans); more antenatal clinic visits (9.9 vs. 7.5 visits); and had more caesarean births (41.6% vs. 26.5% - due to the higher proportion of multiple pregnancies) which contributed to the higher mean total costs of pregnancy in the earlier time period. This is highlighted in Table 6.6c which shows the proportions of the total costs of pregnancy (i.e. the direct costs to the hospital) attributable to each component for each of the two time periods. The biggest cost driver during antenatal care was antenatal/outpatient clinic attendances and during maternal delivery, the biggest cost driver was the mode of delivery. Over 60% of the costs for each period were made up of the following components: labour/delivery bed stay, mode of delivery and postnatal bed stay.

Table 6.6a: Resource components and the number of women who used the resource items for 2001/2002 period

20001/2002 period	TM group (n = 52)	DR group (n = 24)	MR women (n = 245)	LR women (n = 87)	All women (n = 408)
Resource use at district hospital					
1. Antenatal ultrasound scans	52 (100%)	24 (100.0%)	245 (100.0%)	87 (100.0%)	408 (100.0%)
Mean number of scans per patient (SD)	2.8 (1.5)	2.8 (1.3)	3.2 (1.8)	2.2 (1.3)	2.9 (1.7)
Median (Inter-quartile range)	3 (2 to 4)	2 (2 to 4)	3 (2 to 4)	2 (1 to 3)	2 (2 to 4)
2. Antenatal or outpatient clinics	50 (96.2%)	20 (83.3%)	240 (98.0%)	87 (100.0%)	397 (97.3%)
Mean number of clinics per patient (SD)	11.4 (6.7)	10.5 (5.6)	10.1 (4.5)	8.3 (4.2)	9.9 (4.9)
Range	1 to 35	1 to 24	1 to 26	1 to 24	1 to 35
3. Termination of pregnancy	2 (3.9%)	4 (16.7%)	5 (2.0%)	0 (0.0%)	11 (2.7%)
4. Prenatal maternity bed day					
a) During antenatal period	19 (36.5%)	4 (16.7%)	80 (32.7%)	33 (37.9%)	136 (33.3%)
Mean number of days per patient (SD)	3.3 (2.5)	2.5 (1.3)	4.1 (5.3)	2.3 (2.1)	3.5 (4.3)
Median (Inter-quartile range)	2.0 (1.0 to 5.0)	2.5 (1.5 to 3.5)	2.0 (1.0 to 4.0)	2.0 (1.0 to 2.0)	2.0 (1.0 to 4.0)
b) Prior to transfer to labour ward	17 (34.0%)	4 (16.7%)	62 (25.8%)	27 (31.0%)	110 (27.7%)
Mean number of days per patient (SD)	0.9 (0.8)	2.4 (1.9)	3.6 (6.8)	0.7 (0.9)	2.4 (5.3)
Median (Inter-quartile range)	0.4 (0.3 to 1.0)	2.5 (0.8 to 4.0)	1.6 (0.6 to 3.0)	0.4 (0.2 to 0.5)	1.0 (0.4 to 3.0)
*5) Mode of obstetric delivery**					
Normal birth	32 (64.0%)	11 (55.0%)	94 (39.2%)	54 (62.1%)	191 (48.1%)
Forceps birth	0 (0.0%)	0 (0.0%)	11 (4.6%)	2 (2.3%)	13 (3.3%)
Ventouse birth	1 (2.0%)	1 (5.0%)	11 (4.6%)	2 (2.3%)	15 (3.8%)
Caesarean birth (without complications)	16 (32.0%)	8 (40.0%)	117 (48.8%)	24 (27.6%)	165 (41.6%)
Home birth	1 (2.0%)	0 (0.0%)	6 (2.5%)	5 (5.7%)	12 (3.0%)
6) Postnatal maternity bed day	39 (78.0%)	18 (90.0%)	214 (89.2%)	73 (83.9%)	344 (86.6%)
Mean number of days per patient (SD)	4.5 (3.9)	3.9 (2.5)	3.8 (2.6)	2.7 (1.8)	3.7 (2.7)
Median (Inter-quartile range)	4.0 (2.0 to 5.4)	3.0 (2.0 to 6.0)	3.0 (2.0 to 4.7)	2.5 (1.6 to 3.2)	3.0 (2.0 to 4.0)
Resource use at specialist hospital					
1. Specialist ultrasound scans	6 (11.5%)	24 (100.0%)	0 (0.0%)	0 (0.0%)	30 (7.4%)
Mean number of scans per patient (SD)	1.2 (0.4)	1.1 (0.3)	n/a	n/a	1.1 (0.3)
Median (Inter-quartile range)	1 (1 to 1)	1 (1 to 1)			1 (1 to 1)
2. Counselling	5 (9.6%)	9 (37.5%)	0 (0.0%)	0 (0.0%)	14 (3.4%)

* One woman in medium risk group had a water birth

Table 6.6b: Resource components and the number of women who used the resource items for 2005/2006 period

2005/2006 period	TM group (n = 72)	DR group (n = 29)	MR women (n = 90)	LR women (n = 147)	All women (n = 338)
Resource use at district hospital					
1. Antenatal ultrasound scans	72 (100.0%)	29 (100.0%)	90 (100.0%)	147 (100.0%)	338 (100.0%)
Mean number of scans per patient (SD)	2.9 (1.7)	2.0 (1.3)	2.7 (1.9)	2.1 (1.5)	2.4 (1.7)
Median (Inter-quartile range)	3.0 (2.0 to 4.0)	2.0 (1.0 to 2.0)	2.0 (1.0 to 3.0)	2.0 (1.0 to 3.0)	2.0 (1.0 to 3.0)
2. Antenatal or outpatient clinics	72 (100.0%)	27 (93.1%)	86 (95.6%)	147 (100.0%)	332 (98.2%)
Mean number of clinics per patient (SD)	7.3 (3.3)	7.3 (2.4)	9.6 (2.2)	6.5 (1.8)	7.5 (2.6)
Range	1 to 20	1 to 17	2 to 15	1 to 20	1 to 20
3. Termination of pregnancy	0 (0.0%)	2 (6.9%)	4 (4.4%)	0 (0.0%)	6 (1.8%)
4. Prenatal maternity bed day					
a) During antenatal period	42 (58.3%)	16 (55.2%)	21 (23.3%)	61 (41.5%)	140 (41.4%)
Mean number of days per patient (SD)	3.3 (3.8)	2.0 (1.5)	3.2 (3.5)	2.5 (1.7)	2.8 (2.8)
Median (Inter-quartile range)	2.0 (1.0 to 4.0)	1.0 (1.0 to 2.5)	2.0 (1.0 to 3.0)	2.0 (1.0 to 3.0)	2.0 (1.0 to 3.0)
b) Prior to transfer to labour ward	4 (5.6%)	0 (0.0%)	6 (6.9%)	5 (3.4%)	15 (4.5%)
Mean number of days per patient (SD)	1.3 (0.9)	n/a	1.5 (0.5)	1.0 (0.0)	1.3 (0.6)
Median (Inter-quartile range)	1.5 (0.6 to 2.0)		1.5 (1.0 to 2.0)	1.0 (1.0 to 1.0)	1.0 (1.0 to 2.0)
*5) Mode of obstetric delivery**					
Normal birth	37 (51.4%)	17 (63.0%)	43 (50.0%)	96 (65.3%)	193 (58.1%)
Forceps birth	4 (5.6%)	0 (0.0%)	5 (5.8%)	5 (3.4%)	14 (4.2%)
Ventouse birth	5 (6.9%)	1 (3.7%)	6 (7.0%)	8 (5.4%)	20 (6.0%)
Caesarean birth (without complications)	24 (33.3%)	8 (29.6%)	28 (32.6%)	28 (19.0%)	88 (26.5%)
Home birth	2 (2.8%)	1 (3.7%)	4 (4.7%)	10 (6.8%)	17 (5.1%)
6) Postnatal maternity bed day	61 (84.7%)	25 (92.6%)	73 (84.9%)	115 (78.2%)	274 (81.1%)
Mean number of days per patient (SD)	3.0 (2.2)	2.9 (2.6)	2.9 (4.2)	2.3 (2.1)	2.7 (2.9)
Median (Inter-quartile range)	2.8 (1.0 to 4.0)	2.0 (1.0 to 4.0)	2.0 (1.0 to 4.0)	2.0 (1.0 to 3.0)	2.0 (1.0 to 3.0)
Resource use at specialist hospital					
1. Specialist ultrasound scans	9 (12.5%)	29 (100.0%)	0 (0.0%)	0 (0.0%)	38 (11.2%)
Mean number of scans per patient (SD)	1.1 (0.3)	1.4 (1.5)	n/a	n/a	1.3 (1.4)
Median (Inter-quartile range)	1.0 (1.0 to 1.0)	1.0 (1.0 to 1.0)			1.0 (1.0 to 1.0)
2. Counselling	5 (6.9%)	13 (44.8%)	0 (0.0%)	0 (0.0%)	18 (5.3%)

*1 woman in the TM group and 3 women in the direct referral group gave birth in London

Table 6.6c: Total cost of pregnancy and % total (observed cases)

	2001/2002 period with telemedicine (n = 408)		2005/2006 period with telemedicine (n = 338)	
Resource use	**£'s[#]**	**% total**	**£'s[#]**	**% total**
Antenatal ultrasound scans	42,710	2.5%	29,687	2.5%
Antenatal/outpatient clinics	315,305	18.2%	202,434	16.7%
Telemedicine consultations	8,450	0.5%	10,001	0.8%
Specialist scans and counselling	2,447	0.1%	3,724	0.3%
Termination of pregnancy	9,419	0.5%	5,138	0.4%
Prenatal bed stay	176,012	10.2%	99,492	8.2%
Labour/delivery bed stay	164,737	9.5%	132,660	11.0%
Mode of obstetric delivery	714,462	41.3%	550,877	45.5%
Postnatal bed stay	297,254	17.2%	177,121	14.6%
Total	1,730,796	100.0%	1,211,133	100.0%

[#] In 2005/2006 prices

6.2.4.5 Bootstrapped costs results for a sample population

The cost results presented below are for the total events during the second and third trimesters of pregnancy. For the scenario without telemedicine, according to the criteria as described in section 6.2.3.2, all telemedicine women who were seen in each time period, have been reclassified as either women who were assessed by direct referral in London or who were managed in the DGH as 'medium risk' women. One point to note is that the bootstrapped mean costs for the low risk women for all cost scenarios with and without the telemedicine service have stayed the same for each time period. This is because no telemedicine woman would have been reclassified as a low risk woman (the low risk women are not relevant to this comparison, but are shown here as a baseline).

During 2001/2002, of the 52 women who were assessed by telemedicine, if the telemedicine service had not been available, 45 women had risk factors to be assessed directly in London and 7 women would have reclassified as medium risk and their care would have been managed in the DGH. During 2005/2006 of the 72 women who were assessed by telemedicine, if the telemedicine service had not been available, 54 women would have been seen directly in London and the other 18 women would have been managed in the DGH as medium risk patients.

a) Comparison of costs with a telemedicine service

The mean cost per woman for each of the time periods for events from the time of anomaly scan up until just after delivery (or in a few cases after termination of pregnancy) are shown in Table 6.7. The marginal extra cost to the hospital of a teleconsultation over a specialist consultation in 2001/2002 was £87.32 and in 2005/2006 was £50.37. For the 2005/2006 period, the costs for the telemedicine group were slightly lower than the previous period. This is because there were more women during this period that could share the cost of line rental and call charges and also for

this latter period, there were no additional telemedicine and training support costs provided by the telemedicine coordinator from the RBH. The lower cost may also be partly due to fewer antenatal clinic visits now taking place at Medway hospital compared to the earlier time period. The NICE guidelines introduced in October 2003 (updated in March 2008) on antenatal care, provided a recommended guide of the number of antenatal clinic attendances that should take place during antenatal care and what should happen at each appointment [NCCWCH, 2003; NCCWCH, 2008].

Table 6.7: Bootstrapped total mean costs of pregnancy per group in 2005/2006 prices[±]

	Telemedicine group	Direct referral group	Medium risk women	Low risk women	Total for sample
2001/2002 period – with telemedicine					
N	52	24	245	87	408
Mean (SD)	£4,363 (£238)	£3,825 (£311)	£4,446 (£116)	£3,712 (£153)	£4,242 (£85)
95% CI	£3,896-£4,830	£3,215-£4,435	£4,220-£4,672	£3,412-£4,011	£4,075-£4,410
2001/2002 period – without telemedicine					
N		69	252	87	408
Mean (SD)	n/a	£4,055 (£185)	£4,455 (£117)	£3,712 (£153)	£4,229 (£88)
95% CI		£3,692-£4,418	£4,225-£4,685	£3,412-£4,011	£4,057-£4,401
2005/2006 period – with telemedicine					
N	72	29	90	147	338
Mean (SD)	£4,051 (£175)	£3,612 (£233)	£3,705 (£167)	£3,274 (£94)	£3,583 (£76)
95% CI	£3,709-£4,393	£3,155-£4,069	£3,378-£4,032	£3,089-£3,458	£3,435-£3,732
2005/2006 period – without telemedicine					
N		83	108	147	338
Mean (SD)	n/a	£3,877 (£157)	£3,728 (£152)	£3,274 (£94)	£3,567 (£75)
95% CI		£3,570-£4,184	£3,430-£4,026	£3,089-£3,458	£3,420-£3,714

[±] Statistical tests conducted: t tests for all costs when comparing two groups; and F tests for all costs when comparing more than two groups

After costs of other events were taken into account, costs for the telemedicine group were higher, although not significantly so, than the direct referral group (2001/2002: $p = 0.202$; 2005/2006: $p = 0.166$). For the 2001/2002 period, nearly half of the women in the medium risk group had caesarean sections compared to the other three risk groups, as this was the most costly mode of delivery this accounted for the slightly higher mean cost for the medium risk group compared to the other three risk groups. For the latter time period, the total cost may also be a reflection on the type of delivery mode: the rate of caesarean sections for the three risk groups (telemedicine, direct referral and medium risk) was about a third (far fewer multiple pregnancies in this data collection), whereas for the low risk group the rate of caesarean section was only 20%. During 2001/2002, the costs of the medium risk women were approximately £600 greater than direct referral group, a difference which was not significant ($p = 0.111$); however for the 2005/2006 period the costs were approximately similar (£3,612 vs. £3,705: $p = 0.777$). Overall, the total costs for pregnancy for the earlier time period were higher for each risk group and for the observed sample population than the latter time period; this is mainly due to the number of antenatal clinic visits which were taking

place at Medway hospital and the higher percentage of multiple pregnancies in the medium risk group (see Tables 6.4a and 6.4b). When comparing the total costs for events during pregnancy combined for each of the risk groups, the costs for year 2001/2002 were significantly higher than the 2005/2006 period ($p < 0.001$).

b) Comparison of costs without a telemedicine service

Table 6.7 also shows the total mean costs of pregnancy without a telemedicine service. So when comparing the overall observed costs (with a telemedicine service) with the estimated costs (without a telemedicine service) for the 2001/2002 period, they are estimated to be £13 lower and for the 2005/2006 period, they are estimated to be £16 lower; these differences are not statistically significant (2001/2002: $p = 0.915$; 2005/2006: $p = 0.880$). If those patients who were seen by telemedicine are classified as being assessed by direct referral, the mean cost for the direct referral group increases by £231 for the 2001/2002 period and by £265 for the 2005/2006 period, although the differences in mean costs for the direct referral groups at the two different time points are not significant (2001/2002: $p = 0.535$; 2005/2006: $p = 0.379$).

c) Comparison of unadjusted with adjusted costs

The comparison of unadjusted (observed) and adjusted costs are shown in Appendix 8. When adjusting costs for risk factors, the results are in a similar direction and magnitude to those presented in Table 6.7 and it does not change the cost results conclusions.

6.2.4.6 Bootstrapped costs results for total population

Table 6.8 shows the bootstrapped total mean costs of pregnancy for all women delivered at Medway for each cost scenario, after estimating costs for those women (in both the medium and low risk groups) in each time period for whom there were no demographic or resource use data for (see section 6.2.3.4 for more explanation).

Table 6.8: Bootstrapped total costs of pregnancy during each time period in 2005/2006 prices[±]

	2001/2002 period (with telemedicine)	2001/2002 period (without telemedicine)	2005/2006 period (with telemedicine)	2005/2006 period (without telemedicine)
Total costs of pregnancy for all women delivered				
N	5114	5114	5407	5407
Mean (SD)	£3,675 (£7)	£3,689 (£7)	£3,414 (£5)	£3,416 (£5)
95% CI	£3,660 to £3,689	£3,674 to £3,703	£3,404 to £3,423	£3,406 to £3,426

[±] Statistical tests conducted: t tests for all costs when comparing two time periods

Costs have not been presented by risk group, because for the direct referral and telemedicine groups, costs have remained the same as those presented in Table 6.6 and for the medium risk women, the cost magnitudes are very similar. Likewise, for the low risk women, their mean cost is very similar to the total cost for each time period, due to the volume of women who form the low risk category (e.g. about 94% of all mothers for the 2001/2002 period). For both time periods, a service without telemedicine was very similar, although slightly more expensive than a service with telemedicine.

6.3 Addition of patient costs

The literature review in Chapter 3 found that the majority of studies only looked at costs from the healthcare perspective i.e. the direct medical costs, and did not estimate the costs which fall on the patients and their families. Patient costs are important to include in economic evaluations as this gives an indication to policy makers and to individual patients of the likely costs they will face when attending hospitals for treatment or consultations; this is especially important if patients have to attend hospital more than once during a specific time period.

In the next section of this Chapter the various types of costs which may fall under the category of 'patient' costs (both monetary and time costs) and which of these costs should be included in an economic evaluation will be discussed. Also, the main approach to calculating patient costs in economic evaluations will be highlighted. Finally, the analysis presented in the previous section will be extended (and also in Chapter 2) to include both hospital and patient costs of a service with telemedicine compared to a service without telemedicine.

6.3.1 Why are patient costs important?

In Chapter 3, costs were categorised into three categories: 1) direct medical costs are the costs related to the use of resources due to either the disease or treatment, these are the costs to the health service; 2) direct non-medical costs are costs incurred by patients and family members, which contribute to the treatment process; and 3) indirect costs are resources lost due to treating a disease and can reflect two different costs depending on the perspective. If a patient perspective is adopted, this may reflect the loss of time (whether this is work or non-work time) to the individual in attending for treatment. If a societal perspective is adopted, time costs incurred by individuals in receiving treatments reflect the loss of production to society, whether paid or unpaid.

In this section, patient costs which are a sub-category of direct non-medical costs are of interest. Patient costs are important to include in economic evaluations as this gives an indication (i.e. behavioural implications) to policy makers and to individual patients of the likely out-of-pocket expenses they may incur. For example, pregnant women have to attend hospital not only for ultrasound scans, but also for antenatal clinics during the course of their pregnancy and these out-of-pocket expenses such as travel costs can add up. Also, for some screening programmes if patient costs are considerable, this may influence the uptake rate of that specific screening programme [Bryan et al, 1995; Robinson et al, 2007].

In the reference case analysis for NICE, costs relate to resources used for NHS and PSS[9] only. NICE have specifically stated that "Productivity costs and costs borne by patients and carers that are not reimbursed by the NHS or PSS are not included in either the reference-case or non-reference-case analyses" [NICE, 2008]. If patient costs are reimbursed by the NHS or PSS then they can be included in economic evaluations for NICE. When patient costs are included in an economic evaluation they should be included in a non-reference case analysis (explicit methods of valuation are required by NICE) and they should be reported separately to those of NHS/PSS costs. However, patient costs should not be included in the incremental cost-effectiveness ratio, especially if the outcome measure used is QALYs. In this case, including patient costs in an economic evaluation makes them less comparable to other interventions in the context of NICE.

NICE do not explicitly state the reasons why patient costs should not be included in economic evaluations. However, as they are an NHS organisation, they prefer costs and outcomes to be from a healthcare perspective. So for example, if a societal perspective was to be adopted, not only would the costs have to be from a societal viewpoint (i.e. including education, social care costs etc), but the outcomes would also have to be from a societal viewpoint. Although costs are easier to collect and analyse, outcomes are not. However, there may also be a problem of double counting, because some benefits may be classed as both costs and benefits.

6.3.2 What type of patient costs should be included in an economic evaluation?

Costing is an integral part of any economic evaluation and understanding the costs that fall upon patients' (and their families) is clearly important. For each type of cost identified, decisions have to be made as to whether its inclusion in an economic evaluation is relevant. Patients may incur costs whilst attending hospital for treatment

[9] PSS = personal social services

or an intervention and these costs can fall into two main categories: monetary costs and time costs. Monetary costs include: travel costs, and other out-of-pocket expenses such as subsistence costs (i.e. food, drink, and accommodation), and care of children and/or other dependents. If a societal perspective is adopted we would also include loss of pay and loss of productivity. Time costs include: time spent travelling and time spent at hospital (i.e. treatment/consultation time plus waiting time). In some cases, the opportunity cost of patients' time when receiving an intervention or treatment, have also been referred to as indirect costs [Luce and Elixhauser, 1990].

An example of monetary costs is provided by Elford et al (2001) who evaluated a PC-based videoconferencing system for child psychiatry assessments. Thirty children (aged 5-16 years) accompanied by a parent completed a psychiatric assessment using the videoconferencing system. Alternatively, they would have had to travel to a specialist centre (St John's) for assessment, which was approximately 670km away across the island in Newfoundland, Canada. The parents completed a cost questionnaire. The cost analysis looked at what mode of transportation the patients would have taken to St John's (an economy return flight booked at least two weeks in advance, or a return trip in a car, or a Greyhound return bus trip); what type of accommodation they would have stayed in and how many nights they would have stayed (an inexpensive hotel or a hostel, or if they stayed with family, the authors assumed that they would not have to pay anything). Food expenses were taken into account if they stayed at the hotel or the hostel. The authors did not take into account any other costs, such as loss of income from missing work, babysitting costs and any incidentals. Using these figures, each patient's estimated total cost (for travel, accommodation and food) was generated.

The study by Loane and colleagues (2000) who compared the costs of real-time teledermatology with store-and-forward teledermatology looked at patient time costs. Ninety-six patients attended a health centre and in the company of a GP were seen by a hospital dermatologist over a videolink (real-time). Before the videolink took place, the GP took instant photographs of the skin lesion and these photographs were posted along with a standard referral letter to another hospital dermatologist (store-and-forward). For both of these methods the authors looked at the mean total patient time involved which included travel time (to and from appointment), waiting time and consultation time. On average the store-and-forward consultation was shorter than the real-time consultation (41.5 mins vs. 52.2 mins). The authors from this information then estimated the cost of patient time, which was calculated by taking the average

annual income divided by 220 working days, divided by 8 working hours, divided by 60 minutes, and then multiplied by the total patient time.

Another example of time costs is provided by Halvorsen and Kristiansen (1996) who wanted to determine the costs of providing a rural population with radiology services under three different systems. The teleradiology system was compared to the existing system (a small x-ray unit at the remote site and all other examinations at the nearest radiology dept (the host site)) and to all examinations at the host site. The authors looked at the annual direct medical costs, direct non-medical (travel) costs, and indirect costs (lost production) of the three options. With regards to the cost of time the authors claimed that having a radiological examination represents a loss of leisure or of production for those employed. The production loss was assumed to be equivalent to the number of hours absent from work multiplied by the national wage rate. The cost of leisure was assumed to be zero (this was varied in the sensitivity analysis).

6.3.3 Different approaches to calculating patient costs

As mentioned in the Chapter 3, the information provided in the articles varied from detailed cost analyses to simply mentioning some of the costs and that a telemedicine service was cost-effective or cost-saving. Most studies provided very little explanation of the methods used to estimate the costs or what was included in the cost calculation. To the best of my knowledge, I am not aware of any 'gold standard' for estimating patient costs. A UK Working Party on patient costs was set up [Thompson et al, 2001] in 1999 with the specific remit of producing a standard 'patient' cost questionnaire for use in evaluating health care interventions. The aim of the project was to produce a questionnaire for measuring inputs into the health production function that relate directly to patients and informal caregivers, where patients could be used as a source of information. The questionnaire included these categories: single or multiple visits to health care facilities; domiciliary care costs; productivity losses due to illness; medication and medical supplies; and private consultations. The first two categories are relevant to this thesis and similar questions were used in the TelePaed questionnaire and the items included were: patient travel costs; patient time costs; companion costs; and other childcare and other dependent costs.

For example, in the study mentioned earlier by Loane et al (2000), the authors stated that patients completed a questionnaire which detailed the time involved and costs incurred when attending the appointment, along with details of their annual income. They grouped costs into two categories: fixed costs which described the equipment and depreciation costs; and variable costs which estimated the participants' time and

travel costs. However, they did not provide any more detail than this, whereas, Jacklin and colleagues (2003) in their economic evaluation of joint teleconsultation (virtual outreach) compared with conventional outpatient consultation, described in detail how they calculated patient costs. The authors used a postal questionnaire to collect data on the travel costs incurred by patients or anyone accompanying them when they attended their appointment. The questionnaire also recorded the time taken, including travel time, to attend the consultation. There was also information in the questionnaire about the impact on patients and their companions on paid work. If any work time was lost, the questionnaire asked about whether pay was reduced or whether anyone had taken annual leave.

The majority of studies evaluating patient costs in detail have been in the area of screening. For example, Bryan et al (1995) looked at the patient costs associated with abdominal aortic aneurysm screening; and Robinson et al (2007), looked at patient costs associated with a Chlamydia screening programme. Both studies, as mentioned earlier, felt it was important to look at patient costs, as these costs may have an influence on the uptake of screening rates. These two studies were similar in respect to the methods employed to calculate patient costs. Both studies used questionnaires to collect information from patients on travel and time costs at the end of their visit before leaving hospital or GP surgery. The questionnaires asked about travel arrangements for the journey to and from the appointment including the distance travelled, the mode of transport, the time taken and the cost. Costs for travel by car were calculated using published motoring costs (the travel cost was obtained from the journey distance in miles multiplied by an average cost per mile, allowing for fixed costs, depreciation and running costs) and actual return travel costs were used for other modes of transport such as bus or train. In addition, patients were asked about the time spent travelling to and from the appointment, time spent at the clinic/hospital (both waiting time and consultation time) and the activities forgone when attending their appointment. The opportunity cost of time lost from work was estimated from the mean weekly wage rate (minus tax, pension and National Insurance contributions). Other activities such as leisure time were valued at 40% of the mean average wage rate [Bricker et al, 2000]. The study by Bryan et al (1995) also included the travel costs and time costs for companions, assuming that the companion did not have an appointment at the hospital or surgery themselves.

6.3.4 Calculation of patient costs for the sample and the total population

Sample population

In Chapter 2, patient costs were calculated for a sample of women who were sent a questionnaire (including those who were not referred (n = 96) – these results were not presented in Chapter 2). In order to look at the patient costs for all patients who were observed during the study period the following assumptions and calculations were made. It was assumed that a patient would incur a travel cost for a return journey to the district hospital for ultrasound scans (anomaly scan and any further scans later in pregnancy i.e. growth or fetal wellbeing scans), and any antenatal or outpatient visits; and to the specialist hospital for a specialist scan (including counselling where appropriate which was costed as one visit). Travel costs for terminations, prenatal admissions and labour admissions were excluded, as majority of these visits would have been classed as an emergency and not all women would have travelled by car or public transport. Travel costs for telemedicine consultations were also excluded as anomaly scans were videoed and women would not have had to travel again to the hospital for the telemedicine consultation.

For those patients who filled in a questionnaire and indicated which mode of transport they used, this was assumed to be their method of transport throughout their pregnancy. Reflecting the proportions in the TelePaed data, 80% of the women were randomly allocated to travelling by car and the remainder by public transport. Using postcode data from the AA [AA website, 2008a], this source was used to calculate the road distance (in miles) between the home address and the district or specialist hospital. A cost per mile of travel of 44 pence was applied to each mile travelled [AA website, 2008b]. For each hospital visit, the appropriate car parking charge was also added (see next paragraph). For those who travelled by public transport, a mean value was used from the TelePaed data sample, depending on whether the visit was to the district or specialist hospital. All travel costs are in 2005/2006 prices.

Based on the TelePaed data sample, on average women who went to London for an appointment spent 99 minutes in hospital and women who went to the district hospital for an appointment spent 75 minutes in hospital (this included time for the actual appointment and also any waiting time). Using this information, an average estimation was made for each woman for their car parking charge for each visit to the hospital. Car parking charges were obtained from the two hospitals [Medway hospital website, 2008; Personal communication with receptionist at Queen Charlottes Hospital, December 2008] in 2008 prices and were deflated back to 2005/2006 prices [Curtis, 2008].

Loss of pay was also calculated for each woman based on an hourly wage rate for women in 2005/2006 prices as £7.76 (inflated from 2001/2002 prices) using the New Earnings Survey for 2002 [ONS, 2002]. Based on the TelePaed data sample, 40% of the women were randomly allocated to paid employment and the remainder to undertaking household duties or looking after children. For those women who were not in paid employment, their time was valued at 40% of the mean female wage rate [Bricker et al, 2000]. The analysis does not include any additional costs, such as the cost of childcare or care for other dependents for each of these visits, as not enough information was available. Finally, no allowance was made if the patient had a companion accompanying them on their hospital visit.

Total population

To determine patient costs (travel plus additional costs) for the total population, multiple regression models were fitted to the observed caseloads of women for each of the time periods. The costs were adjusted for all risk factors which included: parity; mother's age at anomaly scan; number of fetuses; gestation in weeks at the anomaly scan; and also whether the woman had one of the following risk factors: diabetes; an elevated serum risk; a high risk of Down's syndrome; family history of CHD; or a previous pregnancy with an anomaly. The mean costs were obtained for each time period (and group) and were then allocated to the rest of the women in each group who did not have a patient cost.

6.3.5 Addition of patient costs to sample and total population cost results

Table 6.9 shows the total hospital and patient costs for the observed sample populations for a service with and without telemedicine. Looking at patient costs (this includes patient travel costs and additional costs that patients may have incurred) as expected, the direct referral group incurred more costly journeys than the other risk groups, as these women had to travel to London for a specialist scan(s) (and for some women they also had counselling) for the two time periods when telemedicine was in use. When patient costs were added to the total costs of pregnancy, for the year 2001/2002 with telemedicine, the medium risk group had slightly higher total costs than the other risk groups (again the influence of multiple pregnancies was evident); whereas, for the latter time period when telemedicine was in steady state of use, the telemedicine group had higher total costs than the other three risk groups. For the two scenarios without telemedicine, for 2001/2002 period the medium risk group had higher costs than the direct referral group and for the 2005/2006 period the direct referral group had higher costs than the medium risk group. So on average, during the second

and third trimesters of pregnancy, women would have to pay out approximately £150-£200 out of their own pocket for travel and any other additional costs.

Table 6.9: Bootstrapped total hospital and patient costs for sample population (in £'s in 2005/2006 prices)±

	Telemedicine group	Direct referral group	Medium risk group	Low risk group	Total
2001/2002 period with telemedicine					
Total patient costs (patient travel costs plus additional costs)					
N	52	24	245	87	408
Mean (SD) 95% CI	£199 (£16) £169 to £230	£249 (£26) £197 to £300	£204 (£8) £188 to £219	£163 (£11) £142 to £185	£197 (£6) £186 to £209
Total costs of pregnancy plus patient costs					
Mean (SD) 95% CI	£4,562 (£250) £4,072 to £5,053	£4,073 (£331) £3,424 to £4,723	£4,649 (£118) £4,418 to £4,880	£3,875 (£157) £3,567 to £4,183	£4,439 (£90) £4,264 to £4,615
2001/2002 period without telemedicine					
Total patient costs (patient travel costs plus additional costs)					
N	0	69	252	87	408
Mean (SD) 95% CI	n/a	£233 (£13) £207 to £258	£204 (£8) £188 to £219	£163 (£11) £142 to £185	£199 (£6) £188 to £211
Total costs of pregnancy plus patient costs					
Mean (SD) 95% CI	n/a	£4,287 (£196) £3,904 to £4,671	£4,659 (£121) £4,422 to £4,896	£3,875 (£157) £3,567 to £4,183	£4,429 (£89) £4,254 to £4,604
2005/2006 period with telemedicine					
N	72	29	90	147	338
Total patient costs (patient travel costs plus additional costs)					
Mean (SD) 95% CI	£158 (£10) £139 to £177	£217 (£22) £173 to £260	£172 (£9) £154 to £189	£126 (£5) £117 to £136	£153 (£4) £144 to £162
Total costs of pregnancy plus patient costs					
Mean (SD) 95% CI	£4,209 (£177) £3,862 to £4,555	£3,829 (£243) £3,352 to £4,306	£3,877 (£173) £3,537 to £4,216	£3,400 (£95) £3,214 to £3,587	£3,736 (£76) £3,587 to £3,886
2005/2006 period without telemedicine					
N	0	83	108	147	338
Total patient costs (patient travel costs plus additional costs)					
Mean (SD) 95% CI	n/a	£199 (£11) £177 to £221	£166 (£8) £150 to £181	£126 (£5) £117 to £136	£157 (£5) £148 to £166
Total costs of pregnancy plus patient costs					
Mean (SD) 95% CI	n/a	£4,081 (£161) £3,766 to £4,395	£3,895 (£155) £3,590 to £4,199	£3,400 (£95) £3,214 to £3,587	£3,725 (£76) £3,576 to £3,875

Table 6.10 shows the total costs of pregnancy plus patient costs for the total population. The costs are in the same direction and of similar magnitudes with the results presented in Table 6.8. Overall, a service with telemedicine does not significantly reduce the total costs of pregnancy, even with the addition of patient costs.

Table 6.10: Bootstrapped total hospital and patient costs for total population (in £'s in 2005/2006 prices)±

	2001/2002 period (with telemedicine)	2001/2002 period (without telemedicine)	2005/2006 period (with telemedicine)	2005/2006 period (without telemedicine)
Total costs of pregnancy plus patient costs				
N Mean (SD) 95% CI	5114 £3,862 (£7) £3,847 to £3,877	5114 £3,876 (£7) £3,862 to £3,891	5407 £3,549 (£5) £3,539 to £3,559	5407 £3,551 (£5) £3,541 to £3,562

± Statistical tests conducted: t tests for all costs when comparing two time periods

6.4 Discussion

This chapter set out to determine whether a service with telemedicine is a cost-effective alternative to a service without telemedicine and also to see what the effect is over time. The perspective adopted for the cost analysis was that of the hospital and all costs were presented in pounds sterling (£) in 2005/2006 prices. Patient costs were also presented in a separate analysis. The effectiveness measure was the number of missed cardiac cases.

In terms of costs, for the two periods that used telemedicine compared to the same two periods, when adjusting costs if the telemedicine service was not in use, telemedicine did not add much to the total mean costs of pregnancy for the total population of delivered women. When patient costs were added to the total costs of pregnancy, patient costs did not add much to the overall total mean costs of pregnancy and the overall costs were in the same direction and of similar magnitudes. Finally, in relation to costs, the change over time in antenatal screening protocols, and the fewer antenatal clinic visits, along with fewer multiple pregnancies in the 2005/2006 data collection have meant that the overall costs in the second and third trimesters of pregnancy have fallen slightly.

In terms of effects, when comparing a telemedicine service to a service without telemedicine for both time periods (each time period was assessed separately), there was no change in the effectiveness outcome (the number of missed cases). That is, Specialist A confirmed that none of the telemedicine women (52 women during 2001/2002 and 72 women during 2005/2006) would have had a missed cardiac anomaly if the telemedicine service was not available. So on this basis, a cost-effectiveness analysis was not appropriate and a cost-consequences analysis was sufficient.

However, over time there has been a reduction in the number of missed anomalies amongst the low risk women, partly due to sonographers being more confident in checking fetal heart structures and also due to the additional training they received to carry out more detailed fetal heart examinations. This change in the effectiveness over time may be explained by a number of reasons. Firstly, when telemedicine was introduced in 2001, staff were provided with training, not only on how to operate the telemedicine equipment, but also on how to operate the ultrasound machine more optimally (e.g. the velocity level settings and the Doppler settings to ensure that the videoed images were of an acceptable quality). But, as a prerequisite of using the telemedicine equipment they were also provided with training specifically in checking

the fetal heart structures for cardiac anomalies. Over time, the sonographers have received further training so they are more confident in conducting more detailed examinations to look at the 'extra views: 5 views' (which included the 4-chambers and the position of the vessels ('cross over' views)) and this may have been a factor for the reduction in missed cases. This detailed training in checking the fetal heart structure applied to all women (including low risk women) who had an anomaly scan at 20-22 weeks gestation, and not just to women who were referred to a perinatal cardiologist. Looking at this extra detail in the anomaly scan, isn't an extra burden time wise for the sonographers, as each woman is given a 20 minute slot for their anomaly scan, this is enough time to have a good look at the vessels and the heart [Personal communication with Sonographer A, October 2007]. There may have also been a 'learning curve effect', because over time the sonographers performance in conducting anomaly scans improved the detection rate. This learning curve effect was highlighted in a study by Dumville and colleagues (2006) who looked at different options for female urinary stress incontinence and the authors stated that the less experienced surgeons will have to undergo a learning curve on the first patients they operate and these patients' would be more at risk of post-operative complications.

Secondly, changing practice over time: obstetric staff have become more aware of the telemedicine service. When pregnant women go to the hospital for their booking scan (at approx. 11-13 weeks gestation), doctors are alerted when women are booked whether they are at a 'high risk' of fetal CHD. The hospital has now adopted a protocol for identifying high risk women which is far wider than the criteria set up for the TelePaed project; e.g. the protocol includes women with high body mass index. In addition, since the introduction of nuchal translucency screening in 2007, this has had an impact on the detection rate (as mentioned in Chapter 2, section 2.2, NT has stronger relationship with CHD, than established risk factors). Midwives are pretty good at identifying high risk cases and during the first trimester nuchal scan more high risk women are being identified, so more women are sent for a telemedicine consultation or referred straight away to the specialist if it is urgent. As these women are identified a lot earlier, then plans can be made for them to see a specialist (note that telemedicine is not suitable before 18 weeks gestation). So, telemedicine has helped to achieve a change in practice.

The great majority of costs during pregnancy such as antenatal clinics or ultrasound scans are borne by the health service. When patients attend hospital for consultations they also incur out-of-pocket expenses and time costs. Currently, there is no gold standard in the methods employed to calculate patient costs; however, the

methodologies used to calculate patient costs by various screening studies have been broadly similar.

There are two main methods for estimating productivity costs: 1) human capital approach and 2) friction cost approach. The human capital approach was used in this thesis and has also been used in other screening studies. This method assumes that humans are like machines as they contribute to Gross National Product. When wages paid are equal to the marginal product of labour, this gives an indication of the wage paid [Fox-Rushby and Cairns, 2005]. However, some argue that this method overestimates costs and work is not replaced. The friction cost method is an alternative method which assumes that any output lost is temporary and workers will be replaced [Fox-Rushby and Cairns, 2005]. The thesis did not employ the latter approach as there was not enough information on replacing pregnant women at work and also the thesis focussed on patients' costs and not employers' costs.

One study that estimated patient costs in pregnancy is the study by Henderson and colleagues (2002). The authors estimated resource use and costs associated with antenatal ultrasound screening from both the NHS and women's perspectives. They used questionnaires to assess women's costs of attending hospital for an ultrasound scan (i.e. for one hospital visit). These costs included the opportunity cost of time lost from work (estimated using the gross female weekly wage rate and a mean time away from work); if the patient took annual leave this was estimated at 40% of the mean female wage rate; and if women were not in paid employment, the opportunity cost of them attending the hospital was approximated to cleaning work or informal care. Costs for travel by car were calculated using published motoring costs and actual travel costs were used for other modes of transport. They also costed their companions if they were accompanied on their hospital visit depending on whether they were working or not. Costs to women and their families and friends were estimated £16.59 per scan for one hospital visit (2005/2006 prices - costs inflated from 1998/99 prices) [Curtis, 2007]. The results obtained from the Henderson et al study (2002) are slightly lower than the results obtained from TelePaed study for one hospital visit (£21.69 for both direct referral and telemedicine groups). This can be partly explained by the shorter distances to travel to the hospital: return distance in miles: 14.0 miles [Henderson et al, 2002] vs. 24.8 miles (for both direct referral and telemedicine groups); see Chapter 2 for more details.

The estimates of the cost borne by pregnant women during the second and third trimesters of pregnancy for both the observed samples and the total women delivered involved a number of assumptions:

- Based on the TelePaed sample, it was assumed that 80% of pregnant women travel by car and the remainder (20%) use public transport. It may be that not everyone has a car, and for some pregnant women it may have been more convenient to travel by public transport and/or for some women who live close to the district hospital to actually walk (or maybe cycle?) to the hospital for their appointments, so in this instance for some patients, travel costs may have been overestimated.
- Travel costs were excluded for prenatal admissions, labour stay and terminations and for some of these journeys, patients may have travelled by car to the hospital and also had to pay for car parking for these visits, so in this instance for some patients, travel costs may be underestimated.
- A mean cost for those that have travelled by public transport was used and this may have underestimated the patient costs for these women.
- If the patient was accompanied by a partner or a family member or a friend on each visit, these companions may have also incurred travel costs if they had travelled by public transport.
- With regard to additional costs, for some patients this may have been overestimated. It was assumed that 40% of patients would have been working (based on the TelePaed sample), as the precise number of women who would have been in paid employment were unknown. Of those patients who were paid employment, the following information was unknown: how many received time off with pay, how many had taken annual leave, how many would have had to make time up, how many took time off work with loss of pay and how many went outside work hours. Also, if the woman travelled to the hospital with a companion, their companion may have had to take time off from work with loss of pay.
- With regards to additional costs, in each of the two time periods we had no data on how many women had to pay for someone to care for their children or other dependents whilst they attended hospital; and also whether the pregnant woman (and/or companion) may have experienced other out-of-pocket expenses such as food and drink.

This analysis provides important information about patient costs associated with a woman's pregnancy during the second and third trimesters; and these costs provided here can be used as a guideline. However, to get more comprehensive costs, a full

patient cost survey can be conducted to provide more accurate costs for a sample of patients for each of their visits to the hospital as their circumstances may change throughout pregnancy (i.e. a woman later in her pregnancy may not be able to drive to the hospital for an appointment or may be on maternity leave). This patient survey could also include health status instruments at various time points to find out how the woman's anxiety and depression levels change throughout pregnancy and also for this short time period, some utility data can be collected to enable calculation of QALYs.

One further point to note is that the low risk women have been presented in this Chapter as a baseline; however, they have not really been of interest. Thus, in terms of extending the telemedicine programme to all low risk women (2001/2002 had the most missed cardiac cases; in the same time period, eight of the cardiac cases detected were from women who were in the 'low risk' category – see Table 6.3) and the implications of this will be shown in Chapter 7.

So to summarise the main findings from this chapter: the additional cost of the telemedicine can largely be offset by savings downstream; patient costs did not add much to the total costs of pregnancy; and there were no missed cardiac cases amongst pregnant women who were referred to a perinatal cardiologist. Therefore, the aim of the next chapter is to see what the impact on costs and effects would have been, if telemedicine was offered to all low risk women in order to reduce the number of missed cardiac cases during this same time period. Using decision analytical modelling, a cost-effectiveness analysis comparing a screening programme with telemedicine to a screening programme without telemedicine will be conducted for low risk women. Thus, the next Chapter looks at the third economic issue which has been highlighted throughout this thesis: benefit measures such as QALYs.

CHAPTER 7: MODELLING LIFETIME COSTS AND BENEFITS FOR A TELEMEDICINE SCREENING PROGRAMME USED TO DETECT CHD IN UNBORN CHILDREN

7.1 Introduction

This chapter will look at the third economic issue associated with telemedicine which was highlighted in Chapters 3 and 4: use of benefit measures such as QALYs. There are various health and non-health benefits associated with telemedicine (some of which have been summarised in Chapter 3). For example, some of the health benefits of telemedicine include the possibility of earlier diagnosis which may mean bringing treatment forward in time which may have some benefit. The non-health benefits of telemedicine include the transfer of skills between the specialist and local clinicians and the speed of service. The great majority of studies identified in the literature review in Chapter 3 only looked at the costs, and not at the benefits of telemedicine. Some of the studies which looked at the 'benefits' from telemedicine services classed resource use in terms of transfers, hospitalisations and consultations avoided rather than clinical improvement as a measure of benefit; however, in economic terms these measures are not regarded as benefits, instead they should be classed as 'costs averted'.

The outcomes that are included in an economic evaluation are important as this will in turn determine the type of evaluation which is carried out. Gold et al (1996) suggested the use of the quality-adjusted life year (QALY) as the unit of effectiveness in a cost-effectiveness analysis, as it allows decision makers such as NICE "to compare interventions whose effects on health are qualitatively different, such as prevention of coronary artery disease and treatment of arthritis". This unit of outcome measure has become the norm, as it allows decisions about value for money to be made across different technologies and across different disease areas. In Chapter 3, the thesis looked at what QALYs were and whether they are appropriate for telemedicine. However, in practice it may be difficult to calculate QALYs for telemedicine interventions or services, because the benefits of reduced mortality and morbidity associated with telemedicine are difficult to calculate. That is, QALYs may not be sensitive enough to detect small changes in health outcomes which telemedicine services are most likely to produce.

Chapter 2 compared the costs and effects of two groups of women who were referred directly or via telemedicine to a perinatal cardiologist after a routine anomaly scan for screening of the fetus, or to confirm in suspected cases whether or not there really was a cardiac abnormality. There were no missed cardiac diagnoses among either group.

In Chapter 6, this analysis was extended to look at all women who had undergone an anomaly scan at Medway hospital during the same time period and the Chapter considered conducting a cost-effectiveness analysis comparing a service with telemedicine to one without telemedicine (i.e. 'telemedicine women' were regrouped as direct referral or medium risk women in the without telemedicine scenario). Regarding outcomes for Chapter 6, benefits were measured in terms of 'missed cardiac case avoided'. However, a cost-effectiveness analysis was not appropriate, because no woman in the telemedicine group were assumed to have a 'missed cardiac anomaly' if the service had not been available. Due to this assumption, a cost-consequence analysis was conducted.

In Chapter 6, for the 2001/2002 period we ascertained that the majority of missed cardiac anomalies were amongst the low risk women and there were also two missed cases for the medium risk women. However, for the latter time period there was only one missed case and this woman was in the low risk group. In terms of conducting a cost-effectiveness analysis of a screening programme with telemedicine compared to a screening programme without telemedicine, it would be more appropriate to use the latter data; however, as there was only one missed case it was not feasible. Therefore, the data used in this chapter focuses on the low risk women from the 2001/2002 period. The Chapter does not consider the medium risk women for the 2001/2002 period, because for the 2005/2006 period there were no missed anomalies for this risk group. One further point to note is that as this Chapter considers all low risk women, then there is no issue of 'selection' (see Chapter 5 for more details); hence, costs and effects in this Chapter have not been adjusted for selection bias.

This Chapter explores what the impact on costs and effects would be if telemedicine (this is cheaper and less time consuming in the long-run compared to direct referral) was offered to all low risk women in order to reduce the number of missed cardiac cases. Therefore, by using cost-effectiveness analysis, the main aim of this chapter is to compare a screening programme with telemedicine to a screening programme without telemedicine for low risk women and by comparing these two programmes, the lifetime costs and benefits (QALYs) for children born with and without congenital heart disease can be estimated.

The structure for this chapter is as follows: firstly, to look at various antenatal screening programmes for conditions such as Down's syndrome, to see how they have calculated the longer-term costs and benefits associated with such screening programmes; secondly, to explore the literature on costs and benefits of screening programmes to

help inform the decision model; and in the final part of this chapter, an estimate of the lifetime costs and benefits of introducing a screening service with telemedicine for pregnant women whose unborn babies are at a low risk of CHD compared to a screening service without telemedicine using a decision tree will be provided.

7.2 Background to the longer-term costs and benefits of antenatal screening programmes

Economic analysis of prenatal screening programmes

Literature in the public domain was reviewed on the cost-effectiveness of other antenatal screening programmes to establish what methods had been used, including the calculation of long-term costs and benefits. The three main sets of issues are:

1) what are the averted costs?
2) what are the benefits considered?; and
3) what could otherwise happen?

Over the last thirty years, there have been a growing number of studies looking at the economics of antenatal screening programmes [e.g. Shackley, 1993; Shackley and Cairns, 1993; Karnon et al, 2007]. Antenatal screening programmes focus on the detection of fetal abnormalities, and if a fetal abnormality is detected, a decision can be made as to whether to continue with, or to terminate the pregnancy. In economic terms, if a pregnancy is terminated, cost savings are estimated. However, if the fetal abnormality was not detected through screening and only when the baby is born the anomaly is detected, the costs incurred throughout a child's lifetime to treat such a condition are then estimated. Usually these lifetime costs of a child with an anomaly are more than the cost of replacing an affected pregnancy with a child without the condition [Karnon et al, 1997]. These screening programmes have looked at averted costs such as the extra costs to the health service, the additional education costs and labour market productivity costs. The benefits from such programmes have focused on estimating the future costs which would be avoided by detecting and terminating affected fetuses. Also, most of these studies suggested that if a positive screening result was found, women would elect to terminate their pregnancy. This may not always be the case.

One of the first cost analyses published in this area considered preventing the birth of infants with Down's syndrome and introduced the idea of replacement [Hagard and Carter, 1976]. The authors evaluated the economic benefits (costs averted) resulting from terminating pregnancies affected with Down's syndrome, in both replacement and no replacement situations. Replacement is where the woman becomes pregnant again

after terminating an affected fetus (and the outcome of the new pregnancy is assumed to be normal) and no replacement is where the woman does not become pregnant again after terminating an affected fetus. In terms of the net economic benefit to the community of preventing the birth of a person with Down's syndrome, the costs that arise in both of these situations would then be compared to the resources saved (averted costs) of caring for a person with Down's syndrome (i.e. the cost to the community for their care). The averted costs in this study comprised permanent care costs, education costs and lost maternal income.

Further examples of cost analyses of preventing the births of disabled children followed [Hagard et al, 1976; Henderson, 1982a; Henderson, 1982b; Gill et al, 1997]. In each of the examples, the authors looked at the costs of introducing a programme for the mass screening of pregnancies for the detection and termination of affected fetuses. The issue of replacement was discussed with different assumed rates of replacement and the delay in replacing a pregnancy. Hagard et al (1976) estimated the total costs to the health service of running a mass screening programme for spina bifida for 20 years. They found that spina bifida patients used more resources – medical, educational, social and personal - than patients without spina bifida i.e. 'normal' patients and they calculated the average savings in excess costs of resources used, through births prevented as a result of the programme. However, the authors did not look at any benefits from such screening programmes. Henderson (1982a) conducted an economic appraisal of the costs of a mass screening programme for the prenatal detection of fetuses affected by spina bifida. He compared the net costs to society of the cohort of disabled individuals born in the absence of a screening programme with the cohort of individuals who would be born if the programme was implemented. He estimated the net costs with different assumed rates of replacement for terminated pregnancies (0%, 50%, 100%, 150% and 200%), and the different assumed delays in replacing a pregnancy was between 6 months and 2 years (an estimate of 1 year was used). He used different life expectancies for the disabled and non-disabled children by using age-specific and sex-specific mortality ratios. In a subsequent paper, Henderson (1982b) adopted a similar approach in respect to the costs of screening for open neural tube defects i.e. comparing a cohort of disabled children born if there was no screening programme, with a cohort of 'replacement' non-disabled children born if there was a screening programme. The difference between the two is the cost of the screening programme itself, which has been subtracted to provide an estimate for the expected net benefit [Henderson, 1982b]. The author assumed that: 1) the replacement of pregnancies may be less than 100%; 2) individuals with neural tube defects have a shorter life-expectancy than non-disabled children; and 3) the cohort of

replacement individuals will be born a year later after the termination of an affected fetus, so their costs must be discounted by this extra year. Gill and colleagues (1987) looked at the direct and indirect costs of a screening programme for Down's syndrome taking into account maternal age and serum alpha fetoprotein concentration levels; they did not take into account any benefits associated with the screening programme. The authors looked at the lifetime costs of individuals born with and without Down's syndrome and these costs included: lost parental output and their own output, use of healthcare services and other non-NHS services, education and fostering and adoption costs. The authors also assumed replacement rates for pregnancy after termination of affected fetuses (0%, 50%, and 100%) and they assumed that replacement could occur any time after nine months, although in most cases the authors acknowledged it could be longer.

In summary, from these earlier studies the common themes occurring was the comparison of the costs of a cohort of disabled individuals born in the absence of a screening programme with a cohort of individuals born if a screening programme was implemented. The authors assumed: 1) different rates of replacement pregnancies (anything between 0% and 200%); 2) different delays in replacing a pregnancy (usually an estimate of a year was used); and 3) the outcome of the replacement pregnancy was 'normal'. The authors only looked at the 'averted costs' i.e. the resources used and did not consider any benefits in clinical terms associated with the screening programme.

Antenatal screening programmes using decision analytical models

Next stage was to search for other information available in the public domain on other antenatal screening programmes and how the longer-term costs and benefits were calculated. The earlier studies calculated 'costs' associated with and without screening programmes and did not use decision modelling. The latter studies used decision analytical modelling to see whether a screening programme was cost-effective or not, and the benefits associated with these screening programmes were no longer in terms of averted costs, but instead clinical outcomes such as the number of cases detected and generic utility measures such as QALYs were used. For example, Fletcher et al (1995) used decision analysis to compare different screening policies such as maternal serum testing for Down's syndrome across a broad range of outcome measures including live births with and without Down's syndrome; miscarriages with Down's syndrome; and cases of Down's syndrome detected antenatally. They found that offering serum testing to women of all ages would prevent the birth of approximately one more baby with Down's syndrome per year, than would a screening policy for

women aged 30 years or older. Gilbert et al (2001) conducted a cost-effectiveness analysis of comparing antenatal screening strategies for Down's syndrome. The main outcome measures included the number of babies born with Down's syndrome, miscarriages due to chorionic villus sampling (CVS) or amniocentesis and the additional miscarriages per additional affected live birth prevented by adopting a more effective strategy. They found that compared with no screening, nuchal translucency screening would result in 7.6 fewer babies born with Down's syndrome, at a total cost of £171,000 per 10,000 pregnant women (in 1998 prices). Thus, the additional cost per additional Down's syndrome live birth prevented was £22,000.

Rowley et al (1998) conducted an economic evaluation for prenatal screening for cystic fibrosis carriers using decision analytical modelling. They adopted a societal perspective and considered the impact on the quality of life and earnings of a couple who had a cystic fibrosis child. For each couple, they looked at the outcome of a single pregnancy, unless this pregnancy was terminated due to cystic fibrosis and if the pregnancy was terminated, a replacement pregnancy was also considered (birth outcome for the replacement pregnancy could be with or without cystic fibrosis). Utilities were estimated using the time-trade off method. They found that if a pregnancy was terminated due to cystic fibrosis and was replaced, the marginal cost for prenatal cystic fibrosis carrier screening was estimated to be US$8,290 per QALY (in 1996 prices).

In another study, Harris and colleagues (2004) used cost-utility analysis to compare CVS and amniocentesis versus no invasive testing[10] for women aged 35 and older, or for women who were at high risk of giving birth to an infant with Down's syndrome. The decision model followed women from the 10th week of pregnancy (before any diagnostic testing was undertaken) and then followed the patient through the rest of their pregnancy, including during birth and then throughout the remainder of the woman's life expectancy. The outcomes from diagnostic testing included: a baby born with or without a chromosomal abnormality, miscarriage, termination after positive test results, and whether a future birth occurs after a pregnancy loss. They used data where possible from randomised trials and case registries to populate their model. Preference weights (utilities) for health states were derived using a time-trade off exercise from a large sample of pregnant women (n = 534) from a broad ethnic and socioeconomic mix, aged between 16 and 47 years; where 0 = maternal and fetal death and 1 = perfect health. Where possible, published cost data were used and

[10] One point to be aware of is that there are no invasive risks associated with echocardiography/anomaly ultrasound scans, unlike invasive tests for Down's syndrome such as CVS and amniocentesis.

costs and outcomes were discounted at 3%. They found that compared with no testing, amniocentesis costs less than US$15,000 per QALY gained for women all ages and risk levels (in 2003 prices). CVS was more costly and less effective (i.e. dominated) than amniocentesis for all options.

In relation to congenital heart disease, Odibo and colleagues (2006) wanted to determine whether a policy of universal fetal echocardiography for all pregnant diabetic women was a cost-effective screening tool for congenital heart defects in the second trimester. They used a decision analytical model based on a cohort of 40,000 pregnant diabetic women and compared four strategies: 1) no ultrasound screening; 2) selective fetal echocardiography after abnormal ultrasound results; 3) fetal echocardiography for women with elevated haemoglobin A_{1C} levels; and 4) universal fetal echocardiography for all diabetics. Costs, utilities which were then converted to QALYs and the sensitivity and specificity for each strategy were obtained from a literature search. The authors found that strategy 2 costs less per QALY gained for cardiac defect screening; whereas universal fetal echocardiography was associated with a higher detection rate for cardiac defects, although it was more costly.

However, even though the majority of antenatal screening studies used decision modelling or cost-utility analysis, these type of economic evaluations may not be the most appropriate when it comes to quantifying health benefits. This is because some of the benefits may fall outside the healthcare sector such as process benefits (i.e. speed or reassurance) or it may even be in terms of cost savings to patients. Instead, a cost-consequences analysis may be more relevant. With a cost-consequences analysis, all the different types of outcomes can be taken into account and each of the different outcomes can be presented alongside each other; or perhaps if benefits were to fall outside the healthcare sector, then it maybe more applicable to use a cost-benefit analysis which would look at both costs and outcomes in monetary terms, although this is likely to be unfeasible.

Systematic reviews of antenatal screening programmes

Only one systematic review and critique of studies of economic evaluations of antenatal screening programmes has been published to date. This paper published in 2000 by Petrou and colleagues, which predates some of the studies which used QALYs, identified 566 studies which were published after 1991, of which only 41 were classed as economic evaluations. Fourteen of these 41 studies focused on antenatal screening programmes for Down's syndrome or for routine ultrasonography for the detection of fetal anomalies. The authors found that results varied in terms of the

methodology and reporting of results for the economic evaluations. Of those 14 studies, only three adopted a societal perspective which incorporated direct and indirect costs associated with each screening strategy [see Chapter 3 for more detail on the different types of costs]. Two studies did not provide enough information on costs associated with new screening programmes such as staff training. Two studies on Down's syndrome also looked at the costs attributable to a replacement pregnancy. All 14 studies, if appropriate, discounted future costs into present values; however, five of the 14 studies did not conduct any sensitivity analysis on the key economic parameters. The review also highlighted the narrow definition for benefits which was adopted throughout the literature. For example, five studies used cases detected as the primary outcome measure; another four studies measured outcomes in terms of cases of a particular disorder being prevented by screening and it was assumed that this positively diagnosed fetus would be terminated; and a further two studies looked at outcomes in terms of averted costs.

QALYs for the fetus or the mother?

One of the main issues that arise with measuring benefits for screening programmes is whose QALYs to estimate: 1) those for the pregnant woman; 2) those for the unborn child; 3) or both - the pregnant woman and the unborn child? Shackley and Cairns (1996) raised the issue of whose QALYs are relevant – those of the pregnant woman and/or the unborn child, and suggested that there are potential difficulties in calculating QALYs for both groups of patients. For example, if QALYs are calculated at the point of birth of the child, this would not take into account any QALY losses associated with termination of pregnancy to the mother. However, if QALYs are calculated at some time point during pregnancy i.e. at the anomaly scan, the QALY losses to the mother can be avoided, as aborted fetuses (terminations) are included. The main controversy is whether to include QALYs of the terminated fetus? One argument for including the utility of the pregnant mother into QALY calculations is that some pregnancies do not reach full term (i.e. terminations) and this may still have an impact on the quality of life of the pregnant mother and any future birth that she may have and whether this child has a congenital heart defect or not.

7.3 Modelling lifetime costs and benefits for a telemedicine for a telemedicine screening service

Following on from the literature on antenatal screening programmes, this section aims to calculate the lifetime costs and outcomes for children born with and without CHD when comparing a screening programme with telemedicine to a screening programme without telemedicine using a decision-analytical model. The main aim is to see

whether telemedicine should be offered to all low risk women who have undergone an ultrasound anomaly scan (this group of women had the highest number of missed cardiac anomalies) and to see whether such a screening strategy with telemedicine is cost-effective.

A cost-effectiveness analysis was undertaken and a UK NHS perspective was adopted. This Chapter does not consider patient costs as it was ascertained in Chapter 6 that patient costs did not significantly add to the total costs of pregnancy. Evidence on outcomes and costs were based on patient level data from Chapter 6. Extrapolation beyond the end of the study (just after delivery) was carried out for the lifetime for the children who were born with and without CHD. Expert opinion and data from published sources was used to populate the decision model. The main outcome was QALYs and results were presented as cost per QALY. Future costs and benefits were discounted at an annual rate of 3.5% [HM Treasury, 2003]. All costs are presented in UK pounds sterling in 2005/2006 prices. Various one-way sensitivity analyses were conducted to compare the differences in costs and outcomes for a screening programme with telemedicine compared to a screening programme without telemedicine.

7.3.1 Literature review on costs and benefits and other information for the CHD decision model

This section will provide a brief overview of some of the literature which were available to help populate the decision model in terms of resource use, costs and utilities.

Resource use and unit cost data

A review of the literature found that there was no comprehensive information on the resource use over the lifetime of children who were born with different types of CHD. Mackie and colleagues (2007) quoted “the number of adults with congenital heart disease is increasing. However, rates of health care resource utilization in this population are unknown”; likewise, Moons and colleagues (2001) said that “information on utilization of resources in adults with congenital heart disease is scarce”.

Various studies (see examples below) have looked at resource use for CHD patients of different ages, but not resources used over the lifetime of CHD patients. In addition some studies with the different age groups, including the various stages when resources may have been incurred may not be appropriately applied in the analysis of prenatal screening. Some studies have looked at resource use for CHD patients in different countries, but it is quite hard to generalise to the UK setting the resources

used from other countries. Also, any earlier data would not really be applicable now, as paediatric cardiology over time has changed. For example:

- Garson and colleagues (1994) assessed the cost of CHD among six centres in the United States. They looked at patients from birth to 21 years and from 22 to 40 years of age. Patients were split into three CHD categories: benign disease (mild), acyanotic disease (which includes atrial septal defect and ventricular septal defect) and cyanotic disease (which includes tetralogy of Fallot and hypoplastic left heart syndrome). For each centre, the authors estimated the percentage of patients that fell into each CHD category and for each CHD category the number of clinic visits, hospitalisations and number of years of drug use. The authors said that this data had to be estimated for two reasons: "such detailed data were not available on patients for the last 21 years and the practice of pediatric cardiology has changed markedly during the last 21 years". The paper did not provide a breakdown of resource utilisation for each of the CHD sub-categories. They found that for patients with cyanotic disease, their cost was almost double that of patients who had acyanotic disease (in 1992 prices).
- Garson et al (1996) in a follow-up paper looked at the variations in cost of CHD and clinical practice across nine countries. Five 'typical' patients, each with a different type of CHD were presented to each clinician at each centre for them to estimate the likely course of health care resource utilisation for patients from birth to 21 years of age. Again, this paper did not provide a comprehensive breakdown of resource utilisation for each of the CHD sub-categories.
- Mackie et al (2007) looked at the health care resource utilisation in adults with CHD from 1996 to 2000 in Quebec, Canada. A five-year time period was chosen to capture resource use as patients with mild CHD may receive infrequent health care. They measured the impact of severity of CHD on the use of health care resources using multivariate models to adjust for age, gender, Charlson co-morbidity score and the duration of follow-up. The authors found (as expected) that patients with severe CHD had higher adjusted rates of outpatient visits to cardiologists, emergency department visits, days in hospital and days in critical care units than patients with other congenital cardiac lesions.
- More recently, Knowles et al (2005) developed a decision analytical model based on 100,000 live births to determine the most cost-effective newborn screening strategy for congenital heart defects. The three strategies included: 1) clinical examination alone; 2) clinical examination with pulse oximetry; and 3) clinical examination with screening echocardiography; in making a timely diagnosis before the infant develops life-threatening symptoms of cardiovascular collapse or before death. The authors only provided the costs of screening such as staff and

equipment costs; they did not provide information on other resource use such as the number of bed days a neonate stayed in the hospital or the number of outpatient visits they may have had. The authors concluded that the addition of pulse oximetry to clinical examination is likely to be a cost-effective strategy, whereas screening newborns by echocardiography is unlikely to be cost-effective.

Health State Utilities

A review of the literature found that there was also no comprehensive information on the health state utilities of patients who were born with different types of CHD; instead the health state utilities focused mainly on patients with cardiovascular diseases such as heart failure or coronary heart disease.

In relation to CHD, the studies identified through the literature search reported information on quality of life but not on health state utilities. For example, Latal et al (2009) published a systematic review on the quality of life in children and adolescents following open-heart surgery for CHD. They conducted a review because of the little evidence that existed regarding the long-term health-related quality of life in children with CHD who required open-heart surgery. The literature search looked at studies which were published between 1990 and 2008 and found only 12 studies which focused on the quality of life. Of these 12 studies, 7 various health questionnaires were used to measure quality of life such as the Child Health Questionnaire, PedsQL and the Pediatric Cardiac Quality of Life Inventory; however, none of these instruments are used to estimate or calculate health state utilities. There has been an attempt to develop a new measure of quality of life for children and adolescents with CHD: the ConQol [Macran et al, 2006]. In the paper, Macran and colleagues talk about the process they went through in constructing the questionnaire, the piloting and the development of a weighted scoring system. The questionnaire was used on a sample size of 640 people who were recruited from six regional cardiology centres in the UK. The ConQol has two versions, one designed for use for children aged 8 to 11 years, and the other for adolescents aged 12 to 16 years. However, the ConQol just provides an index score and cannot be used to calculate health state utilities.

The majority of economic evaluations conducted alongside clinical trials for cardiovascular disease (CVD) used health-related quality of life measures alongside clinical outcomes to measure the health status of individual patients. One of the most common measures which has been used to assess the stage of heart failure is the New York Heart Association (NYHA) functional classification system [NYHA Committee, 1994]. This classification system relates symptoms to everyday activities

and the patient's quality of life and has been mainly used for older children and adults with heart failure. Heart failure was categorised into four categories: Class I and II relating to mild heart failure; Class III relating to moderate heart failure; and Class IV relating to severe heart failure. There are generic measures which are routinely being used to measure health-related quality of life in CVD patients such as the EQ-5D. There are also disease-specific instruments such as: the Seattle Angina Questionnaire [Spertus et al, 1995] which is a self-administered 19-item questionnaire measuring five dimensions of coronary artery disease (physical limitation, anginal stability, angina frequency, treatment satisfaction and disease perception); and the MacNew Heart Disease health-related quality of life instrument [Höfer et al, 2004], which is a self-administered questionnaire consisting of 27 items that fall into three domains (physical limitations, emotional function and social function).

A recently published study conducted a structured literature search using keywords relating to CVD and the EQ-5D [Dyer et al, 2010]. The authors identified 147 papers, of which 66 of these met their selection criteria for further review. Sixty studies reported EQ-5D scores (VAS or self-classification) and 10 studies presented evidence on the validity or reliability of the EQ-5D. Overall, the authors concluded that "the published evidence generally supports the validity and reliability of the EQ-5D as an outcome measure within the cardiovascular area" [Dyer et al, 2010]. Even though there was some variation in the results reported, they thought this was due to the differences in CVD in terms of disease stage and patient characteristics. They thought their review provided useful utility estimates across a range of CVD subgroups and would be useful for future modelling of utilities and QALYs in economic evaluations of CVD.

Several studies have attempted to calculate health state utilities for CHD patients who also had other underlying health conditions and the health state utilities from these studies (along with other studies) will be used in the decision model.

<u>a) During pregnancy:</u>

For QALY calculations during pregnancy, the model will consider the mother's utility until birth and when the child is born, the child's utility will be measured (a combined mother and child outcome). This has been the standard norm when looking at mother and child outcomes during pregnancy and after birth.

- Kupperman et al (1999) conducted a cross-sectional study of 72 women aged 35 or older on how they value the outcomes of two prenatal diagnostic tests: CVS and amniocentesis. The authors said that the two tests have different miscarriage risks and in deciding which test to use, women should be aware of the short-term (e.g.

pregnancy loss) and long-term consequences (e.g. whether a termination of pregnancy is followed by a future birth). Preferences for outcomes (utilities) of testing were measured using the standard gamble approach. The authors found no difference in mean utilities assigned to first versus second trimester pregnancy losses. However, utilities for pregnancy losses followed by future birth were higher than for utilities without a future birth [This study is linked to the Harris et al (2004) paper].

b) From birth onwards:

- Caviness et al (2004) looked at the use of bacterial endocarditis prophylaxis for children aged 0 to 24 months who have cardiac lesions and are about to undergo urinary catheterization. Outcomes were based on bacterial endocarditis incidence and QALYs. QALYS were calculated using the Years of Healthy Life measure and the average life expectancy for congestive heart failure. Probabilities were derived from the medical literature and costs were obtained from local and national sources. The authors concluded that the use of bacterial endocarditis prophylaxis for patients with cardiac lesions was not a cost-effective use of resources.
- Brown and colleagues (2009) conducted an economic evaluation of extracorporeal membrane oxygenation (ECMO) as a bridge to transplant for 75 children with end-stage heart failure due to dilated cardiomyopathy. An expert panel established QALY weights for health states using the HUI mark II measure. Results were expressed in terms of cost per QALY. The authors found that bridging was effective but expensive.
- Yount and Mahle (2004) conducted an economic analysis of the use of palivizumab in infants with CHD. Palivizumab is used in the prevention of respiratory syncytial virus infections and is recommended for infants that have CHD. The authors said that "utility data (to evaluate quality of life) in children and adults with CHD is lacking". Therefore, the authors extrapolated utility data from adults with congestive heart failure to the CHD population.
- Kirsch and McGuire (2000) asked 64 respondents aged between 26 and 65 years to provide health state valuations for the four different NYHA classifications of disease progression using the EQ-5D. The authors found consistent mappings between the disease classification and the EQ-5D.

Survival probabilities and life expectancy

The majority of studies looking at the different types of CHD did not provide information on patients' life expectancy. Many studies were more than five years old and survival rates for CHD surgery and the life expectancies for the different types of CHD may now be different.

Knowles et al (2005) conducted a systematic review of the literature and found 104 papers reporting actuarial survival rates for the different types of CHD and this was reported along with the main complications and causes of death for each type of study. For example, for a CHD patient with hypoplastic left heart syndrome survival rates ranged from 33% to 59% (1 year) and for patients with tetralogy of Fallot, survival rates ranged from 75% to 98% (20 years).

Other studies identified which were particularly relevant; although they did not provide comprehensive information, included:

- Williams et al (2000) examined the survival and quality of life of 106 children with CHD who had undergone surgery for hypoplastic left heart syndrome (surgery consists of three stages) between 1990 and 1999. Median age for surgery stages 1, 2 and 3 were 6 days, 9 months and 34 months respectively. They found that the 1-year and 5-year actuarial survival rate using the Kaplan Meier method was 58% and 54% respectively. The authors found these rates comparable to other hypoplastic left heart syndrome survival rates after surgery. The authors quoted survival rates but did not say anything about life expectancy of these CHD patients.
- Walker and colleagues (2002) wanted to look at the survival and quality of life of CHD patients who had undergone repair of tetralogy of Fallot. They quoted that survival after repair of tetralogy of Fallot at 20 years was reported ranging from 84% to 93.7% and with survival at 25 years being 90.9%.
- Mahle et al (2005) conducted a cost-utility analysis for salvage cardiac ECMO therapy in children with CHD. The quality of life status of survivors was determined with the HUI mark II. The authors assumed that life expectancy for children with single-ventricle heart lesions was assumed to be 40 years. The authors found that salvage cardiac ECMO results in reasonable survival and was cost effective (<US$25,000 per QALY saved).
- Garson et al (1994) calculated the expected lifetime of patients with CHD as the age by which 50% of subjects would be expected to die. The expected lifetimes for acyanotic patients was 64.9 years and for cyanotic patients was 49.0 years.

Diagnostic accuracy of telemedicine for CHD

A literature search was undertaken to identify studies which have looked at the diagnostic accuracy of telemedicine in fetal cardiology. However, no studies looked at this issue in detail for pregnant women; instead the diagnostic accuracy focused on fetal echocardiography screening. For example, Stümpflen et al (1996) assessed the prenatal detection of CHD by detailed fetal echocardiography in an unselected,

consecutive group of 3,085 pregnant women in Austria; of which 540 women had maternal risk factors for CHD. Forty-six cases of CHD were detected prenatally and 6 cases had not been detected prenatally. There were no false positives and the diagnostic accuracy of fetal echocardiography was: sensitivity – 85.5% and specificity – 100.0%.

However, a study by Grant et al (2010) seemed the most appropriate (that is, a UK based population cohort) and also used a second opinion to confirm diagnosis (i.e. telemedicine used a screening tool like the TelePaed study) for the information needed for the model. The authors wanted to determine the accuracy of remote diagnosis of CHD in infants by real-time transmission of echocardiographic images via ISDN lines in Northern Ireland. CHD was diagnosed in 84 of the 109 infants (39 infants had major CHD and 45 infants had minor CHD). The initial diagnosis by the paediatrician in the DGH was accurate in 58% cases (63 cases). However, when the echocardiogram was then transmitted via telemedicine to a paediatric cardiologist accuracy significantly increased (sensitivity 97% and specificity 96%).

Similar specificity rates were found in another study (although this was an older study) which used a low cost telemedicine link to diagnose neonatal CHD [Mulholland et al, 1999]. Echocardiographic images from neonates suspected of having CHD were transmitted by a telemedicine link to a regional paediatric cardiology unit for interpretation by a consultant paediatric cardiologist. Sixty-three patients' echocardiographic images were transmitted. CHD was diagnosed in 42 patients (14 patients with major CHD and a further 28 with less serious CHD). Accuracy of diagnosis was improved to 91.0% when transmitted via telemedicine. The mean sensitivity rate was 90.5% and the mean specificity rate was 97.0%.

7.4 Methods

7.4.1 Patient group

Pregnant women who had undergone an anomaly scan at Medway hospital from May 2001 to July 2002 were classified into four categories by Specialist A, these were based on the classification of congenital heart defects by Knowles et al (2005):

1. *No disability* – a fetus or baby who is classified as 'normal'.
2. *Mild CHD disability* – a fetus or baby with mild CHD such as a small ventricular septal defect, which is not significant.
3. *Moderate CHD disability* – a fetus or baby with a moderate CHD such as tetralogy of Fallot.

4. *Severe CHD disability* - a fetus or baby with a rather complex CHD, such as hypoplastic left heart syndrome.

Knowles et al (2005) in their report referred to mild CHD as clinically non-significant; moderate CHD as clinically significant; and severe CHD as life threatening.

7.4.2 Model structure and assumptions

Decision analytical modelling represents the various clinical pathways for alternative treatments and quantifies the probability of a patient following each pathway. For each pathway, the range of possible costs and health-related outcomes can be calculated. A decision tree model was considered to be the most appropriate model for this study, given the short-term nature of the decision problem [Fletcher et al, 1995; Gekas et al, 2009] and was used to compare the different options. The starting point of the model is from the time of the anomaly scan; hence lifetime costs and outcomes for children born with and without CHD were calculated from this point.

Figure 7.1 shows the overall structure for the decision tree for all low risk women. There are two main options:

1) After the anomaly scan is conducted no low risk woman is offered telemedicine (TM). If no anomaly is suspected, the pregnant woman's ultrasound scan is not further reviewed by a specialist and any antenatal care for the patient is provided in the DGH. However, if an anomaly is suspected then a patient is seen by a specialist (direct referral); and
2) After the anomaly scan each patient's scan is further reviewed by a specialist using a store-and-forward approach via the telemedicine service (i.e. this is not a selective use of telemedicine and telemedicine was used for further clarification, that these pregnant women are not carrying a fetus with CHD) and any antenatal care for the patient is provided in the DGH.

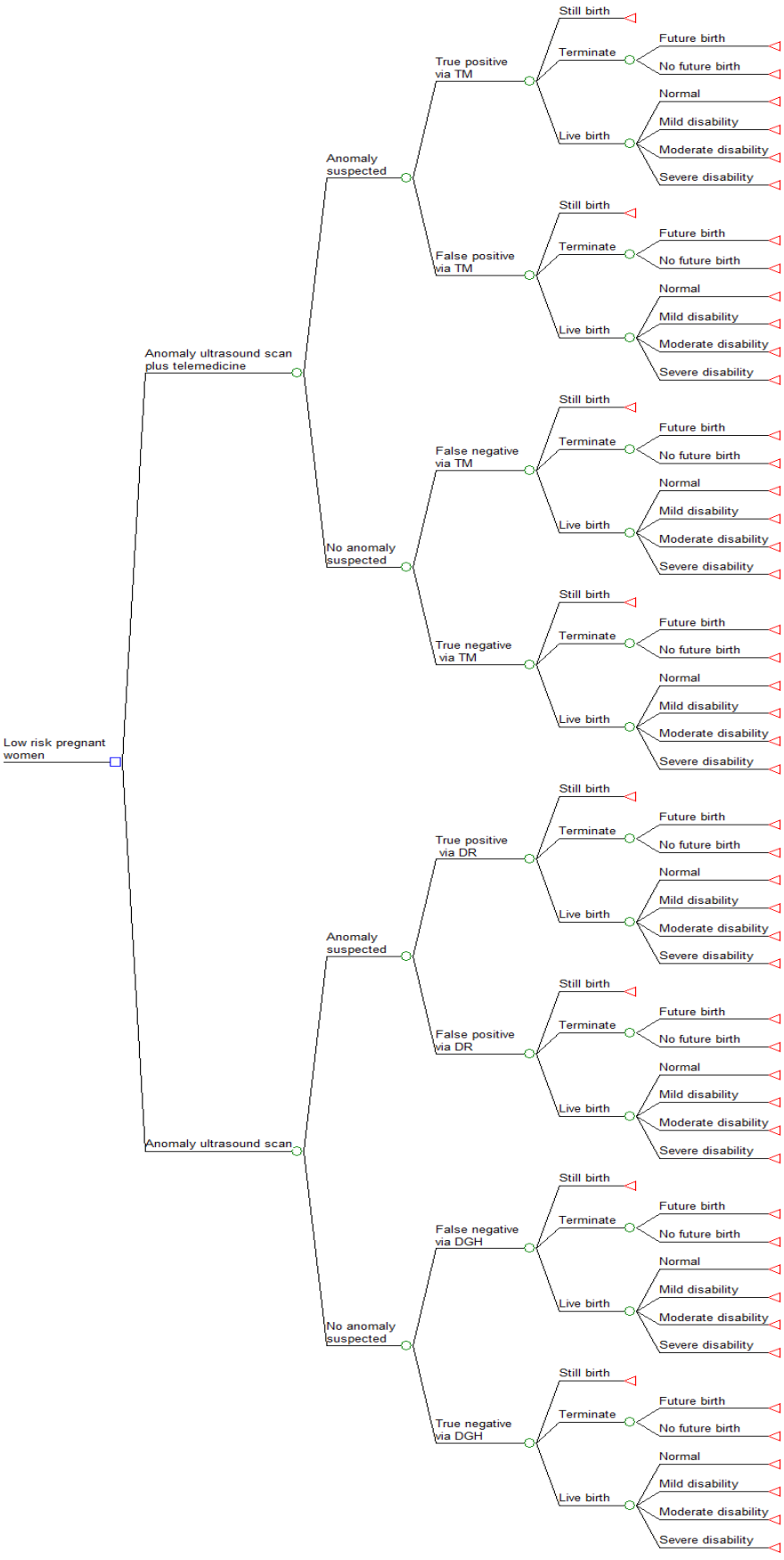


Figure 7.1: Decision tree

No telemedicine for low risk women

Assuming that telemedicine was not offered to any women, at the anomaly ultrasound scan, sonographers may detect a cardiac anomaly. If an anomaly was found, then a woman is seen by a specialist (direct referral) and can then choose to terminate or to continue with the pregnancy. If they continue with the pregnancy, the baby can be born with a heart defect or in a few instances, the defect may have repaired itself so the child would be classified as 'normal'. If no anomaly was found at the scan, the woman would continue with the pregnancy. If they continue with the pregnancy, the baby can be born 'normal' i.e. no heart defect or with a heart defect – in this instance this would be recorded as a 'missed cardiac anomaly'.

Telemedicine for all low risk women

Assuming that telemedicine is offered to all low risk women, telemedicine can then be used as screening tool to confirm whether the patient had a cardiac anomaly or not. If the woman had a fetus with a cardiac anomaly, plans can be made whether to terminate or to continue with the pregnancy.

If telemedicine is offered to all low risk women, there would have to be a team of specially trained staff e.g. sonographers at the specialist hospital to review the anomaly scans that would come through in the telemedicine store-and-forward sessions. For example, if a specialist was on average able to assess 12 women during a telemedicine session which lasted approximately an hour, and assuming that person works 37.5 hours a week, for one person to assess 4,786 low risk women (total number of low risk maternities during this period) by telemedicine this would take up just under 11 working weeks of the person's time over 15-months. Furthermore, extra time would also have to be allocated to the district hospital sonographers for sending all the pre-recorded videoed anomaly scans by telemedicine.

Missed cardiac cases

In Chapter 2, it was established when looking at the referred women (direct referral and telemedicine women) that there were no missed cardiac cases i.e. all cases that were screened by telemedicine or seen by a specialist face-to-face were either a 'true positive' or a 'true negative'. This may be partly due to the sonographers obtaining second opinions for these women i.e. by screening of the fetal heart and to confirm any heart abnormalities. All cardiac cases were detected before birth for these two groups and there was a slightly higher prevalence rate in the direct referral group compared to the telemedicine group, as this group contained the more 'selected' high risk women

who needed a second opinion straightway and any delay in waiting for a 'monthly' scheduled telemedicine clinic was not considered to be appropriate.

However, if telemedicine was used in routine practice and because of the higher throughput, then sensitivity and specificity may actually fall (i.e. there may be a change in the performance levels); on this basis, one would consider that telemedicine would no longer be 100% sensitive and 100% specific. Instead, it was assumed telemedicine has a 97% sensitivity and 96% specificity rate [Grant et al, 2010].

Replacement births

If a woman had a confirmation of a cardiac anomaly and chose to terminate her pregnancy, would she choose to replace this aborted pregnancy with another pregnancy? If a pregnancy is replaced, for simplicity it has been assumed that the outcome of this future pregnancy is normal [Hagard and Carter, 1976; Henderson, 1982a; Gill et al, 1987]. As seen from the earlier papers, there were different rates for replacement of terminated pregnancies and there were also delays in replacing a pregnancy. For these analyses, a 50% replacement of terminated pregnancies[11] has been assumed [Henderson, 1982a; Gill et al, 1987] and the delay in replacing a pregnancy would be one year [Henderson, 1982a; Henderson, 1982b].

7.4.3 Base-case analysis

1. No woman receives telemedicine after the second trimester anomaly scan (*no telemedicine*).
2. All women receive telemedicine after the second trimester anomaly scan and 50% of the affected fetuses detected prenatally are terminated[12] and there is a 50% replacement of all terminated cases (*all with telemedicine*).

7.4.4 Model probabilities

Table 7.1 below shows the model probabilities based on the actual numbers of women who followed each pathway.

[11] A 50% replacement rate was chosen for terminated pregnancies as this seemed much more plausible. In the sensitivity analysis, two other scenarios will be explored: no replacement of terminated cases and 100% replacement of terminated cases.

[12] For the base-case analysis, the termination rate is based on the number of direct referrals and telemedicine women who detected an anomaly prenatally and then subsequently went on to terminate the pregnancy (see Chapter 2 for more details).

Table 7.1: Model probabilities

	Number/Event rate	Base-case probability
Total number of women	4,786	-
Total number of anomalies	8	-
No telemedicine		
Probability of detecting an anomaly	3	0.375
Probability of missing an anomaly	5	0.625
*All telemedicine**		
Probability of detecting an anomaly	7	0.875
Probability of missing an anomaly	1	0.125

*Assuming telemedicine is 97% sensitive & 96% specific

7.4.5 Life expectancy

Table 7.2 below shows the life expectancy at birth. For babies with no CHD disability, this was based on the average life expectancy of a male and a female combined [ONS, 2008]. For babies with a mild CHD, moderate CHD or severe CHD disability these values were obtained from expert opinion [Personal communication with Specialist A, September 2009].

Table 7.2: Life expectancy for normal and CHD babies

Variable	Base case value	Source
No disability	78 years	ONS (2008)
Mild CHD disability	78 years	Specialist A
Moderate CHD disability	50 years	Specialist A
Severe CHD disability	30 years	Specialist A

7.4.6 Resource use and unit cost data

Data for resource use was obtained from expert opinion. With the help of Specialist A, another Specialist (Specialist B) who was based at the Adult Congenital Heart Unit at the RBH and also a representative from the Grown Up Congenital Heart Patients Association (GUCH Patient Representative), provided the relevant resource use for an 'average' CHD patient who was classed as either 'mild, 'moderate' or 'severe'. Table 7.3 shows the amount of resources consumed during a lifetime for an 'average' patient who was classed as either 'normal', 'mild, 'moderate' or 'severe'.

A substantial proportion of patients born with CHD are treated in the first few years of life and do not require regular follow-up in adulthood [Moons et al, 2001]. Based on this assumption, the cost of surgery was only included for moderate and severe CHD patients during the neonatal period. For example, for a patient who has tetralogy of the Fallot (moderate CHD) which is one of the most common types of congenital heart defects, where four heart malformations (pulmonary stenosis, overriding aorta, ventricular septal defect, and right ventricular hypertrophy) present together [Knowles et al, 2005], surgical repair is recommended and is conducted when the child is a neonate. The outcome following surgery is very good and there are relatively few further problems [Knowles et al, 2005]. Whereas, for a patient who has hypoplastic left

heart syndrome (severe CHD), surgery is usually conducted in three stages because of the complexity of the defect. Without surgery, hypoplastic left heart syndrome can lead to early heart failure or even death [Knowles et al, 2005]. The cost of surgery was obtained from the NHS reference costs [Department of Health, 2008] (see Table 7.4).

GUCH patient representative and Specialist B provided information on the most common drugs that either a moderate or a severe CHD patient would be prescribed. A moderate CHD patient would be prescribed about 3 drugs (Digoxin, Frusemide and Warfarin) they are most likely to take every other day, whereas, with severe CHD patients, they would be prescribed about 7 drugs (Digoxin, Frusemide, Warfarin, Amioderone, Bisoprolol, Verapamil, and Ramipiril) and most of these drugs are taken daily. So based on these assumptions and using the British National Formulary [BNF, 2005] as a guide to work out the approximate dosages for both children and adults, the yearly cost of drugs intake for an average moderate and an average severe CHD child and adult was estimated. Furthermore, if a patient had a 'missed cardiac anomaly', both Specialist A and Specialist B asserted that any additional resource use for these patients would be incurred during the neonatal period.

Unit costs for the resource items presented in Table 7.3 are shown in Table 7.4 in 2005/2006 prices. Unit costs taken from other financial years were adjusted to 2005/2006 prices using UK Hospital and Community Health Service indices [Curtis, 2008]. For hospital admissions (inpatient stays and outpatient visits) and tests and investigations, these unit costs were obtained from the NHS reference costs [Department of Health, 2006; Department of Health, 2008]. The cost for the cardiologist was obtained from the Unit Costs of Health and Social Care 2006 [Curtis and Netten, 2006] and the cost of the chest x-ray was provided by Medway finance department. Unit costs for drugs were obtained from the British National Formulary [BNF, 2005]. It was assumed that each patient (normal to severe CHD disability) would incur a cost to the NHS i.e. for prescriptions or a visit to the hospital other than for CHD and this would be a recurring annual cost for these patients in the model. These recurring annual costs were obtained from the report on costing NHS care for pre-term birth of neonates to 18 years of age [Mangham et al, 2009] and these costs vary by age. For simplicity, it was assumed that these costs would also remain the same during their adult lifetime.

Table 7.3: Resource use data

Variable	Resource use	Source
During pregnancy		
From time of anomaly scan until just after delivery or after termination of pregnancy	This includes all costs incurred during pregnancy for low risk women who were seen with or without telemedicine	TelePaed study data
Neonatal care (from birth to 1 year)		
No disability	Normal care in neonatal unit (1.7 days)	Mangham et al (2009)
Mild CHD disability	Normal care (1.7 days) & special care (3.7 days) in neonatal unit	Mangham et al (2009)
	Echocardiogram (outpatient)	Specialist B
Mild CHD disability (missed*)	Same as above plus 2 outpatient visits	Assumption
Moderate CHD disability	Normal care (1.7 days), special care (3.7 days) & high dependency care (5.5 days) in neonatal unit	Mangham et al (2009)
	Surgery	Specialist B
	Echocardiogram (inpatient); Drugs (3 drugs taken every other day)	GUCH PR + Specialist B
Moderate CHD disability (missed*)	Same as above plus additional inpatient stay: special care (2.5 days) in neonatal unit	Assumption
Severe CHD disability	Normal care (1.7 days), special care (3.7 days), high dependency care (5.5 days) & intensive care (6 days) in neonatal unit	Mangham et al (2009)
	Surgery	Specialist B
	Echocardiogram (inpatient); Drugs (7 drugs taken daily)	GUCH PR + Specialist B
Severe CHD disability (missed*)	Same as above plus additional inpatient stay: special care (2.5 days) & high dependency care (5 days) in neonatal unit	Assumption
Children and Adult (2 + years)		
Mild CHD disability	Outpatient clinic every 4 years. Each outpatient clinic includes ECG, chest x-ray and transthoracic echo (plus 30 mins of cardiologists time)	Specialist B + GUCH PR
	Inpatient stay (2 day) every 15 years	Specialist B + GUCH PR
Moderate CHD disability	Outpatient clinic every year. Each outpatient clinic includes ECG and chest x-ray (plus 30 mins of cardiologists time)	Specialist B + GUCH PR
	Inpatient stay (2 day) every 3 years	Specialist B + GUCH PR
	Echo (outpatient) every 2 year; Cardiac MRI scan every 3 years	Specialist B
	Exercise test every 3 years; Catheter every 15 years	Specialist B
	Drugs	GUCH patient representative
Severe CHD disability	Bi-annual outpatient clinic. Each outpatient clinic includes ECG and chest x-ray (plus 30 mins of cardiologists time)	Specialist B + GUCH PR
	Inpatient stay (1 day) every 2 years	Specialist B + GUCH PR
	Echo (outpatient) every 2 years; Cardiac MRI scan every 3 years	Specialist B
	Exercise test every 3 years; Catheter every 10 years	Specialist B
	Drugs	GUCH patient representative

GUCH PR = Grown Up Congenital Heart Patient Representative; * Missed during antenatal period

Table 7.4: Unit costs for resource use (in 2005/2006 prices)

Variable	Base case value (standard error*)		Distribution	Source
Neonatal care				
Normal care	£412		Gamma	DH reference costs 2005/2006
Special care	£412		Gamma	DH reference costs 2005/2006
High dependency care	£726		Gamma	DH reference costs 2005/2006
Intensive care	£1,020		Gamma	DH reference costs 2005/2006
Surgery costs[#]				
Cases detected prenatally				
Moderate CHD disability	£8,738		Gamma	DH reference costs 2006/2007 + Specialist B
Severe CHD disability	£24,406		Gamma	DH reference costs 2006/2007 + Specialist B
Missed cases				
Moderate CHD disability	£7,818		Gamma	DH reference costs 2006/2007 + Specialist B
Severe CHD disability	£31,248		Gamma	DH reference costs 2006/2007 + Specialist B
Cost of health states (for all ages)				
Normal disability	£315 (£59)		Gamma	Mangham et al (2009)
Mild CHD disability	£611 (£95)		Gamma	Mangham et al (2009)
Moderate CHD disability	£660 (£121)		Gamma	Mangham et al (2009)
Severe CHD disability	£1,206 (£237)		Gamma	Mangham et al (2009)
Hospital admissions				
Outpatient visits				
Paediatric cardiology follow-up attendance	£173		Gamma	DH reference costs 2006/2007
Adult cardiology follow-up attendance	£97		Gamma	DH reference costs 2006/2007
Consultant cardiology cost (20 mins)	£50		Gamma	Curtis and Netten (2006)
Inpatient admissions				
Congenital disorders (regular admission)	£361		Gamma	DH reference costs 2005/2006
Tests and investigations				
ECG	£26		Gamma	DH reference costs 2006/2007
Chest X-ray	£16		Gamma	Medway Finance Department
Exercise test	£57		Gamma	DH reference costs 2006/2007
MRI scan	£192		Gamma	DH reference costs 2006/2007
Echocardiogram (outpatient)	£117		Gamma	DH reference costs 2006/2007
Echocardiogram (inpatient)	£2,254		Gamma	DH reference costs 2006/2007
Catheter ≤ 18 years	£199		Gamma	DH reference costs 2006/2007
Catheter ≥ 18 years	£329		Gamma	DH reference costs 2006/2007
Drug (actual net prices from BNF)†	*Children (1 to 18 years)*	*Adults (18+ years)*		
Digoxin	£0.70	£1.04		British National Formulary (Sept 2005)
Frusemide	£12.07	£0.62		British National Formulary (Sept 2005)
Warfarin	£1.47	£1.47		British National Formulary (Sept 2005)
Amioderone	£1.33	£2.43		British National Formulary (Sept 2005)
Bisoprolol	£1.68	£1.68		British National Formulary (Sept 2005)
Verapamil	£1.11	£0.62		British National Formulary (Sept 2005)
Ramipiril	£3.74	£3.74		British National Formulary (Sept 2005)

See section on resource use and unit cost data for more information; † These are the actual net prices from the BNF – however, the dosages and strength for each drug is different; * If standard error was not available, it was assumed to be 0.1 of the mean value [Drummond and McGuire, 2001; Mangham and Petrou, 2008]

7.4.7 Health State Utilities

As the utility data were hard to obtain for patients with different types of CHD, utility values from patients with heart failure were used to populate the model [Personal communication with Specialist B, October 2009]. Table 7.5 below shows the categorisation of patients with different types of CHD into the different heart failure categories [Personal communication with Specialists A and B, October 2009].

Table 7.5: Classification of CHD patients into heart failure categories

Class	Patient Symptoms (NYHA)	Categorisation for model
Class I (Mild)	No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, or dyspnea (shortness of breath).	Mild CHD patients These patients have excellent prognosis, as most of these defects decrease in size or close. They also have no activity restrictions (from birth to 20 years of age).
Class II (Mild)	Slight limitation of physical activity. Comfortable at rest, but ordinary physical activity results in fatigue, palpitation, or dyspnea.	Mild CHD patients These patients may also have some activity restrictions (from 21 years of age). Moderate CHD patients These patients who have undergone surgery will have good to excellent cardiac function with some to no exercise intolerance (from birth to 20 years of age).
Class III (Moderate)	Marked limitation of physical activity. Comfortable at rest, but less than ordinary activity causes fatigue, palpitation, or dyspnea.	Moderate CHD patients These patients after 5-20 years of surgery usually have reduced exercise capacity (from 21 years of age).
Class IV (Severe)	Unable to carry out any physical activity without discomfort. Symptoms of cardiac insufficiency at rest. If any physical activity is undertaken, discomfort is increased.	Severe CHD patients These patients have a good chance of survival but will experience chronic problems for the rest of their lives. They may be advised to limit their physical activities to their own endurance.

In order to estimate QALYs the decision model required utility values. Table 7.6 shows the utility values which were used in the base case analysis. Utility values were obtained for five main categories: for the mother during pregnancy, no disability, mild CHD disability, moderate CHD disability and severe CHD disability.

Table 7.6: Utility values

Variable	Base case value (standard error*)	Distribution	Source
During pregnancy for the mother			
No test, unaffected birth	0.918	Beta	Harris et al (2004); Kuppermann et al (1999)
Test, unaffected birth	0.960 (0.0153)	Beta	Kuppermann et al (1999)
Termination (no future birth)	0.840 (0.0318)	Beta	Kuppermann et al (1999)
Termination (future unaffected birth)	0.910 (0.0259)	Beta	Kuppermann et al (1999)
Fetal death (still birth)	0.070	Beta	Odibo et al (2006)
Defects detected prenatally			
Mild cardiac defect	0.900	Beta	Assumption
Moderate cardiac defect	0.700	Beta	Assumption
Severe (major) cardiac defect	0.500	Beta	Odibo et al (2006)
No disability for the child **			
0 to 5 years	0.940	Beta	Erickson et al (1995)
6 to 24 years	0.940	Beta	Kind et al (1999)
25 to 34 years	0.930	Beta	Kind et al (1999)
35 to 44 years	0.910	Beta	Kind et al (1999)
45 to 54 years	0.850	Beta	Kind et al (1999)
55 to 64 years	0.800	Beta	Kind et al (1999)
65 to 74 years	0.780	Beta	Kind et al (1999)
75 + years	0.730	Beta	Kind et al (1999)
Mild CHD disability for the child **			
0 to 25 years	0.850	Beta	Brown et al (2009)
26 to 45 years	0.834 (0.02705)	Beta	Kirsch and McGuire (2000)
46 to 65 years	0.697 (0.03306)	Beta	Kirsch and McGuire (2000)
66 + years	0.697 (0.03306)	Beta	Assumption
Moderate CHD disability for the child **			
0 to 25 years	0.750 (0.03962)	Beta	Yount and Mahle (2004)
26 to 45 years	0.531 (0.06311)	Beta	Kirsch and McGuire (2000)
46 to 65 years	0.488 (0.06170)	Beta	Kirsch and McGuire (2000)
Severe CHD disability for the child **			
0 to 2 years	0.400	Beta	Caviness at al (2004)
3 to 18 years	0.390	Beta	Brown et al (2009)
19 to 25 years	0.390	Beta	Assumption
26 to 45 years	0.323 (0.06505)	Beta	Kirsch and McGuire (2000)

Assumed test = telemedicine

* If standard error was not available, it was assumed to be equal to 0.1 of the mean value [Fox et al, 2007]

** These utility values are based on a mean value for the males and females combined

During pregnancy, for those women whose fetuses were classed as 'normal', the utility values were based on pregnant women who have undergone second trimester test (telemedicine arm) or have not undergone second trimester test (no telemedicine arm); utility values for women who had a test were higher than those women who did not [Kuppermann et al, 1999; Harris et al, 2004] – this may be because the majority of women who had a test were also carrying 'normal' fetuses. Kupperman et al (1999) also provided utility values for women who had undergone a termination of pregnancy with the view of not having a future pregnancy or a having a future pregnancy (where outcome of future birth is normal). Odibo and colleagues (2006) provided utilities for the pregnant mother for fetal death (i.e. if it was a still birth) and also for a major (severe) cardiac defect. Using this as a basis, assumptions were made for moderate and mild CHD defects which were detected prenatally (see Table 7.6).

For babies and children with no disability aged from 0 to 5 years, utility values were obtained from healthy individuals [Erickson et al, 1995] and for children aged 6 years and older, the utility values were based on the UK general population norms [Kind et al, 1999]. For children and adults with some form of CHD disability, as mentioned before, these values were based on heart failure utilities (see Table 7.6 for more detail on the utility values and sources). For example, Yount and Mahle (2004) used utilities from adults with congestive heart failure for the CHD population (infants and children), whereas for patients aged from 26 to 65 years, heart failure values from Kirsch and McGuire (2000) have been used. One point to note is that for mild CHD patients, the utility values from NHYA Class II were used instead of Class I, because Class I utility values were higher than the EQ-5D norm [Kirsch and McGuire, 2000]. Subsequently, these utility values were used in a Health Technology Assessment report by Fox et al (2007) who looked at the clinical and cost-effectiveness of cardiac resynchronisation therapy for people with heart failure. Unfortunately, there were no utility values for mild CHD patients aged above 66 years, so it was assumed this would be the same as the utility values for patients aged between 46 and 65 years. For patients with moderate CHD disability (aged 26 to 65 years) these were based on NYHA Class III [Kirsch and McGuire, 2000] and for severe CHD disability (aged 26 to 45 years) these were based on NYHA Class IV [Kirsch and McGuire, 2000] as they seemed to be most appropriate for this patient population.

7.4.8 Cost-effectiveness analysis

Using a decision analytical model for the base-case analysis, and focusing on the low-risk women, it was assumed that on average each child would live to their expected lifetime (see Table 7.2). For each year of survival for an average normal child and a

CHD child, the appropriate costs were calculated and the relevant utility values were inputted into the model. Cost-effectiveness was measured in terms of the incremental cost per QALY gained. Discount rates of 3.5% were applied to both costs and QALYs [HM Treasury, 2003]. As both cost and effectiveness data were skewed bootstrapping was used, whereby the distribution of cost-effectiveness ratios are generated by repeated sampling of the data (to stabilise the mean), with replacement and, in the absence of any other data from the population, gives a guide to its distribution [Manly, 1997]. Bootstrapping was performed by taking 1,000 iterations of the data and these bootstrapped iterations were plotted along the cost-effectiveness plane.

7.4.9 Sensitivity analyses

Different one-way sensitivity analyses were undertaken to determine the key determinants of cost-effectiveness:

1) *Changes in replacement rates of terminated pregnancies*. In the base-case, a 50% replacement rate for the terminated pregnancies was assumed. For the sensitivity analyses, the assumptions were: a) no replacement; and b) a 100% replacement of all terminated cases.
2) *Changes in unit costs*. In the base-case, the mean reference costs for: neonatal bed days, surgery, inpatient admissions, outpatient visits and tests and investigations (where available) have been used. For one analysis, the lower quartile costs and in the other, the upper quartile costs have been used.
3) *Changes in discount rates*. In the base-case, a value of 3.5% was used to discount costs and outcomes after the first year. For the sensitivity analyses, two further discount rates were explored: 0% and 6%.
4) *Changes in life expectancy*.

   Base-case values are provided in Table 7.2, for the sensitivity analyses:

   a. It was assumed that the life expectancy of a mild CHD child is 75 years.

   b. For children with moderate CHD, it was assumed that they will live to 40 years of age. It has been quoted that 95% of patients die by the age of 40 years [Patient.Co.UK website, 2010].

   c. For children with severe CHD, it was assumed that they will live to 20 years of age. Life expectancy of patients with severe CHD to date is unknown, but there are many patients already in their early 20s enjoying life [Personal Communication with Specialist A, September 2009].
5) *Changes in utility value assumptions*.

   a. For patients with mild CHD disability aged 66+ years, in the base-case it was assumed they have the same utility value as patients aged 46 to 65 years. For this analysis, a utility value of 0.56 for the patients aged 66+ years was used

(the difference in utility value between the 26-45 years old and the 46-65 years old, which has been subtracted from the utility value used in the base-case).

b. For patients with severe CHD disability aged 19 to 25 years old, it was assumed they have the same utility value as those aged 3 to 18 years old. For this analysis, a utility value of 0.3365 was used (the mean of the utility values for 3-18 years old and the 26-45 years old).

6) *Replacement pregnancy not normal*. In the base-case, if a terminated pregnancy was replaced, it was assumed the outcome would be normal. For this analysis, it was assumed that the replacement pregnancy was not normal and each child would be born with moderate CHD.

7) *Utility for pregnant women for telemedicine and no telemedicine is the same*. In the base-case, the utility value of pregnant women with a 'normal' fetus who were assessed via telemedicine was higher than those who were not assessed via telemedicine. In this sensitivity analysis, it was assumed that the utility value is the same.

8) *Changes in specificity and sensitivity rate*[13]. In the base-case, it was assumed that telemedicine had a 97% sensitivity and a 96% specificity rate [Grant et al, 2010]. For this analysis, it was assumed that telemedicine has a 100% sensitivity and specificity rate i.e. that there will be no missed cases (as was the case in Chapter 2).

9) *Changes in the type of missed cases*. For this analysis, the impact on the cost-effectiveness ratio will be explored if all missed cases were severe or moderate or mild (or any other combination).

10) *Replacement of still birth*s. In the base-case, it was assumed that women who had a still birth they would not replace their pregnancy. In this sensitivity analysis, the impact on the cost-effectiveness ratio is explored when women choose to replace a pregnancy following a still birth.

11) *Changes in termination rates*. In the base-case analysis it was assumed that 50% of affected fetuses would be terminated; however, in reality this figure maybe lower. In this sensitivity analysis, termination rates are varied.

12) *CHD prevalence in England and Wales*. Using data from the Office of National Statistics for the total number of births (including still births) in England and Wales in 2005 [ONS, 2008b] and the estimated number of CHD births (1 in every 145 births), along with the incidence of simple and complex CHD births [Petersen et al, 2003] (see Chapter 2 for more details), the cost-effectiveness of a screening

[13] The sensitivity analysis values went up rather than down to reflect the fact that with an additional screen there will be fewer missed anomalies.

programme with telemedicine compared to a screening programme without telemedicine in detecting congenital heart defects is estimated for England and Wales.

13) *Extrapolating to 2005/2006 data*. Using data from 2005/2006 which was presented in Chapter 6, this analysis explores whether offering telemedicine to all low-risk women was a cost-effective option (in this group of women there was only one missed case and no other cardiac cases were detected).

To further explore the impact of joint uncertainty in resource use (and cost) estimates and the utility values, probabilistic sensitivity analyses were conducted to obtain cost-effectiveness acceptability curves (CEACs). For the probabilistic sensitivity analysis, the gamma distribution was used for costs and the beta distribution was used for utility values [Briggs et al, 2006]. Where only a mean value was provided in the literature, an assumption was made for the standard error in order to calculate the alpha and beta values for the probabilistic sensitivity analysis. For example, for utilities the standard error was assumed to be 0.1 of the mean value [Fox et al, 2007] and for the variation in mean cost, a coefficient of variation of 0.1 of the mean value was used to obtain the standard errors [Drummond and McGuire, 2001; Mangham and Petrou, 2008].

7.5 Results

As reported in Chapter 6, the total number of maternities during the period 1st May 2001 to 31st July 2002 was 5,114. Of these 5,114 women, 52 were assessed by telemedicine, 24 were seen by face-to-face assessment in London and for the rest of the women their care was managed in the DGH (252 women were classed 'medium risk' and 4,786 women were classed as 'low risk'). For this analysis, only the 4,786 low risk women were of interest; as this group of women had the highest number of missed cases (5 or 1.04 per 1,000 women - see Chapter 6).

7.5.1 Base-case analysis results

The results from the base case analysis are shown in Tables 7.7a and 7.7b. In Table 7.7a the overall costs for a screening strategy with telemedicine were slightly lower than a screening strategy without telemedicine. The mean number of cases detected via telemedicine was more than without telemedicine and there were fewer missed cardiac cases with telemedicine compared with no telemedicine i.e. telemedicine dominates. Incremental cost-effectiveness ratios (ICER) are not helpful and therefore have not been presented here.

Table 7.7a: Results from the base-case analyses for low risk women only

Mean results per child's lifetime#	Costs per child's lifetime	Mean no. of cardiac cases detected	Mean no. of missed cases
Deterministic results			
No patients receive telemedicine	£11,451	0.00084	0.00104
All patients receive telemedicine and a 50% replacement	£11,425	0.00233	0.00006

Figures have been rounded up

The incremental cost per QALY gained for a screening strategy with telemedicine compared to screening strategy without telemedicine for children born with and without CHD to low risk women are shown in Table 7.7b. Both the deterministic and probabilistic sensitivity analyses provided similar results. The results from the model show that offering telemedicine to all low risk women is the dominant strategy i.e. that the costs are lower and the QALYs are higher (that is, telemedicine is more effective). In the case of replacement, we are assuming that they would have 'normal child QALYs'; hence, the QALYs for replacement are higher than no replacement (see Table 7.8 for more detail).

Table 7.7b: Results from the base-case analyses for low risk women only (using QALYs)

Mean results per child's lifetime	Costs per child's lifetime	QALYs per child's lifetime	ICER*
Deterministic results			
No patients receive telemedicine	£11,451	23.2400	n/a
All patients receive telemedicine and a 50% replacement	£11,425	23.2822	Dominant
Probabilistic results			
No patients receive telemedicine	£11,483	23.2380	n/a
All patients receive telemedicine and a 50% replacement	£11,457	23.2826	Dominant

*ICER = cost per QALY gained

The uncertainty around the mean estimates is demonstrated in Figure 7.2. For the base-case analysis, the majority of observations were in the bottom right quadrant. So in this case, telemedicine is cheaper and more effective.

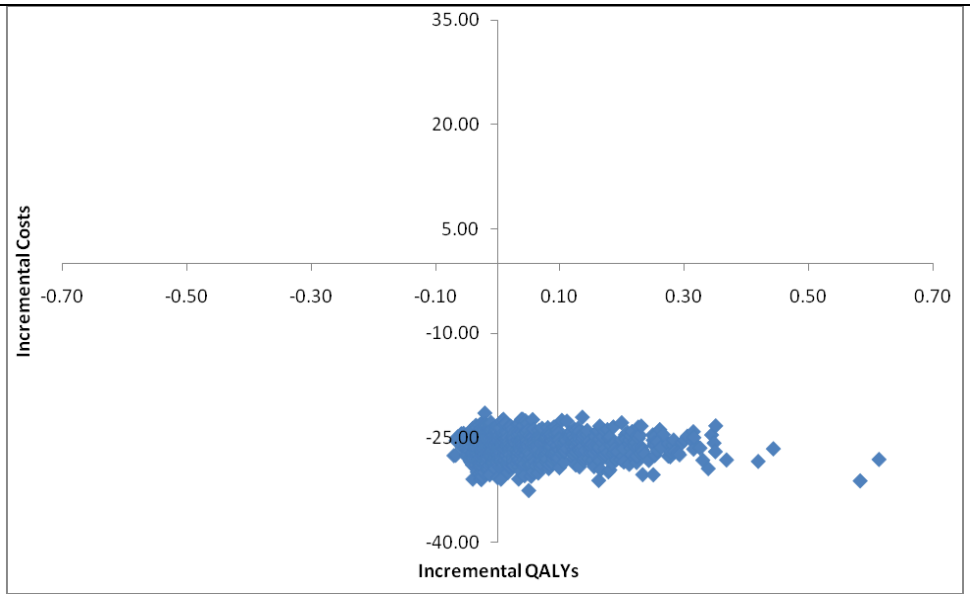


Figure 7.2: Scatterplot of mean bootstrapped ICERs for low risk women only

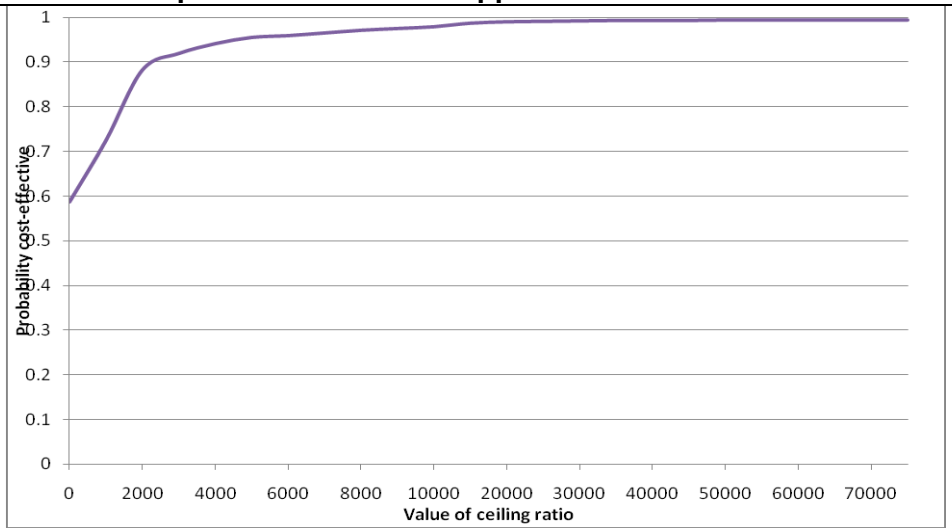


Figure 7.3: Cost-effectiveness acceptability curve for low risk women only

Figure 7.3 shows the CEAC for the base-case analysis and the curve indicates the probability of a screening strategy with telemedicine being more cost-effective than a screening strategy without telemedicine for a range of potential maximum amounts (ceiling ratio) that a decision-maker is willing to pay. For example, if the decision-maker is prepared to pay £2,000 per QALY the probability of a screening strategy with telemedicine being cost-effective is nearly 90%. If the decision-maker is prepared to pay £20,000 per QALY, the probability of a screening strategy with telemedicine being cost-effective is nearly a 100%.

Table 7.8: Results from the sensitivity analyses for low risk women

Deterministic results	Costs per patient lifetime	QALYs per patient lifetime	ICER*
Base-case analysis			
No patients receive telemedicine	£11,451	23.2400	n/a
All patients receive telemedicine and there is a 50% replacement	£11,425	23.2822	Dominant
Changes in replacement rates of terminated pregnancies			
No patients receive telemedicine	£11,451	23.2400	n/a
All patients receive telemedicine and there is no replacement	£11,424	23.2819	Dominant
All patients receive telemedicine and there is a 100% replacement	£11,425	23.2824	Dominant
Changes in unit costs (lower quartile)			
No patients receive telemedicine	£11,324	23.2400	n/a
All patients receive telemedicine and there is a 50% replacement	£11,309	23.2822	Dominant
Changes in unit costs (upper quartile)			
No patients receive telemedicine	£11,548	23.2400	n/a
All patients receive telemedicine and there is a 50% replacement	£11,519	23.2822	Dominant
Changes in discount rates (0% for both costs and outcomes)			
No patients receive telemedicine	£26,613	64.0841	n/a
All patients receive telemedicine and there is a 50% replacement	£26,583	64.1390	Dominant
Changes in discount rates (6% for both costs and outcomes)			
No patients receive telemedicine	£8,490	14.9100	n/a
All patients receive telemedicine and there is a 50% replacement	£8,468	14.9504	Dominant
Changes in life expectancy			
No patients receive telemedicine	£11,447	23.2391	n/a
All patients receive telemedicine and there is a 50% replacement	£11,424	23.2820	Dominant
Changes in utility value assumptions			
No patients receive telemedicine	£11,451	23.2399	n/a
All patients receive telemedicine and there is a 50% replacement	£11,425	23.2821	Dominant
Changes in replacement pregnancy not being normal			
No patients receive telemedicine	£11,451	23.2400	n/a
All patients receive telemedicine and there is a 50% replacement	£11,445	23.2781	Dominant
Utility value is the same for pregnant women with a 'normal' fetus for both telemedicine and no telemedicine arms			
No patients receive telemedicine	£11,451	23.2785	n/a
All patients receive telemedicine and there is a 50% replacement	£11,425	23.2822	Dominant
Changes in telemedicine accuracy (100% for both sensitivity and specificity)			
No patients receive telemedicine	£12,397	25.1718	n/a
All patients receive telemedicine and there is a 50% replacement	£12,373	25.2178	Dominant
Replacement of still births			
No patients receive telemedicine	£11,451	23.2400	n/a
All patients receive telemedicine and there is a 50% replacement	£11,425	23.2833	Dominant

Table 7.8: Results from the sensitivity analyses for low risk women (continued)

Changes in termination rates			
No patients receive telemedicine	£11,451	23.2400	n/a
All patients receive telemedicine and a 50% replacement (25% termination rate)	£11,430	23.2942	Dominant
All patients receive telemedicine and a 50% replacement (75% termination rate)	£11,420	23.2734	Dominant
All patients receive telemedicine and a 50% replacement (33% termination rate)	£11,429	23.2909	Dominant
All patients receive telemedicine and a 50% replacement (66% termination rate)	£11,422	23.2771	Dominant
Changes in CHD prevalence in England and Wales			
No patients receive telemedicine	£11,468	23.1519	n/a
All patients receive telemedicine and there is a 50% replacement	£11,551	23.2412	£923
Using data from 2005/2006 low risk women			
No patients receive telemedicine	£11,189	23.2633	n/a
All patients receive telemedicine and there is a 50% replacement	£11,198	23.3016	£248

7.5.2 Sensitivity analyses results

Several key uncertain parameters within the model have been explored within different sensitivity analyses to assess the robustness of the model (see Table 7.8). Offering telemedicine to all low risk women remained the dominant strategy i.e. cheaper and more effective, as in the base-case analysis for the following: changes in replacement rate of terminated pregnancies, changes in unit costs, changes in discount rates, changes in life expectancies, changes in utility value assumptions, changes in the replacement pregnancy not being normal; utility value for pregnant women for telemedicine and no telemedicine being the same; changes in sensitivity and specificity rate; replacement of still births; and changes in termination rates, implying that the model was robust to the estimates and assumptions which were explored and the new cost and QALY estimates were in the expected direction.

However, a couple of scenarios were not robust to the assumptions explored within the model. Firstly, when looking at the CHD prevalence in England and Wales for the total population, a screening strategy with telemedicine was more costly and more effective, than a screening strategy without telemedicine. However, this is based on the assumption that the additional cost of the telemedicine service would be the same for each hospital, even though this may not be the case as throughput may differ. There may also be other factors such as staff costs and number of clinics for all patient groups (pregnant women, child, adult) which may have an impact on both costs and outcomes and these various factors have not been taken into account. Secondly, using data from 2005/2006, there was only one missed cardiac case amongst the low risk women and all cardiac cases during this time period that had been detected before birth were to women who were in either the direct referral or telemedicine group (even though some of these women may have initially been low risk – see Chapter 6). If telemedicine was offered to all low risk women versus not offering any low risk women telemedicine, telemedicine was slightly more costly and more effective. For low risk women during 2005/2006 period, the costs were lower and the QALYs were higher than 2001/2002 period (cost results are in the same direction as those presented in Chapter 6).

Finally, when looking at changing the type of missed cases (i.e. mild, moderate, and severe) and to see what impact this had on the cost-effectiveness ratio (these results have not been presented in Table 7.8). If all missed cases were severe or moderate, then a screening strategy with telemedicine was still the dominant strategy; however, if the number of mild cases was four or more, then telemedicine was no longer the

dominant strategy – although it was more effective, this additional increase in QALYs came at an extra cost.

7.6 Discussion

This chapter has extended the analysis presented in Chapter 6 by conducting a cost-effectiveness analysis of a screening programme with telemedicine compared to a screening programme without telemedicine, in order to estimate the lifetime costs and QALYs for children born with and without CHD. Thus, the third economic issue: QALYs were explored. For the screening strategy with telemedicine, it was assumed that all low risk pregnant women would have a second opinion via telemedicine and that 50% of affected fetuses detected prenatally would be terminated; and of those terminated pregnancies 50% would be replaced and the delay in replacing a pregnancy was assumed to be a year.

The results from the decision model concluded that offering telemedicine screening to all low risk women was the dominant strategy (cheaper and more effective). For a decision-maker who is willing to pay £20,000 per QALY for a screening strategy with telemedicine, the probability of this strategy being cost-effective is nearly 100%. The results from the various sensitivity analyses concluded that model was robust to the data inputs and assumptions made, and that telemedicine remained the cost-effective strategy.

The main strength from this analysis showed that if telemedicine screening is offered to all women, a telemedicine service can help to reduce the number of missed cardiac cases. If a heart defect is detected in a fetus prenatally (i.e. a true positive result), then decisions can be made as to whether to continue or to terminate the pregnancy, and if parents decide to continue with the pregnancy this allows for earlier treatment. Earlier treatment or intervention can also be associated better quality of life in the future for children born with CHD.

The model created was the most appropriate one using the available data; thus, the data inputs for the model were not perfect. Thus, there are also a few limitations with the model: firstly, the estimates of resource use were based upon expert opinion and the analysis looked at ‘typical’ patients with mild, moderate or severe CHD. However, in practice a CHD patient with the same characteristics, having the same treatment and surgery, may not always have the same outcome. In practice, utilisation of health care resources by patients with CHD disease may vary quite widely. This is due to clinical characteristics, such as cyanosis, type of operation, or occurrence of heart failure,

which are all associated with higher utilisation rates and consequently increased costs [Garson et al, 1994; Moons et al, 2001]. One way of obtaining more accurate resource use estimates may have been to obtain expert opinion via an expert panel or using a Delphi panel method. Unfortunately, due to time constraints with the thesis this was not practical.

Secondly, the analysis presented in this chapter was just restricted to a health service perspective. The Chapter neglects the financial costs effects to patients and their families of attending hospital if they had a disability. For example, if the patient attended hospital say for an outpatient visit, there would not only be travel costs to consider, but other incidental costs and also if the patient is an adult what about the time away from work and the loss of pay? If the patient was a child, what about the loss of pay and time away from work for their parents? In Chapter 6, the addition of patient costs to the total costs of pregnancy did not change the direction of the cost estimates; based on this assumption it is unlikely the direction of the cost estimates would have changed, although the magnitude of cost difference may change. As mentioned in Chapter 6, NICE prefer patient costs not to be included in the cost-effectiveness ratio, especially if the outcome measure is QALYs. Also, the analysis in this Chapter did not take into account the wider effects to families and communities of caring for a disabled child such as carer costs, as these are quite difficult to ascertain; the loss of earnings to parents if they have to give up the work to care for a disabled child, or costs of adaptations to the house (if needed). Thirdly, there may also be some emotional impact, anxiety and burden on families for caring with children and/or adults with disabilities, which has not been taken into account in the cost-effectiveness ratio.

Fourthly, in relation to replacement birth rates, there was no up-to-date information and the older references which have been used in this analysis can be seen as a limitation of the thesis. Fifthly, in terms of utility and expected lifetimes there are also some limitations:

a) Utility values for patients with CHD were based on patients who have heart failure, as utility values were not available for this cohort of patients. More accurate utility estimates are needed for this patient cohort via the use of direct measurements such as the standard gamble or the time trade-off approach;

b) The utility values used in the model have been valued using different instruments: for example, Years of Healthy Life [Caviness et al, 2004], HUI mark II [Brown et al, 2009] and the EQ-5D [Kirsch and McGuire, 2000]. There are various papers that have explored the effects of using different instruments for utility scores and the impact this has on the cost-effectiveness ratio, and generally they have concluded that differences

are evident and caution should be used when comparing health state utility values from different instruments [Brazier et al, 2004; Grieve et al, 2009; McDonough and Grove, 2005]. Bearing this is mind, caution needs to be taken when interpreting the results from this model;

c) It was assumed that each child will live to their expected 'lifetime' according to the CHD status, this may not be the case in practice, as these children can also die of 'other causes' and for those children who have surgery for CHD, some may not survive post-operatively;

d) It was assumed that each child born has a combined utility for a male and a female, and for those who are female they may have a future pregnancy and if this is the case, the prevalence rate of CHD in this population has not been taken into consideration in the model; and

e) Finally, in relation to the health outcomes i.e. whose utility value to use during pregnancy - the mother, the unborn child, or both? There is no consistent approach and it is conceptually difficult to measure and there was not a lot of literature in the public domain about this. For the model, during pregnancy the mother's utility value was used and from birth onwards, the child's utility value was used (a combined mother and child outcome). One thing that that was not taken into account is the disutility for the rest of the mother's life associated with a disabled child. For example, Haberland et al (2002) in their cost-benefit analysis looking at the perinatal screening for Group B Streptococci did not incorporate parental disutilities for losing a child or raising a disabled child as they were difficult to quantify. In any economic evaluation, disutilities are quite difficult to measure, however, they are important.

There are various quality of life instruments as mentioned earlier in the thesis that measure adult's health state utilities such as the EQ-5D, SF-6D and HUI mark 3. It has been argued that we simply cannot use adult utility measures to measure children's utility because of the rapid developmental changes which take place in childhood and adolescence [Petrou, 2003; Griebsch et al, 2005]. Instead, various utility measures have been developed for children, namely, the HUI mark 2 which is used for children 5 years and older as a proxy assessment and 8 years and older for self assessment [Torrance et al, 1996]; and more recently, the EQ-5D-Y (the youth version) for children 8 years and older [Wille et al, 2010]; and the Child Health Utility 9D (CHU-9D) which is used for children between 7 to 11 years for self completion [Stevens, 2010]. However, there is an argument as to whether children can really value their own quality of life and the difficulties that may arise. Griebsch and colleagues (2005) highlighted two main reasons why QALY measurement and valuation is more difficult in children than adults: firstly, children undergo changes in growth and function at different rates and therefore

it is difficult to attribute improvements to health care interventions rather than to normal development; and secondly, young children (under the age of five years) do not have cognitive ability to comprehend and complete valuation or even measurements tasks.

Most screening studies are conducted in the context of a randomised trial and most of them estimate QALYs for adults and not children. Prenatal screening is different as it may not be ethical to conduct a RCT; that is, randomising patients to different screening interventions may have an effect on the health status of the fetus or the mother or both. Prenatal screening is different to other screening programmes for a number of reasons: a) it may be unethical to randomise women to each arm of the trial, because any delay in the consultation, management or treatment will delay the opportunity for antenatal intervention and planned delivery; b) the option to continue with or to terminate a pregnancy, is usually the primary choice after prenatal screening; however, if women wish to continue with their pregnancy, then reassurance is needed for parents of the appropriate treatment or intervention which will be provided immediately after birth; c) if a termination is allowed based on the clinical diagnosis, we would then have to compare the costs and benefits of this life lost with a life which was saved and disabled; d) screening may also raise the number of false positives and in turn, this may increase anxiety in say 90% of the women who are considered to be healthy; e) are the risks of prenatal screening such as using amniocentesis or CVS tests which are considered to be invasive tests, worth the potential benefit?; and finally, f) there are also questions around how people value disabled people in society.

The Chapter did not explore the option of the medium risk women getting a second opinion via telemedicine and only explored the low risk women (this group had the most missed anomalies during the 2001/2002 period). As Chapter 6 had shown that over time the detection of cardiac anomalies at the anomaly scan had improved and for the latter time period (2005/2006) there was only one missed anomaly and this woman was in the low risk group.

In terms of policy implications, is offering telemedicine to all women on a local scale, a good way to identify all congenital heart defects prenatally? This would partly depend on the perspective which is undertaken. For example, if a healthcare perspective is adopted, then both the costs and outcomes to the NHS would have to be considered. Based on the results from the decision model for Medway hospital, a screening strategy with telemedicine is cost-effective for all low risk women. However, when exploring whether the introduction of a telemedicine service for all total births in England and Wales, taking into account CHD prevalence, this analysis was not cost-

effective. Thus, there were a number of issues which were not looked into further such as staff numbers, cost of setting up the telemedicine service in each hospital etc. Therefore, if telemedicine was to be implemented in other hospitals in the area, listed below are some of the factors which should be taken into account:

- Are the clinicians in favour of obtaining second opinions for all women via a store-and-forward telemedicine service?
- How would the introduction of a telemedicine service in one hospital impact on other services such as cancer care provided by the hospital?
- What about the impact on the other services provided by the specialist hospital?
- If more than one DGH set up such screening services, what would be the impact on specialist(s), as there are very few perinatal cardiologists across England and Wales.
- Who would pay for the initial telemedicine equipment and set-up costs?
- Who would pay for the telemedicine equipment maintenance costs and the ongoing costs (i.e. line rental and call charges)?
- Would a telemedicine co-ordinator have to be employed?
- Extra staff would have to be employed at both ends of the telemedicine link to carry out this service. Would extra staff be employed or would current staff have to conduct this extra work in addition to the current duties?
- Providing staff with training in using the telemedicine equipment and providing extra fetal heart training.

A few studies have looked at whether a screening strategy for all women is cost-effective in identifying congenital defects prenatally and results are similar to the model presented in this Chapter; that is, offering telemedicine to all low risk women is cost-effective. Odibo and colleagues (2006) wanted to find out whether all pregnant diabetic women should be offered universal fetal echocardiography as a screening tool for congenital heart defects. Their results showed that option 2 (selective fetal echocardiography after abnormal ultrasound results) was the most cost-effective strategy compared to option 4 (universal fetal echocardiography); although option 4 was associated with a higher detection rate for cardiac defects, it was more expensive. However, they argue that if the "financial implications of missing these cases exceed this net saving (using option 2), then a policy of universal fetal echocardiogram should be continued". The authors found that the ideal policy for screening cardiac defects could be influenced by various factors, for example, if the emphasis was on utilities or on QALYs, then option 2 was preferred; however, if the emphasis of the policy was societal i.e. to prevent cardiac defects, then option 4 was preferred. They also noted that depending on which strategy was seen as cost-effective would depend on the

thresholds that were suggested as reasonable for the cost per QALY. The authors also noted the limitations of their model (some of which mirror the limitations in the model presented in this Chapter): they only took into account direct medical costs; they had difficulty obtaining reliable and generalisable point estimates for the model probabilities; and they found that there was an absence of utility estimates which were specific to cardiac malformations in infants of diabetic mothers.

In the paper by Harris and colleagues (2004), they used a decision model to assess the cost-utility of CVS and amniocentesis compared to no diagnostic testing for pregnant women of all ages and risk levels. The authors found that prenatal testing for chromosomal disorders is cost-effective irrespective of maternal age or risk of carrying an affected fetus and they argued that there was no economic evidence to support the existing guidelines in the USA that recommend offering testing to women above a certain age (35 years or older) or to women of a similar risk as determined by maternal serum screening or ultrasonography, or both. They concluded that universal prenatal diagnostic testing should be offered to all pregnant women on economic grounds.

Finally, Buskens et al (1997) developed a decision analytical model to assess the potential impact of fetal ultrasound screening of pregnant women at low risk for CHD in their unborn child. The authors stated the option of screening for high risk women was not included in the model, as its merits had already been established. The decision model for a Dutch population (1 million pregnancies) took into account the prevalence and history of CHD, characteristics of ultrasound, pregnancy terminations and a literature search was undertaken to obtain probabilities for the model. The results suggested that screening programs may prevent the birth of approximately 1,300 severely affected newborns per million second-trimester pregnancies. However, this meant that over 2,000 terminations of pregnancy would be required, and a further, 9,900 false-positive screening results would occur which required referral. They concluded that the impact of routine screening for CHD was relatively small and further data were needed to fully assess the prenatal screening programmes. Although this study provided some useful information on screening for low risk women, it did not look at the impact on the short-term (during screening) and the long-term (life-time) costs and benefits such as QALYs for children born with and without CHD.

In summary, even though the data used in the model was not perfect, the findings from the model suggest that offering telemedicine to all low risk women is a cost-effective strategy. The next chapter will bring together the three economic issues identified in

this thesis, and the strengths and limitations of the thesis will be discussed. Finally, some concluding remarks and the implications for further research will be provided.

CHAPTER 8: GENERAL DISCUSSION AND CONCLUDING REMARKS

8.1 Introduction

Ever increasing pressures on healthcare budgets have made it necessary to show that healthcare technologies such as telemedicine, not only demonstrate their safety and efficacy, but also to show that they are a cost-effective use of resources. When markets are not 'perfectly competitive' such as within healthcare, economic evaluations can help provide information on whether healthcare technologies are an efficient use of resources by comparing the costs and benefits of one healthcare technology, to the costs and benefits of another healthcare technology. So when healthcare budgets are limited, scarce healthcare resources should be allocated towards those technologies where the incremental benefits outweigh the incremental costs. Thus, this thesis aimed to find out whether the use of telemedicine was a cost-effective use of resources in obtaining specialist advice for fetal cardiology.

Starting from the TelePaed study which was presented in Chapter 2, the thesis looked at how the case study analysis can be improved in the light of the literature review and some of the issues which arose. The primary aim of this thesis was to address these economic issues which came to light from the literature review on costs and benefits of telemedicine and to see how these issues could be addressed in the context of an economic evaluation. These issues were: selection bias; patient costs and benefit measures such as QALYs. These concerns are not specific to telemedicine, but are perceived as a particular problem for telemedicine technologies. This final chapter will summarise: the main contributions of the thesis; the limitations of the thesis; discuss the policy implications of the results; summarise issues for future research; and some concluding comments will be provided.

8.2 Overview of thesis and contributions made to literature

This thesis has aimed to fill a gap in the literature by providing new evidence as to whether telemedicine should be used in a hospital to provide specialist advice in fetal cardiology to all low risk women. Furthermore, the thesis has also addressed some of the economic issues which were associated with telemedicine. Currently, telemedicine is only used routinely for 'high risk' women, but it is within the low risk population where many cases of CHD occur [Simpson, 2009].

The thesis began by looking at the data and results from the TelePaed case study (Chapter 2), that is, comparing the costs (total costs of second and third trimesters of pregnancy) and the effects (detection of cardiac cases before birth) for two groups of

women who were referred for specialist cardiac advice for fetal cardiology (direct referral women who saw a specialist face-to-face in London or telemedicine women whose assessment was conducted in their absence via a store-and-forward telemedicine service). The results suggested that telemedicine was slightly more expensive than direct referral and that the direct referral group had a higher prevalence of cardiac anomalies than the telemedicine group. This was due to the selection and eligibility criteria that the obstetricians and sonographers used at Medway hospital; that is, women with suspected fetal abnormalities were seen mainly by direct referral (unless a telemedicine clinic was scheduled to take place in the next few days). Telemedicine was mainly used for screening purposes. This is important as the telemedicine link reduces the time taken to assess a woman and screen the fetal heart. There were also no missed cardiac cases amongst these two groups of referred women (that is, all cardiac anomalies were detected during the antenatal period).

In Chapter 3, a literature review was conducted to see what existing evidence there was on the costs and benefits of telemedicine. Even though more studies have been added, the results from the literature review were consistent with previous findings [Whitten et al, 2002; Bergmo, 2009] and the conclusions have remained the same. Despite this, the findings from the literature review made a contribution to the literature, as there were some specific economic issues which may be of concern which were identified when conducting economic evaluations of telemedicine. For example, one of the issues which arose was that the majority of studies did not give enough details about the study design or they had weak methodologies and for this reason there was no consistency in the cost analyses across the studies. If studies reported appropriate methodology on how costs (and/or outcomes) are collected, calculated and reported, then it would be easier for the reader to understand how the author(s) came to their results and conclusions. Also, one thing that became apparent from the literature review is that if investigators were more transparent on how they conducted and reported findings on telemedicine interventions this may lead to different conclusions. Transparency would also ensure that all the relevant statistical information are included in each study, so that a meta-analysis on costs (and/or cost-effectiveness) of telemedicine can be conducted.

In Chapter 4, following on from the findings in the literature review, a critique of the TelePaed study and a reflection on some of the problems with the design of the study which was determined prior to my involvement was provided. The design of the study in part influenced the uptake and usage of the telemedicine service and the information which was collected for the economic evaluation. Within this Chapter, the ‘Drummond

et al checklist' [Drummond et al, 2005] was applied to the case study and the analysis which was presented in Chapter 2 was in line with economic evaluation guidelines [Bergmo, 2009]. However, the three economic issues identified in the literature review showed how the study could be improved and these three issues were highlighted in this Chapter: selection bias, repeated measurement of patient costs and measures of benefits such as QALYs.

The next step for the thesis was to address these three economic issues. In Chapter 5, the first economic issue, selection bias was examined. Selection bias existed within the dataset, because pregnant women were selected for referral for further assessment from a perinatal cardiologist due to the nature of their presenting demographic characteristics and risk factors. Thus, the observed costs and effects which were presented in Chapter 2 are said to be biased. Various methods were identified in the literature review to reduce selection bias and these methods were applied to the dataset. However, the analysis cannot conclude which method was the most accurate for this dataset due to the nature and size of the dataset; only some direction on which method may be the best way forward could be provided. Propensity score matching was a more reliable way of obtaining cost and effect estimates, because after matching the groups were similar in terms of background characteristics (i.e. 'balanced'). Both the regression and the Heckman models may not have explicitly balanced the covariates among the groups; therefore the two groups may not be similar. The analysis suggested that both regression analyses and propensity scoring methods (the Heckman selection model was not appropriate for the dataset) when applied to the dataset may have reduced the observed selection bias between the two groups, because after adjustment, the differences between the two groups were smaller, thereby increasing the homogeneity and reducing the variance in the adjusted costs and effects. The majority of studies which have used these methods to reduce selection bias have been conducted on large sample sizes. In Chapter 5, these methods were explored on a small sample size, but it was difficult to tell which method is the most appropriate and if only one method was used, then the results should be interpreted with caution. Chapter 5 has contributed to the literature, because it is the first study, I believe, which has applied these methods for reducing selection bias in a fetal cardiology dataset given a small sample size.

In the previous chapters (Chapters 2, 4 and 5), the thesis showed that the data had some limitations. As mentioned earlier the two groups (referral methods) were not strictly comparable (unless adjustments are made) because of the selection criteria for each referral method. Therefore, the next step in Chapter 6 was not only to look at the

two groups of women who were referred to a perinatal cardiologist, but at all pregnant women (this included women who were not referred to a specialist and were managed in the DGH: both medium risk and low risk women) who had an anomaly scan at Medway hospital during the same time period (May 2001 to July 2002) to see what would happen to costs and effects if the telemedicine service had not been available. By looking at all women during the same time period, we are not artificially introducing 'selection'. It was ascertained that if telemedicine service had not been available, then those women who were in the 'telemedicine group' would have been categorised as either 'direct referral' or 'medium risk' women, and there would have been no 'missed' cardiac cases amongst these women. As there was no difference in the number of missed cases avoided when comparing the same time periods (with and without a telemedicine service), a cost-effectiveness analysis was not appropriate, and results for a cost-consequence study were presented in Chapter 6 and the results showed that a service with telemedicine was slightly cheaper than a service without telemedicine.

In Chapter 6, additional data were presented to look at changes over time. There was no change in the effectiveness data when comparing a service with telemedicine to a service without telemedicine for the latter time period, so again only costs were compared. Changes in antenatal screening protocols, fewer antenatal clinic visits and fewer multiple pregnancies (in the data collection) have meant that the overall costs in the second and third trimesters of pregnancy have fallen slightly over time. In addition Chapter 6 showed that over time (from 2001/2002 to 2005/2006) the number of missed cardiac cases had fallen amongst the low risk women and this was partly due to sonographers being confident in checking fetal heart structures, after receiving additional training to carry out more detailed fetal heart examinations.

Within Chapter 6, the thesis addressed the second economic issue: patient costs. Patient costs are important to include in economic evaluations as this gives an indication to individual patients of the likely costs they will face when attending hospitals for treatment or consultations; this is important if patients have to attend hospital more than once during a specific time period. There is currently no 'gold standard' for estimating patient costs. When patient costs (travel costs and loss of pay) were added to the total costs of pregnancy for all women who were seen during this time period (total population of women delivered), patient costs did not add much to the total costs of pregnancy; that is, the total costs were in the same direction and were of similar magnitudes. Chapter 6 has made a contribution to the literature, as it is the first study to have calculated patient costs for more than one time point during the second and third trimesters of pregnancy.

Chapter 7 first explored how other antenatal screening programmes had calculated the lifetime costs and benefits associated with such programmes (because short-term screening costs are not helpful and longer-term analyses are needed) and literature searches were undertaken to determine the data inputs for the decision model. A hypothetical decision model was developed to address this final economic issue: QALYs. Thus, the cost-effectiveness analysis compared a screening programme with telemedicine to a screening programme without telemedicine for all low risk women and by comparing these two programmes, the lifetime costs and QALYs for children born with and without CHD were estimated. This chapter only looked at low risk women, because in Chapter 6, it was ascertained that the majority of missed cardiac anomalies were amongst the low risk women. Consequently, the model aimed to show that for this cohort, if an additional screen is provided, whether this upfront telemedicine cost is justified. The deterministic results from the decision model suggested that offering telemedicine to all low risk women was the dominant strategy (cheaper and more effective). For a decision-maker who is willing to pay £20,000 per QALY for a screening strategy with telemedicine, the likelihood this strategy being cost-effective is nearly 100%. The model in Chapter 7 was a simple decision analytical model that allowed for extensive sensitivity analyses to be undertaken so that the robustness of the model could be rigorously tested and priorities for future data collection could be determined. Given the available data, the model is the most appropriate. Although, future research priorities can look at refining these data inputs.

Chapter 7 has filled a gap in research and contributed to the literature, because it is the first study which has compared a screening programme with telemedicine to a screening programme without telemedicine in order to calculate the lifetime costs and QALYs for children born with congenital heart disease (mild, moderate, severe) and without congenital heart disease (normal). The Chapter provided new evidence that QALYs can be calculated for telemedicine services.

8.3 Limitations of the thesis

Whilst this thesis has contributed to the existing literature, it also has a number of shortfalls.

To determine whether providing specialist advice via telemedicine compared with direct referral was a cost-effective use of resources, ideally the data for this economic evaluation should have been collected alongside a RCT as they produce the least biased estimates of costs and effects. However, for the TelePaed study it was deemed

unethical to randomise individual patients because of the urgency of cases of suspected abnormalities and any delay in the specialist assessment was not appropriate for these patients. As individual patients could not be randomised (in theory, medium risk women could have been randomised, it was the women with suspected abnormalities which were considered urgent), the use of a cluster RCT may have been a more appropriate study design. This is where different hospitals or units are randomised instead of individual patients, and this type of trial is said to be the least biased. However, there was always a question as to whether the TelePaed study as mentioned in Chapter 4 was set-up a cluster randomised trial; because even by study design two clusters in each arm is questionable, especially when you look at it in terms of statistical power which are required to detect statistical differences in RCTs. Due to time constraints and the slow uptake of the telemedicine service for the TelePaed study, a cluster randomised trial of the project was not feasible and instead an observational study was conducted. Steps were undertaken in Chapter 5 to minimise selection bias between the two referral groups.

Results from the literature review concluded that there is a need for telemedicine studies to be transparent in reporting their methodologies and results. Cost-effectiveness of telemedicine depends not only on the service being evaluated, its comparator, the perspective of the analysis, patient group and sample size, type of economic analysis and how the costs and outcomes were measured and valued, but also on the take-up rate and the usage of the service. Decision makers and readers must be cautious as to the degree to which they can apply the results of such assessments to their own circumstances. Further research needs to be done in which telemedicine interventions and their comparators are conducted in accordance with general standards for health economic evaluations.

The results from the meta-analysis of telemedicine cost studies which was presented in Appendix 3 was for illustration purposes only and was not really informative; as 100 of the 109 studies had insufficient statistical information and could not be included in the meta-analysis. A further limitation in Chapter 3 was that a meta-analysis of costs (and/or cost-effectiveness) of telemedicine studies could not be conducted for each health area separately as again there was not enough statistical information in each of the studies. These differences were primarily due to the different studies reporting variations in patient populations, study design, intervention and comparator groups and not enough appropriate information was reported in each of the studies.

The limitations of the analysis which was presented in Chapter 5 to reduce selection bias between the two referral groups included: 1) all the methods explored were applied to a small sample; however, in practice most of these models are usually conducted on bigger sample sizes; 2) the exclusion of variables such as income and smoking which are routinely observed but were not recorded in the audit dataset, as well as unobserved variables which are not routinely recorded such as a patient's preference for referral method or the quality of care which the patient receives at the hospital; and 3) the Heckman model was not suited to the dataset and this may be partly due to the small number of observations (women) and also partly due to the identifier variable (instrument) which had a low correlation with referral mode. Each of these limitations may have affected the precision of the costs and effects estimates. Nevertheless, there is some confidence in the adjusted costs and effect estimates, as the regression and propensity scoring methods reduced the difference (incremental) in costs and effects between the two groups.

The resource use data in Chapters 2 and 5 was based on the TelePaed project and data were collected prospectively during 2001/2002 for patients who were referred to a perinatal cardiologist. The analysis for this section is based on a small number of referred women (n = 76) and this might increase the level of uncertainty in the cost and effect calculations and limits the generalisation to other patients who may need a specialist opinion. However, in Chapter 6 data were extrapolated from the sample of patients (referred women, medium and low risk women) for whom resource use data was available for, to all maternities over a 15-month period (by looking at the whole population we are not introducing 'selection'). Extrapolating the cost data was undertaken by the use of multiple regression models which were fitted to observed caseloads of women to predict total costs of pregnancy for women for whom there was no resource use data. A mean estimate of the total costs of pregnancy was assigned to each woman who had missing data; this may in some cases have under- or over-estimated the total costs of pregnancy for the cohort (this also applies to the 2005/2006 data, as total costs of pregnancy were extrapolated from the sample to the total population).

In Chapter 6, patient costs were estimated for each woman for each antenatal visit during the second and third trimesters of their pregnancy to see what the total additional out-of-pocket expenses would be and these estimated costs can be used as a guideline. As only a sample of patients provided patient costs for one visit, assumptions were made so that extrapolation of costs could take place beyond this one visit for the remainder of their pregnancy, as well as for the total population for

whom there was no patient cost information for. The assumptions which were made including the mode of travel to hospital (i.e. by car or public transport) and the number of women who may have been in employment during this period, may have led to patient costs being under- or over-estimated. However, to get more comprehensive costs, a full patient cost survey could be conducted for a sample of patients for each of their visits to the hospital as their circumstances may change throughout pregnancy (i.e. in the later months, a woman may not be able to drive to the hospital for an appointment or may be on maternity leave).

There is a large amount of uncertainty in the data inputs used in the decision model in Chapter 7, as it was largely based on assumptions and expert opinion. Firstly, the data assumptions for resources used by children and adults with CHD were based on limited expert opinion. One way of obtaining more accurate resource use estimates may have been to obtain expert opinion via an expert panel or using a Delphi panel method. Secondly, utility values for patients with CHD were based on patients who have heart failure, as directly elicited utility values were not available for this cohort of patients. More accurate utility estimates are needed for this patient cohort (patients with different types of CHD, as well as for those that undergo surgery for a CHD) via the use of direct measures such standard gamble or the time trade-off methods or indirect measures such as the EQ-5D or SF-6D measures. Thirdly, accurate estimates are also needed on the life expectancy of patients with CHD. Ideally, a prospective study estimating lifetime costs and QALYs for patients with different types of CHD is needed.

8.4 Implications of findings for researchers

Given that various studies on telemedicine from the literature review have provided conflicting results about whether telemedicine is cost-effective or not, researchers should strive towards a consistent approach in the information that is collected, reported and analysed for telemedicine evaluation. Telemedicine evaluations need to be more transparent in reporting their methodologies and results, especially in terms of the economics issues which were identified in Chapter 3 namely: study design and methodologies; time frame; study perspective; sample size justification; choice of alternatives; economic importance of question; type of economic evaluation conducted; costs and benefits how they are measured, valued and reported; and whether an incremental approach, discounting and sensitivity analyses are reported. By following guidelines such as the Drummond checklist [Drummond et al, 2005], results from economic evaluations of telemedicine studies will enable readers to make comparisons in their own settings and ensure that results are transferable and generalisable. If

these economic issues are adhered to, then robust economic evaluations of telemedicine services can be reported; and in turn, by having the appropriate statistical information a meta-analysis of telemedicine costs (and/or cost-effectiveness) evidence can be performed.

The literature review also identified a few points which journal editors could bear in mind when assessing an article for publication:

- Research teams don't always have health economists on their team for the costing studies;
- Papers published are too restrictant on word length, therefore all methodology and findings can't be reported; and
- Publication relies on the data analysis being completed and reported, and this may be a few years out of date since the fieldwork was actually completed.

In Chapter 5, after adjusting for selection bias, the results varied greatly between each of the different analytical methods; this means that a researcher cannot simply use the results from one method of correcting for selection bias to represent all methods especially for a dataset of this kind, rather an argument has to be made concerning which method is best to accept. Also, due to the size and nature of this dataset, the bias adjustment still does not provide a clearer picture of how much selection bias remains in the dataset and what the direction of selection bias is. In addition, reviewing the literature from non-randomised telemedicine studies where these studies have not been adjusted for selection bias, the results from these studies should be interpreted for caution. This is because we don't know how much selection bias was in their datasets and what the direction of the bias was.

8.5 Implications of findings for policy makers

The results from the decision model showed that a screening strategy with telemedicine for all low risk women is more likely to be cost-effective than for a screening strategy without telemedicine. In addition, as women are getting a 'second opinion', the number of missed cardiac cases is also likely to fall and if a cardiac anomaly is detected before birth then the appropriate plans can be made by parents as to whether to terminate or to continue with the pregnancy. These findings provide some important implications for policy makers if telemedicine was to become part of routine antenatal screening.

For example as highlighted in Chapter 7, in terms of policy implications, is offering telemedicine to all women on a local scale, a good way to identify all congenital heart

defects prenatally and to minimise the number of missed cardiac cases? If the number of missed cardiac cases is reduced, over the longer-term this will generate savings for both the NHS and patients (the costs of telemedicine are offset by savings downstream). If telemedicine were to be introduced in other hospitals in the area it may have a considerable effect on the organisation of health services (this is especially true for the specialist hospital, where the specialists will have to assess all patients' anomaly scans). That is, an increase in the demand for the service may lead to an increased cost to the provider. Decision makers would then have to decide on the level of telemedicine technology and in how many hospitals it should be implemented. Some factors which should be taken into account include:

- The introduction of telemedicine services may have consequences for resource allocation for other public sector services in the local region;
- What impact would the introduction of a telemedicine service have on other services provided by both hospitals (DGH and the specialist)?
- Would the telemedicine service complement or substitute existing services in a district hospital?
- The telemedicine service should be both feasible and acceptable to clinicians;
- Each hospital would have to decide on what telemedicine service they want: 1) diagnostic work – ad hoc 'live scanning' (which would be more costly and time consuming) or 2) 'store-and-forward' scanning.
- Who pays for the equipment and the maintenance costs, along with the line rental and call charges?
- Would a telemedicine co-ordinator have to be employed to organise the telemedicine activities?
- Would extra staff be employed or would current staff have to carry out this extra work in addition to the current duties? That is, would additional sonographers be needed to send anomaly scans via the telemedicine store-and-forward link, and likewise, perinatal cardiologists are needed to interpret and provide a diagnosis for the anomaly scans.
- Sonographers at the district hospitals would have to be trained properly in conducting and interpreting anomaly scans. There are training implications as there are so few perinatal cardiologists at present.
- In addition, extra training for staff would be required for using the telemedicine equipment.

Karnon et al (2007) provided a critique of methods for model-based cost-utility analysis of screening programmes including antenatal screening programmes. The report provides guidance to policy makers regarding the development of screening

programmes and the guidelines were grouped into seven categories: research question; general modelling approach; model structure; modelling technique; model population; validation and calibration; and issues specific to antenatal screening. The latter point included issues such as termination of affected pregnancies as an option and the complex issues around the estimation of QALYs (these issues were looked at in this thesis).

Finally, Broens et al (2007) conducted a qualitative literature review of 45 conference papers to find out what important determinants had influenced successful telemedicine implementations. They classified the findings into categories: 1) technology - which includes support and training that was given to users and the technical quality of the technology; 2) acceptance – by both clinician and patients; 3) financing – who pays the costs associated with the implementation of technology including the maintenance and operating costs; and 4) organisation – telemedicine may influence the structure of the current healthcare organisation and this may lead to changes with staffing structures.

8.6 Issues for future research

This thesis has highlighted some of the economic issues which surround telemedicine interventions. However, future research is important and listed below are ways to improve and address these economic issues.

Non-randomised studies and selection bias

If sufficient resources are available for the evaluation, researchers should ideally collect data for economic evaluations alongside randomised clinical trials. Sometimes, individual patients or subjects cannot be randomised, and as mentioned earlier, a cluster randomised trial may be a more appropriate study design for telemedicine services, were different hospitals or units are randomised instead of individual patients and this type of trial would aim to produce the least biased estimates for costs and effects. However, as shown in the thesis, it is not always possible to conduct a randomised trial. When such data were not available, then the use of the data obtained from non-randomised studies should be used appropriately and its limitations should also be acknowledged. Adjustments should be made when there are known biases and the appropriate data has not been collected.

It is worth pointing out that economic evaluations of telemedicine services in general require more consistency in their conduct, so that the methodology and results are appropriate, reliable and comparable with other studies, and thus, a meta-analysis of telemedicine studies can also be conducted.

Patient costs

NICE states that patient costs should not be included in economic evaluations for the reference case analyses (and non-reference case analyses) [NICE, 2008]. However, patient costs are particularly important if patients have to travel more than once to a hospital and if the specialist hospital is not in the local area. It would be more useful, if NICE in their 'Methods of Technology Appraisal' Guide, suggested specific circumstances in which patient costs should be included, because of the impact these costs may have on patients and their families and of course, patient costs may be relevant for non-NICE economic evaluations.

For any clinical trial especially screening studies, it would be useful to include instruments such as questionnaires to collect information on patient costs. The results from these questionnaires could inform other future providers or users of telemedicine of what the time and cost implications might be of setting up and using a telemedicine service. Maybe some sort of 'gold standard' questionnaire to collect this information is needed; therefore making it easier for other researchers who need to calculate patient costs and also to help compare across different screening programmes.

Measures of benefits such as QALYs

Cost-benefit analysis is a type of economic evaluation which requires both costs and outcomes to be measured and valued in monetary units. Here, benefits are valued in monetary terms using valuations of peoples' observed (choices that people make in practice) or stated preferences (choices that people make in hypothetical situations). Two methods for eliciting stated preferences are: contingent valuation and discrete choice experiments.

Contingent valuation asks people the maximum they are willing to pay for a good or service [Donaldson, 1990]. This approach asks people to attach to healthcare, the amount they would be willing to pay to obtain the benefits or to avoid certain events. The value that an individual places on a health care intervention or service will depend on the perceived benefits to them. For example, Ryan et al (1997) used willingness to pay to value alternative models of antenatal care. Discrete choice experiments (DCEs) are also used to elicit preferences for healthcare services, and have the ability to take into account the non-health benefits to a patient associated with a particular service or intervention. McIntosh and Cairns (1997) claimed "by using a technique which can measure changes in utility arising through changes in the 'process' of care, any benefit arising from telemedicine can be assessed in terms of how people value the service".

DCEs are based on the notion that any good or service can be described by its characteristics (or attributes), and it is peoples' preferences for these attributes that determine the overall preference for a good. The technique can show how individuals are willing to trade between attributes of services, to estimate the relative importance of the different attributes, to estimate whether an attribute is important, and to predict the demand for a given service given its attributes [Ryan and Farrar, 2000; Drummond et al, 2005; Lancsar E and Louviere J, 2008]. There are five key steps in undertaking a DCE: 1) characteristics of the service must be identified; 2) levels need to be assigned to each of the characteristics; 3) scenarios need to be drawn up that describe all the possible outcomes of the characteristics and levels chosen; 4) preferences for the scenarios included in the survey are elicited by discrete choices (respondents given various choices and asked for their preferred choice i.e. do they prefer A or B); 5) data are analysed using regression techniques [Ryan and Farrar, 2000]. DCEs can look at the important aspects around telemedicine, but not necessarily QALYs; that is, it is conceptually difficult to measure both the quantity and quality of life. However, the aim of this thesis was to focus on QALYs.

The QALY was used as an outcome measure in Chapter 7 to look at the benefits from a screening programme with telemedicine compared to a screening programme with telemedicine. However, utilities which were used to derive the QALYs were based on literature for heart failure patients and these utilities have not captured all the benefits associated with telemedicine (including the non-health benefits such as speed of service). Hence, DCEs can be used instead to take into account the health and non-health benefits associated with telemedicine, while utility measures such as the SF-6D and EQ-5D can be used to derive QALYs. Furthermore, the QALY is the preferred outcome measure for NICE and can help decision makers to make decisions when using the incremental cost per QALY ratio to make comparisons across different interventions or treatments or across different disease areas.

One recent study which used DCEs for telemedicine was in the area of endoscopy services [van der Pol and McKenzie, 2010]. The authors looked at the costs and benefits for two clinics: 1) a tele-endoscopy clinic and 2) a conventional, mainland clinic in Scotland. The benefits from the two clinics were estimated from a sample of general public using a DCE survey, where a monetary value was derived. The relevant attributes were: type of service or clinic, one-way drive time, waiting time, and cost. Based on these attributes, 16 hypothetical scenarios were developed. The authors found that if more than 27 patients were seen in each year, the average cost per

patient was lower for the tele-endoscopy clinic (£353) than for the mainland clinic (£381). Assuming equal waiting times, individuals preferred the tele-endoscopy clinic to the mainland clinic. The net benefits were larger for tele-endoscopy clinics as long as the additional waiting time was not longer than four weeks.

Other issues for future research

Further research is needed to refine the model assumptions and data inputs in order to generate 'good cost-effectiveness estimates'. One could use value of information analysis [Claxton, 1999] to help provide a framework for analysing uncertainty within an economic model, by focusing on the value of reducing uncertainty in terms of which intervention is cost-effective through further information or additional research. For example, value of information may be able to highlight whether the uncertainty around resource use estimates for children with mild, moderate and severe CHD would have an impact on the cost-effectiveness ratio and whether this would mean that another study should be undertaken to collect such information needed for a model.

8.7 Final thoughts

The purpose of this thesis was to see how the TelePaed study analysis could have been improved in light of the economic issues which were highlighted in the literature review and to address some of these issues associated with telemedicine: selection bias, patient costs and measures of benefits, namely QALYs.

There have been increasing pressures on healthcare technologies to show that not only they are safe and efficacious, but that they are also an efficient use of resources. The thesis looked at the use telemedicine in providing specialist advice to pregnant women who were at risk of a fetal cardiac anomaly and how these three economic issues can be addressed in terms of an economic evaluation. Overall, with the available data and the decision model which was developed, the results demonstrated that if telemedicine is implemented as part of a screening programme for all low risk women in one hospital, it can be a cost-effective use of resources and help to prevent future 'missed' cardiac anomalies. In the future, with better model data inputs, telemedicine as part of antenatal screening programme maybe cost-effective.

References

Agha Z, Weinstein RS, Dunn BE (1999) Cost minimization analysis of telepathology. American Journal of Clinical Pathology. 112(4): 470-47.

Agha Z, Schapira RM, Maker AH (2002) Cost effectiveness of telemedicine for the delivery of outpatient pulmonary care to a rural population. Telemedicine Journal & E-Health. 8(3): 281-291.

Aoki N, Dunn K, Fukui T, Beck JR, Schull WJ, Li HK (2004) Cost-effectiveness analysis of telemedicine to evaluate diabetic retinopathy in a prison population. Diabetes Care. 27(5): 1095-1101.

Armstrong AW, Dorer DJ, Lugn NE, Kvedar JC (2007) Economic evaluation of interactive teledermatology compared with conventional care. Telemedicine Journal & E-Health. 13(2): 91-99.

Auerbach H, Schreyogg J, Busse R (2006) Cost effectiveness analysis of telemedical devices for pre-clinical traffic accident emergency rescue in Germany. Technology & Health Care. 14(3): 189-97.

Automobile Association (AA) (2008a) Route planner. http://www.theaa.com/route-planner/index.jsp. Last accessed 26/07/2010.

Automobile Association (AA) (2008b) Your guide to motoring costs. http://www.theaa.com/motoring_advice/motoring_costs.html. Last accessed 26/07/2010.

Bailes JE, Poole CC, Hutchinson W, Maroon JC, Fukushima T (1997) Utilization and cost savings of a wide-area computer network for neurosurgical consultation. Telemedicine Journal. 3(2): 135-139.

Barker G, McNeill KM, Krupinski EA, Weinstein RS (2004) Clinical encounters costing for telemedicine services. Telemedicine Journal & E-Health. 10(3): 381-8.

Barnett TE, Chumbler NR, Vogel WB, Beyth RJ, Qin H, Kobb R (2006) The effectiveness of a care coordination home telehealth program for veterans with diabetes mellitus: a 2-year follow-up. The American Journal of Managed Care. 12(8): 467-474.

Baser O (2006) Too much ado about propensity score models? Comparing methods of propensity score matching. Value in Health. 9(6):377-385.

Becker SO, Ichino A (2002) Estimation of average treatment effects based on propensity scores. The Stata Journal. 2:358-377.

Begg L, Chan FY, Edie G, Hockey R, Wootton R (2001) Minimum acceptable standards for digital compression of a fetal ultrasound video-clip. Journal of Telemedicine & Telecare. 7(Suppl 2): 88-90.

Bergmo TS (1996) An economic analysis of teleradiology versus a visiting radiologist service. Journal of Telemedicine & Telecare. 2(3): 136-142.

Bergmo TS (1997) An economic analysis of teleconsultation in otorhinolaryngology. Journal of Telemedicine & Telecare. 3(4): 194-199.

Bergmo TS (2000) A cost-minimization analysis of a realtime teledermatology service in northern Norway. Journal of Telemedicine & Telecare. 6(5): 273-7.

Bergmo T (2009) Can economic evaluation in telemedicine be trusted? A systematic review of the literature. Cost Effectiveness and Resource Allocation. 7:18.

Berman M, Fenaughty A (2005) Technology and managed care: patient benefits of telemedicine in a rural health care network. Health Economics. 14(6): 559-573.

Bird KT (1972) Cardiopulmonary frontiers: quality health care via interactive television. Chest. 61(3): 204-205

Bishai DM, Ferris DG, Litaker MS (2003) What is the least costly strategy to evaluate cervical abnormalities in rural women? Comparing telemedicine, local practitioners, and expert physicians. Medical Decision Making. 23(6): 463-70.

Bjorvig S, Johansen MA, Fossen K (2002) An economic analysis of screening for diabetic retinopathy. Journal of Telemedicine & Telecare. 8(1): 32-35.

Borenstein M, Hedges L, Higgins J, Rothstein H (2005) Comprehensive Meta-analysis Version 2, Biostat, Englewood NJ.

Bowes A, McColgan G (2006) Smart technology and community care for older people: innovation in West Lothian, Scotland. Edinburgh: Age Concern Scotland.

Bracale M, Cesarelli M, Bifulco P (2002) Telemedicine services for two islands in the Bay of Naples. Journal of Telemedicine & Telecare. 8(1): 5-10.

Brazier J, Roberts J, Tsuchiya A, Busshbach J (2004) A comparison of the EQ-5D and SF-6D across seven patient groups. Health Economics. 13(9): 873-884.

Breslow MJ, Rosenfeld BA, Doerfler M, Burke G, Yates G, Stone DJ et al (2004) Effect of a multiple-site intensive care unit telemedicine program on clinical and economic outcomes: an alternative paradigm for intensivist staffing. Critical Care Medicine. 32(1): 31-38.

Bricker L, Garcia J, Henderson J, Mugford M, Neilson J, Roberts T, Martin MA (2000) Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views. Health Technology Assessment. 4(16).

Briggs A, Clark T, Wolstenholme J, Clarke P (2003) Missing presumed at random: cost-analysis of incomplete data. Health Economics. 12(5): 377-392.

Briggs A, Claxton K, Sculpher M (2006) Decision modelling for health economic evaluation. Oxford: Oxford University Press.

British National Formulary 50 (2005) British Medical Association and the Royal Pharmaceutical Society of Great Britain. London: BMJ Publishing Group.

Broens TH, Huis in't Veld RM, Vollenbroek-Hutten MM, Hermens HJ, van Halteren AT, Nieuwenhuis LJ (2007) Determinants of successful telemedicine implementations: a literature study. Journal of Telemedicine & Telecare. 13(6): 303-309.

Brooks R (1996) EuroQol: the current state of play. Health Policy. 37(1): 53-72.

Brown KL, Wray J, Wood TL, McMahon AM, Burch M, Cairns J (2009) Cost utility evaluation of extracorporeal membrane oxygenation as a bridge to transplant for

children with end-stage heart failure to dilated cardiomyopathy. The Journal of Heart and Lung Transplantation. 28(1): 32-38.

Brumage MR, Chinn S, Cho K (2001) Teleradiology in a military training area. Journal of Telemedicine & Telecare. 7(6): 348-352.

Brunicardi BO (1998) Financial analysis of savings from telemedicine in Ohio's prison system. Telemedicine Journal. 4(3): 49-54

Bryan S, Buxton M, McKenna M, Ashton H, Scott A (1995) Private costs associated with abdominal aortic aneurysm screening: the importance of private travel and time costs. Journal of Medical Screening. 2(2): 62-66.

Bull C (1999) Current and potential impact of fetal diagnosis on prevalence and spectrum of serious congenital heart disease at term in the UK. The Lancet. 354(9186): 1242-1247.

Burgiss SG, Julius CE, Watson HW, Haynes BK, Buonocore E, Smith GT (1997) Telemedicine for dermatology care in rural patients. Telemedicine Journal. 3(3): 227-233.

Buskens E, Steyerberg EW, Hess J, Wladimiroff JW, Grobbee DE (1997) Routine prenatal screening for congenital heart disease: what can be expected? A decision-analytical approach. American Journal of Public Health. 87(6): 962-967.

Bynum AB, Irwin CA, Cranford CO, Denny GS (2003) The impact of telemedicine on patients' cost savings: some preliminary findings. Telemedicine Journal & E-Health. 9(4): 361-7.

Callahan CW, Malone F, Estroff D, Person DA (2005) Effectiveness of an internet-based store-and-forward telemedicine system for pediatric subspecialty consultation. Archives of Pediatrics & Adolescent Medicine. 159(4): 389-93.

Carvalho JS, Mavrides E, Shinebourne EA, Campbell S, Thilaganathan B (2002) Improving the effectiveness of routine prenatal screening for major congenital heart defects. Heart. 88(4): 387-391.

Castillo-Riquelme MC, Lord J, Moseley MJ, Fielder AR, Haines L (2004) Cost-effectiveness of digital photographic screening for retinopathy of prematurity in the United Kingdom. International Journal of Technology Assessment in Health Care. 20(2): 210-213.

Caviness AC, Cantor SB, Allen CH, Ward MA (2004) A cost-effectiveness analysis of bacterial endocarditis prophylaxis for febrile children who have cardiac lesions and undergo urinary catheterisation in the emergency department. Pediatrics. 113(5): 1291-1296.

Cepeda MSRB, Farrar JT, Strom BL (2003) Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. American Journal of Epidemiology. 158:280–287.

Chan FY, Soong B, Watson D, Whitehall J (2001) Realtime fetal ultrasound by telemedicine in Queensland. A successful venture? Journal of Telemedicine & Telecare. 7(Suppl 2): 7-11.

Chan HH, Woo J, Chan WM, Hjelm M (2000) Teledermatology in Hong Kong: a cost-effective method to provide service to the elderly patients living in institutions. International Journal of Dermatology. 39(10): 774-778.

Chodroff PH (1999) A three-year review of telemedicine at the community level - clinical and fiscal results. Journal of Telemedicine & Telecare. 5(Suppl 1): 28-30.

Chua R, Craig J, Wootton R, Patterson V (2001) Cost implications of outpatient teleneurology. Journal of Telemedicine & Telecare. 7(Suppl 1): 62-64.

Claxton K (1999) The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. Journal of Health Economics. 18(3): 342-364.

Cochrane Collaboration Handbook (2004) Alderson P, Green S, Higgins JPT, eds. Chapter 6: Assessment of study quality. In: The Cochrane Library, Issue 1. Chichester: Wiley.

Confidential Enquiry into Maternal and Child Health (2005) Pregnancy in Women with Type 1 and Type 2 Diabetes 2002-2003, England, Wales and Northern Ireland. London: CEMACH. http://www.cmace.org.uk/getdoc/dea8d48b-3dc1-4185-8a57-e7a6efad663f/Pregnancy-in-women-with-type-1-and-type-2-diabetes.aspx. Last accessed 26/07/2010.

Coyte PC, Young W, Croxford R (2000) Costs and outcomes associated with alternative discharge strategies following joint replacement surgery: analysis of an observational study using a propensity score. Journal of Health Economics. 19(6): 907-929.

Crowe BL, Hailey DM, de Silva M (1996) Teleradiology at a children's hospital: a pilot study. Journal of Telemedicine & Telecare. 2(4): 210-216.

Crown WH, Obenchain RL, Englehart L, Lair T, Buesching DP, Croghan T (1998) The application of sample selection models to outcomes research: the case of evaluating the effects of antidepressant therapy on resource utilization. Statistics in Medicine. 17(17): 1943-1958.

Crown WH (2001) Antidepressant selection and economic outcome: a review of methods and studies from clinical practice. The British Journal of Psychiatry 179(S2): S18-S22.

Curtis L (2007) Pay and Prices Index. Unit Costs of Health and Social Care 2007, Personal Social Services Research Unit, University of Kent at Canterbury, Canterbury.

Curtis L (2008) Pay and Prices Index. Unit Costs of Health and Social Care 2008. Personal Social Services Research Unit, University of Kent at Canterbury, Canterbury.

Curtis L, Netten A (2006) Pay and Prices Index. Unit Costs of Health and Social Care 2006. Personal Social Services Research Unit, University of Kent at Canterbury, Canterbury.

D'Agostino RB (1998) Propensity score methods for bias reduction in the comparison of a treatment to a non-randomised control group. Statistics in Medicine. 17(19): 2265-2281.

Darkins A, Dearden CH, Rocke LG, Martin JB, Sibson L, Wootton R (1996) An evaluation of telemedical support for a minor treatment centre. Journal of Telemedicine & Telecare. 2(2): 93-99.

Daucourt V, Sicotte C, Pelletier-Fleury N, Petitjean ME, Chateil JF, Michel P (2006) Cost-minimization analysis of a wide-area teleradiology network in a French region. International Journal for Quality in Health Care. 18(4): 287-93.

Davis MC (1997) Teleradiology in rural imaging centres. Journal of Telemedicine & Telecare. 3(3): 146-153.

Davis P, Howard R, Brockway P (2001) Telehealth consultations in rheumatology: cost-effectiveness and user satisfaction. Journal of Telemedicine & Telecare. 7(Suppl 1): 10-11.

de la Torre A, Hernandez-Rodriguez C, Garcia L (2004) Cost analysis in telemedicine: empirical evidence from sites in Arizona. Journal of Rural Health. 20(3):253-7.

Deeks JJ, Dinnes J, D'Amico, Sowden AJ, Sakarovitch C, Song F, et al (2003) Evaluating non-randomised intervention studies. Health Technology Assessment. 7(27).

Dehejia RH, Wahba S (2002) Propensity score-matching methods for non-experimental causal studies. The Review of Economics and Statistics 84(1): 151-161.

Delgado-Rodríguez M, Llorca J (2004) Bias. Journal of Epidemiology and Community Health. 58(8): 635-641.

Deodhar J (2002) Telemedicine by email-experience in neonatal care at a primary care facility in rural India. Journal of Telemedicine & Telecare. 8(Suppl 2): 20-21.

Department of Health (1997) The new NHS Modern and Dependable white paper. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4008869. Last accessed 26/07/2010.

Department of Health (1998) Information for Health: An Information Strategy for Modern NHS. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4007832. Last accessed 26/07/2010.

Department of Health (1999) Modernising Government white paper. London: Department of Health. http://archive.cabinetoffice.gov.uk/moderngov/index.htm. Last accessed 26/07/2010.

Department of Health (2001a) Building the Information Core: Implementing the NHS Plan. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4005249. Last accessed 26/07/2010.

Department of Health (2001b) Telemedicine Website Launched. Press Release. London: Department of Health. http://www.gov-news.org/gov/uk/news/telemedicine_website_launched/26536.html Last accessed 26/07/2010.

Department of Health (2002a) Advanced letter (MD) 1/02: Pay and conditions of service of hospital medical and dental staff and doctors in public health medicine and the community health service. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Lettersandcirculars/Advancedletters/index.htm. Last accessed 26/07/2010.

Department of Health (2002b) Delivering 21st century IT support for the NHS. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4008227. Last accessed 26/07/2010.

Department of Health (2003) NHS Maternity Statistics, England: 2001-02. Bulletin 2003/09. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/StatisticalWorkAreas/Statisticalhealthcare/DH_4086520. Last accessed 26/07/2010.

Department of Health (2005) Building Telecare in England. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4115303. Last accessed 26/07/2010.

Department of Health (2006) National Schedule of Reference Costs 2005–06. Appendix NSRC1 (NHS Trusts). London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_062884. Last accessed 26/07/2010.

Department of Health (2007) NHS Maternity Statistics, England: 2005-06. NHS Information Centre, London. http://www.ic.nhs.uk/pubs/maternity0506. Last accessed 26/07/2010.

Department of Health (2008) National Schedule of Reference Costs 2006–07. Appendix NSRC1 (NHS Trusts). London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_082571. Last accessed 26/07/2010.

Department of Health (2009) Whole Systems Demonstrators. An overview of Telecare and Telehealth. London: Department of Health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_100946. Last accessed 26/07/2010.

Dolan P, Gudex C, Kind P, Williams A (1996) The time-trade off method: results from a general population study. Health Economics. 5(2): 141-154.

Dolan P (1997) Modelling valuations for EuroQol Health States. Medical Care. 35(11): 1095-1108.

Donaldson C (1990) Willingness to pay for publicly-provided goods: a possible measure of benefit? Journal of Health Economics. 9(1): 103-118.

Doolittle GC, Williams A, Harmon A, Allen A, Boysen CD, Wittman C, et al (1998) A cost measurement study for a tele-oncology practice. Journal of Telemedicine & Telecare. 4(2): 84-88.

Doolittle GC (2000) A cost measurement study for a home-based telehospice service. Journal of Telemedicine & Telecare. 6(Suppl 1): 193-195.

Doolittle GC, Williams AR, Cook DJ (2003) An estimation of costs of a pediatric telemedicine practice in public schools. Medical Care. 41(1): 100-109.

Doolittle GC, Williams AR, Spaulding A, Spaulding RJ, Cook DJ (2004) A cost analysis of a tele-oncology practice in the United States. Journal of Telemedicine & Telecare. 10(Suppl 1): 27-29.

Dowie R, Young T, Mistry H, Weatherburn G on behalf of the TelePaed project team (2003) Economic evaluation of the role of telemedicine in paediatric cardiology. First report: Paediatric cardiology outpatient services. Final Report to the Department of Health, Brunel University, Uxbridge.

Dowie R, Mistry H, Young T, Weatherburn G on behalf of the TelePaed project team (2004) Economic evaluation of the role of telemedicine in paediatric cardiology. Second report: Fetal cardiology services. Final Report to the Department of Health, Brunel University, Uxbridge.

Dowie R, Mistry H, Young TA, Weatherburn G, Gardiner HM, Rigby M, Rowlinson G, Franklin RCG (2007) Telemedicine in pediatric and perinatal cardiology: economic evaluation of a service in English hospitals. International Journal of Technology Assessment and Health Care. 23(1): 116-125.

Dowie R, Mistry H, Young TA, Franklin R, Gardiner HM (2008) Cost implications of introducing a telecardiology service to support fetal ultrasound screening. Journal of Telemedicine & Telecare. 14(8): 421-426.

Doze S, Simpson J, Hailey D, Jacobs P (1999) Evaluation of a telepsychiatry pilot project. Journal of Telemedicine & Telecare. 5(1): 38-46.

Drummond M, McGuire A (2001) Economic Evaluation in HealthCare: Merging Theory with Practice. Oxford: Oxford University Press.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005) Methods for the Economic Evaluation of Health Care Programmes. 3rd Edition. New York: Oxford University Press.

Dumville JC, Manca A, Kitchener HC, Smith ARB, Belson L, Torgerson DJ on behalf of the COLPO study group (2006) Cost-effectiveness analysis of open colposuspension versus laparoscopic colposuspension in the treatment of urodynamic stress incontinence. British Journal of Obstetrics and Gynaecology. 113(9): 1014-1022.

Dusheiko M, Gravelle H, Jacobs R (2004) The effect of practice budgets on patient waiting times: allowing for selection bias. Health Economics. 13: 941-958.

Dyer MT, Goldsmith KA, Sharples LS, Buxton MJ (2010) A review of health utilities using the EQ-5D in studies of cardiovascular disease. Health and Quality of Life Outcomes. 8(13):1-12.

Egger M, Smith GD, Song F, Sheldon TA (1993) Making sense of meta-analysis. Pharmacoepidemiology and drug safety. 2(1): 65-72.

Elford DR, White H, St John K, Maddigan B, Ghandi M, Bowering R (2001) A prospective satisfaction study and cost analysis of a pilot child telepsychiatry service in Newfoundland. Journal of Telemedicine & Telecare. 7(2): 73-81.

Endnote (2002) Endnote Windows Version 6. Thomson ISI. Philadelphia, PA.

Erickson P, Wilson R, Shannon I (1995) Years of health life. Hyattsville (MD): US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics [Healthy People 2000 Statistical Note No 7; DHHS Publication No (PHS) 95-1237]

Estrada CA, Unterborn JN, Price J, Thompson D, Gibson L (2000). Judging the effectiveness of clinical pathways for pneumonia: the role of risk adjustment. Effective Clinical Practice 3(5): 221-228.

Feinstein AR (1985) Clinical Epidemiology: the Architecture of Clinical Research. WB Saunders Company, Philadelphia, USA.

Ferrer-Roca O, Sosa-Iudicissa MC (1998) Handbook of Telemedicine (Studies in Health Technology and Informatics). IOS Press. Washington. Volume 54.

Ferris DG, Bishai DM, Litaker MS, Dickman ED, Miller JA, Macfee MS (2004) Telemedicine network telecolposcopy compared with computer-based telecolposcopy. Journal of Lower Genital Tract Disease. 8(2): 94-101.

Finley JP, Sharratt GP, Nanton MA, Chen RP, Bryan P, Wolstenholme J, et al (1997) Paediatric echocardiography by telemedicine-nine years' experience. Journal of Telemedicine & Telecare. 3(4): 200-204.

Fisk NM, Sepulveda W, Drysdale K, Ridley D, Garner P, Bower S, Kyle P, Dhillon H, Carvalho JS, Wootton R (1996) Fetal telemedicine: six month pilot of real-time ultrasound and video consultation between the Isle of Wight and London. British Journal of Obstetrics and Gynaecology. 103(11): 1092-1095.

Fletcher J, Hicks NR, Kay JDS, Boyd PA (1995) Using decision analysis to compare policies for antenatal screening for Down’s syndrome. British Medical Journal. 311(7001):351-356.

Fortney J, Booth B, Zhang M, Humphrey J, Wiseman E (1998) Controlling for selection bias in the evaluation of alcoholics anonymous as aftercare treatment. Journal of Studies on Alcohol. 59(6): 690-7.

Foster EM (2003) Propensity score matching - an illustrative analysis of dose response. Medical Care 41(10): 1183-1192.

Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, Taylor RS (2007) The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model. Health Technology Assessment. 11(47).

Friedman RH, Kazis LE, Jette A, Smith MB, Stollerman J, Torgerson J, Carey K (1996) A telecommunications system for monitoring and counselling patients with hypertension. Impact on medication adherence and blood pressure control. American Journal of Hypertension. 9(4 Pt 1): 285-92.

Garson A, Allen HD, Gersony WM, Gillette PC, Hohn AR, Pisky WW, Mikhail O (1994) The cost of congenital heart disease in children and adults: A model for multicenter assessment of price and practice variation. Archives of Pediatric and Adolescent Medicine. 148(10): 1039-1045.

Garson A, Davignon A, Deanfield J, Demoor M, Kachener J, Kallfelz C, Squarcia U, Sumitomo N, Wilkinson J (1996) Payment and practice variation in congenital heart

disease – an international study. Cardiovascular Journal of Southern Africa. 86(Suppl 1): 25-29.

Gekas J, Gagné G, Bujold E, Douillard D, Forest JC, Reinharz D, Rousseau F (2009) Comparison of different strategies in prenatal screening for Down's syndrome: cost-effectiveness analysis of computer simulation. British Medical Journal. 338:b138; doi: 10.1136/bmj.b138.

Gershon-Cohen J and Cooley AG (1950) Telognosis. Radiology. 55(4): 582-587.

Gibbs JL, Monro JL, Cunningham D, Rickards A (2004) Survival after surgery or therapeutic catheterisation for congenital heart disease in children in the United Kingdom: analysis of the central cardiac audit database for 2000–1. British Medical Journal. 328(7440): 611-615.

Gilbert RE, Augood C, Gupta R, Ades AE, Logan S, Sculpher M, van der Muelen JHP (2001) Screening for Down's syndrome: effects, safety, and cost effectiveness of first and second trimester strategies. British Medical Journal. 323(7310): 423-428.

Gill M, Murday V, Slack J (1987) An economic appraisal of screening for Down's syndrome in pregnancy using maternal age and serum alpha fetoprotein concentration. Social Science and Medicine. 24(9): 725-731.

Glaser V (2009) Telehealth for managing COPD. Telemedicine Journal & E-Health. 15(8): 739-741.

Gold MR, Siegel JE, Russell LB, Weinstein MC (1996) Cost-effectiveness in Health and Medicine. New York: Oxford University Press.

Grant B, Morgan GJ, McCrossan BA, Crealey GE, Sands AJ, Craig B, Casey FA (2010) Remote diagnosis of congenital heart disease: the impact of telemedicine. Archives of Disease in Childhood. 95(4): 276-280.

Griebsch I, Coast J, Brown J (2005) Quality-adjusted life years lack quality in pediatric care: a critical review of published cost-utility studies in child health. Pediatrics. 115(5): 600-608.

Grieve R, Grishchenko M, Cairns J (2009) SF-6D versus EQ-5D: reasons for differences in utility scores and impact on reported cost-utility. European Journal of Health Economics. 10(1): 15-23.

Gum PA, Thamilarasan M, Watanabe J, Blackstone EH, Lauer MS (2001) Aspirin use and all-cause mortality among patients being evaluated for known or suspected coronary artery disease: a propensity analysis. Journal of the American Medical Association. 286(10): 1187-1194.

Haberland CA, Benitz WE, Sanders GD, Pietzsch JB, Yamada S, Nguyen L, Garber AM (2002) Perinatal screening for Group B Streptococci: Cost-Benefit Analysis of Rapid Polymerase Chain Reaction. Pediatrics. 110(3): 471-480.

Hagard S, Carter FA (1976) Preventing the birth of infants with Down's syndrome: a cost-benefit analysis. British Medical Journal. 1(6012): 753-756.

Hagard S, Carter F, Milne RG (1976) Screening for spina bifida cystic – a cost-benefit analysis. British Journal of Preventive and Social Medicine. 30(1): 40-53.

Hailey D, Roine R, Ohinmaa A (2002) Systematic review of evidence for the benefits of telemedicine. Journal of Telemedicine & Telecare. 8(Suppl 1): 1-7.

Hailey D, Ohinmaa A, Roine R (2004) Study quality and evidence of benefit in recent assessments of telemedicine. Journal of Telemedicine & Telecare. 10(6): 318-324.

Hailey D (2005) The need for cost-effectiveness studies in telemedicine. Journal of Telemedicine & Telecare. 11(8): 379-383.

Håkasson S, Gavelin C (2000) What do we really know about the cost-effectiveness of telemedicine? Journal of Telemedicine & Telecare. 6(Suppl 1): 133-136.

Halvorsen PA, Kristiansen IS (1996) Radiology services for remote communities: cost minimisation study of telemedicine. British Medical Journal. 312(7042): 1333-1336.

Harley J (2006) Economic evaluation of a tertiary telepsychiatry service to an island. Journal of Telemedicine & Telecare. 12(7): 354-57.

Harno KS (1999) Telemedicine in managing demand for secondary-care services. Journal of Telemedicine & Telecare. 5(3): 189-192.

Harno K, Paavola T, Carlson C, Viikinkoski P (2000) Patient referral by telemedicine: effectiveness and cost analysis of an Intranet system. Journal of Telemedicine & Telecare. 6(6): 320-329.

Harno K, Arajarvi E, Paavola T, Carlson C, Viikinkoski P (2001) Clinical effectiveness and cost analysis of patient referral by videoconferencing in orthopaedics. Journal of Telemedicine & Telecare. 7(4): 219-225.

Harris RA, Washington AE, Nease RF, Kuppermann M (2004) Cost utility of prenatal diagnosis and the risk-based threshold. The Lancet. 363(9405): 276-282.

Hassall S, Wootton R, Guilfoyle C (2003) The cost of allied health assessments delivered by videoconference to a residential facility for elderly people. Journal of Telemedicine & Telecare. 9(4): 234-237.

Heckman J (1979) Sample selection bias as a specification error. Econometrica. 47(1): 153-161.

Henderson J (1982a) An economic appraisal of the benefits of screening for open spina bifida. Social Science and Medicine. 16(5): 545-560.

Henderson J (1982b) Measuring the benefits of screening for open neural tube defects. Journal of Epidemiology and Community Health. 36(3): 214-219.

Henderson J, Bricker L, Roberts T, Mugford, Garcia J, Neilson J (2002) British National Health Service's and women's costs of antenatal ultrasound screening and follow-up tests. Ultrasound in Obstetrics & Gynecology. 20(2): 154-162.

Heron J, O'Connor TG, Evans J, Golding J, Glover V, and the ALPSPAC Study Team (2004) The course of anxiety and depression through pregnancy and the postpartum in a community sample. Journal of Affective Disorders. 80(1): 65-73.

Hersh W, Helfand M, Wallace J, Kraemer D, Patterson P, Shapiro S, Greenlick M (2002) A systematic review of the efficacy of telemedicine for making diagnostic and management decisions. Journal of Telemedicine & Telecare. 8(4): 197-209.

Hersh WR, Hickam DH, Severance SM, Dana TL, Krages KP, Helfand M (2006) Diagnosis, access and outcomes: update of a systematic review of telemedicine services. Journal of Telemedicine & Telecare. 12(Suppl 2): 3-31.

Higgins JPT, Green S (2009) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.2 (Updated September 2009) The Cochrane Collaboration. Available from http://www.cochrane-handbook.org. Last accessed 20/01/2011.

Himpens B (2003) The Pentalfa project. 2: profile of participants and economic aspects. Journal of Telemedicine & Telecare. 9(2): 104-108.

HM Treasury (2003) The Green Book: Appraisal and Evaluation in Central Government. http://www.hm-treasury.gov.uk/d/green_book_complete.pdf Last accessed 26/07/2010.

Höfer S, Lim L, Guyatt G, Oldridge N (2004) The MacNew Heart Disease health-related quality of life instrument: A summary. Health and Quality of Life Outcomes. 2(3):1-8.

Hogan JW, Lancaster T (2004) Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. Statistical Methods in Medical Research. 13(1): 17-48.

Horsman J, Furlong W, Feeny D, Torrance G (2003) The Health Utilities Index (HUI®): concepts, measurement properties and applications. Health and Quality of Life Outcomes. 2003. 1:54.

Hosmer DW, Lemeshow S (1989) Applied Logistic Regression. First Edition. New York: Wiley.

Hui E, Woo J (2002) Telehealth for older patients: the Hong Kong experience. Journal of Telemedicine & Telecare. 8(Suppl 3): 39-41.

Hunter S, Heads A, Wyllie J, Robson S (2000) Prenatal diagnosis of congenital heart disease in the northern region of England: benefits of a training programme for obstetric ultrasonographers. Heart. 84(3): 294-298.

Hyett J, Perdu M, Sharland G, Snijders R, Nicolaides KH (1999) Using fetal nuchal translucency to screen for major congenital heart defects at 10-14 weeks of gestation: Population based cohort study. British Medical Journal. 318(7176): 81-85.

Hyler SE, Gangure DP (2003) A review of the costs of telepsychiatry. Psychiatric Services. 54(7): 976-980.

Imbens GW (2000) The role of the propensity score in estimating dose-response functions. Biometrika 87(3): 706-710.

Institute of Medicine (1996) Telemedicine: a guide to assessing telecommunications in health: Committee on evaluating clinical applications of telemedicine. Editor: Field MF. Washington: National Academy Press.

Ippolito A, De Falco M, Triassi M, Di Lieto A (2003) A cost study of prenatal telemedicine. Journal of Telemedicine & Telecare. 9(5): 288-291.

Jacklin PB, Roberts JA, Wallace P, Haines A, Harrison R, Barber JA, Thompson SG, Lewis L, Currell R, Parker S, Wainwright P and the Virtual Outreach Project Group (2003) Virtual outreach: economic evaluation of joint teleconsultations for patients

referred by their general practitioner for a specialist opinion. British Medical Journal. 327(7406): 84-88.

Jin AJ, Martin D, Maberley D, Dawson KG, Seccombe DW, Beattie J (2003) Evaluation of a mobile diabetes care telemedicine clinic serving Aboriginal communities in Northern British Columbia, Canada. International Journal of Circumpolar Health. 63(Suppl 2): 124-8.

Johnson A, Berg G, Fleegler E, Sauerbrun M (2005) A matched-cohort study of selected clinical and utilization outcomes for an asthma care support program. Disease Management. 8(3): 144-154.

Johnston K, Kennedy C, Murdoch I, Taylor P, Cook C (2004) The cost-effectiveness of technology transfer using telemedicine. Health Policy & Planning. 19(5): 302-9.

Jong M (2004) Managing suicides via videoconferencing in a remote northern community in Canada. International Journal of Circumpolar Health. 63(4): 422-8.

Karnon J, Brown J, Zeuner D, Briggs A (1997) Issues in the economic evaluation of antenatal screening programmes for genetic disorders. Brunel University: Health Economics Research Group. Discussion paper 18.

Karnon J, Goyder E, Tappenden P, McPhie S, Towers I, Brazier J, Madan J (2007) A review and critique of modelling in prioritising and designing screening programmes. Health Technology Assessment. 11(52).

Kenny D, Stuart AS (2009) Long-term outcome of the child with congenital heart disease. Paediatrics and Child Health. 19(1): 37-42.

Kesler C, Balch D (1995) Development of a telemedicine and distance learning network in rural eastern North Carolina. Journal of Telemedicine & Telecare. 1(3): 178-182.

Kielhorn A, Graf von der Schulenburg JM (2000) The Health Economics Handbook. Adis International Ltd, Chester, England.

Kildemoes HW, Kristiansen IS (2004) Cost-effectiveness of interventions to reduce the thrombolytic delay for acute myocardial infarction. International Journal of Technology Assessment in Healthcare. 20(3): 368-74.

Kind P, Hardman G, Macran S (1999) UK Population Norms for EQ-5D. University of York: Centre for Health Economics. Discussion paper 172.

Kirsch J, McGuire A (2000) Establishing health state valuations for disease specific states: an example from heart disease. Health Economics. 9(2): 149-158.

Kitt SM, Clayton L (2002) The cost-effectiveness of telehealth in metropolitan hospitals. Journal of Telemedicine & Telecare. 8(Suppl 3): 42-3.

Knapp M, Windmeijer F, Brown J, Kontondimas S, Tzivelekis S, Haro JM, Ratcliffe M, Hong J, Novick D on behalf of the SOHO Study Group (2008) Cost-utility analysis of treatment with olanzapine compared with other antipsychotic treatments in patients with schizophrenia in the Pan-European SOHO study. Pharmacoeconomics. 26(4): 341-358.

Knowles R, Griebsch I, Dezateux C, Brown, Bull C, Wren C (2005) Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. Health Technology Assessment. 9(44).

Kristo DA, Eliasson AH, Poropatich RK, Netzer CM, Bradley JP, Loube DI, et al (2001) Telemedicine in the sleep laboratory: feasibility and economic advantages of polysomnograms transferred online. Telemedicine Journal & E-Health. 7(3): 219-224.

Kumar S, Tay-Kearney ML, Chaves F, Constable IJ, Yogesan K (2006) Remote ophthalmology services: cost comparison of telemedicine and alternative service delivery options. Journal of Telemedicine & Telecare. 12(1): 19-22.

Kunz RA, Oxman AD (1995) Empirical evidence of selection bias in studies of the effects of health care: a systematic review. Presented at the Cochrane Colloquium, Oslo, 5-8 October 1995.

Kunz R, Vist G, Oxman AD (2005) Randomisation to protect against selection bias in healthcare trials (review). The Cochrane Collaboration. Issue 4.

Kuppermann M, Feeny D, Gates E, Posner SF, Blumberg B, Washington AE (1999) Preferences of women facing a prenatal diagnostic choice: long-term outcomes matter most. Prenatal Diagnosis. 19(8): 711-716.

Lamminen H, Tuomi ML, Lamminen J, Uusitalo H (2000) A feasibility study of realtime teledermatology in Finland. Journal of Telemedicine & Telecare. 6(2): 102-107.

Lamminen H, Lamminen J, Ruohonen K, Uusitalo H (2001) A cost study of teleconsultation for primary-care ophthalmology and dermatology. Journal of Telemedicine & Telecare. 7(3): 167-173.

Lancsar E, Louviere J (2008) Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 26(8): 661-677.

Landrum MB, Ayanian JZ (2001) Causal effect of ambulatory speciality care on mortality following myocardial infarction: a comparison of propensity score and instrumental variable analyses. Health Services and Outcomes Research Methodology. 2(3-4): 221-245.

Latal B, Helfricht S, Fischer JE, Bauersfeld U, Landolt MA (2009) Psychological adjustment and quality of life in children and adolescents following open-heart surgery for congenital heart disease: a systematic review. BMC Pediatrics. 9(6):1-10.

Lee LF (2003) Generalized econometric models with selectivity. Econometrica. 51(2): 507-512.

Leuven E, Sianesi B (2003) "psmatch2: Propensity Score and Mahalanobis Matching" http://ideas.repec.org/c/boc/bocode/s432001.html. Last accessed 26/07/2010.

Loane MA, Bloomer SE, Corbett R, Eedy DJ, Gore HE, Hicks N, et al (1999) Patient cost-benefit analysis of teledermatology measured in a randomized control trial. Journal of Telemedicine & Telecare. 5(Suppl 1): 1-3.

Loane MA, Bloomer SE, Corbett R, Eedy DJ, Hicks N, Lotery HE, et al (2000) A comparison of real-time and store-and-forward teledermatology: a cost-benefit study. British Journal of Dermatology. 143(6): 1241-1247.

Loane MA, Bloomer SE, Corbett R, Eedy DJ, Evans C, Hicks N, et al (2001a) A randomized controlled trial assessing the health economics of realtime teledermatology compared with conventional care: an urban versus rural perspective. Journal of Telemedicine & Telecare. 7(2): 108-118.

Loane MA, Oakley A, Rademaker M, Bardford N, Fleischl P, Kerr P, et al (2001b) A cost-minimization analysis of the societal costs of real-time teledermatology compared with conventional care: results from a randomized controlled trial in New Zealand. Journal of Telemedicine & Telecare. 7(4): 233-238.

Lobley D (1997) The economics of telemedicine. Journal of Telemedicine & Telecare. 3(3):117-125

Luce BR, Elixhauser A (1990) Estimating costs in the economic evaluation of medical technologies. International Journal of Technology Assessment in Health Care. 6(1): 57-75.

Luellen JK, Shadish WR, Clark MH (2005). Propensity scores. An introduction and experimental test. Evaluation Review 29(6): 530-558.

Mackie AS, Pilote L, Ionescu-Ittu R, Rahme E, Marelli AJ (2007) Health care resource utilization in adults with congenital heart disease. The American Journal of Cardiology. 99(6): 839-843.

Macran S, Birks Y, Parsons J, Sloper P, Hardman G, Kind P, van Doorn C, Thompson D, Lewin R (2006) The development of a new measure of quality of life for children with congenital cardiac disease. Cardiology in the Young. 16(2):165-172.

Mahle WT, Forbess JM, Kirshbom PM, Cuadrado AR, Simsic JM, Kanter KR (2005) Cost-utility analysis of salvage cardiac extracorporeal membrane oxygenation in children. The Journal of Thoracic and Cardiovascular Surgery. 129(5): 1084-1090.

Mair F, Whitten P (2000) Systematic review of studies of patient satisfaction with telemedicine. British Medical Journal. 320(7248): 1517-1520.

Mangham L, Petrou S (2008) Modelling the long-term costs of pre-term birth. Tommy's, The Baby Charity: London.

Mangham L, Petrou S, Doyle LW, Draper ES, Marlow N (2009) The cost of preterm birth throughout childhood in England and Wales. Pediatrics. 123(2): 312-327.

Manificat S, Dazord A, Langue J, Danjou G, Bauche P, Bovet F, Cubells J, Luchelli R, Tockert E, Conway K (2000) Évaluation de la qualité de vie du nourrisson et du très jeune enfant: validation d'un questionnaire. Étude multicentrique européenne. Archives de Pédiatric. 7(6):605-614.

Manly BFJ (1997) Randomisation, Bootstrap and Monte Carlo Methods in Biology (2nd Edition), Texts in Statistical Science, Chapman and Hall, London.

Malone FD, Athanassiou A, Craigo SD, Simpson LL, D'Alton ME (1998) Cost issues surrounding the use of computerized telemedicine for obstetric ultrasonography. Ultrasound in Obstetrics & Gynaecology. 12(2): 120-124.

Marcin JP, Nesbitt TS, Struve S, Traugott C, Dimand RJ (2004) Financial benefits of a paediatric intensive care unit-based Telemedicine program to a rural adult intensive care unit: Impact of keeping acutely ill and injured children in their local community. Telemedicine Journal & E-Health. 10(Suppl 2): S1-S5.

McClellan M, McNeil BJ, Newhouse JP (1994) Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. Journal of the American Medical Association. 272(11): 859-866.

McCue MJ, Mazmanian PE, Hampton C, Marks TK, Fisher E, Parpart F, et al (1997) The case of Powhatan Correctional Center/Virginia Department of Corrections and Virginia Commonwealth University/Medical College of Virginia. Telemedicine Journal. 3(1): 11-17.

McCue MJ, Mazmanian PE, Hampton CL, Marks TK, Fisher EJ, Parpart F, et al (1998) Cost-minimization analysis: A follow-up study of a telemedicine program. Telemedicine Journal. 4(4): 323-327.

McCue MJ, Hampton CL, Malloy W, Fisk KJ, Dixon L, Neece A (2000) Financial analysis of telecardiology used in a correctional setting. Telemedicine Journal & E-Health. 6(4): 385-391.

McDonough CM, Grove MR (2005) Comparison of EQ-5D, HUI, and SF-36-Derived Societal Health State Values among Spine Patient Outcomes Research Trial (SPORT) Participants. Quality of Life Research. 14(5): 1321-1332.

McIntosh E, Cairns J (1997) A framework for the economic evaluation of telemedicine. Journal of Telemedicine & Telecare. 3(3):132-139.

McIntosh WA, Alston LT, Booher JR, McCaughan W, Segura CB, Sykes D, et al (2003) Time spent with patients and charges to patients for specialty consultations using telemedicine. Telemedicine Journal & E-Health. 9(4): 345-50.

Medway hospital (2008) Car parking charges. http://www.medway.nhs.uk/Visiting/Car_parking/index.html. Last accessed 26/07/2010.

Mielonen ML, Ohinmaa A, Moring J, Isohanni M (2000) Psychiatric inpatient care planning via telemedicine. Journal of Telemedicine & Telecare. 6(3): 152-157.

Millimet D (2001) What is the difference between “endogeneity” and “sample selection bias”? http://www.stata.com/support/faqs/stat/bias.html

Mistry H, Dowie R, Young T, Gardiner H (2007) The costs of maternity care for women with multiple pregnancy compared with high-risk and low-risk singleton pregnancy. British Journal of Obstetrics and Gynaecology. 114(9): 1104-1112.

Mitra N, Indurkhya A (2005) A propensity score approach to estimating the cost-effectiveness of medical therapies from observational data. Health Economics. 14(8): 805-815.

Modai I, Jabarin M, Kurs R, Barak P, Hanan I, Kitain L (2006) Cost effectiveness, safety, and satisfaction with video telepsychiatry versus face-to-face care in ambulatory setting. Telemedicine Journal & E-Health. 2(5): 515-520.

Moons P, Siebens K, De Geest S, Abraham I, Budts W, Gewilling M (2001) A pilot study of expenditures on, and utilization of resources in, health care in adults with congenital heart disease. Cardiology in the Young. 11(3): 301-313.

Morris S, Sutton M, Gravelle H (2005) Inequity and inequality in the use of health care in England: an empirical investigation. Social Science and Medicine. 60(6): 1251-1266.

Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, et al (2006) Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. Journal of Neurology, Neurosurgery and Psychiatry. 77(2):193–198.

Mulholland HC, Casey F, Brown D, Corrigan N, Quinn M, McCord B, Rogers J, Craig BG (1999) Application of a low cost telemedicine link to the diagnosis of neonatal congenital heart defects by remote consultation. Heart. 82(2): 217-221.

Mulligan JA, Fox-Rushby J (2005) Transferring cost-effectiveness data across time and space. In Fox-Rushby J and Cairns J. Economic Evaluation. Oxford: Open University Press.

Mulrow CD (1994) Rationale for systematic reviews. British Medical Journal. 309(6954): 597-599.

National Collaborating Centre for Women's and Children's Health (2003) Antenatal care: routine care for the healthy pregnant woman. Clinical guideline 6, October 2003. Commissioned by the National Institute for Health and Clinical Excellence. Royal College of Obstetricians and Gynaecologists, London.

National Collaborating Centre for Women's and Children's Health (2008) Antenatal care: routine care for the healthy pregnant woman. Clinical guideline 62, March 2008, partial update of 2003 guidelines. Commissioned by the National Institute for Health and Clinical Excellence. Royal College of Obstetricians and Gynaecologists, London.

National Institute for Health and Clinical Excellence (NICE) (2008) Guide to methods of technology appraisal. London, June 2008.

Navein J, Arose D, Pietermich A (1999) A business model for telemedicine. Journal of Telemedicine & Telecare. 5(Suppl 1): 76-78.

Netten A, Curtis L (2002) Unit Costs of Health and Social Care 2002, Personal Social Services Research Unit, University of Kent at Canterbury, Canterbury.

Neophytou E (2000) *TELE-REMEDY.* Remote Diagnosis, Management and Education in Congenital Heart Disease: a cost-effective Telemedicine Solution based on EURO-ISDN technology, Business Development Section, ERGO SA, Athens, Greece.

Newhouse JP, McClellan M (1998) Econometrics in outcomes research: the use of instrumental variables. Annual Review in Public Health. 19(17-34).

Nguyen LT, Massman NJ, Franzen BJ, Ahrenholz DH, Sorensen NW, Mohr WJ 3rd, et al (2004) Telemedicine follow-up of burns: lessons learned from the first thousand visits. Journal of Burn Care & Rehabilitation. 25(6): 485-90.

Noble SM, Coast J, Benger JR (2005) A cost-consequences analysis of minor injuries telemedicine. Journal of Telemedicine & Telecare. 11(1): 15-9.

Norris AC (2001) Essentials of telemedicine and telehealth. Chichester: John Wiley and Sons Ltd.

Norum J, Bergmo TS, Holdo B, Johansen MV, Vold IN, Sjaaeng EE, et al (2007) A tele-obstetric broadband service including ultrasound, videoconferencing and cardiotocogram. A high cost and a low volume of patients. Journal of Telemedicine & Telecare. 13(4): 180-4.

Odibo AO, Coassolo KM, Stamilio DM, Ural SH, Macones GA (2006) Should all pregnant diabetic women undergo a fetal echocardiography? A cost-effectiveness analysis comparing four screening strategies. Prenatal Diagnosis. 26(1): 39-44.

OECD (2008) Main economic indicators. March 2008. France: Paris. http://www.oecd.org/department/0,3355,en_2649_34357_1_1_1_1_1,00.html. Last accessed 26/07/2010.

Office of National Statistics (2002) New Earnings Survey. London: Office for National Statistics. http://www.statistics.gov.uk/downloads/theme_labour/NES2002_GB/NES2002_Streamlined_analyses.pdf Last accessed 26/07/2010.

Office of National Statistics (2005) Live births, stillbirths and infant deaths by ONS cause groups, 2005, Health Statistics Quarterly 32. London: Office for National Statistics. http://www.statistics.gov.uk/StatBase/Expodata/Spreadsheets/D9519.xls. Last accessed 26/07/2010.

Office of National Statistics (2008) Interim Life Tables, England & Wales, 1980-82 to 2005-07. London: Office for National Statistics.

Office of National Statistics (2008b) Birth Statistics, England & Wales (Series FM1), No.37. London: Office for National Statistics.

Ohinma A, Scott R (2006) A costing model for videoconferencing in Alberta. Journal of Telemedicine & Telecare. 12(7): 363-369.

Ohinmaa A, Vuolio S, Haukipuro K, Winblad I (2002) A cost-minimization analysis of orthopaedic consultations using videoconferencing in comparison with conventional consulting. Journal of Telemedicine & Telecare. 8(5):283-289.

Patient.Co.UK website (2010) Fallot's Tetralogy: Prognosis. http://www.patient.co.uk/doctor/Fallot's-Tetralogy.htm. Last accessed: 26/07/2010.

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology. 49:1373–1379.

Peng PW, Stafford MA, Wong DT, Salenieks ME (2006) Use of telemedicine in chronic pain consultation: a pilot study. Clinical Journal of Pain. 22(4): 350-52.

Persaud DD, Jreige S, Skedgel C, Finley J, Sargeant J, Hanlon N (2005) An incremental cost analysis of telehealth in Nova Scotia from a societal perspective. Journal of Telemedicine & Telecare. 11(2): 77-84.

Petersen S, Peto V, Rayner (2003) Congenital heart disease statistics 2003, British Heart Foundation Health Promotion Research Group, Department of Public Health, University of Oxford. http://www.heartstats.org. Last accessed 26/07/2010.

Petrou S, Henderson J, Roberts T, Martin MA (2000) Recent economic evaluations of antenatal screening: a systematic review and critique. Journal of Medical Screening. 7(2): 59-73.

Petrou S (2003) Methodological issues raised by preference-based approaches to measuring the health status of children. Health Economics. 12(8): 697-702.

Preston J (1995) Texas Telemedicine Project: a viability study. Telemedicine Journal. 1(2): 125-132.

Redlick F, Roston B, Gomez M, Fish JS (2002) An initial experience with telemedicine in follow-up burn care. Journal of Burn Care & Rehabilitation. 23(2): 110-115.

Rendina MC, Downs SM, Carasco N, Loonsk J, Bose CL (1998) Effect of telemedicine on health outcomes in 87 infants requiring neonatal intensive care. Telemedicine Journal. 4(4): 345-51.

Rendina MC, Carrasco N, Wood B, Cameron A, Bose C (2001) A logit model for the effect of telecardiology on acute newborn transfers. International Journal of Technology Assessment in Healthcare. 17(2): 244-249.

Robins JM, Hernan MA, Brumback B (2000) Marginal structural models and causal inference in epidemiology. Epidemiology. 11(5): 550-560.

Robinson S, Roberts T, Barton P, Bryan S, Macleod J, McCarthy A, et al (2007) Healthcare and patient costs of a proactive chlamydia screening programme: the Chlamydia Screening Studies project. Sexually Transmitted Infections. 83(4): 276-281.

Roine R, Ohinmaa A, Hailey D (2001) Assessing telemedicine: a systematic review of the literature. Canadian Medical Association Journal. 165(6): 765-771.

Rosenbaum PR (1987) Model-based direct adjustment. Journal of the American Statistical Association. 82(398): 387-394.

Rosenbaum PR, Rubin DB (1984) Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association. 79(387): 516-524.

Rowley PT, Loader S, Kaplan RM (1998) Prenatal screening for cystic fibrosis carriers: an economic evaluation. American Journal of Human Genetics. 63(4): 1160-1174.

Rubin DB (1979) Using multivariate matched sampling and regression adjustment to control bias in observational studies. Journal of the American Statistical Association. 74(366): 318-328.

Rubin DB (1997) Estimating causal effects from large data sets using propensity scores. Annals of Internal Medicine. 127(8S): 757-763.

Rumpsfeld M, Arild E, Norum J, Breivik E (2005) Telemedicine in haemodialysis: a university department and two remote satellites linked together as one common workplace. Journal of Telemedicine & Telecare. 11(5): 251-55.

Ruskin PE, Silver-Aylaian M, Kling MA, Reed SA, Bradham DD, Hebel JR et al (2004) Treatment outcomes in depression: comparison of remote treatment through telepsychiatry to in-person treatment. American Journal of Psychiatry. 161(8): 1471-1476.

Ryan M, Ratcliffe J, Tucker J (1997) Using willingness to pay to value alternative models of antenatal care. Social Science and Medicine. 44(3): 371-380.

Ryan M, Farrar S (2000) Using conjoint analysis to elicit preferences for health care. British Medical Journal. 320(7248): 1530-1533.

Sackett D (1979) Bias in analytic research. Journal of Chronic Diseases. 32(1-2): 51-63.

Santamaria N, Carville K, Ellis I, Prentice J (2004) The effectiveness of digital imaging and remote expert wound consultation on healing rates in chronic lower leg ulcers in the Kimberley region of Western Australia. Primary Intention. 12(2): 62-70.

Schaafsma J, Pantazi SV, Moehr JR, Anglin CR, Grimm NA (2007) An economic evaluation of a telehealth network in British Columbia. Journal of Telemedicine & Telecare. 13(5): 251-6.

Scuffham PA, Steed M (2002) An economic evaluation of the Highlands and Islands teledentistry project. Journal of Telemedicine & Telecare. 8(3): 165-177.

Sezeur A, Degramount A, Touboul E, Mosnier H (2001) Teleconsultation before chemotherapy for recently operated on patients. American Journal of Surgery. 182(1): 49-51.

Shackley P (1993) An economic appraisal of alternative prenatal screening programmes for Down's syndrome. Journal of Public Health Medicine. 15(2): 175-184.

Shackley P, Cairns J (1996) Evaluating the benefits of antenatal screening: an alternative approach. Health Policy. 36(2): 103-115.

Shah BR, Laupacis A, Hux JE, Austin PC (2005) Propensity score methods gave similar results to traditional regression modelling in observational studies: a systematic review. Journal of Clinical Epidemiology. 58(6): 550-559.

Sharland G (2004) Routine fetal cardiac screening: what are we doing and what should we do? Prenatal Diagnosis. 24(13): 1123-1129.

Sharma S, Parness IA, Kamenir SA, Ko H, Haddow S, Steinberg LG, Lai WW (2003) Screening fetal echocardiography by telemedicine: Efficacy and community acceptance. Journal of American Society for Echocardiography. 16(3): 202-208.

Shih YCT, Kauf TL (1999) Reconciling decision models with the real world: an application to anaemia of renal failure. Pharmacoeconomics. 15(5): 481-493.

Sicotte C, Lehoux P, Van Doesburg N, Cardinal G, Leblanc Y (2004) A cost-effectiveness analysis of interactive paediatric telecardiology. Journal of Telemedicine & Telecare. 10(2): 78-83.

Simpson JM (2009) Impact of fetal echocardiography. Annals of Pediatric Cardiology. 2(1): 41-50.

Simpson J, Doze S, Urness D, Hailey D, Jacobs P (2001a) Telepsychiatry as a routine service - the perspective of the patient. Journal of Telemedicine & Telecare. 7(3): 155-160.

Simpson J, Doze S, Urness D, Hailey D, Jacobs P (2001b) Evaluation of a routine telepsychiatry service. Journal of Telemedicine & Telecare. 7(2): 90-98.

Sisk JE, Sanders JH (1998) A proposed framework for economic evaluation of telemedicine. Telemedicine Journal. 4(1):31-37.

Smith AC, Scuffham P, Wootton R (2007) The costs and potential savings of a novel telepaediatric service in Queensland. BMC Health Services Research 7(35): 1-7.

Specht JK, Wakefield B, Flanagan J (2001) Evaluating the cost of one telehealth application connecting an acute and long-term care setting. Journal of Gerontological Nursing. 27(1): 34-9.

Spertus JA, Winder JA, Dewhurst TA, Deyo RA, Prodzinski J, McDonell M, Fihn SD (1995) Development and evaluation of the Seattle Angina Questionnaire: A new functional status measure for coronary artery disease. Journal of the American College of Cardiology. 25(2):333-41.

Stalfors J, Bjorholt I, Westin T (2005) A cost analysis of participation via personal attendance versus telemedicine at a head and neck oncology multidisciplinary team meeting. Journal of Telemedicine & Telecare. 11(4): 205-210.

StataCorp (2007) Stata Statistical Software: Release 10.0. Stata Corporation, College Station, Texas, USA.

Steer P (2005) The epidemiology of preterm labour. British Journal of Obstetrics and Gynaecology. 112(1): 1-3.

Stensland J, Speedie SM, Ideker M, House J, Thompson T (1999) The relative cost of outpatient telemedicine services. Telemedicine Journal. 5(3): 245-256.

Stevens K J (2010) Assessing the performance of a new generic measure of health related quality of life for children and refining it for use in health state valuation. Health Economics and Decision Science Discussion Paper 10/02.

Stoeger A, Strohmayr W, Giacomuzzi SM, Dessl A, Buchberger W, Jaschke W (1997) A cost analysis of an emergency computerized tomography teleradiology system. Journal of Telemedicine & Telecare. 3(1): 35-39.

Stone RA, Obrosky DS, Singer DE, Kapoor WN, Fine MJ (1995) Propensity score adjustment for pretreatment differences between hospitalized and ambulatory patients with community acquired pneumonia. Medical Care 33(4): AS56-AS66.

Stümpflen I, Stümpflen A, Wimmer M, Bernaschek G (1996) Effect of detailed fetal echocardiography as part of routine prenatal ultrasonographic screening on detection of congenital heart disease. The Lancet. 348(9031): 854-857.

Sturm R, Jackson CA, Meredith LS, Yip W, Manning WG, Rogers WH, Wells KB (1995) Mental health care utilization in prepaid and fee-for-service plans among depressed patients in the medical outcomes study. Health Services Research. 30(2): 319-40.

Takizawa M, Sone S, Takashima S, Feng L, Maruyama Y, Hasegawa M, et al (1998) The mobile hospital-an experimental telemedicine system for the early detection of disease. Journal of Telemedicine & Telecare. 4(3): 146-151.

Thompson S, Wordsworth S, on behalf of the UK Working Party on Patient Costs (2001) An annotated cost questionnaire for completion by patients. HERU Discussion Paper (03/01). University of Aberdeen, Aberdeen.

The Criteria Committee of the New York Heart Association (1994) Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels (9th edition). Boston: Little, Brown & Co. 253-256.

Tiny Tickers Charity (2008) Fetal Heart Q&A, Echo UK. http://www.tinytickers.org/sonog5vprotocol.html. Last accessed 26/07/2010.

Torrance GW (1986) Measurement of health state utilities for economic appraisal. Journal of Health Economics. 5(1): 1-30.

Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q (1996) Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. 34(7): 702-722.

Trott P, Blignault I (1998) Cost evaluation of a telepsychiatry service in northern Queensland. Journal of Telemedicine & Telecare. 4(Suppl 1):66-68.

Tsitlakidis C, Mylonakis J, Niakas D (2005) Economic evaluation of telemedicine for a remotely located population: The case of two Greek islands. International Journal of Electronic Healthcare. 1(3): 243-260.

Tuulonen A, Ohinmaa T, Alanko HI, Hyytinen P, Juutinen A, Toppinen E (1999) The application of teleophthalmology in examining patients with glaucoma: a pilot study. Journal of Glaucoma. 8(6): 367-373.

Uebing A, Steer PJ, Yentis SM, Gatzoulis MA (2006) Pregnancy and congenital heart disease. British Medical Journal. 332(7538):401–6.

Van der Pol M, McKenzie L (2010) Costs and benefits of tele-endoscopy clinics in a remote location. Journal of Telemedicine and Telecare. 16(2): 89-94.

Varni JW, Seid M, Kurtin PS (2001) PedsQL™ 4.0: Reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient populations. Medical Care. 39(8): 800-812.

Vincent JA, Cavitt DL, Karpawich PP (1997) Diagnostic and cost-effectiveness of telemonitoring the pediatric pacemaker patient. Pediatric Cardiology. 18(2): 86-90.

Vogel WB, Rittman M, Bradshaw P, Nissen D, Anderson L, Bates B, Marshall C (2002) Outcomes from stroke rehabilitation in Veterans Affairs rehabilitation units: detecting and correcting for selection bias. Journal of Rehabilitation Research and Development. 39(3): 367-384.

Vuletic S (2001) Teleradiology in Croatia. Journal of Telemedicine & Telecare. 7(Suppl 1): 73-75.

Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM (2003) First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS). Health Technology Assessment. 7(11).

Walker WT, Temple IK, Gnanapragasam JP, Goddard JR, Brown EM (2002) Quality of life after repair of tetralogy of Fallot. Cardiology in the Young. 12(6): 549-553.

Whited JD, Datta S, Hall RP, Foy ME, Marbrey LE, Grambow SC, et al (2003) An economic analysis of a store and forward teledermatology consult system. Telemedicine Journal & E-Health. 9(4): 351-60.

Whited JD, Datta SK, Aiello LM, Aiello LP, Cavallerano JD, Conlin PR, et al (2005) A modeled economic analysis of a digital tele-ophthalmology system as used by three federal health care agencies for detecting proliferative diabetic retinopathy. Telemedicine Journal & E-Health. 11(6): 641-51.

Whitten P, Kingsley C, Grigsby J (2000) Results of a meta-analysis of cost-benefit research: is this a question worth asking? Journal of Telemedicine & Telecare. 6(Suppl 1): 4-6.

Whitten PS, Mair FS, Haycox A, May CR, Williams TL, Hellmich S (2002) Systematic review of cost effectiveness studies of telemedicine interventions. British Medical Journal. 324(7351): 1434-1437.

Whitten P, Johannessen LK, Soerensen T, Gammon D, Mackert M (2007) A systematic review of research methodology in telemedicine studies. Journal of Telemedicine & Telecare. 13(5): 230-235.

Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, Egmar AC, Greiner W, Gusi N, Herdman M, Jelsma J, Kind P, Scalone L, Ravens-Sieberer U (2010) Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Quality of Life Research. 19(6): 875-886.

Williams DL, Gelijns AC, Moskowitz AJ, Weinberg AD, Ng JH, Crawford E, Hayes CJ, Quaegebeur JM (2000) Hypoplastic left heart syndrome: valuing the survival. The Journal of Thoracic and Cardiovascular Surgery. 119(4): 720-731.

Williams TL, May CR, Esmail A (2001) Limitations of patient satisfaction studies in telehealthcare: a systematic review of the literature. Telemedicine Journal & E-Health. 7(4): 293-316.

Wittson CL, Benschoter R (1972) Two-way television: Helping the Medical Center Reach Out. American Journal of Psychiatry. 129(5): 624-627.

Wong HT, Poon WS, Jacobs P, Goh KYC, Leung CHS, Lau FL, et al (2006) The comparative impact of video consultation on emergency neurosurgical referrals. Neurosurgery. 59(3): 607-612.

Wootton R, Bloomer SE, Corbett R, Eedy DJ, Hicks N, Lotery HE, et al (2000) Multicentre randomised control trial comparing real time teledermatology with conventional outpatient dermatological care: societal cost-benefit analysis. British Medical Journal. 320(7244): 1252-1256.

World Bank - International Monetary Fund (2006) World Economic Outlook Database. September 2006. USA: Washington DC.

World Health Organisation (1998) Information technology in support of health care. Geneva: WHO. http://www.who.int/eht/en/InformationTech.pdf. Last accessed 26/07/2010.

Wu J, Kessler DK, Chakko S, Kessler KM (1995) A cost-effectiveness strategy for transtelephonic arrhythmia monitoring. American Journal of Cardiology. 75(2): 184-5.

Young TL, Ireson C (2003) Effectiveness of school-based telehealth care in urban and rural elementary schools. Pediatrics. 112(5): 1088-94.

Yount LE, Mahle WT (2004) Economic analysis of palivizumab in infants with congenital heart disease. Pediatrics. 114(6): 1606-1611.

Zincone LH Jr, Doty E, Balch DC (1997) Financial analysis of telemedicine in a prison system. Telemedicine Journal. 3(4): 247-255.

Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale. Acta Psychiatrica, Scandinavia. 67(6): 361-370.

Zollo S, Kienzle M, Loeffelholz P, Sebille S (1999) Telemedicine to Iowa's correctional facilities: initial clinical experience and assessment of program costs. Telemedicine Journal. 5(3): 291-301.

Zuckerman IH, Lee E, Wutoh AK, Xue Z, Stuart B (2006) Application of regression-discontinuity analysis in pharmaceutical health services research. Health Services Research. 41(2): 550-563.

Zundel KM (1996) Telemedicine: history, applications, and impact on librarianship. Bulletin of the Medical Library Association. 84(1): 71-79.

APPENDIX 1: SEARCH STRATEGY

MEDLINE search strategy

1. exp Cost-Benefit Analysis/ (41359)
2. (cost effective$ or cost-effective$).mp. (41927)
3. (cost utility$ or cost-utilit$).mp. (1123)
4. (cost benefit$ or cost-benefit$).mp. [mp=ti, ot, ab, nm, hw] (43973)
5. (willingness to pay or wtp or willingness-to-pay or willingness to accept or willingness-to-accept or net benefit or net-benefit or contingent valuation).mp. (1630)
6. (Pharmacoeconomic$ or pharmaco-economic$ or Economic analy$ or Economic evaluation$).mp. (7047)
7. (economic adj2 (evaluation$ or analy$ or study or studies)).mp. (6148)
8. (cost adj2 (evaluation$ or analy$ or study or studies or effective$ or benefit$ or utili$ or consequence$)).mp. (105992)
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (109709)
10. exp quality adjusted life year/ (3105)
11. quality adjusted life year.mp. (1131)
12. (QALY or QALYs).mp. (1872)
13. utilit$.mp. (63539)
14. (EuroQol or Euro Qol or Euro-Qol or EQ 5D or EQ5D or EQ-5D).mp. (1158)
15. (health utilities index or health-utilities-index or HUI).mp. (551)
16. (SF 6D or SF6D or SF-6D).mp. (83)
17. (quality of wellbeing or quality of well-being or QWB).mp. (216)
18. (health years equivalent or hyes or hye).mp. (48)
19. (time trade off or time trade-off or time-trade-off or TTO).mp. (518)
20. (standard gamble or standard-gamble or SG).mp. (3670)
21. (15 D or 15D).mp. (2182)
22. ((willing$ adj2 pay) or WTP).mp. (1256)
23. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (74290)
24. exp telemedicine/ (8670)
25. (telehealth or telecare or telemonitoring).mp. (954)
26. exp telecommunications/ (33471)
27. tele$.mp. (88576)
28. 24 or 25 or 26 or 27 (95719)
29. 9 or 23 (177685)
30. 28 and 29 (2907)
31. limit 30 to english language (2743)
32. limit 31 to humans (2279)
33. remove duplicates from 32 (2252)

CINAHL search strategy

1. exp Cost-Benefit Analysis/ (5920)
2. (cost effective$ or cost-effective$).mp. (7377)
3. (cost utility$ or cost-utilit$).mp. (165)
4. (cost benefit$ or cost-benefit$).mp. [mp=ti, ot, ab, nm, hw] (6483)
5. (willingness to pay or wtp or willingness-to-pay or willingness to accept or willingness-to-accept or net benefit or net-benefit or contingent valuation).mp. (220)
6. (Pharmacoeconomic$ or pharmaco-economic$ or Economic analy$ or Economic evaluation$).mp. (1703)
7. (economic adj2 (evaluation$ or analy$ or study or studies)).mp. (1761)
8. (cost adj2 (evaluation$ or analy$ or study or studies or effective$ or benefit$ or utili$ or consequence$)).mp. (16307)
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (17376)
10. exp quality adjusted life year/ (0)
11. quality adjusted life year.mp. (172)
12. (QALY or QALYs).mp. (232)
13. utilit$.mp. (4917)
14. (EuroQol or Euro Qol or Euro-Qol or EQ 5D or EQ5D or EQ-5D).mp. (488)
15. (health utilities index or health-utilities-index or HUI).mp. (152)
16. (SF 6D or SF6D or SF-6D).mp. (21)
17. (quality of wellbeing or quality of well-being or QWB).mp. (626)
18. (health years equivalent or hyes or hye).mp. (1)
19. (time trade off or time trade-off or time-trade-off or TTO).mp. (75)
20. (standard gamble or standard-gamble or SG).mp. (214)
21. (15 D or 15D).mp. (65)
22. ((willing$ adj2 pay) or WTP).mp. (170)
23. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (6515)
24. exp telemedicine/ (1740)
25. (telehealth or telecare or telemonitoring).mp. (1243)
26. exp telecommunications/ (24416)
27. tele$.mp. (20509)
28. 24 or 25 or 26 or 27 (30951)
29. 9 or 23 (23075)
30. 28 and 29 (843)
31. limit 30 to english language (838)
32. limit 31 to humans (838)
33. remove duplicates from 32 (838)

EMBASE search strategy

1. ('cost benefit'/exp OR 'cost benefit') AND ('analysis'/exp OR 'analysis') (17647)
2. ('cost effective' OR 'cost-effective') (16973)
3. ('cost utility' OR 'cost-utilit') (1820)
4. ('cost benefit' OR 'cost-benefit') (18066)
5. ('willingness to pay' OR 'WTP' OR 'willingness-to-pay' OR 'willingness to accept' OR 'willingness-to-accept' OR 'net benefit' OR 'net-benefit' OR ('contingent valuation'/exp OR 'contingent valuation')) (1132)
6. (pharmacoeconomic* OR 'pharmaco-economic' OR 'economic analy' OR 'economic evaluation') (34250)
7. (economic AND 2 AND (evaluation* OR analy* OR ('study'/exp OR 'study') OR studies)) (16423)
8. (('cost'/exp OR 'cost') AND 2 AND (evaluation* OR analy* OR ('study'/exp OR 'study') OR studies OR effective* OR benefit* OR utili* OR consequence*)) (61648)
9. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 (101419)
10. ('quality adjusted life year'/exp OR 'quality adjusted life year') (2950)
11. (('QALY'/exp OR 'QALY') OR QALYs) (3132)
12. utilit* (34913)
13. (EuroQol OR 'Euro Qol' OR 'Euro-Qol' OR 'EQ 5D' OR EQ5D OR 'EQ-5D') (992)
14. (('health'/exp OR 'health') AND utilities AND index OR health-utilities-index OR HUI) (2867)
15. (('SF'/exp OR 'SF') AND 6D OR SF6D OR 'SF 6D') (97)
16. (quality AND of AND ('wellbeing'/exp OR 'wellbeing') OR quality AND of AND ('well being'/exp OR 'well being') OR QWB) (5714)
17. (('health'/exp OR 'health') AND years AND equivalent OR hyes OR hye) (1472)
18. (('time'/exp OR 'time') AND trade AND off OR ('time'/exp OR 'time') AND 'trade off' OR time-trade-off OR TTO) (617)
19. (('standard'/exp OR 'standard') AND gamble OR 'standard gamble' OR SG) (15978)
20. ('15-D' OR '15D') (2830)
21. ((willing* AND 2 AND pay) OR WTP) (651)
22. 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 (66223)
23. ('telemedicine'/exp OR 'telemedicine') (1615)
24. (('telehealth'/exp OR 'telehealth') OR telecare OR ('telemonitoring'/exp OR 'telemonitoring')) (1010)
25. ('telecommunications'/exp OR 'telecommunications') (3502)
26. tele* (33618)
27. 23 OR 24 OR 25 OR 26 (33619)
28. 9 OR 22 (160784)
29. 27 AND 28 (2470)
30. 29 remove duplicates (2464)

Web of Science search strategy

1. TI = Cost-Benefit Analysis (1058)
2. TI = (cost effective* or cost-effective*) (8202)
3. TI = (cost utility* or cost-utilit*) (424)
4. TI = (cost benefit* or cost-benefit*) (2253)
5. TI =-(willingness to pay or wtp or willingness-to-pay or willingness to accept or willingness-to-accept or net benefit or net-benefit or contingent valuation). (1263)
6. TI = (Pharmacoeconomic* or pharmaco-economic* or Economic analy* or Economic evaluation*) (6166)
7. TI = (economic AND (evaluation* or analy* or study or studies)) (5625)
8. TI = (cost AND (evaluation* or analy* or study or studies or effective* or benefit* or utili* or consequence*)) (13740)
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (20952)
10. TI = quality adjusted life year (9)
11. TI = (QALY or QALYs) (98)
12. TI = utility* (15453)
13. TI = (EuroQol or Euro Qol or Euro-Qol or EQ 5D or EQ5D or EQ-5D) (156)
14. TI = (health utilities index or health-utilities-index or HUI) (158)
15. TI = (SF 6D or SF6D or SF-6D) (20)
16. TI = (quality of wellbeing or quality of well-being or QWB) (373)
17. TI = (health years equivalent or hyes or hye) (12)
18. TI = (time trade off or time trade-off or time-trade-off or TTO) (283)
19. TI = (standard gamble or standard-gamble or SG) (390)
20. TI = (15-D or 15D) (84)
21. TI = ((willing* AND pay) or WTP) (749)
22. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (17604)
23. TI = telemedicine (1079)
24. TI = (telehealth or telecare or telemonitoring) (307)
25. TI = telecommunications (2926)
26. TI = tele* (44461)
27. 24 or 25 or 26 or 27 (44461)
28. 9 or 23 (37319)
29. 27 and 28 (180)

APPENDIX 2: LITERATURE REVIEW RESULTS

No.	Authors	Publication Year	Country	Study Design	Sample size	Technology and comparators	Clinical area or application	Method	Viewpoint
1	Agha Z et al	1999	USA	Non-randomised	NS	TM vs. CC (on-site care or transfer by courier)	Pathology	CMA	Healthcare system
2	Agha Z et al	2002	USA	Non-randomised	65 patients	TM vs. CC (patient travel or on-site care)	Outpatient pulmonary care	CEA	Societal
3	Aoki N et al	2004	USA	Non-randomised	10,000 patients	TM vs. CC (on-site care)	Diabetes	CEA	Healthcare system
4	Armstrong AW et al	2007	USA	Non-randomised	451 patients	TM vs. CC (face-to-face clinic)	Dermatology	CMA	Healthcare system
5	Auerbach H et al	2006	Germany	Non-randomised	NS	TM (2 applications) vs. CC (no-TM)	A&E	CEA	Societal
6	Bailes JE et al	1997	USA	Non-randomised	100 patients	TM vs. CC (patient transfer)	Neurology	CA	NS
7	Barker G et al	2004	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Different health areas	Costing model	Healthcare system
8	Bergmo TS	1996	Norway	Non-randomised	6,000 patients per year	TM vs. CC (visiting specialist)	Radiology	CA	Healthcare system
9	Bergmo TS	1997	Norway	Non-randomised	100 patients per year	TM vs. CC (visiting specialist or patient travel)	Ear, nose & throat	CA	Societal
10	Bergmo TS	2000	Norway	Non-randomised	TM = 375; CC = 100	TM vs. CC (visiting specialist & patient travel or patient travel or on-site care)	Dermatology	CMA	Healthcare system
11	Berman M & Fenaughty A	2005	USA	Non-randomised	TM = 240; CC = 836	TM vs. CC (patient travel)	Ear, nose & throat	Costing model	NS
12	Bishai DM et al	2003	USA	Non-randomised	90 patients	TM vs. CC (on-site care or visiting specialist)	Cancer	CA	Societal
13	Bjorvig S et al	2002	Norway	Non-randomised	42 patients	TM vs. CC (face-to-face clinic)	Diabetes	CMA	NS
14	Bracale M et al	2002	Italy	Non-randomised	NS	TM vs. CC (patient transfer)	Different health areas	CCA	NS
15	Breslow MJ et al	2004	USA	Non-randomised	1,240 patients	Before and after TM	Intensive care	CCA	NS
16	Brumage MR et al	2001	USA	Non-randomised	TM = 323; CC = 33	TM vs. CC (patient transfer)	Radiology	CA	NS
17	Brunicardi BO	1998	USA	Non-randomised	NS	TM vs. CC (patient transfer)	Prison service	CA	Prison health service
18	Burgiss SG et al	1997	USA	Non-randomised	87 patients	Before and after TM	Dermatology	CA	NS
19	Bynum AB et al	2003	USA	Non-randomised	236 patients	Before and after TM	Different health areas	CA	Patient
20	Callahan CW et al	2005	USA	Non-randomised	267 patients	TM vs. CC (patient transfer)	Paediatrics	CA	NS
21	Castillo-Riquelme MC et al	2004	UK	Non-randomised	235 patients	TM vs. CC (visiting specialist or transfer of images)	Ophthalmology	CEA	Healthcare system
22	Chan HH et al	2000	Hong Kong	Non-randomised	74 patients	TM vs. CC (patient transfer or visiting specialist)	Dermatology	CA	NS
23	Chodroff PH	1999	USA	Non-randomised	329 patients	TM vs. CC (patient transfer)	Neurology	CA	NS
24	Chua R et al	2001	Northern Ireland	Randomised	TM = 76; CC = 65	TM vs. CC (face-to-face clinic)	Neurology	CA	NS
25	Crowe BL et al	1996	Australia	Non-randomised	NS	TM vs. CC (patient transfer)	Radiology	CA	NS
26	Darkins A et al	1996	UK and Ireland	Non-randomised	51 TM patients	TM vs. CC (on-site care)	Minor injuries	CA	NS
27	Daucourt V et al	2006	France	Non-randomised	NS	TM vs. CC (patient transfer)	Radiology	CMA	Healthcare system
28	Davis MC	1997	USA	Non-randomised	2,000 patients per year	TM vs. CC (transfer by courier)	Radiology	CA	NS
29	Davis P et al	2001	Canada	Non-randomised	52 patients	TM vs. CC (face-to-face clinic)	Rheumatology	CA	NS
30	de la Torre A et al	2004	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Different health areas	CA	Societal
31	Deodhar J	2002	India	Non-randomised	182 patients	TM vs. CC (patient transfer)	Neonatal care	CA	NS

32	Doolittle GC et al	1998	USA	Non-randomised	TM = 103; CC: patient travel = 2,400 & on-site care = 81	TM vs. CC (face-to-face clinic or visiting specialist)	Cancer	CA	NS
33	Doolittle GC	2000	USA	Non-randomised	NS	TM vs. CC (on-site care)	Hospice care	CA	NS
34	Doolittle GC et al	2003	USA	Non-randomised	286 patients	TM vs. CC (on-site care)	Paediatrics	CA	NS
35	Doolittle GC et al	2004	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Cancer	CA	NS
36	Doze S et al	1999	Canada	Non-randomised	90 patients	TM vs. CC (visiting specialist)	Psychiatry	CA	NS
37	Elford DR et al	2001	Canada	Non-randomised	30 patients	TM vs. CC (face-to-face clinic)	Psychiatry	CA	NS
38	Ferris DG et al	2004	USA	Non-randomised	264 patients	TM vs. CC (on-site care)	Cancer	CA	NS
39	Finley JP et al	1997	Canada	Non-randomised	135 patients	TM vs. CC (patient transfer)	Cardiology	CA	NS
40	Halvorsen PA & Kristiansen IS	1996	Norway	Randomised	597 patients	TM vs. CC (on-site care or patient travel)	Radiology	CMA	Societal
41	Harley J	2006	UK	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Psychiatry	CA	NS
42	Harno KS	1999	Finland	Non-randomised	NS	TM vs. CC (patient transfer or on-site care)	Different health areas	CA	NS
43	Harno K et al	2000	Finland	Non-randomised	292 patients	TM vs. CC (face-to-face clinic)	Different health areas	CMA	NS
44	Harno K et al	2001	Finland	Non-randomised	TM = 57; CC = 362	TM vs. CC (face-to-face clinic)	Orthopaedics	CMA	NS
45	Hassall S et al	2003	Australia	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Different health areas	CA	NS
46	Himpens B	2003	Belgium	Non-randomised	NS	TM vs. CC (patient travel)	Administration, education & clinical	CA	NS
47	Hui E & Woo J	2002	Hong Kong	Non-randomised	NS	TM vs. CC (face-to-face clinic or on-site care)	Care for elderly	CA	NS
48	Ippolito A et al	2003	Italy	Non-randomised	TM = 162; CC = 2,710 deliveries.	TM vs. CC (on-site care)	Cardiology	CA	NS
49	Jacklin PB et al	2003	UK	Randomised	TM = 1,051; CC = 1,043	TM vs. CC (face-to-face clinic)	Patients who visited the GP	CCA	Societal
50	Jin AJ et al	2004	Canada	Non-randomised	339 patients	TM vs. CC (patient travel)	Diabetes	CEA	NS
51	Johnston K et al	2004	UK & South Africa	Non-randomised	90 patients	Use of TM in 2 countries	Ophthalmology	CEA	NS
52	Jong M	2004	Canada	Non-randomised	71 patients	TM vs. CC (face-to-face clinic)	Mental health	CA	NS
53	Kesler C & Balch D	1995	USA	Non-randomised	NS	TM vs. CC (patient transfer)	Prison service	CA	NS
54	Kildemoes HW & Kristiansen IS	2004	Denmark	Non-randomised	NS	TM vs. CC (public campaign only)	Cardiology	CEA	Healthcare system
55	Kitt SM & Clayton L	2002	Australia	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Different health areas	CA	NS
56	Kristo DA et al	2001	USA	Non-randomised	54 patients	TM vs. CC (on-site care)	Sleep studies	CA	NS
57	Kumar S et al	2006	Australia	Non-randomised	118 patients	TM vs. CC (face-to-face clinic or on-site care or patient transfer)	Ophthalmology	CA	Healthcare system
58	Lamminen H et al	2000	Finland	Non-randomised	25 patients	TM vs. CC (patient travel)	Dermatology	CA	Healthcare system
59	Lamminen H et al	2001	Finland	Non-randomised	TM = 42; CC = 249	TM vs. CC (patient travel)	Ophthalmology & dermatology	CA	NS
60	Loane MA et al	2000	Northern Ireland	Non-randomised	96 patients	TM vs. CC (transfer by courier)	Dermatology	CA	Societal
61	Loane MA et al	2001a	Northern Ireland	Randomised	TM = 126; CC = 148	TM vs. CC (face-to-face clinic)	Dermatology	CA	Societal
62	Loane MA et al	2001b	New Zealand	Randomised	TM = 109; CC = 94	TM vs. CC (face-to-face clinic)	Dermatology	CMA	Societal
63	Malone FD et al	1998	USA	Non-randomised	300 patients	TM vs. CC (transfer by courier)	Pregnant women	CA	NS
64	Marcin JP et al	2004	USA	Non-randomised	TM = 47; CC = 90	TM vs. CC (patient transfer)	Paediatrics	CA	NS
65	McCue MJ et al	1997	USA	Non-randomised	165 patients	TM vs. CC (patient transfer)	HIV	CA	NS

66	McCue MJ et al	1998	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Different health areas	CMA	Prison health service
67	McCue MJ et al	2000	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Cardiology	CA	NS
68	McIntosh WA et al	2003	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Different health areas	CA	NS
69	Mielonen ML et al	2000	Finland	Non-randomised	14 patients	TM vs. CC (face-to-face clinic)	Psychiatry	CA	Societal
70	Modai I et al	2006	Israel	Non-randomised	TM = 39; CC = 42	TM vs. CC (face-to-face clinic)	Psychiatry	CA	NS
71	Navein J et al	1999	USA and Europe	Non-randomised	2000 air referrals	TM vs. CC (patient transfer)	Different health areas	CA	NS
72	Nguyen LT et al	2004	USA	Non-randomised	294 patients	TM vs. CC (face-to-face clinic)	Burns	CA	NS
73	Noble SM et al	2005	UK	Randomised	TM = 191; CC = 62	TM vs. CC (face-to-face clinic)	Minor injuries	CCA	Societal
74	Norum J et al	2007	Norway	Non-randomised	130 patients	TM vs. CC (face-to-face clinic)	Pregnant women	CA	Societal
75	Ohinma A & Scott R	2006	Canada	Non-randomised	NS	TM for various activities	Administration, education & clinical	Costing model	Healthcare system
76	Ohinmaa A et al	2002	Finland	Randomised	TM = 69; CC = 76	TM vs. CC (face-to-face clinic)	Orthopaedics	CMA	NS
77	Peng PW et al	2006	Canada	Non-randomised	8 patients	TM vs. CC (face-to-face clinic)	Chronic pain	CA	NS
78	Persaud DD et al	2005	Canada	Non-randomised	TM = 86; CC = 129	TM vs. CC (face-to-face clinic)	Psychiatry and dermatology	CA	Societal
79	Preston J	1995	USA	Non-randomised	NS	Before and after TM	Different health areas	CA	NS
80	Redlick F et al	2002	Canada	Non-randomised	NS	TM vs. CC (patient travel)	Burns	CA	NS
81	Rendina MC et al	1998	USA	Non-randomised	TM = 48; CC = 39	TM vs. CC (transfer by courier)	Cardiology	CA	Healthcare system
82	Rendina MC et al	2001	USA	Non-randomised	2,142 infants admitted to NICU	TM vs. CC (face-to-face clinic)	Cardiology	CA	NS
83	Rumpsfeld M et al	2005	Norway	Non-randomised	9 patients	TM vs. CC (patient travel)	Dialysis	CMA	Societal
84	Ruskin PE et al	2004	USA	Randomised	TM = 59 ; CC = 60	TM vs. CC (on-site care)	Psychiatry	CCA	Healthcare system
85	Santamaria N et al	2004	Australia	Randomised	93 patients	TM vs. CC (on-site care)	Wound care	CCA	NS
86	Schaafsma J et al	2007	Canada	Non-randomised	NS	TM vs. CC (patient travel)	Administration, education & clinical	CA	Societal
87	Scuffham PA & Steed M	2002	UK	Non-randomised	18 patients	TM vs. CC (on-site care vs. patient travel)	Dentistry	CMA	Societal
88	Sezeur A et al	2001	France	Non-randomised	16 patients	TM vs. CC (patient transfer)	Cancer	CA	NS
89	Sicotte C et al	2004	Canada	Non-randomised	78 patients	TM vs. CC (on-site care or patient travel)	Cardiology	CEA	Societal
90	Simpson J et al	2001a	Canada	Non-randomised	230 patients	TM vs. CC (patient travel)	Psychiatry	CA	Patient
91	Simpson J et al	2001b	Canada	Non-randomised	379 patients	TM vs. CC (on-site care)	Psychiatry	CA	Healthcare system
92	Smith AC et al	2007	Australia	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Paediatrics	CA	Healthcare system
93	Specht JK et al	2001	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Chronic pain	CMA	Societal
94	Stalfors J et al	2005	Sweden	Non-randomised	TM = 45; CC = 39	TM vs. CC (face-to-face clinic)	Cancer	CMA	NS
95	Stensland J et al	1999	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Orthopaedic & dermatology	CA	Societal
96	Stoeger A et al	1997	Austria	Non-randomised	116 patients	TM vs. CC (patient transfer or transfer by courier)	Radiology	CA	NS
97	Takizawa M et al	1998	Japan	Non-randomised	TM = 18; CC = 20	TM vs. CC (face-to-face clinic)	Lung disease	CA	NS
98	Trott P & Blignault I	1998	Australia	Non-randomised	50 patients	TM vs. CC (patient transfer)	Psychiatry	CA	NS
99	Tsitlakidis C et al	2005	Greece	Non-randomised	38 patients	TM vs. CC (patient transfer)	General health advice	CMA	NS
100	Tuulonen A et al	1999	Finland	Non-randomised	TM = 29; CC = 41	TM vs. CC (face-to-face clinic)	Ophthalmology	CA	NS
101	Vincent JA et al	1997	USA	Non-randomised	96 patients	TM vs. CC (face-to-face clinic or on-site care)	Cardiology	CEA	NS
102	Vuletic S	2001	Croatia	Non-randomised	442 patients	TM vs. CC (patient transfer)	Radiology	CA	NS
103	Whited JD et al	2003	USA	Randomised	TM = 135; CC =	TM vs. CC (face-to-face clinic)	Dermatology	CA &	Healthcare

					140			CEA	system
104	Whited JD et al	2005	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Ophthalmology	CEA	Healthcare system
105	Wong HT et al	2006	China	Randomised	TM = 239; CC = 471	TM vs. CC (face-to-face clinic or a video-consultation)	Neurology	CEA	NS
106	Wootton R et al	2000	Northern Ireland	Randomised	TM = 102; CC = 102	TM vs. CC (face-to-face clinic)	Dermatology	CBA	Societal
107	Young TL & Ireson C	2003	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Acute care	CA	NS
108	Zincone LH Jr et al	1997	USA	Non-randomised	NS	TM vs. CC (face-to-face clinic)	Prison service	CA	Societal
109	Zollo S et al	1999	USA	Non-randomised	274 patients	TM vs. CC (patient transfer)	Different health areas	CA	Prison health service

APPENDIX 2: RESULTS FROM THE LITERATURE REVIEW (CONTINUED)

No.	Authors	Publication Year	Time frame	Costs	Patient costs	Benefits	Discounting	Sensitivity analyses	Cost results
1	Agha Z et al	1999	1 year	DMC	None	None	Yes	Yes	Courier method < TM & TM < on-site care
2	Agha Z et al	2002	1 year	DMC, DNMC, IC	Travel and accommodation costs	Healthcare access for patients who received outpatient consultations	Yes	Yes	TM < CC
3	Aoki N et al	2004	NS	DMC	None	QALYs calculated using a Markov model	Yes	Yes	TM < CC
4	Armstrong AW et al	2007	18 months	DMC	None	None	No	Yes	TM < CC
5	Auerbach H et al	2006	11 years	DMC, DNMC	Personal injury costs i.e. hospital admissions and rehabilitation	Life years gained	Yes	Yes	Both TM methods > CC
6	Bailes JE et al	1997	1 year	DMC	None	None	No	No	TM < CC
7	Barker G et al	2004	1 year	DMC	None	None	No	No	TM > CC
8	Bergmo TS	1996	NS	DMC	None	None	No	Yes	TM < CC
9	Bergmo TS	1997	1 year	DMC, DNMC	Travel costs, subsistence costs and cost of care to look after child or adult.	None	No	No	TM cost saving above certain patient throughput
10	Bergmo TS	2000	1 year	DMC, DNMC	Travel costs	None	No	Yes	TM < CC
11	Berman M & Fenaughty A	2005	2 years	DMC, DNMC	Travel, food & accommodation costs	None	No	No	TM < CC
12	Bishai DM et al	2003	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM > CC
13	Bjorvig S et al	2002	1 year	DMC	None	None	No	No	TM < CC
14	Bracale M et al	2002	2 years	DMC	None	None	No	No	TM < CC
15	Breslow MJ et al	2004	18 months	DMC	None	Mortality and length of stay	No	No	TM < CC
16	Brumage MR et al	2001	2 months	DMC, PSav	None	Transfers and hospitalisations avoided	No	Yes	TM < CC
17	Brunicardi BO	1998	1 year	DMC	None	None	No	No	TM cost saving above certain patient throughput
18	Burgiss SG et al	1997	17 months	DMC	None	None	No	No	TM < CC
19	Bynum AB et al	2003	2 years	DNMC, IC	Travel costs	None	No	No	TM reduces patients costs
20	Callahan CW et al	2005	1 year	DMC, DNMC, PSav	Travel and accommodation costs	None	No	No	TM < CC
21	Castillo-Riquelme MC et al	2004	16 months	DMC	None	QALYs based on survival and utility weights	Yes	Yes	TM > CC
22	Chan HH et al	2000	NS	DMC	None	None	No	No	TM < CC
23	Chodroff PH	1999	35 months	DMC	None	None	No	No	TM < CC
24	Chua R et al	2001	NS	DMC, DNMC	Travel costs	None	No	No	TM > CC
25	Crowe BL et al	1996	3 months	DMC	None	None	No	No	TM has potential to be cost saving
26	Darkins A et al	1996	1 year	DMC	None	None	No	No	TM < CC
27	Daucourt V et al	2006	1 year	DMC, PSav	None	Number of transfers, hospitalisations & consultations avoided	No	Yes	TM has potential to be cost saving
28	Davis MC	1997	2 years	DMC	None	None	No	No	TM < CC
29	Davis P et al	2001	NS	DMC	None	None	No	No	TM costs were similar to CC

30	de la Torre A et al	2004	1 year	DMC, DNMC, IC	Travel costs	None	No	No	TM > CC
31	Deodhar J	2002	18 months	DMC	None	Number of transfers, hospitalisations and consultations avoided	No	No	TM < CC
32	Doolittle GC et al	1998	1 year	DMC	None	None	No	No	Patient travel< TM & TM < on-site care
33	Doolittle GC	2000	NS	DMC	None	None	No	No	TM < CC
34	Doolittle GC et al	2003	9 months	DMC, DNMC	Travel costs & out-of-pocket expenses	None	No	No	TM cost saving above certain patient throughput
35	Doolittle GC et al	2004	NS	DMC	None	None	No	No	TM has potential to be cost saving
36	Doze S et al	1999	9 months	DMC	None	None	No	Yes	TM has potential to be cost saving
37	Elford DR et al	2001	3 months	DMC, DNMC	Travel, food & accommodation costs	None	No	No	TM < CC
38	Ferris DG et al	2004	23 months	DMC	None	None	No	No	TM > CC
39	Finley JP et al	1997	2 years	DMC, DNMC, PSav	Travel and accommodation costs	None	No	No	TM < CC
40	Halvorsen PA & Kristiansen IS	1996	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM > CC
41	Harley J	2006	6 months	DMC, DNMC	Travel costs	None	No	Yes	TM < CC
42	Harno KS	1999	1 year	DMC	None	None	No	No	TM has potential to be cost saving
43	Harno K et al	2000	8 months	DMC, DNMC	Travel costs	None	No	No	TM < CC
44	Harno K et al	2001	8 months	DMC, DNMC	Travel costs	None	No	No	TM < CC
45	Hassall S et al	2003	3 months	DMC	None	None	No	Yes	TM cost saving above certain patient throughput
46	Himpens B	2003	3 years	DMC, DNMC	Travel costs	None	No	No	TM reduces patients costs
47	Hui E & Woo J	2002	1 year	DMC	None	None	No	No	TM < CC
48	Ippolito A et al	2003	1 year	DMC	None	None	No	No	TM costs were similar to CC
49	Jacklin PB et al	2003	6 months	DMC, DNMC, IC	Travel costs	SF-12 as a utility measure	No	Yes	TM > CC
50	Jin AJ et al	2004	1 year	DMC, DNMC	Travel, food & accommodation costs	None	No	No	TM < CC
51	Johnston K et al	2004	1 year	DMC	None	DALYs	No	Yes	TM has potential to be cost saving
52	Jong M	2004	1 year	DMC, DNMC	Travel, food & accommodation costs	None	No	No	TM < CC
53	Kesler C & Balch D	1995	NS	DMC	None	None	No	No	TM has potential to be cost saving
54	Kildemoes HW & Kristiansen IS	2004	Lifetime	DMC	None	Life years gained	Yes	Yes	TM > CC
55	Kitt SM & Clayton L	2002	NS	DMC	None	None	No	No	TM < CC
56	Kristo DA et al	2001	7 months	DMC	None	None	No	No	TM < CC
57	Kumar S et al	2006	1 year	DMC	None	None	No	Yes	TM < CC
58	Lamminen H et al	2000	8 months	DMC, DNMC	Travel costs	None	No	No	TM costs were similar to CC
	Lamminen H et al	2001	9 months	DMC,	Travel costs	None	No	No	TM cost saving above certain

59				DNMC					patient throughput
60	Loane MA et al	2000	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM has potential to be cost saving
61	Loane MA et al	2001a	2 years	DMC, DNMC, IC	Travel costs	None	No	Yes	TM > CC
62	Loane MA et al	2001b	10 months	DMC, DNMC, IC	Travel costs	None	No	No	TM < CC
63	Malone FD et al	1998	2 months	DMC	None	None	No	No	TM has potential to be cost saving
64	Marcin JP et al	2004	2 years	DMC	None	None	No	No	TM has potential to be cost saving
65	McCue MJ et al	1997	7 months	DMC	None	None	No	Yes	TM < CC
66	McCue MJ et al	1998	1 year	DMC	None	None	No	No	TM < CC
67	McCue MJ et al	2000	3 years	DMC	None	None	No	No	TM cost saving above certain patient throughput
68	McIntosh WA et al	2003	2 years	DNMC, IC, PSav	Cost to patient for specialty consultation	None	No	No	TM > CC
69	Mielonen ML et al	2000	11 months	DMC	None	None	No	No	TM < CC
70	Modai I et al	2006	1 year	DMC, DNMC	Travel costs	Patient satisfaction was measured	No	No	TM > CC
71	Navein J et al	1999	NS	DMC, DNMC	Travel and accommodation costs	None	No	No	TM has potential to be cost saving
72	Nguyen LT et al	2004	5 years	DMC	None	None	No	No	TM < CC
73	Noble SM et al	2005	NS	DMC, DNMC	Travel and subsistence costs	Safety and clinical effectiveness at 7 days after presentation	No	Yes	TM > CC
74	Norum J et al	2007	8 months	DMC, DNMC, IC	Travel costs	None	No	No	TM > CC
75	Ohinma A & Scott R	2006	1 year	DMC	None	None	Yes	Yes	TM > CC
76	Ohinmaa A et al	2002	NS	DMC, DNMC, IC	Travel costs and cost of home help	None	No	Yes	TM cost saving above certain patient throughput
77	Peng PW et al	2006	NS	DNMC	Travel, food & accommodation costs	None	No	No	TM < CC
78	Persaud DD et al	2005	NS	DMC, DNMC	Travel costs, out-of-pocket and accommodation costs	None	No	No	TM > CC
79	Preston J	1995	NS	DMC, PSav	None	None	No	No	TM has potential to be cost saving
80	Redlick F et al	2002	1 year	DMC, DNMC	Travel, food, childcare and accommodation costs	Patient satisfaction was measured	No	No	TM < CC
81	Rendina MC et al	1998	NS	DMC	None	Mortality, length of stay and transfer rate	No	No	TM > CC
82	Rendina MC et al	2001	3 years	PSav	None	None	No	No	TM has potential to be cost saving
83	Rumpsfeld M et al	2005	8 months	DMC, DNMC, PSav	Travel costs	Travel and hospitalisations avoided	No	No	TM > CC
84	Ruskin PE et al	2004	6 months	DMC	None	Scales for depression, anxiety, functioning and SF-12.	No	No	TM > CC
85	Santamaria N et al	2004	12 months	DMC, DNMC	Travel costs	Healing rate difference	No	No	TM < CC
	Schaafsma J et al	2007	17 months	DMC,	Travel, food &	None	No	Yes	TM has potential to be cost

86				DNMC, IC	accommodation costs				saving
87	Scuffham PA & Steed M	2002	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM > on-site care & TM < patient travel
88	Sezeur A et al	2001	NS	DMC	None	None	No	No	TM < CC
89	Sicotte C et al	2004	4 years	DMC, DNMC	Travel costs	Avoidance of patient journeys	No	Yes	TM > CC
90	Simpson J et al	2001a	NS	DNMC, IC	Travel costs	Patient satisfaction was measured	No	No	TM < CC
91	Simpson J et al	2001b	2 years	DMC	None	None	No	Yes	TM cost saving above certain patient throughput
92	Smith AC et al	2007	5 years	DMC, DNMC	Travel and accommodation costs	Travel avoided	No	Yes	TM < CC
93	Specht JK et al	2001	NS	DMC, DNMC, PSav	Travel costs	None	No	No	TM < CC
94	Stalfors J et al	2005	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM < CC
95	Stensland J et al	1999	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM cost saving above certain patient throughput
96	Stoeger A et al	1997	13 months	DMC	None	None	No	No	TM < patient transfer & TM > courier transfer
97	Takizawa M et al	1998	11 months	DMC	None	None	No	No	TM < CC
98	Trott P & Blignault I	1998	1 year	DMC	None	None	No	No	TM < CC
99	Tsitlakidis C et al	2005	1 year	DMC, DNMC, IC	Travel and accommodation costs	None	No	Yes	TM cost saving above certain patient throughput
100	Tuulonen A et al	1999	NS	DMC	None	None	No	No	TM costs were similar to CC
101	Vincent JA et al	1997	3 years	DMC	None	None	No	No	TM < CC
102	Vuletic S	2001	2 years	PSav	None	None	No	No	TM < CC
103	Whited JD et al	2003	1 year	DMC, DNMC	Travel costs	Effectiveness measured in days to intervention	No	Yes	TM > CC
104	Whited JD et al	2005	1 year	DMC	None	No. of patients who needed treatment & number of cases of severe vision loss averted	No	Yes	TM > CC
105	Wong HT et al	2006	3 years	DMC	None	Range of outcomes including mortality	No	No	TM costs were similar to CC
106	Wootton R et al	2000	1 year	DMC, DNMC, IC	Travel costs	None	No	Yes	TM > CC
107	Young TL & Ireson C	2003	2 years	DMC, DNMC, PSav	Travel costs	None	No	No	TM reduces patients costs
108	Zincone LH Jr et al	1997	NS	DMC	None	None	No	Yes	TM cost saving above certain patient throughput
109	Zollo S et al	1999	1 year	DMC	None	None	No	No	TM cost saving above certain patient throughput

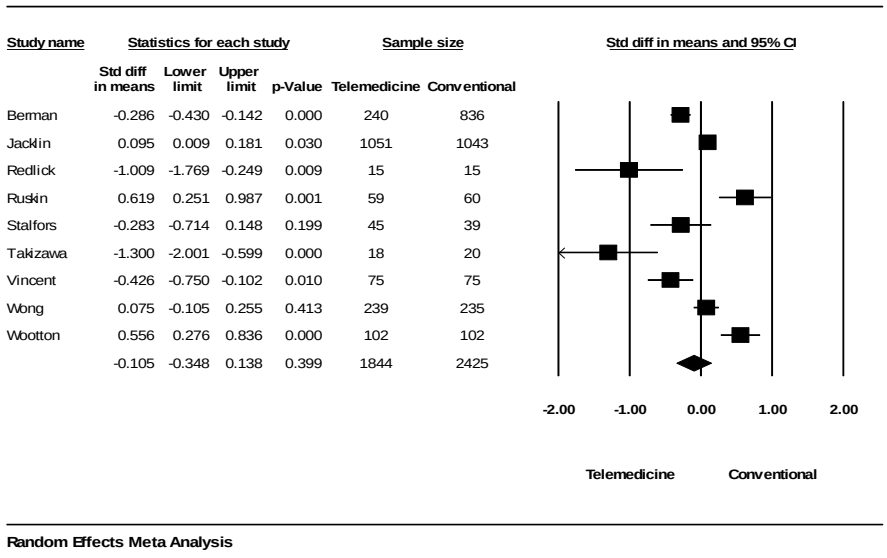
Key:
Method: CA = cost analysis; CCA = cost-consequences analysis; CBA = cost-benefit analysis; CMA = cost-minimisation analysis; CEA = cost-effectiveness analysis
Technology & comparators: TM = telemedicine; CC = Conventional Care
Conventional Care: on-site care = care at local hospital or in an outreach clinic; patient travel or face-to-face clinic = patient travels to specialist hospital for assessment; patient transfer = patient is transferred to specialist hospital for treatment/care; visiting specialist = visiting specialist visits local hospital
Costs: DMC = direct medical costs; DNMC = direct non-medical costs; IC = indirect costs; PSav = potential savings
NS = not stated

APPENDIX 3: RESULTS FROM THE META-ANALYSIS

Of the 109 studies, 100 studies (91.7%) had insufficient statistical information in order for a meta-analysis to be conducted. Of those studies which were excluded the number of studies reporting only: 1) total costs, mean costs and sample size; and 2) means costs and sample size, for the telemedicine arm were: 37 (37.0%) and 20 (20.0%), and for the conventional arm were: 28 (28.0%) and 23 (23.0%) respectively. However, when looking at both the telemedicine and control arms, these figures were: 28 (28.0%); and 18 (18.0%). If those studies which reported mean costs and sample sizes, also included either standard deviations or p- or t- values from statistical tests, then these studies could have been included in the meta-analysis.

Only nine studies could be included in the meta-analysis. Of these 7 studies [Jacklin et al, 2003; Redlick et al, 2002; Ruskin et al, 2004; Stalfors et al, 2005; Takizawa et al, 1998; Vincent et al, 1997; Wong et al, 2006] reported information on mean costs, sample sizes and p-values. A further two studies [Berman and Fenaughty, 2005; Wootton et al, 2000] reported information on mean costs, standard deviations and sample sizes, and using this information a p-value was computed.

Figure A1: Random effects meta-analysis of telemedicine cost studies (using means, samples sizes and p-values)



Figures A1 above shows the random effects meta-analysis results from the 9 studies. Overall the random effects meta-analysis concludes that telemedicine is cheaper than conventional care, although differences are not significant (-10.5% (95% CI: -34.8% to

13.8%), p = 0.399). Four studies were significantly ($p \leq 0.05$) in favour of telemedicine [Berman and Fenaughty, 2005; Redlick et al, 2002; Takizawa et al, 1998; Vincent et al, 1997]; whereas three studies were significantly in favour of the conventional arm [Jacklin et al, 2003; Ruskin et al, 2004; Wootton et al, 2000]. The meta-analysis for the telemedicine cost studies using a random effects procedure most closely reflects the health services being analysed as the sample sizes of the various studies were different and the procedure has the ability to control for between study variations. The results also showed that there was significant heterogeneity between the studies (heterogeneity test: $Q = 74.73$, $df = 8$, $p < 0.001$) and that there was also evidence of funnel plot asymmetry.

Conclusions cannot be made precisely from these nine studies as one can't be 100% sure that the costs included in each study have been collected appropriately i.e. that each of the studies have defined costs in the same way; that they used similar methods to collect and analyse the cost data; and finally, the same valuation technique was used for the costs. As telemedicine has been used in a different context for each study, it makes it quite difficult to come to some generalisable conclusions.

APPENDIX 4: USE OF TELEMEDICINE IN SPECIFIC HEALTH AREAS

Detailed below are studies of telemedicine interventions in four specific health areas which were identified in the literature review: ten studies in dermatology, nine studies in psychiatry and eight studies each in radiology and cardiology.

a) Dermatology. The first study by Armstrong and colleagues (2007) was a cost-minimisation analysis of interactive teledermatology compared with conventional care (outpatient face-to-face dermatology clinic) in the USA. The study was conducted for a period of 18 months and 451 new patient and follow-up visits were made to the interactive teledermatology clinic. The two hour clinic sessions occurred weekly and in each session, four patients were seen in an hour. The costs of teledermatology included the costs of the videoconferencing equipment, clinic space, personnel and overhead costs. For conventional clinics, the costs included room space, overheads and staff costs. Total hourly operating costs for teledermatology were lower than conventional clinic (US$274 vs. US$346). The lower cost of teledermatology was due to the low cost of the technology and reduced rental charge for clinic space in rural areas. The authors concluded that from a healthcare provider perspective, teledermatology can be economically viable means of providing dermatological care to remote regions.

Bergmo (2000) compared real-time teledermatology to three alternative methods of providing dermatology care to patients living in remote areas in northern Norway. The alternatives were a combination of visiting service and patient travel; patient travel to the nearest secondary care centre; and a locally employed dermatologist. Three hundred and seventy five patients were seen by teledermatology and 100 patients visited the outpatient clinic during the course of the year. The costs for teledermatology included a videoconferencing unit including two monitors, connected by ISDN lines, video camera, a phototherapy machine and staff costs. Costs for conventional care included travel costs and salary costs for the specialist, GP and nurse. Costs common to all methods, such as drugs costs and hospitalization costs were excluded. The cost-minimisation analysis from the healthcare perspective found that for a workload of 375 patients, teledermatology cost NKr470,780; while the visiting service and patient travel to hospital cost NKr880,530; patient travel to the nearest secondary-care centre cost NKr1,635,075; and a locally employed dermatologist cost NKr958,660. The authors found that generalisation of the study results, depends on the workload, whether the visiting specialist service is an available option and the actual travel costs. The cost analysis shows that telemedicine was the cheapest method for workloads above 195 per year.

The effect of teledermatology consultations on the cost of care for patients in rural areas in the USA was assessed by Burgiss and colleagues (1997). The study was a before and after comparison of teledermatology. In total, 119 visits were made by the 87 patients who were referred for teledermatology over a 17-month period. The telemedicine costs included the costs of the video equipment and special cameras for observing dermatologic lesions. Other costs included the hospital room costs, medication costs, procedure and test costs and staff costs. The cost analysis found that the average cost of care for all patients during an average period of 8 months prior to teledermatology was US$294, compared with US$141 for the 6 months after diagnosis by teledermatology. The authors stated that teledermatology can be effective for dermatology consultation in new patients and that the data indicated teledermatology can decrease the cost of care for the diagnosed conditions.

A pilot study to look at the costs of establishing a teledermatology centre for a community in Hong Kong was conducted by Chan and colleagues (2001). Teledermatology at an elderly person's centre was compared to sending patients to a speciality clinic and monthly visits from a dermatologist. Seventy-four patients were recruited into the study. The teledermatology costs included the costs of the equipment, maintenance costs, cost of the ISDN lines and installation and staff costs. The conventional costs included the costs of the transportation, escort costs and staff costs. The cost analysis found that the cost per patient for teledermatology was HK$57.70; cost per patient for sending patients to a specialty clinic was HK$322.80; and total cost per patient of sending a dermatologist to a centre was HK$445.90. Therefore, the total saving per patient seen via teledermatology was HK$265.10. The authors found that teledermatology is a cost-saving means for providing a service to elderly persons.

Lamminen et al (2000) conducted a study from a health service perspective of real-time teledermatology using low cost equipment compared to conventional consultations in Finland. Costs for the real-time teledermatology included the cost of the videoconferencing equipment, annual maintenance cost, document camera, a dermatoscope, telecommunication costs and staff costs. For the conventional consultations the costs included cost of specialist and GP, and cost of travel. The average time the patient spent in travelling for the conventional consultations was 24 minutes (one-way), whereas the mean duration of a teleconsultation was 15 minutes. Over the eight months, the cost analysis found that the cost of teleconsultations for the 18 patients who avoided travel was FM18,627 and the cost of the 18 conventional

consultations would have been FM18,034. The authors concluded that the equipment was generally reliable and easy to use and that the main economic benefits of videoconferencing were attributable to the reduced travelling and hospital costs.

A cost analysis comparing real-time versus store-and-forward teledermatology in Northern Ireland was conducted by Loane et al (2000). During the course of the year, 96 patients attended their GP health centre and were seen by a hospital dermatologist over the video link (real-time teledermatology). Before the video link started, the GP took photographs and posted them with a referral letter to a different hospital dermatologist (store-and-forward). Fixed costs included the cost of the videoconferencing unit, installation of ISDN lines, and a video camera for real-time consultations and a camera for store-and-forward consultations. The variable costs were participant's time and travel costs. The average time for real-time consultations was 15.7 minutes and for store-and-forward consultations was 1.6 minutes. From a societal perspective, the net societal cost of the initial real-time consultation was £132.10 per patient compared with £26.90 per patient for the initial store-and-forward consultation. The authors concluded that the store-and-forward consultation was cheaper, but less clinically efficient, compared with the real-time consultation.

Loane and colleagues (2001a) within their randomised controlled trial assessed the costs of real-time teledermatology compared with conventional dermatology care for patients from urban and rural areas in Northern Ireland from a societal perspective. Over the two years, 126 patients (46%) were randomised to telemedicine consultation and 148 (54%) to a conventional hospital outpatient consultation. The costs for teledermatology included videoconferencing units, installation of ISDN lines, line rental costs, call costs, video camera and staff costs. The conventional care costs included the cost of clinician's time, patient's time, and patient travel. The observed marginal cost per patient of the initial real-time teledermatology consultation was £52.85 (urban areas) and £59.93 per patient (rural areas). The observed marginal cost of the initial conventional consultation was £47.13 for urban patients and £48.77 for rural patients. The total observed costs of teledermatology were higher than the costs of conventional care in both urban and rural areas, mainly because of the fixed equipment costs. The authors concluded from a patient perspective, telemedicine was cheaper than conventional care, as it involved less travel and time costs; however, from a healthcare perspective telemedicine was not the cheaper option.

A cost-minimisation analysis of real-time teledermatology compared with conventional hospital care in New Zealand was conducted by Loane et al (2001b). Over 10 months,

109 (54%) patients were randomised to teledermatology and 94 (46%) patients randomised to conventional care. Teledermatology costs included ISDN line installation and rental, cost of capital equipment, call costs and staff costs. Conventional care costs included cost of clinician's time, patient's time and patient travel. The total cost of the 123 teledermatology consultations was NZ$34,346 and the total cost of the 106 conventional hospital consultations was NZ$30,081. The average societal cost of the teledermatology consultation was NZ$279.23 compared with NZ$283.79 for the conventional hospital consultation. The marginal cost of seeing an additional patient was NZ$135 via teledermatology and NZ$284 via conventional hospital appointment. From a societal viewpoint, and assuming an equal outcome, teledermatology was a more cost-efficient use of resources than conventional hospital care.

Whited et al (2003) conducted a cost analysis and cost-effectiveness analysis of a store-and-forward teledermatology service compared to usual care in USA from a healthcare perspective. One hundred and forty patients were randomised to usual care (patients assessed in a dermatology clinic) and 135 patients to teledermatology. The data for these patients was then extrapolated to the total population served by the dermatology clinic in 2001 (n = 5,440). The costs for the teledermatology included the equipment (a digital camera, laptop, desktop computers, and printer), maintenance, training and line installation, and also labour, overheads, space and travel costs. Usual care costs include staff and travel costs. Teledermatology was not cost-saving when compared to usual care using observed costs and outcomes. The decision model estimated that the base-case annual total cost of serving 5,440 dermatology patients was US$198,016 with teledermatology and US$116,416 with usual care. Expected costs were US$36.40 per teledermatology patient and US$21.40 per usual care patient. Total incremental cost of teledermatology was US$81,600 and per patient incremental cost was US$15.00. Benefits were measured in terms of days to the intervention. The incremental cost-effectiveness ratio of teledermatology was US$0.17 per patient per day of time to initial intervention saved (the incremental cost of teledermatology (US$15.00) divided by the incremental effectiveness of teledermatology (87.5 days)). Sensitivity analyses indicated that teledermatology has the potential to be cost-saving if clinic visit costs, travel costs, or averted clinic visits were higher than observed in the study.

As part of a multicentre randomised trial, real-time teledermatology was compared with conventional dermatological care in terms of clinical outcomes, cost-benefits and patient re-attendance in Northern Ireland [Wootton and colleagues, 2000]. During the

year, 102 patients were randomised to telemedicine and 102 patients randomised to a conventional hospital outpatient care. Fixed costs included the cost of the videoconferencing units, installation of ISDN lines, line rental costs and video camera. Variable costs were the costs of the clinician's time, patient's time, patient travel and call costs. No major differences were found in the reported clinical outcomes. Of the patients randomised to teledermatology, 55 (54%) were managed within primary care and 47 (46%) required at least one hospital appointment. Of the patients randomised to conventional care, 46 (45%) required at least one further appointment, 15 (15%) required general practice review, and 40 (39%) required no follow-up visits. The net societal cost of the initial consultation was £132.10 per patient for teledermatology and £48.73 for conventional consultation. Real-time teledermatology was clinically feasible, but not cost-saving compared with conventional care. However, the authors stated that if the equipment were purchased at current prices and the travelling distances greater, teledermatology would be cost-saving.

In the ten articles mentioned above on dermatology: three studies were conducted in Northern Ireland and another three studies were from the USA. Five studies were cost analyses, three were cost-minimisation analyses, one was a cost analysis combined with a cost-effectiveness analysis and the final study was a cost-benefit analysis. All nine studies which reported a time-frame, were conducted in a period of 2 years or less; six studies conducted some sort of sensitivity analyses; and no studies undertook discounting. Four studies were from the healthcare perspective and four from a societal perspective, and the other two studies did not state a perspective. Telemedicine was cheaper than conventional care in five studies, for three studies telemedicine was more expensive, in one study telemedicine had the potential to be cost-saving and in the final study, telemedicine produced similar costs results to conventional care.

b) Radiology. Bergmo (1996) conducted a cost analysis to see whether teleradiology was more or less expensive than a visiting radiologist in Norway. The study was conducted from a health service perspective, with an average workload of 25 patients per day or 6,000 patients (8,000 examinations) per year. The journey takes two and a half hours for a visiting radiologist and each session typically lasts five hours. Costs of the teleradiology service included the costs of the equipment plus maintenance, personnel costs and call charges; whereas for the visiting radiologist the costs included cost of senior radiologist, travel costs, costs of an x-ray processor and emergency patient transfer costs. Costs common to both arms were not included, such as electricity and the radiographer who assisted in taking radiographs. Total cost of

teleradiology was NKr646,900 per year; compared with visiting radiologist cost NKr1,069,000 per year. Teleradiology cost NKr108 per patient, in comparison with NKr178 per patient for the visiting radiologist. For teleradiology to be the cheaper option, the workload had to exceed 1,576 patients per year.

A cost analysis of the use of teleradiology at a military training area in Hawaii or sending patients to a hospital which was 400km away for radiological diagnosis (direct referral to hospital – ground evacuation) was conducted by Brumage and colleagues (2001). They hypothesised that a teleradiology system would reduce the number of patient transfers to the hospital, thereby leading to savings. In total, 323 patients were seen by teleradiology, and 33 patients were transferred to hospital during the 2 months training exercise. Teleradiology costs included the cost of the teleradiology system, operating costs, and personnel costs. Direct referral costs included the cost of transportation and fuel, driver, a medical attendant, costs of emergency room and tests. Direct referral to hospital cost was US$1,261. When teleradiology was provided, in total 29 evacuations were avoided and the savings amounted to US$36,569. Over the course of the year, the teleradiology system would cost US$167,203 and assuming a constant rate of evacuations over the year which would be 140, the expected cost would have been US$176,540, and hence the teleradiology system saved US$9,337. Over five years, the costs and savings were estimated to be US$349,940 and US$882,700, respectively. The authors stated the benefits of teleradiology included: (1) a reduction in air and ground transportation costs; (2) a reduction in hospital costs; (3) man-hours saved during medical evacuations; (4) less use of materials (e.g. X-ray film); (5) better quality of care (determined by satisfaction questions of providers); and (6) better access to care (determined by time to diagnosis and treatment). The authors concluded that not only did the teleradiology system prove to be cost-effective during the exercise, but there was also an improvement in the quality of care at the point of need.

Crowe and colleagues (1996) wanted to find out whether a 3-month pilot teleradiology project at a children's hospital in Australia was cost-saving compared to patient transfers to a hospital. In total, 575 images were transmitted and the mean transmission time per image was 3.26 minutes. Costs were calculated in terms of transmission, equipment, maintenance, ISDN lines and staff components. The cost analysis found that the cost per image transmitted would vary from A$80.14 for 2,500 images to A$34.00 for 10,000 images per year. The study estimated that 20 emergency transfers between A$7,000 to A$10,000 would be needed to break-even each year. The authors found that the adoption of teleradiology would have potential

improvements in patient management due to quicker diagnosis and earlier intervention, and also potential savings through avoiding transfer of some emergency cases.

A cost-minimisation analysis from the healthcare perspective of a wide-area teleradiology network in a French region was undertaken by Daucourt and colleagues (2006). They compared the teleradiology service with direct transfers over the course of a year for all patients for whom a remote consultation in radiology was required. Teleradiology costs included equipment costs per workstation, maintenance costs and staff costs. Direct referral costs included costs related to transfers and hospitalisations, consultations and staff costs. In total, 664 teleconsultations took place during the year and 309 transfers were avoided. Hospitalisations were avoided for 27 patients. Therefore, the savings due to proportions of transfers, hospitalisation and consultations avoided were estimated to be €102,779. The authors concluded that the teleradiology network enabled savings despite the large capital outlay.

Davis (1997) conducted a cost analysis of teleradiology in a rural imaging centre in the USA compared with a courier service for radiological transfers over the course of two years. Just over 8,000 teleradiology examinations were performed and the average transmission time for a typical case of 50 images was between 6 and 8 minutes. For the teleradiology service the costs included: the cost of the capital, maintenance, floor-space rental, line rental charges, call charges, ISDN transmission and personnel costs. For the off-site radiology, the cost of the courier was also included. So for 2,000 cases, a mid-field magnetic resonance imaging unit was predicted to cost US$470 per case using teleradiology (total costs US$940,515) and US$544 per case using film and a courier service (total costs US$1,087,975). The author found that teleradiology reduced travelling time for both patients and physicians, and had an overall positive effect on patient care. The results of the survey indicated that the teleradiology system was cheaper than the more traditional courier service.

A cost-minimisation analysis from a societal perspective of teleradiology services for remote communities in Norway was conducted by Halvorsen and Kristiansen (1996). The teleradiology system (most examinations at site and more advanced examinations at the host site) was compared to the existing system (a small x-ray unit at the remote site and all other examinations at the nearest radiology dept (the host site)) and to all examinations at the host site. Of the 1,793 identified referrals for radiological examinations in 1993, one third (n = 597) were randomly selected. For the three options, the annual direct medical costs were the equipment costs and staff costs; the direct non-medical costs were the travel costs; and indirect costs were the lost

production (number of hours absent from work). During the 12-month study, the direct medical, direct non-medical, and indirect costs of the three options were: £9,000, £51,000, and £31,500 (total £91,500) for the existing system; £108,000, £2,000, and £13,500 (total £123,500) for the teleradiology option; and £0, £75,000, and £42,000 (£117,000 in total) for the "all at host" option, respectively. The teleradiology option did not seem to be cost-saving and an increase in the annual number of teleradiology referrals to 590, would make the teleradiology system cost-saving.

A cost analysis of an emergency computerized tomography teleradiology system compared to transporting films by taxi or transporting patients to nearest central hospital for scanning in Austria was undertaken by Stoeger and colleagues (1997). During the 13-month study period, 121 emergency examinations of 116 patients took place and the average transmission time per examination was 15 minutes. The costs of teleradiology included the ISDN connection and rental, whereas for the alternative options this included the transportation (courier service only) and staff costs. The capital cost of the scanners were excluded as they were installed in both hospitals. The fixed cost of teleradiology was DM926 and average cost of one emergency computerized tomography examination by teleradiology was DM372. Transporting the films by taxi was cheaper, estimated cost DM156. Transporting the patient to the nearest hospital was more expensive: DM524 by road or DM4,667 by helicopter ambulance. Overall, teleradiology was cheaper than transporting the patient to the nearest central hospital for scanning; however, more expensive than transporting the films by taxi, but taxi option would have been much slower.

Vuletic (2001) wanted to determine the costs of teleradiology versus patient referral in Croatia. A teleradiology system connecting 35 workstations in 27 hospitals was established and the teleconsultations which took place were transmitted in real-time and store-and-forward. Eighty patients were transported urgently, a further 181 were transported non-urgently and an additional 181 patients remained in their local hospital. Over two years, 2,071 consultations took place which represented an average rate of four neurosurgical teleconsultations per day. By avoiding unnecessary patient transfers, over 3.5 million kunas per year was saved. The author felt the benefits of teleradiology included: decreased travelling time and expenses for both patients and specialists; decreased hospital expenses; shorter waiting times; increased efficiency; and better medical education. The author concluded that the avoidance of unnecessary patient transfers led to significant cost savings.

So in the eight articles mentioned above in radiology, two studies were conducted in USA and further two studies were from Norway. Six studies conducted cost analyses and the other two articles conducted cost-minimisation analyses. All studies that reported a time frame were conducted within 2 years. Five studies did not state their viewpoint, two were from the healthcare perspective and one from a societal perspective. Four studies undertook some sort of sensitivity analyses and none undertook discounting. Teleradiology was cheaper than conventional care in four studies; only one study indicated that teleradiology was more expensive than conventional care; in another study teleradiology was either more expensive or cheaper depending on the alternative; and two studies stated that telemedicine can enable savings as long as the patient throughput is above a certain number.

c) Psychiatry. Doze and colleagues (1999) conducted a cost analysis of a telepsychiatry pilot project in which a psychiatric hospital was linked with five general hospitals mental health clinics in Canada and this was compared to the costs of a travelling psychiatrist. Ninety psychiatric patients were seen during the 9 months and 109 telepsychiatry consultations took place. The telepsychiatry costs included a videoconferencing unit, dual monitors, document camera, video recorder, the installation of ISDN lines, call charges and staff costs. For the travelling psychiatrist, staff costs and travel costs were included. The cost analysis indicated that at 396 consultations per year, the service cost the same as providing a travelling psychiatrist (C$610 per consultation); with more consultations, telepsychiatry was cheaper. In the sensitivity analysis, the authors found that a reduction of 10% in equipment cost reduced the break-even point from 396 to 368 telepsychiatry consultations a year. For patients, the authors emphasised that telepsychiatry had greater benefits in terms of reduced travel time, less stress, less absence from work, less delays in accessing a psychiatrist, more patient choice and improvement in quality of life. The authors concluded that telepsychiatry was a more expensive option than in-person psychiatry at a low volume of service, but less expensive at a higher volume.

A cost analysis of a 3-month pilot telepsychiatry service for children in Canada was undertaken by Elford and colleagues (2001). Thirty patients accompanied by a parent, completed a psychiatric assessment using the videoconferencing system compared to conventional consultation at a referral centre. The costs of telepsychiatry included the videoconferencing unit, installation of ISDN lines, line rental, call charges and personnel costs. The conventional care costs included the cost of travel, accommodation, food expenses and staff costs. Excluded costs included the loss of income for parents, babysitting costs and other incidental costs. The cost analysis

estimated the total travel cost for the 30 patients was C$12,849, an average of C$428.30 per patient, whereas the total cost of the telepsychiatry service was C$12,575, or C$419.17 per patient, this was based on a patient volume of 10 patients per month. The authors indicated that telepsychiatry advantages included decreased travel costs, decreased travel time, greater convenience, improved access to psychiatrist and children felt more comfortable. The authors stated that the pilot project was a success and with more patients the costs for telepsychiatry would decrease.

Harley (2006) conducted a cost analysis of a telepsychiatry service compared with direct travel to a specialist psychiatric hospital in London. During the course of the six months pilot study, 11 videoconferences took place between Jersey in the Channel Islands to London, England. The costs for videoconferencing included the cost of the equipment, installation of ISDN lines, annual line rental, ISDN call charges, maintenance cost and personnel costs. For direct referral the costs included travel costs and personnel costs. Family travel costs and subsistence costs were excluded. The total cost of videoconferencing was £3,483 and the traditional service was £12,975. The study suggested that telemedicine is cost-saving in providing tertiary mental health services.

A cost analysis of psychiatric inpatient care-planning consultations to remote areas using videoconferencing, instead of the conventional face-to-face consultations at a hospital in Finland from a societal perspective was conducted [Mielonen et al, 2000]. The study duration was 11 months and during this time 14 videoconferences and 20 conventional consultations took place. The videoconferencing costs included the equipment costs which consisted of the video and monitor, adjustable camera, an audio unit; installation and monthly line rental of the ISDN lines and personnel costs. For conventional care the costs included personnel and travel costs. The results from the costs analysis found at a workload of 20 patients per year, the cost of the videoconferences was FM2,510 per patient and the cost of the conventional alternative was FM4,750 per patient. At 50 consultations per year, a remote area would save about FM117,000. The authors concluded that consultations via videoconferencing in the long run are cost-saving compared with conventional practice.

Modai et al (2006) looked at the cost-effectiveness of video telepsychiatry versus face-to-face care for psychiatric treatment for mental health patients who lived in remote communities in Israel. Data was collected for 1 year, 39 patients formed the telepsychiatry group and 42 patients formed the comparison group. Costs included the costs of the salaries, psychiatrist travel expenses, equipment and running costs, phone

expenses, and patient travel expenses. During the study, satisfaction was measured at the 3-, 6-, 9-, and 12-month visits using the Patient Satisfaction Questionnaire. For the telepsychiatry group, total operating costs were US$94.40 compared with the comparison group, were total operating costs were US$84.70. One hour of telepsychiatry treatment was more expensive than face-to-face care and the authors found patients were generally satisfied. The authors also commented on the limited sample size which meant it was harder to draw definitive conclusions and further studies are needed involving a larger population and a longer study duration.

Ruskin et al (2004) in a cost-consequences analysis alongside a randomised trial, examined treatment outcomes of patients with depressive disorders who were treated remotely by means of telepsychiatry compared to in-person treatment from the healthcare perspective in the USA. Fifty-nine veterans were randomised to the telepsychiatry group and 60 veterans were seen conventionally. Psychiatric treatment lasted 6 months (8 sessions over a 6 month period, with each session lasting 20 minutes) and consisted of psychotropic medication, psychoeducation, and brief supportive counselling. Telepsychiatry costs included the computer based videoconferencing equipment, cameras, ISDN lines, line charges, maintenance fees and staff costs. For the conventional consultation, staff costs and travel costs were included. Telepsychiatry was more expensive per treatment session than onsite care (US$86.16 vs. US$63.25) and the cost was lower if the psychiatrists had to travel more than 22 miles. The authors concluded that telepsychiatry and in-person treatment for depression have comparable outcomes and equivalent levels of patient adherence, patient satisfaction, and healthcare cost.

An evaluation of a telepsychiatry service compared to a visiting psychiatrist following the completion of a pilot project in rural areas of a Canadian province was undertaken by Simpson and colleagues (2001b). During the 24 months of routine operation, a total of 546 consultations for 379 patients were completed at the five participating general hospitals, at an average rate of 23 consultations per month. The telemedicine equipment was also used for administrative and clinical meetings. The telemedicine costs included the cost of the equipment, installation, annual line charges, call charges and staff costs. For the visiting psychiatrist costs included the cost of the travel and time for the psychiatrist. The total fixed costs for telepsychiatry were C$169,800 per year, the total variable costs for telepsychiatry were C$140 per consultation and for a travelling psychiatrist (for time, travel and subsistence) were C$630 per consultation. The cost analysis found that the break-even point was at 348 consultations. If the cost of the other meetings were included, the break-even point would be 224 consultations

per year. From the health service perspective, telepsychiatry proved to be an appropriate use of resources and was a useful addition to existing mental health services.

Simpson and colleagues (2001a) evaluated the telepsychiatry service versus patient travel to a referral centre in Canada. Information on patient's time and travel costs were gathered through self-reported questionnaires and telephone interviews. Patient satisfaction with the service was also obtained from the questionnaires. The authors found that the availability of telepsychiatry led to an estimated cost-saving of C$210 per consultation for patient who would otherwise have had to travel and thus, lose time at work and pay for child care expenses. Patients overall were very satisfied with the service. The main advantages for the patient were reduction in waiting times for consultation and the avoidance of the expense and inconvenience in travel. From the patient's perspective, telepsychiatry was an acceptable technique that increased access to services and produced cost savings.

A cost analysis of delivering a mental health service either by telepsychiatry or conventional methods (patient transfer to the specialist centre) in Australia was undertaken by Trott and Blignault (1998). Telemedicine costs included: videoconferencing equipment, camera, ISDN connection and line rental, call charges and staff costs. Conventional costs included travel costs, accommodation, meals, and staff costs. Fifty patients including both adults and children were seen during the 12 months. Total telemedicine costs would be A$82,340 in the first year and A$54,960 in subsequent years and for the conventional service the cost was A$167,750. The savings were estimated to be A$85,380 in the first year and A$112,790 in subsequent years, not allowing for maintenance and equipment upgrading. The authors estimated there was a 40% reduction in transfers due to the introduction of telemedicine. Based on the previous year of 27 transfers at A$8,920 each, this would produce an annual saving of A$96,336 for the Royal Flying Doctor Service. The results of the study showed considerable savings from reduced travel by patients and health care workers.

In the nine articles mentioned above on psychiatry: four of these studies were from Canada; all nine studies compared telemedicine to one method of conventional referral; all studies except one were cost analyses; and eight of the nine studies which stated a timeframe were conducted within a year, apart from one study which lasted two years. Only four studies stated a viewpoint: one was a societal perspective, two were from the healthcare perspective and the final study was from the patient's perspective. Only three studies undertook some sort of sensitivity analyses and no

studies undertook discounting. In three studies, benefits were measured in terms of patient satisfaction. In conclusion, telepsychiatry was cheaper than conventional care in five studies (one was from the viewpoint of the patient), two studies found telepsychiatry to be more expensive than conventional care, and the final two studies stated that telepsychiatry had the potential to be cost-saving.

d) Cardiology. The experience of real-time telemedicine to transmit paediatric echocardiographic images in Canada was reported by Finley and colleagues (1997). During the 24 months, telemedicine was used for 135 paediatric patients: 69 (51%) were urgent examinations of newborns; 30 (22%) were urgent examinations of older children; and 36 (27%) of the examinations were for repeat or post-operative checks. The cost analysis found that the cost of the telemedicine network (equipment leasing costs and telecommunications costs including connection costs) was C$90,000 for the two years (annual cost of C$45,000). Use of the telemedicine network saved unnecessary patient transfers in 31 cases, at an average of C$8,500 per flight and C$1,200 per ambulance trip for transportation avoided. The cost of the transportations avoided was between C$100,000 and C$118,000. The authors found telemedicine to be cost-saving and provides a service comparable in availability and accuracy to that provided in their paediatric cardiology division.

To determine the costs over a year of using a store-and-forward telemedicine system for cardiotocographic recording of fetal heart rate, Ippolito and colleagues (2003) compared the telemedicine system to conventional referral in Italy. The cohort of patients included women with both high and low risk pregnancies: 162 patients were seen by telemedicine and a control group consisted of 2,710 deliveries. Equipment costs and operating costs of each of the five centres included equipment, maintenance, electricity charges, staff costs, telephone charges and modem. Hospital costs were calculated according to length of hospitalisation and the resources used. The cost analysis found that the total cost of telemedicine was €344,796. In the intervention group, 11 of the 87 high-risk patients were admitted to hospital and in the control group, 203 of the 813 women in the high-risk group were admitted to hospital and stayed on average of 20 days. If the women in the control group had been seen by telemedicine there would have been a saving, through avoided bed days, amounting to €358,280, which was similar to the cost of the telemedicine system. The study suggested that the use of telemedicine in cardiotocographic monitoring improves the quality of prenatal care.

Kildemoes and Kristiansen (2004) used decision analysis to estimate costs and health benefits of a public awareness campaign compared to a campaign combined with telemedicine which was aimed at shortening the delay for thrombolytic therapy in patients with acute myocardial infarction in Denmark. The authors adopted a healthcare perspective. Costs of the public awareness campaign, the cost of the telemedicine equipment, the cost of running the telemedicine diagnostics in ambulances and the cost of hospital stays were included. The results found that the campaign only will translate into five fewer fatal acute myocardial infarctions (62 life years gained) and a cost per life year of DKK283,300. When combining the public campaign with prehospital telemedicine diagnostics, the incremental cost per life year gained was DKK854,700. The authors concluded that whether such programs can be considered cost-effective will depend on how life year gains are valued by society.

McCue and colleagues (2000) evaluated the cost savings of 3 years of telecardiology use in a prison in the USA. The study compared the cost per visit of providing cardiology services by telemedicine to patients in prison versus the cost of providing traditional cardiology services at the cardiology clinic. Telemedicine costs included the cost of the equipment, maintenance costs, transmission lines, telephone charges and personnel costs and the non-telemedicine costs included staff costs and travel costs. The number of telemedicine consultations for the 3 years were: 24 (1996), 78 (1997) and 86 (1998). The authors found that the lower use of telecardiology services in 1996 resulted in higher cost per visit of US$189. This was US$45 more than the cost of traditional cardiology in the clinic (US$144). In 1997 and 1998, however, the higher utilization of telecardiology services decreased the cost per visit to US$135 and US$132, respectively. This resulted in a cost saving with telecardiology of US$15 per visit in 1997 (clinic cost = US$150) and US$46 per visit in 1998 (clinic cost = US$178). The authors found that because the vast proportion of telemedicine operating costs are fixed, increased utilization reduces the cost per visit and results in cost savings.

Rendina et al (1998) looked the effect of the use of telemedicine to transmit echocardiograms digitally for immediate interpretation versus sending videotapes via an overnight courier in the USA. The study was conducted from a healthcare perspective. In total, 48 infants were in the telemedicine group and 39 infants were in the control group. The telemedicine costs included desktop videoconferencing unit, video recorder, lines and installation, maintenance and communication charges. The control costs included the averted costs of shipping videotapes by overnight courier, materials and sonographer's time. The cost analysis found that the total telemedicine costs were US$6,405 and the total control costs were US$2,750. The cost per

echocardiogram transmitted was calculated at US$33 (difference between two methods (US$3,955) divided by the number of echocardiograms transmitted during the intervention period (110)), compared to the previous method of sending videotapes via an overnight courier (US$25). In terms of benefits the authors looked at the mortality rates, length of stay in the neonatal intensive care unit and the rate of transfer. Even though telemedicine was more expensive, over a longer time period and given a bigger patient throughput, telemedicine has the potential to be cost-saving. The authors found that while the sample size was inadequate to demonstrate improvements in health outcomes, the magnitude of change and the low costs of the system suggest that this intervention is practical for obtaining rapid diagnostic and treatment support.

To investigate the effect of immediate echocardiogram interpretation via telemedicine on rates of neonatal transfer to academic medical centres, Rendina and colleagues (2001) developed a logit model to predict the probability of transfer from two regional level three neonatal intensive care units to academic medical centres in the USA. One unit implemented a telecardiology program and the other acted as a comparison site. Telecardiology intervention began 18 months into the 36-month study period. The authors estimated that there would be a 58% reduction in transfers and approximately 30 transfers (1 transfer = US$5,000) were eliminated during the study period, resulting in the elimination of approximately US$150,000 in hospital charges. Introduction of telecardiology can lead to elimination of transfers and ultimately to cost savings.

Sicotte and colleagues (2004) conducted a cost-effectiveness analysis of interactive paediatric telecardiology to see whether the service was a worthwhile alternative to conventional referral (patient travel and 6 monthly outreach cardiology clinics) in Canada. The study was conducted from a societal perspective and during the 4 year study period, 78 children suffering from cardiac pathologies were seen in 129 consultations, at an average of 32 consultations per year. Telemedicine costs included the cost of the equipment and system installation, equipment maintenance, ISDN telecommunication fees and staff costs. The conventional care costs included patient travel expenses and consultant fees. The outcome measure was avoidance of patient journeys and with telemedicine the number of patient journeys was reduced by 42%. The total cost of telecardiology was C$272,327 and the total cost of conventional care would have been C$157,212. Telemedicine represented a supplementary cost of C$1,500 per patient. The incremental cost-effectiveness ratio of teleconsultation was estimated to C$3,488 per patient journey avoided. The authors felt that telemedicine can be reliable method of reducing patient journeys and the delay in consulting specialists.

Vincent et al (1997) wanted to examine the costs of a routine monthly and anytime emergency telemedicine monitoring in children and young adults compared to routine outpatient clinic visits or comparable emergency room treatment in the USA. Ninety-six patients were followed for 3 years after pacemaker implant. Costs included the cost of telemedicine equipment, outpatient clinic charges, emergency room charges and pacemaker analysis charges. Monthly charges for use of telemedicine including emergency transmission was US$70 and total charges for 75 emergency telemedicine transmissions were US$5,250. Outpatient clinic charges including pacemaker analysis were approximately US$200, whereas the standard emergency room charge without pacemaker analysis was US$260. For conventional care, the total charges for either alternative were US$19,500. Charges for use of telemedicine were significantly less ($p < 0.01$) than comparable outpatient visits.

So in summarising the impact of telemedicine in cardiology, four of the eight articles were from USA and a further two studies were conducted in Canada. Five studies conducted simple cost analyses and three were cost-effectiveness analyses. All studies except one were conducted for the duration of a year or more. One study was conducted from a societal viewpoint, two studies were from a healthcare perspective, and the other five studies did not state their viewpoint. Three studies had some outcome measure in their analyses; only two studies undertook sensitivity analyses; and only one study undertook discounting, even though the majority of the studies were equal to or greater than a year in length. Telecardiology was cheaper than conventional care in two studies, three studies indicated that telecardiology was more expensive than conventional care, one study found that the costs of telecardiology were similar to conventional care, and finally, two studies stated that telemedicine had the potential to be cost-saving.

APPENDIX 5: SUMMARY OF SYSTEMATIC REVIEWS OF TELEMEDICINE

Whitten and colleagues (2002) as mentioned in Chapter 3, section 3.2.1 undertook a systematic review of the cost-effectiveness of telemedicine studies which were published until June 2000. The authors separated the articles into those without (n = 557) and with (n = 55) cost data and established a checklist of criteria for assessing the quality of economic evaluations in healthcare. Of the 55 cost studies, only 24 articles met their selection criteria and were subject to a full review. For those 24 articles which reported cost data they presented quantitative and qualitative analyses. Twenty of these 24 studies were limited to simple cost comparisons; no study used cost-utility analysis; 11 studies stated the viewpoint of the cost analyses and only four of these were from a societal perspective; and most of the studies provided no details of sensitivity analyses. They found that studies assessing the cost-effectiveness of telemedicine are generally small, of a short-term duration and of poor quality. They concluded that "there was no good evidence that telemedicine is a cost-effective means of delivering healthcare compared to standard healthcare delivery".

Roine and colleagues (2001) as mentioned earlier examined the evidence for the effectiveness or cost-effectiveness of telemedicine in order to clarify the current status of the technology. They identified telemedicine studies that reported information on patient outcomes, administrative changes or economic assessments. From more than 1,100 abstracts which were surveyed, they reviewed 133 full text articles for further inspection and of these 133 articles, only 50 met their inclusion criteria for their review. Thirty-four of these articles assessed at least some clinical outcomes, and the remaining 16 articles were economic analyses. They found that most of the available literature referred only to pilot studies, short-term outcomes and most studies were of low quality. They felt that a systematic comparison of the costs and more work on the effects should be done in the future. They concluded that the "evidence regarding the effectiveness or cost-effectiveness of telemedicine is still limited. Based on the current scientific evidence, only a few telemedicine applications can be recommended for broader use".

Hailey and colleagues (2002) in their systematic review of the evidence for the benefits of telemedicine identified 66 studies that included a comparison with a non-telemedicine alternative (mentioned earlier in section 3.2.1). The main problems they encountered were that: most of the available literature referred only to pilot projects and to short-term outcomes, poor measurement of effectiveness, identification and timing of costs (and consequences), the absence of incremental analysis and only a few papers considered the long-term or routine use of telemedicine. They found that

the most convincing evidence on the efficacy and effectiveness of telemedicine was in the area of radiology, where savings were through avoidance of travel and associated delays. Over half the studies (56.1%, 37 articles) suggested that telemedicine had advantages over the alternative approach, 24 studies (36.4%) were unclear whether telemedicine had advantages; and the final five studies (7.6%) found the alternative approach had advantages over telemedicine. The authors concluded that "although useful clinical and economic outcomes data have been obtained for some telemedicine applications, good quality studies are still scarce and the generalisability of most assessment findings is rather limited".

In a follow-up to this review, Hailey and colleagues (2004) developed a simple approach to the measurement of quality for telemedicine studies that takes into account both study design and study performance. For example, for study design a large RCT was given a higher score than a retrospective comparative study; and for study performance such as patient selection and description of the interventions, if no information was missing this was given a higher score than if relevant information was missing or was given in little detail. They also assessed the quality of the economic evaluation by using a checklist, for each study a score of 1 was given to each of the criterion that was fulfilled (score range: 0 to 10). From the literature review, 48 papers met the selection criteria and were included in the review. Twenty-five papers (52.1%) included some kind of economic analysis and only two of these papers included health related quality-of-life measures. Thirteen of the 25 papers met five or more of the economic criteria and were rated as good, or good to fair, although the authors mentioned that the scope of the economic analysis was limited. The authors felt the results confirmed previous findings and that good quality studies are still scarce.

Mair and Whitten (2000) provided a systematic review of 32 studies on patient satisfaction with telemedicine, involving real-time videoconferencing. The main outcome measures examined included patients' satisfaction and patients' willingness to use telemedicine in the future. The studies used videoconferencing for a variety of purposes ranging from specialist consultations to home nursing. Most of the studies included in the review were pilot and feasibility studies with small sample sizes. Twenty-six studies used simple survey instruments, five did not specify any methods and one study used qualitative methods. All studies reported good levels of patient satisfaction with telemedicine and the main advantages from teleconsultations were noted as: increased accessibility of specialist expertise, less travel required, and reduced waiting times. The authors concluded that "methodological deficiencies (low

sample sizes, context, and study designs) of the published research limit the generalisability of the findings".

Williams et al (2001) provided a systematic review of 93 studies on patient satisfaction with telemedicine and this review differed to the one by Mair and Whitten (2000) as it had more emphasis on patient satisfaction measures and included all forms of telemedicine rather than just videoconferencing. Telepsychiatry and telepsychology studies represented the largest proportion of studies which used patient satisfaction measures (24.7%) and real-time videoconferencing alone was used in 71.0% of the studies. Ninety-four percent of studies on patient satisfaction in telemedicine used uniquely designed questionnaires and the most of them (86.0%) did not report on the validity or reliability of the questionnaires. Almost a quarter of studies did not provide enough detail to determine the type of questions patients were asked. Sixty-four studies used Likert-type questions in which the patient responds to statements about telemedicine using a scale of agreement to disagreement. Aspects of patient satisfaction which were assessed included: professional-patient interaction, patient's feeling about the consultation such as comfort and convenience, overall satisfaction, technical aspects and physical environment. Preferences between telemedicine and face-to-face consultation were assessed in 28 studies. Consistent with the earlier review, more than 80.0% of patients were satisfied with the service they received. One of the main criticisms the authors stated was that "patient satisfaction data may be biased due to self-selection, as those who are "less satisfied" are more likely to opt out of satisfaction studies". Overall, the authors concluded that "the current evidence concerning patient satisfaction with telemedicine is rather limited."

Other systematic reviews of telemedicine also found no conclusive evidence that their findings could be generalised to other telemedicine applications. Hersh and colleagues (2002) conducted a systematic review of the literature to evaluate the efficacy of telemedicine for making diagnostic and management decisions for three different types of telemedicine applications: office/hospital based, store-and-forward and home-based. The authors wanted to determine whether telemedicine can provide at least equally good diagnostic and management decisions as face-to-face care. A total of 58 studies were included in the review and the articles were summarised and graded for quality and direction of the evidence. The authors stated that the overall methodological quality of most studies was low. Typical problems they encountered with studies included: small sample sizes, the use of the same clinician for both telemedicine and face-to-face care, and lack of inter-observer agreement within modalities. The strongest evidence for the efficacy of telemedicine for diagnostic and management

decisions came from the areas of psychiatry and dermatology. There was also good evidence that general medical history and physical examinations performed via telemedicine had relatively good sensitivity and specificity. The authors concluded "despite the widespread use of telemedicine in most major medical specialties, there is strong evidence in only a few of them that diagnostic and management decisions provided by telemedicine are comparable to face-to-face care". Hersh et al (2006) provided an update to the systematic review published in 2002 [Hersh et al, 2002]; 106 studies were included in the review. However, the methodology of many studies was still weak such as small sample sizes and most studies lacked statistical power to detect significant differences between the two groups. The authors concluded that "there are still significant gaps in the evidence base between where telemedicine is used and where its use is supported by high-quality evidence".

Whitten et al (2007) conducted a systematic review of 15 relevant databases for articles about telemedicine and telehealth relating to research methodology and after the exclusion criteria was applied, 1,615 articles remained for analysis. Only 85 studies (5%) of the telemedicine articles mentioned any theory or paradigmatic approach. Most studies reported the overall study aim (96%), however only a few papers (11%) provided hypotheses or a research question. Randomised selection of the subjects was reported in 11% of patient studies and the majority of studies were based on small sample sizes (100 patients or less). Only 26% of the studies reported a time frame i.e. start and end points of a study. The authors concluded that "until the telemedicine field adheres to agreed standards of reporting methodological details it will be difficult to draw firm conclusions from review studies".

Finally, the results produced in Chapter 3 are consistent with a recently published and up to date review by Bergmo (2009) who found that "the majority of the economic evaluations were not in accordance with standard evaluation techniques". The author referred to telemedicine as "technologies used in direct patient care such as real-time videoconferencing and store-and-forward applications". Only articles published in journals and written in English from the period 1990 to 2007 were analysed. In total, the literature search identified 779 abstracts, and after review only 33 of them were full economic evaluations where both costs and outcomes were measured. Studies which only looked at costs were excluded and studies which looked at outcome measures, for example, in terms of travel avoided or hospitalisations avoided were also excluded (these have been included in the literature review). All 33 studies were also identified during this literature search, of which only 15 were included. Seventeen of these studies evaluated telemedicine in a home setting and were excluded from the literature

review (see definition for ‘telecare’ in Chapter 1, section 1.2) and the other excluded study was Dowie et al study (2007), which forms part of the case study for the thesis. The author said that only eight studies had addressed all the key issues: a clear study objective, adequate comparison(s), reporting of study perspective and design, transparent measurements and valuation of costs and outcomes, reporting of data sources, addressing of uncertainty and clear presentation of the results. Most studies used multiple outcome measures and a cost-consequences approach was adopted for most economic evaluations. Overall, the objectives, study design and choice of comparator were mostly well reported. However, most studies lacked information on the study perspective; the cost methods were not reported in detail; and very few studies provided information on statistical and sensitivity analysis to assess the validity and robustness of the results.

APPENDIX 6: SELECTION BIAS

A6.1 Literature search to identify methods to reduce selection bias in healthcare

The aim of the literature search was to identify all methods which have been used to reduce selection bias in non-randomised studies. The following electronic databases outlined below were searched from their inception date to 30th January 2009: MEDLINE, CINAHL, EconLit, Bids IBSS and Scopus. Search terms were entered as free text and/or as MESH terms and they included:

"selection bias OR bias selection OR susceptibility bias OR selectivity bias"

AND

"non-random* OR observational OR quasi-experiment* OR cohort OR case-control OR cross-sectional OR case series OR case report OR before and after"

Once all abstracts were retrieved they were exported into the citation software package, Endnote [Endnote, 2002]; and any duplicates identified from the different databases and non-English language articles were deleted within Endnote. The search retrieved 2,442 abstracts. All titles, bibliographic data and abstracts of the results from these searches were then read and scanned. Of these 2,442 abstracts, 246 of these abstracts indicated that they used some method to reduce selection bias. If the method was explicitly stated in the abstract, a note was made of the method used; if the method was not explicitly stated in the abstract, then the full-text paper was obtained to find out the method which was used to reduce selection bias. Some abstracts (or papers) identified more than one method. Both, theoretical and practical abstracts (papers) were identified, which were not just related to healthcare, but were also related to the labour market. Table A6.1 below lists the type and frequency of each method identified within the 246 abstracts.

Table A6.1: Methods identified in the literature search to reduce selection bias and/or endogeneity

Method	Frequency
Matching	13
Stratification	5
Regression analysis i.e. logistic or linear or multivariate or advanced	46
Propensity analysis	29
Propensity score matching	41
Propensity scores used in stratified analysis (stratification method)	9
Propensity scores used as a covariate in regression models (regression adjustment)	10
Inverse probability weighting	8
Instrumental variables analysis	26
Sample selection models i.e. Heckman; Lee; Greene	65
Two-part models	21
Regression discontinuity	2
Decomposition technique	1
Difference-in-differences method	5
Total	**281**

A6.2 Brief overview of the other methods identified in the literature review

A6.2.1 Matching

Matching is a method to remove bias in observational studies. For example, in case-control studies, individuals who have the disease under investigation (cases) are matched to individuals who do not have the disease (controls), but who are thought to be comparable in other respects. Matching takes place to ensure that each case is matched to a control that has the same (or similar) values of the matching variables. Matching variables include risk factors such as age and gender; and sometimes there are two or more such controls for each case. Matching ensures that any difference between cases and controls cannot be a result of differences in the matching variables, and any difference between the two groups is the treatment effect plus a random element.

The main advantage of matching is that it is simple and can balance confounders (risk factors). The main disadvantage to matching is that the more variables that are matched on, the more difficult it may be to find such controls. If this is the case, a large population of potential controls from which to draw is also required. There is also a possibility of overmatching, which can reduce statistical power. In a large study with many variables, it is easier to take an unmatched control group and use regression methods.

A6.2.2 Stratification

The concept of stratification is very similar to matching. Confounding variables (risk factors) are identified and subgroups are created using these variables. Stratification is the process of dividing members of the population (i.e. cases and controls) into different strata (or subgroups) before comparisons are made. For example, patients with similar ages (i.e. 15-24 years) form one group, and another group is formed with patients with similar ages (i.e. 25-34 years) and so forth. Even though each stratum must contain patients from both groups, they may appear in unequal numbers. The strata should be mutually exclusive: that is, every element in the population must be assigned to only one stratum. Once subgroups are defined, analysis is then performed on each subgroup separately. The main advantage of stratification is that it is simple; however, the main disadvantage is that it can be difficult to interpret with many subgroups.

A6.2.3 Inverse probability weighting[14]

When treatment is offered over multiple time points, marginal structural models are used [Robins et al, 2000]. That is, marginal structural models are useful when there are time dependent treatments and/or time dependent confounders (i.e. observed covariates that are affected by treatment and relevant to the outcome of interest). Marginal structural models are estimated using a method called the inverse probability of weighting. Inverse probability weighting has also been used for non-randomised studies where treatment is offered at a single time point, where there are two or more groups [Imbens, 2000]. Inverse probability weighting can only control for observed variables.

The inverse probability weighting approach uses observed variables to estimate a treatment selection probability i.e. that is to calculate the probability of an individual receiving the treatment they actually received, conditional on their observed covariates (this is similar to the propensity score method) and the inverses of these treatment selection probabilities are used as observation weights [Rosenbaum, 1987]. That is, an individual is assigned a weight, equal to the inverse of the conditional probability of receiving his or her own treatment. Specifically, the weight is the inverse of the propensity score for treated subjects and it is the inverse of one minus the propensity score for control subjects. The weights (these are usually not known) can be calculated using logistic regression and by obtaining the predicted values.

For example, Hogan and Lancaster (2004) used data from women in the HIV Epidemiologic Research Study and they wanted to evaluate the difference in CD4 cell count for an individual who receives the new therapy, versus if that same individual did not receive therapy. They used inverse probability weighting, whereby observed confounders were used to estimate treatment selection probabilities which corresponded to their observed treatment histories and the inverse of these treatment selection probabilities were used as observation weights. The authors used ordinary least squares regression analyses, where each unit was weighted relative to the inverse of its probability of being sampled (e.g. in a random sample (size = n), the sampling probability is 1/n, so the relative weights are equal to 1). The weights then can be interpreted to quantify the number of non-sampled members of the population that are being represented by the sample unit (e.g. if the weight for an observed unit is ¼, then this represents information from four people). They found that the inverse probability weighting provided estimates close to ordinary least squares estimates.

[14] Inverse probability weighting or marginal structural model is a variant of regression adjustment for propensity score.

A6.2.4 Instrumental variables

Instrumental variables analysis accounts for both observed and unobserved covariates. The idea is to find one or more variables (instruments) that have two properties. First, they should be related to the choice of treatment; the higher the correlation between the instrument and treatment variable, the more reliable the instrument. Second, the instrument should be uncorrelated (have no direct effect) with outcome. If the instrument(s) satisfies both properties, then the instruments are said to be good and consistent [Newhouse and McClellan, 1998].

An instrumental variable is a device that aims to achieve pseudo randomisation [Newhouse and McClellan, 1998]. That is, to mimic a randomisation process whereby the instrument(s) predicts treatment allocation and these values are then used as a covariate in the outcome model (instrumental variable regression).

The instrumental variables analysis is conducted in two steps. In the first stage a logistic regression model is conducted, using both the independent variables and the instrument(s) in order to obtain predicted probabilities. In the second stage, the instrumental variables regression is conducted using the predicted probabilities obtained from the first equation instead of actual (observed) values. By using these predicted probabilities rather than the actual values one should get unbiased estimates. As the groups differ in terms of likelihoods for treatment and not in the treatments they receive, this method estimates an incremental effect of treatment only over a range of variation in the treatment across the instrumental variable groups. In other words, the coefficients from the instrumental variable model are interpreted in terms of “mean” differences between the two groups.

If the instrument(s) are weakly correlated with treatment, then instrumental variables analysis has less statistical power than standard regression analysis; and even a weak correlation between instrument(s) and the outcome variable may result in the coefficients being biased or at least more biased than the ordinary least squares estimates and this can also lead to large standard errors (this can also happen if the sample size is not large enough) [Fortney et al, 1998; Newhouse and McClellan, 1998].

In the example by McClellan and colleagues (1994), they used instrumental variables analysis to find out whether more intensive treatment of acute myocardial infarction in the elderly reduces mortality. The instrument used here was differential distance, where differential distance is defined as the additional distance, if any, beyond the

distance to the nearest hospital to reach a specialist hospital (e.g. catheterisation hospital). The method estimated the incremental effect of invasive management for all patients who are 'marginal' (that is, who undergo invasive treatment in the 'relatively near' group but not in the 'relatively far' group) given that the groups are balanced in observable characteristics and that there are no other treatment differences between the groups. The authors found that observable health characteristics after instrumental variables estimation between the two groups were more similar than with the standard statistical comparison. They also concluded that there was a lower mortality rate among elderly patients who received catheterisation than among those treated more conservatively.

A6.2.5 Two-part models

A two-part model uses a two-stage approach to control for observed covariates. In the first part either a probit or logistic regression is conducted, to find out the probability of treatment selection i.e. separating those treated from those who are not treated. In the second part, another type of regression is conducted.

Two-part models can take various forms. For example, one type of two-part model is the censored Tobit regression model. This model is applied when the dependent variable is censored at some upper or lower bound. In the first stage, the selection equation is estimated by a Tobit regression. In the second stage, an additional variable estimated by Tobit regression is included in the equation to correct for possible endogeneity.

Sturm and colleagues (1995) wanted to compare mental health utilization in prepaid and fee-for-service plans. They used two-part models because of the presence of both non-use and skewness of use. In the first part of the model, a logit equation was used to estimate the probability of any outpatient mental healthcare, this separated out users from non-users and addressed the large number of zero use. The second part of the model used a multivariate regression model whereby the natural logarithm of the number of mental health visits on explanatory variables where used to analyse the level of use for patients with one or more visits. The logarithmic transformation of the number of visits for users in the second equation alleviates the skewness displayed by the data. The authors concluded that depressed pre-paid patients obtained substantially fewer mental health services than similar patients in fee-for-service.

A6.2.6 Regression discontinuity design

Regression discontinuity design method only controls for observed variables that affect whether a subject is assigned to a treatment or control group. The idea is that members of the study group are compared with themselves; instead of to a control group. That is, observations on the variable of interest are collected from the study group prior to the intervention and after the intervention. The design is characterised by its method of assigning subjects i.e. a cut-off score on an assignment measure, rather than by random assignment. All subjects who are on one side of the cut-off are assigned to the intervention group, while those scoring on the other side of the cut-off are assigned to a control group. The difference or ‘‘discontinuity’’ in the two regression lines (one for the control and the other for the intervention) at the cut-off provides an estimate of the intervention effect.

Zuckerman et al (2006) used the regression discontinuity design to analyse data from a Mid-Atlantic state Medicaid drug utilization intervention with the aim of improving the pharmacologic management of paediatric asthma. For the simple analysis, they used repeated measure ANOVA to compare the differences in monthly short-acting b2-agonist inhalers (SAB) prescription fills for the 333 ‘‘high-user’’ children (intervention group) for a period of five months pre-intervention and five months post-intervention. For the regression discontinuity analysis, the authors compared the pre–post experience of the intervention group (n=333) to that of 3,306 children with asthma who had at least one prescription for a SAB and were below the cut-off value of an average of one SAB canister per month during the matched pre-intervention period. They used ordinary least squares regression for the regression discontinuity model. Both analyses indicated that the intervention significantly reduced SAB use among the high users.

A6.2.7 Decomposition technique

The ‘decomposition technique’ is where the observed outcome effects in the treatment and control groups are divided into two components: treatment and population effects. The treatment effect is the difference in relative risk applied to the treatment groups, whereas the population effect is an adjustment for differences in the two study populations.

The decomposition technique considers the benefits of a policy in terms of ‘avoided costs’. After decomposition, the treatment effect answers the question, ‘how much money will the new treatment save (cost) compared with the existing treatment?’ In contrast, the population effect addresses the question, ‘how much money would have

been spent by a 'potential' treatment group receiving the existing treatment compared with the control groups?' [Shih and Kauf, 1999]. In other words, the treatment effect is the incremental cost (benefit) resulting from the treatment itself, and the population effect is the incremental cost (benefit) resulting from population differences.

In the example by Shih and Kauf (1999), individuals who were diagnosed with end-stage renal disease, either had to have a kidney transplant or be on maintenance dialysis in order to survive. Treatment consists of blood transfusions and the new treatment includes erythropoietin which increases haematocrit levels and prevents the onset of anaemia symptoms, and this treatment is given at every dialysis session. In the context of Medicare coverage of erythropoietin and to quantify the cost impact a decision model of anaemia treatment in end-stage renal disease patients was built. Outcome effects are decomposed into a treatment effect and a population effect. Logistic and multiple regression analysis were used to estimate branch probabilities and payoffs (outcomes), respectively, for the two treatment options within a decision model. Under standard methods of decision analysis, the authors found that an increase of US$7,032 per patient following erythropoietin coverage is observed. With the decomposition technique, the policy effect is estimated to be less, US$6,172, the difference coming from the population effect.

A6.2.8 Difference-in-differences method

The difference-in-differences approach seeks to measure a treatment effect while accounting for any pre-treatment differences between the treatment and control groups. Before the intervention, the difference between the treatment and control groups measures any existing (intrinsic) difference between the two groups; and after the intervention, the difference between the treatment and control groups measures the treatment effect plus intrinsic difference. So, to calculate the treatment effect alone, one must subtract the difference between the treatment and control groups before the intervention from the difference between the treatment and control groups following the intervention [Barnett et al, 2006]. Hence, the treatment effect is measured as the difference between two differences (i.e. the term difference-in-differences). By obtaining the treatment effect alone and by eliminating any existing difference between the treatment and control groups, one can control for observed differences between the groups.

Barnett et al (2006) assessed the health care use among veterans with diabetes mellitus who were enrolled in a care co-ordination telehealth program compared to veterans with diabetes mellitus who were not enrolled in a program over 24 months.

Health care utilization such as length of stay and outpatient visits were assessed at baseline and at 24 months after intervention for the treatment (n = 400) and control (n = 400) groups. The authors used a difference-in-differences approach and found that the program reduced avoidable healthcare services (such as hospitalizations) for patients with diabetes mellitus.

APPENDIX 7: DATA EXTRACTION FORM

Section 1: Personal Details

ID		
Postcode		
Date of birth		
Number of fetuses		
Parity	Primaparous	☐
	Multiparous	☐
Risk factors	Diabetes	☐
(tick all that apply)	Anti-epileptic therapy	☐
	Lithium	☐
	Down's risk	☐
	Elevated serum	☐
	Family history of CHD	☐
	Previous pregnancy with abnormality	☐
	High BMI	☐
	Any other factor _____	☐
	No risk	☐
Was the patient high risk according to Medway criteria?	Yes	☐
	No	☐
Was the patient high risk according to TelePaed criteria?	Yes	☐
	No	☐
Was the patient low risk but an abnormality was suspected?	Yes	☐
	No	☐
Any other comment(s)?		

Section 2: DGH Anomaly Scan at 20 weeks

Type of scan	Anomaly scan Other scan e.g. _____	☐ ☐
Date of scan		
Gestation in weeks		
Visual heart satisfactory	Yes No	☐ ☐
Visual heart satisfactory	Normal heart Abnormal heart Abnormal excluding heart Multi-organ abnormality including heart	☐ ☐ ☐ ☐
Cardiac abnormalities		
Was the scan videoed?	Yes No	☐ ☐

Section 3: Referral to specialist via Telemedicine

Date of TM consultation			
When did they decide the patient was a suitable candidate for TM?			
Type of consult Live transmission Store-and-forward			 ☐ ☐
Outcome of consult Normal heart Abnormal heart Abnormal excluding heart Multi-organ abnormality including heart			 ☐ ☐ ☐ ☐
Severity Mild Moderate	 ☐ ☐	 Severe Unknown	 ☐ ☐
Provisional Diagnosis			
Any further TM consults (please provide the dates)			

Section 4: Referral to specialist – direct referral to London

Date seen by specialist in London			
Specialist scan done		Yes	☐
		No	☐
Outcome of scan			
Normal heart		☐	
Abnormal heart		☐	
Abnormal excluding heart		☐	
Multi-organ abnormality including heart		☐	
Counselling by specialist			
Yes		☐	
No		☐	
Severity			
Mild	☐	Severe	☐
Moderate	☐	Unknown	☐
Provisional Diagnosis			
Any further specialist scans (please provide the dates)			
Any further counselling by specialist (please provide the dates)			

Section 5: Antenatal scans and clinic resource use at Medway Hospital from 20 weeks

Antenatal Scans from 20 weeks			
Date	***Gestation in weeks***	***Type of scan:*** 1) anomaly; 2) repeat anomaly; 3) detailed; or 4) other	***Outcome of scan:*** 1) normal heart; 2) abnormal heart; 3) abnormal exc. heart; or 4) multi-organ abnormality inc. heart

Antenatal clinics from 20 weeks	
Date	***Type of clinic:*** 1) peripheral antenatal; 2) hospital antenatal; or 3) other

Section 6: Prenatal Admissions from 20 weeks

Prenatal stays		
Admission date	***Discharge date***	***Outcome:*** 1) Discharge home 2) Discharged to London 3) Transferred to labour ward

Section 7: Terminations (if applicable)

Date admitted	
Date discharged	
Type of procedure	
Fetocide	☐
Induction/vaginal delivery	☐
Intercardiac KLC inject	☐
Medical TOP	☐
Surgical TOP	☐
Other	☐
Fetal outcome	
Normal heart	☐
Abnormal heart	☐
Abnormal excluding heart	☐
Multi-organ abnormality including heart	☐
Reason for termination – diagnosis for abnormal heart	

Section 8: Labour information

Labour admission date and time			
Labour discharge date and time			
Gestation in weeks at delivery			
Mode of delivery		***Delivery place***	
Normal	☐	Medway Hospital	☐
Forceps	☐	London Hospital	☐
Ventouse	☐	Home	☐
Caesarean – elective	☐	Other ----------	☐
Caesarean - emergency	☐	----------	
Water	☐		

Section 9: Birth outcome

Number of births			
Birth outcome	Baby 1	Baby 2	Baby 3
Live birth	☐	☐	☐
Still birth	☐	☐	☐
Miscarriage	☐	☐	☐
Unknown	☐	☐	☐
Baby's cardiac status	Baby 1	Baby 2	Baby 3
Normal heart	☐	☐	☐
Abnormal heart	☐	☐	☐
Abnormal excluding heart	☐	☐	☐
Multi-organ abnormality including heart	☐	☐	☐
Date of birth(s)			
Any other comment(s)?			

Section 10: Postnatal admission

Postnatal stays		
Admission date	***Discharge date***	***Outcome:*** 1) Discharge home 2) Other

Section 11: Any other information or comments

APPENDIX 8: REGRESSION RESULTS

Table A8.1: Conditional costs adjusting for all risk factors

	Unadjusted costs	Adjusted costs for all risk factors
Total costs of pregnancy		
2001/2002 period with telemedicine	£4,242.15	£4,118.01
2001/2002 period without telemedicine	£4,229.04	£4,104.90
2005/2006 period with telemedicine	£3,583.24	£3,733.25
2005/2006 period without telemedicine	£3,567.07	£3,717.09

Table A8.1a: Unadjusted costs (observed) and adjusted costs for all risk factors from time of anomaly scan up until after delivery for all time periods

Variable	Unadjusted costs		Adjusted costs for all risk factors	
	N = 1492, R^2 = 0.860		N = 1492, R^2 = 0.884	
	Coefficient SE)	p-value	Coefficient (SE)	p-value
1) 2001/2002 (with TM)	4242.15 (86.79)	< 0.001	3473.84 (808.53)	< 0.001
2) 2001/2002 (without TM)	4229.04 (86.76)	< 0.001	3460.73 (808.39)	< 0.001
3) 2005/2006 (with TM)	3583.24 (75.76)	< 0.001	3089.08 (791.04)	< 0.001
4) 2005/2006 (without TM)	3567.07 (75.51)	< 0.001	3072.92 (790.98)	< 0.001
Mother's age			23.37 (9.57)	0.015
Number of fetuses			1498.43 (220.50)	< 0.001
Parity			-368.94 (114.65)	0.001
Gestation at anomaly scan			-61.56 (34.25)	0.073
Diabetes			1807.06 (301.80)	< 0.001
Downs			25.26 (288.22)	0.930
Elevated serum			-67.89 (138.65)	0.624
Family history of CHD			267.16 (172.37)	0.122
Previous pregnancy with anomaly			18.43 (265.30)	0.945
Tests:				
1 = 2	$F_{1, 745}$ = 54.23, p < 0.001		$F_{1, 745}$ = 53.91, p < 0.001	
3 = 4	$F_{1, 745}$ = 82.12, p < 0.001		$F_{1, 745}$ = 81.62, p < 0.001	

Note: Standard errors (SE) are robust standard errors

Adjusting for all risk factors:
Period 1: 3473.84+ (23.37*29.448)+(1498.43*1.135)+(-368.94*1.540)+(-61.56*21.038)+ (1807.06*0.054) +(25.26*0.063)+(-67.89*0.150)+(267.16*0.107)+(18.43*0.051) = £4,118.01
Period 2: 3460.73+ (23.37*29.448)+(1498.43*1.135)+(-368.94*1.540)+(-61.56*21.038)+ (1807.06*0.054)+ (25.26*0.063)+(-67.89*0.150)+(267.16*0.107)+(18.43*0.051) = £4,104.90
Period 3: 3089.08+ (23.37*29.448)+(1498.43*1.135)+(-368.94*1.540)+(-61.56*21.038)+ (1807.06*0.054)+ (25.26*0.063)+(-67.89*0.150)+(267.16*0.107)+(18.43*0.051) = £3,733.25
Period 4: 3072.92+ (23.37*29.448)+(1498.43*1.135)+(-368.94*1.540)+(-61.56*21.038)+ (1807.06*0.054)+(25.26*0.063)+(-67.89*0.150)+(267.16*0.107)+(18.43*0.051) = £3,717.09

Table A8.2: Conditional cost results adjusting for all risk factors by risk group

	Telemedicine group	Direct referral group	Medium risk women	Low risk women
Total costs of pregnancy				
2001/2002 period with TM	£4,429.02	£3,937.85	£4,338.22	£3,945.55
2001/2002 period without TM	n/a	£4,099.53	£4,375.21	£3,911.03
2005/2006 period with TM	£3,876.18	£3,535.66	£3,651.49	£3,423.49
2005/2006 period without TM	n/a	£3,638.62	£3,709.99	£3,422.34

Table A8.2a: Unadjusted costs (observed) and adjusted costs for all risk factors from time of anomaly scan up until after delivery (or for a few cases after termination of pregnancy) for period 2001/2002 with TM

Variable	**Unadjusted costs**		**Adjusted costs for all risk factors**	
	N = 408, R^2 = 0.859		**N = 408, R^2 = 0.885**	
	Coefficient (SE)	**p-value**	**Coefficient (SE)**	**p-value**
Telemedicine (TM)	4363.31 (240.02)	< 0.001	3122.79 (1480.26)	0.036
Direct referral (DR)	3824.78 (353.29)	< 0.001	2631.62 (1564.91)	0.093
Medium risk (MR)	4445.67 (110.58)	< 0.001	3031.99 (1523.83)	0.047
Low risk (LR)	3711.72 (185.56)	< 0.001	2639.32 (1451.26)	0.070
Parity			-494.35 (170.25)	0.004
Number of fetuses			1221.40 (309.01)	< 0.001
Mother's age			46.88 (14.84)	0.002
Gestation at anomaly scan			-35.60 (63.37)	0.575
Diabetes			1526.87 (364.58)	< 0.001
Elevated serum			-390.48 (310.58)	0.209
Downs			184.00 (473.17)	0.698
Family history of CHD			-39.79 (296.50)	0.893
Previous pregnancy with anomaly			-571.10 (447.68)	0.203
Tests:				
TM = DR	$F_{1, 404}$ = 1.59, p = 0.208		$F_{1, 395}$ = 1.50, p = 0.221	
TM = MR	$F_{1, 404}$ = 0.10, p = 0.756		$F_{1, 395}$ = 0.10, p = 0.757	
TM = LR	$F_{1, 404}$ = 4.61, p = 0.032		$F_{1, 395}$ = 2.33, p = 0.127	
DR = MR	$F_{1, 404}$ = 2.81, p = 0.094		$F_{1, 395}$ = 1.22, p = 0.271	
DR = LR	$F_{1, 404}$ = 0.08, p = 0.777		$F_{1, 395}$ = 0.00, p = 0.985	
MR = LR	$F_{1, 404}$ = 11.55, p = 0.001		$F_{1, 395}$ = 1.54, p = 0.215	

Adjusted for all risk factors:
TM = 3122.79 + (-494.35*1.600)+(1221.40*1.218)+(46.88*29.260)+(-35.60*20.973)+ (1526.87*0.064)+(-390.48*0.203)+(184.00*0.059)+(-39.79*0.137)+(-571.10*0.069) = £4429.02
DR = 2631.62 + (-494.35*1.600)+(1221.40*1.218)+(46.88*29.260)+(-35.60*20.973)+ (1526.87*0.064)+(-390.48*0.203)+(184.00*0.059)+(-39.79*0.137)+(-571.10*0.069) = £3937.85
MR = 3031.99 + (-494.35*1.600)+(1221.40*1.218)+(46.88*29.260)+(-35.60*20.973)+ (1526.87*0.064)+(-390.48*0.203)+(184.00*0.059)+(-39.79*0.137)+(-571.10*0.069) = £4338.22
LR = 2639.32 + (-494.35*1.600)+(1221.40*1.218)+(46.88*29.260)+(-35.60*20.973)+ (1526.87*0.064)+(-390.48*0.203)+(184.00*0.059)+(-39.79*0.137)+(-571.10*0.069) = £3945.55

Table A8.2b: Unadjusted costs (observed) and adjusted costs for all risk factors from time of anomaly scan up until after delivery (or for a few cases after termination of pregnancy) for period 2001/2002 without TM

Variable	Unadjusted costs		Adjusted costs for all risk factors	
	N = 408, R^2 = 0.858		N = 408, R^2 = 0.884	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Direct referral (DR)	4054.91 (208.14)	< 0.001	3032.23 (1476.31)	0.041
Medium risk (MR)	4455.31 (108.91)	< 0.001	3307.91 (1483.29)	0.026
Low risk (LR)	3711.72 (185.37)	< 0.001	2843.73 (1424.88)	0.047
Parity			-482.86 (169.54)	0.005
Number of fetuses			1147.53 (297.41)	< 0.001
Mother's age			45.70 (14.83)	0.002
Gestation at anomaly scan			-41.01 (62.63)	0.513
Diabetes			1547.67 (364.06)	< 0.001
Elevated serum			-464.92 (300.13)	0.122
Downs			113.00 (475.86)	0.812
Family history of CHD			-43.09 (294.01)	0.884
Previous pregnancy with anomaly			-584.47 (440.59)	0.185
Tests:				
DR = MR	$F_{1, 405}$ = 2.91, p = 0.089		$F_{1, 396}$ = 1.12, p = 0.291	
DR = LR	$F_{1, 405}$ = 1.52, p = 0.219		$F_{1, 396}$ = 0.39, p = 0.534	
MR = LR	$F_{1, 405}$ = 11.96, p = 0.001		$F_{1, 396}$ = 2.32, p = 0.128	

Adjusted for all risk factors:
DR = 3032.23 + (-482.86*1.600)+(1147.53*1.218)+(45.70*29.260)+(-41.01*20.973)+ (1547.67*0.064)+(-464.92*0.203)+(113.00*0.059)+(-43.09*0.137)+(-584.47*0.069) = £4099.53
MR = 3307.91 + (-482.86*1.600)+(1147.53*1.218)+(45.70*29.260)+(-41.01*20.973)+ (1547.67*0.064)+(-464.92*0.203)+(113.00*0.059)+(-43.09*0.137)+(-584.47*0.069) = £4375.21
LR = 2843.73 + (-482.86*1.600)+(1147.53*1.218)+(45.70*29.260)+(-41.01*20.973)+ (1547.67*0.064)+(-464.92*0.203)+(113.00*0.059)+(-43.09*0.137)+(-584.47*0.069) = £3911.03

Table A8.2c: Unadjusted costs (observed) and adjusted costs for all risk factors from time of anomaly scan up until after delivery (or for a few cases after termination of pregnancy) for period 2005/2006 with TM

Variable	Unadjusted costs		Adjusted costs for all risk factors	
	N = 338, R^2 = 0.875		N = 338, R^2 = 0.889	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Telemedicine (TM)	4050.74 (160.88)	< 0.001	4441.89 (1406.90)	0.002
Direct referral (DR)	3612.22 (253.49)	< 0.001	4101.37 (1347.32)	0.003
Medium risk (MR)	3705.30 (143.89)	< 0.001	4217.20 (1420.37)	0.003
Low risk (LR)	3273.80 (112.59)	< 0.001	3989.20 (1373.29)	0.004
Parity			-279.86 (148.59)	0.061
Number of fetuses			1426.85 (369.99)	< 0.001
Mother's age			4.45 (11.89)	0.709
Gestation at anomaly scan			-87.41 (60.51)	0.150
Diabetes			1762.39 (383.26)	< 0.001
Elevated serum			-142.68 (297.32)	0.632
Downs			-45.28 (308.19)	0.883
Family history of CHD			173.51 (320.60)	0.589
Previous pregnancy with anomaly			367.02 (450.09)	0.415
Tests:				
TM = DR	$F_{1, 334}$ = 2.13, p = 0.145		$F_{1, 325}$ = 1.01, p = 0.317	
TM = MR	$F_{1, 334}$ = 2.56, p = 0.111		$F_{1, 325}$ = 0.56, p = 0.454	
TM = LR	$F_{1, 334}$ = 15.65, p < 0.001		$F_{1, 325}$ = 3.63, p = 0.058	
DR = MR	$F_{1, 334}$ = 0.10, p = 0.750		$F_{1, 325}$ = 0.13, p = 0.718	
DR = LR	$F_{1, 334}$ = 1.49, p = 0.223		$F_{1, 325}$ = 0.15, p = 0.701	
MR = LR	$F_{1, 334}$ = 5.58, p = 0.019		$F_{1, 325}$ = 1.03, p = 0.311	

Adjusted for all risk factors:
TM = 4441.89+(-279.86*1.467)+(1426.85*1.036)+(4.45*29.675)+(-87.41*21.115)+ (1762.39*0.041)+(-142.68*0.086)+(-45.28*0.068)+(173.51*0.071)+ (367.02*0.030) = £3876.18
DR = 4101.37+(-279.86*1.467)+(1426.85*1.036)+(4.45*29.675)+(-87.41*21.115)+ (1762.39*0.041)+(-142.68*0.086)+(-45.28*0.068)+(173.51*0.071)+ (367.02*0.030) = £3535.66
MR = 4217.20+(-279.86*1.467)+(1426.85*1.036)+(4.45*29.675)+(-87.41*21.115)+ (1762.39*0.041)+(-142.68*0.086)+(-45.28*0.068)+(173.51*0.071)+ (367.02*0.030) = £3651.49
LR = 3989.20+(-279.86*1.467)+(1426.85*1.036)+(4.45*29.675)+(-87.41*21.115)+ (1762.39*0.041)+(-142.68*0.086)+(-45.28*0.068)+(173.51*0.071)+ (367.02*0.030) = £3423.49

Table A8.2d: Unadjusted costs (observed) and adjusted costs for all risk factors from time of anomaly scan up until after delivery (or for a few cases after termination of pregnancy) for period 2005/2006 without TM

Variable	Unadjusted costs		Adjusted costs for all risk factors	
	N = 338, R^2 = 0.874		N = 338, R^2 = 0.888	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Direct referral (DR)	3877.06 (150.02)	< 0.001	4036.65 (1342.30)	0.003
Medium risk (MR)	3728.00 (131.52)	< 0.001	4108.02 (1386.78)	0.003
Low risk (LR)	3273.80 (112.73)	< 0.001	3820.37 (1341.27)	0.005
Parity			-265.85 (147.96)	0.073
Number of fetuses			1412.36 (370.75)	< 0.001
Mother's age			2.48 (11.81)	0.834
Gestation at anomaly scan			-77.12 (59.04)	0.192
Diabetes			1876.26 (387.93)	< 0.001
Elevated serum			-192.44 (290.62)	0.508
Downs			-98.77 (304.00)	0.745
Family history of CHD			248.01 (329.12)	0.452
Previous pregnancy with anomaly			409.82 (449.64)	0.363
Tests:				
DR = MR	$F_{1, 335}$ = 0.56, p = 0.456		$F_{1, 326}$ = 0.08, p = 0.775	
DR = LR	$F_{1, 335}$ = 10.33, p = 0.001		$F_{1, 326}$ = 0.97, p = 0.326	
MR = LR	$F_{1, 335}$ = 6.88, p = 0.009		$F_{1, 326}$ = 1.91, p = 0.168	

Adjusted for all risk factors:
DR = 4036.65+(-265.85*1.467)+(1412.36*1.036)+(2.48*29.675)+(-77.12*21.115)+ (1876.26*0.041)+(-192.44*0.086)+(-98.77*0.068)+(248.01*0.071)+ (409.82*0.030) = £3638.62
MR = 4108.02+(-265.85*1.467)+(1412.36*1.036)+(2.48*29.675)+(-77.12*21.115)+ (1876.26*0.041)+(-192.44*0.086)+(-98.77*0.068)+(248.01*0.071)+ (409.82*0.030) = £3709.99
LR = 3820.37+(-265.85*1.467)+(1412.36*1.036)+(2.48*29.675)+(-77.12*21.115)+ (1876.26*0.041)+(-192.44*0.086)+(-98.77*0.068)+(248.01*0.071)+ (409.82*0.030) = £3422.34

Appendix 9: Published paper – Dowie et al (2007)

International Journal of Technology Assessment in Health Care, 23:1 (2007), 116–125.

Telemedicine in pediatric and perinatal cardiology: Economic evaluation of a service in English hospitals

Robin Dowie, Hema Mistry
Brunel University

Tracey A. Young
University of Sheffield

Gwyn C. Weatherburn
Buckinghamshire Chilterns University College

Helena M. Gardiner
Imperial College and *Royal Brompton & Harefield NHS Trust*

Michael Rigby
Imperial College and *Royal Brompton & Harefield NHS Trust*

Giselle V. Rowlinson
Royal Brompton & Harefield NHS Trust

Rodney C. G. Franklin
National Heart & Lung Institute and *Royal Brompton & Harefield NHS Trust*

Objectives: Pediatric cardiology has an expanding role in fetal and pediatric screening. The aims of this study were to observe how district hospitals use a pediatric telecardiology service, and to compare the costs and outcomes of patients referred to specialists by means of this service or conventionally.
Methods: A telemedicine service was set up between a pediatric cardiac center in London and four district hospitals for referrals of second trimester women, newborn babies, and older children. Clinicians in each hospital decided on the role for their service. Clinical events were audited prospectively and costed, and patient surveys were conducted.
Results: The hospitals differed in their selection of patient groups for the service. In all, 117 telemedicine patients were compared with 387 patients seen in London or in outreach clinics. Patients selected for telemedicine were generally healthier. For all patients, the mean cost for the initial consultation was £411 for tele-referrals and £277 for conventional referrals, a nonsignificant difference. Teleconsultations for women and children were

Funding for this study was from the Department of Health and the Charitable Funds Committee of the Royal Brompton and Harefield NHS Trust. The Health Economics Research Group at Brunel University receives funding from the Department of Health Policy Research Programme. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health. We thank the staff for their generous support in Basildon Hospital, Colchester General Hospital, Southend General Hospital, Medway Maritime Hospital in Gillingham, and the Royal Brompton Hospital in London. Liz Evans, Val Asker, Sulie Anstis, Daphne Brown, and Nuala Brady-Murphy, and Tom Carter facilitated the project in these hospitals.

significantly more expensive because of technology costs, whereas for babies, ambulance transfers were much more costly. After 6-months follow-up, the difference between referral methods for all patients was nonsignificant (telemedicine, £3,350; conventional referrals, £2,172), and nonsignificant within the patient groups.
Conclusions: Telemedicine was perceived by cardiologists, district clinicians, and families as reliable and efficient. The equivocal 6-month cost results indicate that investment in the technology is warranted to enhance pediatric and perinatal cardiology services.

Keywords: Costs and cost analysis, Heart defects, Congenital, Outpatients, Perinatal care, Telemedicine

Although the annual incidence of congenital heart disease per 1,000 live births in the United Kingdom appears constant at 1.5 cases for complex abnormalities and 4.5 cases for simple conditions (27), pressures are mounting on the nation's 15 pediatric cardiology (congenital cardiac) units as substantial improvements in survival rates for complex cases have resulted in a sustained expansion in cohorts of children requiring long-term monitoring. Pediatric cardiologists have an expanding role as they follow up these children and triage other children with murmurs (which usually prove to be innocent [23]). They also assess fetuses with suspected congenital heart disease and fetuses in high risk groups of pregnant women (11).

Pediatric cardiologists from most congenital cardiac units hold outreach clinics in district hospitals on a monthly, bimonthly, or quarterly basis (21), thus waiting times for new nonurgent appointments may be many weeks (33). Tertiary fetal medicine centers, where cardiologists specializing in fetal echocardiology hold sessions (11), do not usually provide outreach services, so women in their second trimester often make lengthy journeys for an assessment. Early recognition of congenital heart disease in babies is essential, because deterioration may be sudden and some treatable defects may cause death if diagnosis is delayed. Pediatricians in district hospitals can have difficulty in arranging the urgent transfer of a sick baby if the nearest cardiac center has no suitable cot available. Another receiving center has to be found, causing further delays in diagnosis and treatment.

In the late 1990s, telemedicine and telecare were seen as having a key role in the British government's plans to modernize the National Health Service (NHS) by helping to eliminate unnecessary travel and delay for patients (25). Reliability of the technology for sharing cardiac information between clinicians about children, neonates, and unborn babies had been demonstrated (2;5;9;10;26), but there was no robust information on the cost-effectiveness of pediatric telemedicine services (13), and the situation was unchanged in 2003 (12).

In England in 2001, the Royal Brompton Hospital, which already had pediatric telecardiology links with hospitals in Greece and Portugal (30), introduced a telemedicine service for district hospitals in southeast England that was designed for use in pediatric departments, neonatal units, and obstetric departments. The clinicians and managers decided on the precise role for this service within their hospital. An independent observational study was conducted to see how the service would be used in different hospitals, and to compare the costs and outcomes of patients referred to specialists by means of the service or by conventional methods.

METHODS

Setting

The Royal Brompton Hospital in west London and four hospitals in the towns of Basildon, Colchester, Gillingham, and Southend, between 35 and 65 miles from central London, participated in the project. Pediatric cardiologists from the Royal Brompton held outreach clinics in the hospitals: monthly in Gillingham, and every 3 or 4 months in the other towns. The hospitals recorded between 3,100 and 3,800 deliveries annually. Gillingham had a Level III neonatal intensive care unit providing comprehensive medical neonatal care, and the units in the other hospitals provided Level II high dependency care and short-term intensive care (3). According to the hospitals' established referral patterns, pregnant women were referred to three fetal medicine centers in London, one being linked to the Royal Brompton; babies were referred to three pediatric cardiology units including the Royal Brompton; and infants and children usually attended outreach clinics.

The telemedicine equipment packages installed in the district hospitals included a Tandberg video conferencing system for use with six integrated services digital network (ISDN-6) lines, additional monitors, a video recorder, an object camera visualizer, and an electronic stethoscope sender. Staff were trained to use the equipment. Clinicians could hold face-to-face teleconsultations with specialists at the Royal Brompton with the patients being present and live ultrasound images transmitted as necessary, or prerecorded video ultrasound images could be transmitted (the "store and forward" approach) in the absence of patients. The specialists provided advanced tutoring in fetal heart scanning and pediatric echocardiography.

The telemedicine service complemented the existing outreach services. Basildon and Gillingham hospitals were randomized to begin using the new service 6 months before

Table 1. Mean Cost per Patient for the Components of the Telemedicine Service in the District Hospitals

Mean cost per telemedicine referral (£)	Basildon $N = 38$	Gillingham $N = 61$	Colchester $N = 11$	Southend $N = 7$
Telemedicine equipment	93.56	58.29	323.24	507.93
ISDN-6 line installation, and equipment maintenance contract	11.69	9.49	37.05	68.12
Telemedicine training and service support	12.04[a]	8.00[a]	23.94[b]	36.20[b]
ISDN line rental and call charges	77.00[a]	55.90[a]	35.17[b]	21.65[b]
Total mean cost per referred patient	194.29	131.68	519.40	833.90

[a] Costs and charges pro rated over 12 months.
[b] Costs and charges pro rated over 6 months.

the other two hospitals to undertake a comparative evaluation of intervention versus control sites. Multicenter and local research ethics committees approved the project.

Patients

Three patient groups were considered for teleconsultations: pregnant women referred for ultrasound examination of the fetal heart after an anomaly scan (performed usually between 18 and 22 weeks gestation); newborn babies with a suspected heart problem; and older infants and children referred for cardiac assessment. Project facilitators in the hospitals identified all eligible patients over a 15-month period, including 3 months when the telemedicine equipment was being installed. Babies and children were followed up for a maximum of 12 months. Women were followed up until delivery.

Evaluation

The economic evaluation adopted a cost consequences approach from the dual viewpoints of NHS acute service providers, and patients, and their families. Clinical outcomes after the patients' initial consultation with the specialists were recorded. Postal surveys conducted over 10 months assessed the health-related quality of life of women and children after their initial consultation and the costs incurred by families on hospital visits.

Hospital resource use events were audited by the project facilitators. The items covered babies' and childrens' hospital attendances and admissions relating to their heart problems, women's antenatal attendances and prenatal admissions, the patients' clinical care (tests, investigations, cardiac procedures and cardiac drugs), and the status of NHS personnel who were consulted. Details of ambulance journeys and teleconsultations were recorded, and fieldwork was undertaken in the outreach clinics to estimate mean times for conventional consultations.

Health Service Costs

Hospital Unit Costs. Finance departments in the hospitals supplied unit costs, including overhead, for the relevant resource items at 2001–02 financial year prices. A pharmacy department priced the pharmaceutical items. As there were wide interhospital variations both in the submitted costs and the caseloads of patients, weighted unit costs rather than mean costs were applied to all district items for which information had been supplied by two or more finance departments (17).The weights were derived according to the total number of referrals in each patient group for each hospital.

Telemedicine Service Costs. An annual equivalent cost for the telemedicine items in each hospital, including installation of the ISDN-6 lines and 17.5 percent value added tax (VAT), was calculated, with an expected lifetime for the equipment of 5 years (15), and an annual discount rate of 3.5 percent (14). Telephone bills provided details of ISDN-6 line rental, call charges, and VAT. A mean components cost per telemedicine patient was then derived (Table 1).

Other Cost Components. Staff time was costed using NHS salary scales (6–8) and Netten and Curtis (24), and pro rated according to the mean number of minutes for completing relevant tasks. NHS ambulance trusts provided costs for transferring babies, taking account of distances traveled when making return journeys between hospitals, and time spent waiting (16). A hospital with a retrieval team provided staffing and equipment costs for a neonatal transfer team. Because the distances from the hospitals to London varied by 30 miles, a weighted cost for an ambulance transfer was derived. Postcode data were used to calculate mileages of car journeys made by patients when attending hospital (22), and motoring costs were applied to the mileage (1). The main resource items with costs and patient utilization on the day of the initial specialist consultation are shown in Table 2.

Analytical Perspective

A cohort approach was adopted for the economic analysis, whereby the costs (mean and 95 percent confidence intervals [CI]) of patients referred by means of telemedicine from all four hospitals were compared with the costs of patients referred conventionally over 15 months. Three sets of mean costs per patient were generated: the initial consultation with a specialist, 14 days inclusive of the initial consultation, and a maximum period of 6 months or, for women, until admission before delivery.

Table 2. Resource Items with Costs, and Utilization on the Day of the Initial Consultation with a Specialist

Resource item and mean times in minutes (min)	No. of patients		Cost (£)	
	Telemedicine	Conventional	Telemedicine	Conventional
Pregnant women	$N = 52$	$N = 196$		
Ultrasound attendance	4	196	19.79	42.63
DGH clinician (5 min)	52	NA	3.12	–
Specialist				
Telemedicine (5 min)	52	NA	3.23	–
London (20 min)	NA	196	–	12.73
Counselling (15 min)	NA	39	–	7.50
TM coordinator (5 min)	52	NA	1.17	–
Antenatal clinic attendance	NA	NA	36.00–47.71	36.00–47.71
Clinic staff (10.7–12.5 min)	NA	NA	2.33–7.80	2.33–7.80
Termination	NA	NA	644.01–883.00	644.01–883.00
Prenatal maternity bed day	NA	NA	185.27	185.27
Newborn babies	$N = 17$	$N = 23$		
DGH cot day				
Ventilated intensive care	6	7	690.60	690.60
High dependency care	3	3	441.14	441.14
Special care	6	9	286.92	286.92
Pediatric ward	2	1	228.70	228.70
Specialist cot day				
Neonatal intensive care	1	15	1,020.00	1,020.00
Specialist outpatient clinic	NA	7	–	118.00
DGH neonatologist (20 min)	17	NA	12.48	–
Specialist (20 min)	17	NA	12.92	–
TM coordinator (20 min)	17	NA	4.68	–
Echocardiogram	13	4 DGH 9 London	18.25	18.25 DGH 133.00 London
Ambulance transfer London	1	18	1,476.23	1,476.23
Older children	$N = 48$	$N = 168$		
Outpatient attendance				
Tele-clinic	48	NA	128.45	–
Outreach clinic	NA	156	–	128.00
London clinic	NA	10	–	118.00
DGH consultant				
Telemedicine (15 min)	48	NA	9.36	–
Outreach clinic (11.5 min)	NA	156	–	7.18
Specialist				
Telemedicine (15 min)	48	NA	9.69	–
Outreach (11.5 min)	NA	156	–	7.32
London (9 min)	NA	10	–	5.73
TM coordinator (15 min)	48	NA	3.52	–
Echocardiogram	42	117 outreach 6 London	18.25	28.68 outreach 133.00 London
Resting ECG	25	28 outreach 1 London	14.02	14.02 outreach 45.80 London
Chest X-ray	16	24 outreach 1 London	10.21	10.21 outreach 22.00 London
Specialist bed day	NA	2	–	631.00
DGH bed day	NA	NA	281.16	281.16

DGH, district general hospital; ECG, electrocardiogram; NA, not applicable on day of initial specialist consultation; TM, telemedicine.

Two NHS cost analyses are presented. The primary analysis compares the alternative referral methods for all patients, and for each of the three patient groups. The secondary analysis focuses on the two hospitals that had access to the telecardiology service for 12 months and established regular, although different, usage patterns. Family costs associated with hospital visits are also presented. A sensitivity analysis assesses the likely impact on costs if a telemedicine service in a hospital is shared with other users.

The statistical packages of S-PLUS and Stata Version 8 (28;29) were used to explore differences between referral methods using Kruskal–Wallis test, *t*-tests, Chi-squared tests, and Fisher's exact tests. All statistical tests were two-sided. A p value of $\leq .05$ was considered to be statistically

Table 3. Specialist Referrals from the District Hospitals over 15 Months

	District hospital				
Caseloads over 15 months	Basildon	Gillingham	Colchester	Southend	Total referrals
All referrals					
Pregnant women	34	76	11	127	248
Newborn babies	8	17	7	8	40
Older children	69	54	59	34	216
Total	111	147	77	169	504
Telemedicine					
Duration of access	12 months	12 months	6 months	6 months	
Tele-service used					
Pregnant women	X	✓	X	X	52
Newborn babies	✓	✓	X	✓	17
Older children	✓	X	✓	✓	48
All tele-referrals	38	61	11	7	117

✓, telemedicine referral service used; X, telemedicine referral service not used.

significant. As the distributions of the patient costs were skewed, bias adjusted nonparametric bootstrapping, taking 5,000 iterations of the data, were performed to generate CIs around the means (20).

RESULTS

Over the 15-month period, 504 new patients were assessed by specialists, of whom 117 (23.2 percent) were referred by means of the telemedicine service. However, during the periods when the service was available to the individual hospitals (Table 3), more than half of the 206 patients who became eligible for the new service had a teleconsultation (56.8 percent, 117 of 206). Within the patient groups over 15 months, telemedicine was used for 52 of 248 women, 17 of 40 newborn babies, and 48 of 216 older children.

Demographic and Clinical Attributes

Statistically significant differences were observed only among the women. The telemedicine women were younger by an average of 3.4 years, no one was pregnant with twins, and most had a high risk of conceiving a fetus with congenital heart disease (11) (78.8 percent, 41 of 52: Chi-squared test, $p < .001$). The purpose of most referrals (telemedicine 90.4 percent, 47 of 52; conventional 81.1 percent, 159 of 196) was to screen the fetus rather than to confirm a suspected anomaly. With the babies, although no statistically significant difference was observed, 34.8 percent (8 of 23) of the London transfers had symptoms suggestive of critical congenital heart disease compared with 11.8 percent (2 of 17) of the telemedicine babies ($p = .234$). The two groups of children were similar in age, mean 4.4 years (standard deviation [SD] 5.2) for telemedicine users and 5.1 (4.5) for clinic attendees, and most children were asymptomatic (telemedicine 79.2 percent, 38 of 48; clinic referrals 67.9 percent, 114 of 168; $p = .344$).

Outcome of the Specialist Assessment

A fetal diagnosis of severe or moderately severe congenital heart disease was made for four telemedicine women and thirty-three women seen in London, and there was no statistically significant difference in the referral methods used for the women diagnosed in this way ($p = .126$). There was no significant difference either in the outcomes for the children: three quarters were assessed for heart murmurs (40 seen by means of telemedicine and 126 in clinics; $p = .227$) and most were normal or had self-correcting congenital heart lesions (telemedicine, 90.0 percent (36 of 40); outreach, 81.7 percent (103 of 126); $p = .218$). Not surprisingly, 41.7 percent (20 of 48) of the telemedicine children and 44.6 percent (75 of 168) of the clinic attendees were discharged immediately. Patterns of care for the newborn babies were significantly different. Among the twenty-three babies transferred directly to a cardiac center, nine (39.1 percent) were returned to their referring hospital for medical management, whereas fifteen (88.2 percent) of the seventeen telemedicine babies remained in the district units (Fisher's exact test, $p = .007$).

Health-Related Quality of Life

The EuroQol EQ-5D instrument (4) was completed by twenty-six women assessed by means of telemedicine and eleven who traveled to London from the hospitals in Gillingham and Colchester. The EQ-5D mean (SD) tariff of .72 (.22) for the London travelers, derived from five physical and psychological dimensions, was significantly lower than the tariff for telemedicine women of .86 (.14) (Kruskal–Wallis, $p = .031$). Parents of children referred from the four hospitals completed either an English translation of the QUALIté de vie du Nourrisson (QUALIN) instrument for infants between 4 and 24 months of age (18;19) or the Pediatric Quality of Life Questionnaire (PedsQL™ Generic Core Scales version 4.0 [31;32]) for older children. Twelve telemedicine children had a slightly better, although not significantly better, quality of life than forty-six clinic attendees.

Table 4. Bootstrapped Mean Cost per Patient for Telemedicine and Conventional Referrals over Three Time Periods

Patient groups and time periods		Referral method: Telemedicine	Referral method: Conventional	*p* value
All referred patients		$N = 117$	$N = 387$	
Initial consultation day	Mean (SD)	£411 (£355)	£277 (£862)	.107
	95% CI	£352–£481	£212–£389	
14 days	Mean (SD)	£1,437 (£3,753)	£863 (£3,329)	.114
	95% CI	£888–£2,305	£582–£1,269	
6 months	Mean (SD)	£3,350 (£9,725)	£2,172 (£6,736)	.141
	95% CI	£2,035–£6,020	£1,670–£3,132	
Pregnant women		$N = 52$	$N = 196$	
Initial consultation day	Mean (SD)	£143 (£11)	£59 (£11)	<.001
	95% CI	£141–£147	£58–£61	
14 days	Mean (SD)	£190 (£162)	£167 (£372)	.668
	95% CI	£159–£263	£128–£238	
6 months	Mean (SD)	£925 (£539)	£714 (£728)	.052
	95% CI	£800–£1,097	£632–£849	
Newborn babies		$N = 17$	$N = 23$	
Initial consultation day	Mean (SD)	£917 (£465)	£2,449 (£2,323)	.012
	95% CI	£755–£1,237	£1,782–£4,009	
14 days	Mean (SD)	£6,962 (£7,018)	£11,206 (£8,378)	.099
	95% CI	£4,019–£10,671	£7,771–£14,532	
6 months	Mean (SD)	£17,121 (£20,937)	£20,156 (£18,066)	.624
	95% CI	£9,401–£30,776	£14,292–£30,040	
Older children		$N = 48$	$N = 168$	
Initial consultation day	Mean (SD)	£523 (£245)	£235 (£556)	.001
	95% CI	£463–£604	£175–£377	
14 days	Mean (SD)	£834 (£2,255)	£265 (£620)	.004
	95% CI	£487–£2,187	£199–£420	
6 months	Mean (SD)	£1,103 (£2,721)	£1,423 (£3,575)	.572
	95% CI	£554–£2,388	£959–£2,062	

Note. Statistical tests: *t*-tests were performed on non-bootstrapped mean costs.
SD, standard deviation; CI, confidence interval.

Analyses of NHS Costs (Table 4)

All Referred Patients. No statistically significant difference was found in the mean NHS cost for all 504 patients using the alternative referral options in any time period, although the telemedicine service was always more costly.

Pregnant Women. Assessments of women by means of the fetal telecardiology service, which was only used in Gillingham, were significantly more costly, but the cost of the technology of £131.68 per referral (see Table 1) accounted for most of the difference. The comparative 14-day mean costs were relatively similar. Over the months until delivery, care experienced by the telemedicine cohort bordered on being significantly more costly ($p = .052$). However, there were wide variations in the frequencies (and numbers) of antenatal visits made by women and prenatal inpatient admissions in all four hospitals, which resulted in large variations in their costs as reflected in the large standard deviations at 6 months.

Newborn Babies. The mean cost for teleconsultations involving babies was significantly cheaper, because only one baby incurred the additional cost of an ambulance transfer later the same day. The London-referred babies spent an average of 5.5 days receiving specialist care, at a cost per neonatal cot day of £1,020, in the 13 days after their transfer. Nevertheless, the mean cost per telemedicine baby over 14 days was not significantly lower. After 6 months, telemedicine remained the cheaper option.

Older Children. Teleconsultations for children were significantly more costly than clinic attendances because of the costs of the technology. The mean 14-day cost for the telemedicine cohort was also significantly higher, even though nearly 90 percent of the children assessed remotely were either discharged immediately or booked for a follow-up appointment. This mean cost of £834 (95 percent CI, £487–£2,187) included an outlier who received emergency treatment valued at over £15,000, but even when the patient was excluded, the 14-day cost for the cohort remained significantly higher (£508, 95 percent CI, £454–£581; $p = .009$). After 6 months, there was relatively little difference between the referral strategies, telemedicine being cheaper.

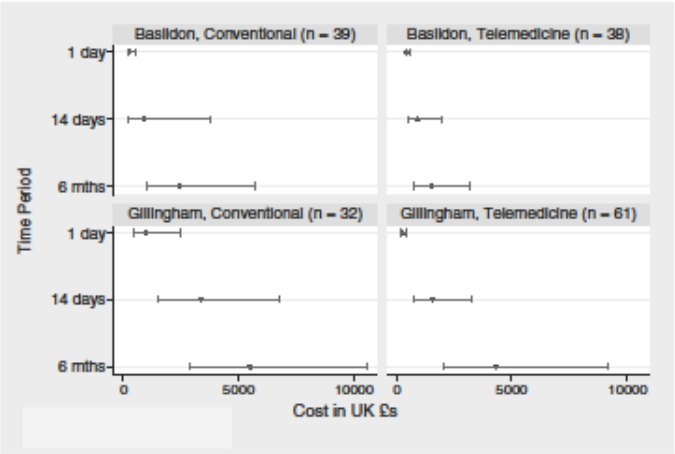


Figure 1. Bootstrapped mean cost per patient for conventional referrals and telemedicine referrals in Basildon hospital, covering newborn babies and older children, and Gillingham hospital, covering pregnant women and newborn babies. Dots, mean costs; brackets, 95 percent confidence intervals.

Telemedicine Service Costs in Two Hospitals

The telemedicine service was used over 12 months for babies and children in Basildon and for women and babies in Gillingham. Figure 1 indicates the magnitude of the hospitals' mean patient costs for the three time periods. In each hospital after 6-months follow-up, the telemedicine mean was lower than the mean for conventional referrals, although the cost differences were not statistically significant. The higher costs for Gillingham generally were attributable to neonatal case severity in the Level III unit.

Family Costs

The median cost of hospital visits, inclusive of travel, incidental expenses, and any loss of income, for six Gillingham women who journeyed 35 miles to London was significantly higher than the median cost for twenty-six women who attended the local hospital where their ultrasound scan was recorded for telemedicine transmission: £50.36 (interquartile range [IQR], £38.00–£77.20) versus £12.59 (IQR, £2.52–£15.60; Kruskal–Wallis, $p = .002$). Local visits were mostly completed within 2.5 hours compared with 5.5 hours for London visits. Children usually traveled by car to the four district hospitals for either an outreach clinic appointment or a teleconsultation, so the median costs were similar: £8 (IQR, £5–£12) for sixteen telemedicine families and £6 (IQR, £3–£16) for fifty clinic families. A median of 2.5 hours was spent on the visits by both groups of families.

Sensitivity Analysis

In Colchester hospital, the telemedicine equipment was installed in a central suite and made available to other users. In 2004–05, the suite was used once or twice a week for cancer network teleconferences, with approximately 10 cases discussed each session (R. Emslie, personal communication). For the sensitivity analysis, the costs of the telemedicine service were shared among the eleven cardiac children who used the service (Table 3) and 300 cancer cases. After adding the telephone charges for the cancer network to the observed costs of setting up and operating the telemedicine service, the re-attributed service cost per child was £18.94. According to this scenario, the mean cost of the eleven teleconsultations was slightly lower than the mean for 48 consultations in the hospital's outreach clinics: £240 (95 percent CI, £178–£515) versus £268 (95 percent CI, £168–£712; $p = .901$).

DISCUSSION

The economic evaluation was designed as an observational study with four hospitals randomized as "early" or "delayed" users of the telecardiology service. Uptake of the service was slower than anticipated and for fewer numbers of referrals. So, instead of comparing the telemedicine cases from the "early" hospitals with the conventionally referred cases from the "delayed" hospitals as originally intended, a cohort approach was adopted for the primary analysis whereby the attributes and costs of patients referred by means of the telemedicine service were compared with patients referred

conventionally over 15 months. Williams et al. (38) observed that conventional methods for conducting biomedical research, such as randomized controlled trials, may not be suitable for evaluating telehealthcare systems because of the emergent nature of these systems and their uncertain impact on organizational and professional structures.

Telemedicine patients were generally in a better state of health, although no statistically significant difference in the presenting clinical circumstances was observed. The patient cost results over 6 months for babies and children also indicated that the telemedicine cohorts had lower utilization levels of NHS hospital resources, although not significantly so, than the conventionally referred cohorts. The district clinicians, when deciding on the methods of referral for their patients, may have been influenced by the relatively short distances of 35 to 65 miles between their hospitals and London, even though emergency teleconsultations were easily arranged, particularly on weekdays. Our equivocal 6-month cost results differ from those of a randomized controlled trial of a telemedicine (virtual outreach) service for routine outpatient consultations in two NHS hospitals, in which the virtual outreach 6-month mean cost was significantly greater than the standard outpatient clinic cost (15). The virtual outreach system linked primary care physicians with consultants in eight specialties by means of personal computers and ISDN-2 lines (15;34).

In both projects, the equivalent annual costs of the video conferencing systems and ISDN line rental and call charges were key resource components for the initial consultation, while the key variable for assigning these costs to individual patients was the number of consultations conducted using the systems (15). However, the extent to which the cost of the technology impacted on the overall mean teleconsultation cost for an individual hospital in our project depended upon the case mix of patients being referred in this manner. For ambulatory patients (children or pregnant women), teleconsultations were more costly. With sick babies, the converse applied: transferring a baby by ambulance to a specialist center at a cost of £1,476 was a far more expensive strategy. Sharing the technology with other users in one hospital reduced the mean teleconsultation cost for a child to £240, but the volume of additional users needed to achieve this target cost over 6 months was considerable: 300 patients over 30 hours.

For this single-specialty project, the equipment packages had to be of sufficient quality to transmit color Doppler ultrasound images of the heart for diagnostic purposes, which necessitated ISDN-6 lines. The clinical benefit of this capital investment was apparent in the outcome data from the initial consultations. Discharge rates were similar for the two cohorts of cardiac children. Only two of the forty-seven telemedicine women whose prerecorded fetal images were assessed as normal were followed up by a specialist, and all the women gave birth to healthy babies. In the area of neonatal care, the district doctors relied on the telemedicine service for problem solving; that is, when they were uncertain about the diagnosis of a heart problem or the management of a baby who was failing to thrive.

The district clinicians found learning to use the telemedicine system was less difficult than acquiring sufficient expertise in scanning the heart for remote diagnosis. It was important that trust and professional respect existed between specialists and district staff, who valued the educational benefits (36). Parents commended the telemedicine service because of its potential for reducing waiting times for appointments for children referred for screening, and its convenience for pediatricians requiring advice for patients about whom they were particularly concerned, so reducing the need for families to travel to London. As for the technology itself, patients and parents found remote consultations acceptable as long as transmission difficulties did not arise, which happened very infrequently (35).

The role adopted for the telecardiology service in the district hospitals was to supplement, rather than substitute, existing services provided from the Royal Brompton Hospital. The three hospitals that used the service for pediatric referrals continued as before to host outreach clinics. By allowing the district clinicians as providers (37), in collaboration with the specialists, to determine the roles locally for the telemedicine service, patterns in the use of the technology for perinatal and pediatric care of new and review patients were permanently established in all four hospitals.

POLICY IMPLICATIONS

For asymptomatic children, our results confirm that there are numerous patients requiring assessment with a low yield of abnormality. In the future, this workload will be best handled by training local pediatricians in the use of echocardiography as a screening tool with the back up of a telemedicine link up when uncertainty over the presence of pathology arises. Consideration should be given to introducing remote diagnostic facilities in neonatal units for triaging babies for specialist care. Teleconsultations would provide diagnostic confirmation in infants with questionable heart disease, thus avoiding costly ambulance journeys. Most women referred for perinatal echocardiography have a normal fetus. As standards of second trimester fetal heart imaging improve, higher detection rates of equivocal and unequivocal cardiac anomalies will lead to higher rates of women journeying to fetal medicine centers for evaluation. Transmitting prerecorded ultrasound images during telemedicine sessions is an efficient and economic method of providing this important service.

CONTACT INFORMATION

Robin Dowie, PhD (robin.dowie@brunel.ac.uk), Senior Research Fellow, **Hema Mistry**, BA, MSc (hema.mistry@brunel.ac.uk), Research Fellow in Health Economics, Health Economics Research Group, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK

Tracey A. Young, BSc, MSc, (CStat) (t.a.young@sheffield.ac.uk), Lecturer, Health Economics and Decision Science, School of Health and Related Research (ScHARR), University of Sheffield, 30 Regent Street, Sheffield S1 4DA, UK

Gwyneth C. Weatherburn, PhD (gweath01@bcuc.ac.uk), Reader in Medical Imaging, Research Centre for Health Studies, Buckinghamshire Chilterns University College, Gorelands Lane, Chalfont St Giles, Buckinghamshire, HP8 4AD, UK

Helena M. Gardiner, PhD (helena.gardiner@imperial.ac.uk), Senior Lecturer and Director of Perinatal Cardiology, Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College, Hammersmith Campus, Du Cane Road, London W12 0NH, UK; Honorary consultant in Perinatal Cardiology, Department of Pediatric and Fetal Cardiology, Royal Brompton & Harefield NHS Trust, Sydney Street, London SW3 6NP, UK

Michael Rigby, MD (m.rigby@rbht.nhs.uk), Senior Lecturer, Department of Pediatrics, National Heart and Lung Institute/Imperial College, Cale Street; Consultant Pediatric Cardiologist, Department of Pediatric Cardiology, Royal Brompton Hospital, Sydney Street, London SW3 6NP, UK

Giselle V. Rowlinson, MBBS (giselle72@doctors.org.uk), Specialist Registrar, Department of Pediatric Cardiology, Royal Brompton Hospital, Sydney Street, London SW3 6NP, UK

Rodney C. G. Franklin, MD (r.franklin@rbht.nhs.uk), Honorary Senior Lecturer in Paediatric Cardiology, Department of Pediatric Cardiology, National Heart & Lung Institute, Dovehouse Street; Consultant Pediatric Cardiologist, Pediatric Cardiology Directorate, Royal Brompton Hospital, Sydney Street, London SW3 6NP, UK

REFERENCES

1. Automobile Association website. *Petrol car running costs*. Available at: www.theaa.com/allaboutcars/advice/advice_rcosts_home.html. Accessed March 13, 2003.
2. Belmont JM, Mattioli LF, Goertz KK, Ardinger RH Jr, Thomas CM. Evaluation of remote stethoscopy for pediatric telecardiology. *Telemed J*. 1995;1:133-149.
3. British Association of Perinatal Medicine. *Standards for hospitals providing neonatal intensive and high dependency care*. 2nd ed. London: British Association of Perinatal Medicine; 2001.
4. Brooks R. EuroQol: The current state of play. *Health Policy*. 1996;37:53-72.
5. Casey F, Brown D, Craig BG, Rogers J, Mulholland HC. Diagnosis of neonatal congenital heart defects by remote consultation using a low-cost telemedicine link. *J Telemed Telecare*. 1996;2:165-169.
6. Department of Health. *Advanced letter (MD) 1/02: Pay and conditions of service of hospital medical and dental staff and doctors in public health medicine and the community health service*. London: DoH; 2002.
7. Department of Health. *Advanced letter (NM) 1/2002: Nurses' and midwives' pay for 2002/03*. London: DoH; 2002.
8. Department of Health. *Advanced letter (PTA) 1/2002: Professions allied to medicine and related grades of staff – pay for 2002/2003*. London: DoH; 2002.
9. Finley JP, Sharratt GP, Nanton MA, et al. Paediatric echocardiography by telemedicine – nine years' experience. *J Telemed Telecare*. 1997;3:200-204.
10. Fisk NM, Sepulveda W, Drysdale K, et al. Fetal telemedicine: Six month pilot of real-time ultrasound and video consultation between the Isle of Wight and London. *Br J Obstet Gynaecol*. 1996;103:1092-1095.
11. Gardiner HM. Access to perinatal cardiology in the United Kingdom. *Postgrad Med J*. 2001;77:1-3.
12. Hailey D, Ohinmaa A, Roine R. *Evidence for the benefits of telecardiology applications: A systematic review*. Edmonton: Alberta Heritage Foundation for Medical Research; 2004.
13. Hersh WR, Wallace JA, Patterson PK, et al. *Telemedicine for the Medicare population: Pediatric, obstetric and clinician-indirect home interventions*. Evidence Report/Technology Assessment: Number 24, Supplement. (AHRQ Publication No. 01-E060). Rockville: Agency for Healthcare Research and Quality; 2001. Available at: www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat1.chapter.34386.
14. HM Treasury. *The Green Book: Appraisal and evaluation in central government*. London: The Stationery Office; 2003. Available at: www.hm-treasury.gov.uk/media/f/05553/Green_Book_03.pdf
15. Jacklin PB, Roberts JA, Wallace P, et al. Virtual outreach: Economic evaluation of joint teleconsultations for patients referred by their general practitioner for a specialist opinion. *BMJ*. 2003;327:84-88.
16. Leslie A, Stephenson T. Neonatal transfers by advanced neonatal nurse practitioners and paediatric registrars. *Arch Dis Child Fetal Neonatal Ed*. 2003;88:F509-F512.
17. Longworth L, Young T, Buxton MJ, et al. Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. *Liver Transpl*. 2003;9:1295-1307.
18. Manificat S, Dazord A, Langue J, et al. A new instrument to evaluate infant quality of life. *QoL Newsletter*. 1999;23:7-8.
19. Manificat S, Dazord A, Langue J, et al. Évaluation de la qualité de vie du nourrisson et du très jeune enfant: Validation d'un questionnaire. Étude multicentrique européenne. *Arch Pédiatr*. 2000;7:605-614.
20. Manly BFJ. *Randomisation, bootstrap and Monte Carlo methods in biology*. 2nd ed. London: Chapman & Hall Texts in Statistical Science; 1997.
21. Maynard L, Franklin RCG, Wray J. A survey of professionals delivering secondary care regarding their requirements for paediatric cardiac services as provided by specialists. *Cardiol Young*. 2005;15:489-492.
22. Multimap. Multimap Com. *London;* 2003. Available at: www.multimap.com/.
23. Murugan SJ, Thompson J, Parsons JM, et al. New outpatient referrals to a tertiary paediatric cardiac centre: Evidence of increasing workload and evolving patterns of referral. *Cardiol Young*. 2005;15:43-46.
24. Netten A, Curtis L. *Unit costs of health and social care 2002*. Canterbury: University of Kent at Canterbury, Personal Social Services Research Unit; 2002.

25. NHS Executive. *Information for health: An information strategy for the modern NHS 1998-2005*. Leeds: NHS Executive; 1998.
26. Nitzkin JL, Zhu N, Marier RL. Reliability of telemedicine examination. *Telemed J*. 1997;3:141-157.
27. Petersen S, Peto V, Rayner M. *Congenital heart disease statistics 2003*. Oxford: University of Oxford, British Heart Foundation Health Promotion Research Group, Department of Public Health; 2003. Available at: www.heartstats.org/uploads/documents/%5CCoHDstatistics.pdf.
28. S-PLUS. *Guide to statistics, volumes 1 and 2*. Seattle, WA: Mathsoft; 2000.
29. StataCorp. *Stata statistical software: Release 8.0*. College Station, TX: Stata Corporation; 2003.
30. Tsilimigaki A, Maraka S, Tsekoura T, et al. Eighteen months' experience with remote diagnosis, management and education in congenital heart disease. *J Telemed Telecare*. 2001;7:239-243.
31. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: Reliability and validity of the Pediatric Quality of Life Inventory™ version 4.0 generic core scales in healthy and patient populations. *Med Care*. 2001;39:800-812.
32. Varni JW, Burwinkle T, Seid M, Zellner J. The PedsQL™ as a population health measure: Implications for states and nations. *QoL Newsletter*. 2002;28:4-6.
33. Wagstaff MH, Rigby ML., Redington AN. Increasing workload and changing referral patterns in paediatric cardiology outreach clinics: Implications for consultant staffing. *Heart*. 1998;79:223-224.
34. Wallace P, Barber J, Clayton W, et al. Virtual outreach: A randomised controlled trial and economic evaluation of joint teleconferenced medical consultations. *Health Technol Assess*. 2004;8(50):1-106.
35. Weatherburn G, Dowie R, Mistry H, Young T. An assessment of parental satisfaction with mode of delivery of specialist advice for pediatric cardiology: Face-to-face versus teleconference. *J Telemed Telecare*. 2006:12(Suppl 1):57-59.
36. Weatherburn G, Dowie R, Mistry H, Young T. Specialist pediatric cardiology advice by telemedicine. Proceedings of the Second IASTED Conference TELEHEALTH. July 3-5, 2006, Banff, AB, Canada. Available at: http://www.actapress.com/SearchResults.aspx#Papers.
37. Whitten PS, Mackert MS. Addressing telehealth's foremost barrier: Provider as initial gatekeeper. *Int J Technol Assess Health Care*. 2005:21:517-521.
38. Williams T, May C, Mair F, Mort M, Gask L. Normative models of health technology assessment and the social production of evidence about telehealth care. *Health Policy*. 2003;64:39-54.

Appendix 10: Published paper – Dowie et al (2008)

Cost implications of introducing a telecardiology service to support fetal ultrasound screening

Robin Dowie*, Hema Mistry*, Tracey A Young†, Rodney CG Franklin‡ and Helena M Gardiner‡§

*Health Economics Research Group, Brunel University, Uxbridge; †Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield; ‡Brompton Fetal Cardiology, Royal Brompton and Harefield NHS Trust, London; §Institute of Reproductive and Developmental Biology, Imperial College, London, UK

Summary

A district hospital in south-east England used a telecardiology service for fetal cardiac diagnosis alongside an existing arrangement for referring pregnant women directly to perinatal cardiologists in London for detailed fetal echocardiography. Women were identified for referral according to local protocols when having a second trimester anomaly scan. For the telemedicine referrals, the sonographers video-recorded images from the anomaly scans for transmission during monthly videoconferences. The cost of the women's antenatal care was calculated from the specialist assessment until delivery, while family costs were collected in a postal survey. Over 15 months, telemedicine was used in 52 cases, while 24 women were seen in London. The London women were more likely to have had an ultrasound abnormality (29% *v* 10%, $P = 0.047$). A telemedicine assessment of 5 min duration was more costly than an examination in London (mean cost per referral of £206 *v* £74, $P < 0.001$). However, the telecardiology service was cost neutral after 14 days and for the extended period until delivery. Travel costs for London women averaged £37 compared with £5.50 for the telemedicine referrals. Telemedicine may be useful to support perinatal cardiologists in the UK whose workloads are expanding in response to improved standards in antenatal ultrasound screening.

Introduction

Almost all pregnant women in the UK are offered an anomaly scan at 18–22 weeks gestation, which is usually performed by an obstetric sonographer in the obstetric unit of a district hospital. The anomaly scan typically examines the baby's head, spine, heart, abdominal shape and content, and limbs. The UK National Screening Committee has recommended replacing the minimum imaging standard of a four-chamber view of the heart with one that includes the outflow tracts,[1] thus detecting major lesions, such as transposition of the great arteries and tetralogy of Fallot that require major surgery. Despite local training initiatives the proportion of obstetric units who report that they routinely assess outflow tracts has remained at about 60% for the last five years,[1,2] and detection rates of major abnormalities of the heart remain low overall;[3,4] the estimated rate for England and Wales in the mid 1990s was about 23%, but with marked variation between regions.[5]

In the National Health Service (NHS) the major determinants of referral to a fetal cardiologist for detailed echocardiography are the local protocols of the referring hospitals that emphasize referral of women whose pregnancies are thought to be at an increased risk for congenital heart disease (CHD). While it is accepted that the risk of CHD in these women is about 3%, the majority of babies with CHD are not born to at-risk women. Thus, this protocol-driven strategy results in a referral pattern where only a small proportion of all the babies examined by fetal cardiologists have heart defects.[6]

Diagnostic fetal echocardiography may be undertaken in paediatric cardiology departments, isolated from obstetric services, or in tertiary fetal medicine centres. Both are sited in major city hospitals and many pregnant women make lengthy journeys to attend the regional centres. In the case of a suspected abnormality only a few days elapse between the local decision to refer and the specialist appointment, so travel arrangements have to be made quickly. For all women, there are travel costs to consider as well as incidental costs, such as child-minding and possible loss of earnings for themselves or partners.

Telemedicine offers an alternative method of delivering a fetal echocardiography service. The digitized images from obstetric ultrasound machines are suitable for electronic transmission,[7] and the structures of a fetal heart can be visualized with clarity after 18 weeks gestation.

Accepted 25 September 2008
Correspondence: Robin Dowie, Health Economics Research Group, Brunel University, Uxbridge UB8 3PH, UK (*Fax: +44 1895 269 708; Email: robin.dowie@brunel.ac.uk*)

DOI: 10.1258/jtt.2008.080401

Teleconsultations can be conducted either in realtime with a sonographer examining a heart while a remote specialist jointly views the images,[8–10] or via the transmission of pre-recorded images (i.e. store-and-forward mode).

In London, the Royal Brompton Hospital (RBH) installed a telemedicine system in the 1990s. Paediatric telecardiology links were formed with hospitals in Greece and Portugal.[11] In 2001, the RBH set up a telecardiology network with four district general hospitals (DGHs) in south-east England for the provision of specialist advice to clinicians in obstetric and paediatric departments. The local clinicians and managers decided on the precise role for this service in their hospital. An economic analysis covering all three telemedicine applications (fetal, neonatal and paediatric) showed that the telecardiology network was cost neutral when compared with conventional referral practice once patients were followed up over six months.[12]

In one DGH, the telemedicine system was used for fetal cardiac diagnosis alongside the existing arrangement for referring women directly to London specialist centres for fetal echocardiography. A pattern of monthly 'virtual outreach' clinics was established during which obstetric sonographers transmitted to a specialist, ultrasound images of fetal hearts that had been recorded when women were screened for anomalies in the second trimester of their pregnancy. The present paper describes the impact of the new service on NHS costs of antenatal care received by the women and on their personal costs. It also identifies cost-related factors that could influence the wider introduction of fetal telecardiology.

Methods

The Medway Maritime Hospital in Gillingham, where the fetal telecardiology service was used, is located 55 km from central London. The fetal ultrasound department routinely referred women to perinatal cardiologists based at the RBH. The average number of women who were delivered in Gillingham was 3810 per year.

The hospital was supplied with a videoconferencing system (model 2500, Tandberg) and connected to the ISDN network at 384 kbit/s. Training in using the telemedicine equipment was provided, as well as advanced training in fetal heart scanning. At the RBH, the existing telemedicine suite contained a videoconferencing system (model 2500, Tandberg) that was used frequently by hospital staff.

Patients

Over a 15-month period in 2001–2002, demographic details were recorded of all women referred for an assessment of their baby's heart and the women's clinical events from the time of the anomaly scan until they were delivered. The personal information included any antenatal risk factors for fetal abnormalities. The factors were of two kinds: traditional CHD risk factors (a maternal or family history of CHD; maternal dependency on insulin, anti-epilepsy or lithium therapies; or current multifetal pregnancy[13]); and risk factors for Down's syndrome (notably an adjusted risk based on a quadruple serum test result or a nuchal translucency measurement ≥3.5 mm). Diagnostic information on fetal heart abnormalities was extracted from ultrasound and echocardiography reports. The project was approved by the appropriate ethics committees.

The referred women formed two groups for the analyses: a telemedicine-referrals group and a London group as a comparator. The telemedicine group consisted of women who, at the time of their anomaly scan, were identified for echocardiographic referral according to the obstetric department's protocol for high-risk women, and women whose fetal heart images during the anomaly scan were abnormal in some respect or were poorly visualized. The London group consisted of women who were referred for the same reasons in the five months before the telecardiology service entered regular use and those who travelled to London during the following 10 months, usually during the intervening weeks between the scheduled teleconferencing sessions. It also included women whose fetuses were strongly suspected of having a heart defect. A consultation with a fetal cardiologist was held according to the hospital protocols, usually within two working days.

Health service and women's costs

Detailed resource use information was collected for each woman from the anomaly scan until delivery. Unit costs in pounds sterling (£1 ≈ €1.3 ≈ US$ 2.0) for the resource items were obtained from hospital finance departments in Gillingham and specialist centres in London. The items covered ultrasound scans, antenatal and outpatient review consultations, and prenatal inpatient admissions. The unit cost for an outpatient or antenatal clinic attendance included nursing supervision and clerical staff time, consumables and equipment, overheads and capital charges. The status of the clinician (midwife or doctor) who examined each woman at each attendance was recorded separately. The mean durations of the consultations were estimated to allow the time of these clinicians to be costed and the staff cost was added to the unit cost in the analysis. The cost for an ultrasound examination incorporated the sonographer's time as well as consumables, an administrative cost, overheads and capital charges. Originally, the costs applied to the 2001–2002 financial year, but for the present paper they were inflated to 2005–2006 prices (Table 1).[14] When deriving the fetal telemedicine system cost, an annual equivalent cost covering the equipment, the installation of the ISDN lines and 17.5% tax (VAT) was calculated, with an expected lifetime for the equipment of 5 years[15] and an annual discount rate of 3.5%.[16] The monthly invoices from the telephone company provided details of the ISDN line rental

Table 1 Antenatal resource items with unit costs (2005–2006 prices) and frequencies of the clinical events over the period between the specialist assessment and delivery

		Frequency of events in total period until delivery			
		Telemedicine referrals ($n = 52$)		London referrals ($n = 24$)	
Antenatal resource items	Unit cost (£)*	Events	Patients (n)	Events	Patients (n)
Antenatal ultrasound scan (anomaly or other scans)	36.22	52 anomaly, 74 other scans	52	39 other scans	20
Antenatal or outpatient clinic attendance (review)	75.95	580 attendances	50	206 attendances	20
Clinician consulted in clinic (mean time 15 min)					
Consultant	10.94				
Doctor in training	6.85				
Midwife	3.83				
Teleconsultation clinicians (mean time 5 min)					
Specialist, obstetric sonographer and co-ordinator	8.79	55 teleconsultations	52	–	–
Prenatal maternity bed-day	240.71	57 days	19	8 days	3
Termination of pregnancy	857.68	2 procedures	2	4 procedures	4
Ultrasound scan at the specialist hospital (inclusive of specialist's mean time 20 min)	64.69	7 scans	6	27 scans	24

*Costs were provided by Gillingham hospital for 2001–02 and inflated to 2005–06

and call charges. A mean component cost per telemedicine patient was then derived (Table 2).

Information on travel arrangements and expenditure when visiting hospital was obtained in a 10-month postal survey in Gillingham of women who were referred to the specialists and women with either CHD or Down's syndrome risk factors who were not referred. Postcode data were used to calculate mileages of car journeys[17] and motoring costs were applied to the mileage.[18] For other travel methods, the actual costs recorded by the women were used.

Statistical analysis

After attributing unit costs to the clinical events recorded for all women (Table 1), a mean cost per woman was estimated for telemedicine referrals and London referrals for each of three time periods: the events relating to the specialist consultation; antenatal care over 14 days inclusive of the specialist consultation; and antenatal care spanning the second and third trimesters until delivery. These follow-up periods were adopted for the evaluation of all three applications of the telecardiology service in the district hospitals.[12] The short-term period covered emergency events arising from the specialist assessment (such as a termination of pregnancy following prenatal diagnosis). Fourteen days duration was selected to minimize the weighting effect on a patient's costs of any complications associated with an emergency event. The longer-term period covering the women's care until delivery had also been used in a companion paper from this project on the costs of NHS maternity care for women with high-risk and low-risk pregnancies.[13]

As the patient total costs were skewed, a simulation technique known as bootstrapping[19] was used, taking 5000 iterations of the data, to calculate 95% confidence intervals around the mean costs. Statistical analyses were conducted using the software Stata version 10[20] and S-PLUS.[21]

Table 2 Mean costs for the fetal telemedicine system in the district hospital (2005–2006 prices)

Components of the fetal medicine system	Mean cost (£) $n = 52$ referred women
Telemedicine equipment	60.99
ISDN line installation, and equipment maintenance contract	9.56
Telemedicine training and service support	10.96*
ISDN line rental and call charges (mean time 5 min per woman)	76.62*
Total mean cost per woman (inclusive of value added tax)	158.14

*Costs and charges pro rated over 12 months

Results

Over 15 months 76 Gillingham women were referred for a fetal cardiac assessment, a referral rate of 16 per 1000 maternal deliveries. Fifty-two were assessed via telemedicine and 24 travelled to London. Thirteen London women were referred before the telecardiology service entered regular use and 11 women over the following 10 months. The two cohorts were similar in terms of mean age, parity and duration of gestation (Table 3). One woman was pregnant with twins. The proportion of women with CHD-risk factors was higher in the telemedicine cohort (79%) than in the cohort referred directly to London (46%, $P = 0.016$). The London women were more likely to have had an ultrasound abnormality (7/24 versus 5/52, Fisher's exact test, $P = 0.047$). Of the 64 women referred for screening according to the departmental protocol, 47 (73%) were assessed during videoconferencing sessions. As well as following up the pregnancies until delivery, admissions to the neonatal unit were monitored and no missed diagnosis was detected among the screened babies.

The Gillingham ultrasound department adopted the store-and-forward method for using the telemedicine service. During 10 months, 10 pre-arranged videoconferencing sessions were held. Obstetric sonographers video-recorded fetal

Table 3 Demographic characteristics of women referred to fetal cardiology specialists over 15 months from Gillingham hospital

Characteristics of referred women	Referral method	
	Telemedicine service $n = 52$	Direct to London $n = 24$
Mean age (SD)	29 (7)	30 (7)
Parity		
Primaparous	19 (37%)	12 (50%)
Multiparous	33 (64%)	12 (50%)
Risk factors for fetal abnormalities		
CHD factors	41 (79%)	11 (46%)
Down's syndrome factors	5 (10%)	7 (29%)
No apparent factor	6 (12%)	6 (25%)
Timing of anomaly scan		
Gestation ≤ 20 weeks	16 (31%)	5 (21%)
Gestation = 21 weeks	28 (54%)	12 (50%)
Gestation ≥ 22 weeks	8 (15%)	7 (29%)

CHD = congenital heart disease
Statistical tests: Age: t-test, $P = 0.56$; Parity: chi-square test, $P = 0.27$; Gestation week: chi-square test, $P = 0.33$; and CHD-risk women assessed via telemedicine: Fisher's exact test, $P = 0.016$

heart images during the anomaly scans of women and the recordings were transmitted during the videoconferences. On one occasion an urgent teleconsultation was arranged, and on three other occasions, a serious heart abnormality was diagnosed during an anomaly scan performed within 48 hours of a videoconference session, so the images were transmitted for verification. About five video recordings were transmitted per session (range 2 to 8, median 5) and the mean time spent discussing a case was five minutes. Diagnostic confirmation was provided for the five cases where an abnormality had already been detected; the cases were hypoplastic left heart (3), coarctation of the aorta with mild aortic stenosis (1) and mid-muscular ventricular septal defect (1). (Four of these mothers then travelled to London for a diagnostic scan and counselling about the diagnosis, planned management and prognosis.) Five women with normal babies were re-assessed by the specialists either personally (two women) or by telemedicine (three women) to exclude progressive valvular disease because of a family history of pulmonary or aortic stenosis.

Antenatal costs of the referral services

A telemedicine assessment was significantly more costly than an examination in London (bootstrapped mean costs per referral of £206 versus £74, $P < 0.001$) (Table 4). This was a predictable result since the mean cost per woman of the telemedicine service in its first year of operation was £158 (Table 2). However, there was no significant difference in the antenatal mean costs for the two cohorts in each follow-up period. Six women had a late termination of pregnancy; two in the telemedicine group (4%) and four among the directly referred women (17%). These procedures were assigned a unit cost of £858, and this outcome for the pregnancies of four London-referred women contributed to the higher mean cost for the London cohort within the 14-day period. Among all the referred women, 25 were under treatment for diabetes or epilepsy; 21 in the telemedicine group (40%) and four (17%) in the directly-referred group (chi-square test, $P = 0.041$). A separate analysis of the study's pregnancy data-set found that the antenatal care for women with underlying medical conditions was significantly more expensive than antenatal care for singleton women in general.[13] Thus the differing risk profiles of the telemedicine and London cohorts would account for the higher mean cost for the telemedicine women for the extended period until delivery.

Family costs

Details of travel arrangements when attending Gillingham hospital were included in 23 questionnaires from women whose anomaly scans were video-recorded for a telemedicine assessment and 98 questionnaires from women managed locally. The mean travel costs for the two groups in 2005–2006 prices were similar: £5.49 (SD 5.03) for telemedicine referrals and £7.01 (SD 5.51) for local care. Six women were seen in London and their mean travel cost of £37.33 (SD 12.06) was significantly greater (F test, $P < 0.001$). Twenty-one (17%) of the Gillingham hospital attenders reported a loss of income or an expenditure on babysitting, bringing their total mean cost to £59.96 (SD 45.53).

Sensitivity analysis

A sensitivity analysis was used to examine the effect on the initial specialist assessment when personal travel costs of the referred women were added to the NHS resource costs. (Multiple imputation[22] was performed to estimate the costs of women who had not been surveyed.) The bootstrapped mean cost for the telemedicine group remained significantly greater than the mean cost for the London group: £213 (SD 13) versus £108 (SD 27), t-test, $P < 0.001$. Sharing the telemedicine system with other users was the scenario with the greatest potential for moderating the cost of a teleconsultation.[12,15] So, as the Gillingham telecardiology service was also used in the neonatal unit,

Table 4 Bootstrapped mean NHS cost per patient for telemedicine and London referrals over three time periods (2005–2006 prices)

Time periods for antenatal care		Telemedicine referrals $n = 52$	London referrals $n = 24$	*P* value*
Day 1: Initial specialist assessment (inclusive of the anomaly scan for the telemedicine referrals)	Mean (SD) 95% CI	£206 (£12) £204 to £213	£74 (£20) £68 to £86	<0.001
14-day period (inclusive of Day 1)	Mean (SD) 95% CI	£267 (£180) £232 to £345	£348 (£516) £198 to £657	0.46
Total period from assessment on Day 1 until delivery	Mean (SD) 95% CI	£1469 (£831) £1280 to £1743	£1181 (£584) £983 to £1447	0.093

CI = confidence interval
*t-tests conducted on non-bootstrapped means

the NHS resource costs for 52 women and nine babies seen by telemedicine were compared with the costs for 24 women and eight babies who travelled to London. The telemedicine initial assessment cost was significantly less than the London assessment: £335 (SD 396) versus £1131 (SD 2708), t-test, $P = 0.027$. This was due to the high cost of neonatal ambulance transfers. However, at 14 days and six months, the alternative referral services were cost neutral: the six-month mean costs were £5303 (SD 14,176) for the telemedicine group and £6650 (SD 11,500) for the London group, t-test, $P = 0.65$).

Discussion

In the 1980s, as obstetric ultrasound techniques advanced to enable most forms of CHD to be detected in fetal life, a pattern of practice developed in the UK for hospital obstetric departments to refer women with an increased risk for fetal CHD to tertiary units for fetal echocardiography. For instance, in 1987, almost 1000 high-risk patients were referred to a tertiary centre at Guy's Hospital in London.[23] A national report on heart disease in 2002 noted that referrals for fetal echocardiography were increasing as obstetric departments' routine anomaly scans (performed at 18 to 22 weeks) had become more searching.[24] Perinatal or fetal cardiology consultations in England are mostly held in the 12 regional congenital cardiac units (or their associated fetal medicine centres). When the telecardiology service for this project was introduced in the four DGHs in 2001, only the Gillingham obstetric department, which was already referring high-risk women to the RBH cardiac unit, took up the service primarily for obtaining assurance of normality in ultrasound images of the fetal heart from high-risk women who would otherwise have travelled to London for echocardiography. The referral arrangements for the obstetric departments in two of the other hospitals were already established with other congenital cardiac units in London; these were unchanged and a total of 161 women were referred during our study period. In the fourth hospital, detailed fetal ultrasonography was performed within the obstetric department, so in general, relatively few women were referred (11 cases over the period).[12]

From the perspective of a district hospital, the most important telecommunications factor is the annual volume of use. The expected lifetime for the telemedicine equipment was five years,[15] so the mean patient cost was based on the first year of the equipment's use plus the ISDN line rental and call charges for the same period. When the expense of the videoconferencing equipment in Gillingham was shared among the 52 women and nine babies, the adjusted mean cost (in 2005–2006 prices) for the service in the first year was reduced to £135 per patient. Senior ultrasonographers discussed pre-recorded videoed images of fetal hearts with the fetal cardiologists, spending five minutes on each case. If the women had been referred directly to the London clinics, as new patients they would have been booked for a 45-minute appointment to allow time for counselling should an abnormality be detected.

The magnitude of the financial savings to women who avoided travelling to London was determined primarily by the distances of the hospitals from their homes. The average return journey to the London clinics was 145 km compared with 21 km to the local hospital.

By assessing comprehensively the antenatal care received by the surveyed women during the second and third trimesters of their pregnancy, we observed how a teleconsultation was a relatively modest cost component of the total package of care provided by the NHS. Moreover, the telemedicine referral service was cost neutral in the extended period until delivery. Analyses of the study's data-sets showed that the total antenatal mean cost for a cohort of women could be markedly affected by costs associated with their pre-existing or pregnancy-related medical conditions.[13] Thus, in the telemedicine group, almost 80% of the 52 women had CHD risk factors, such as diabetes, epilepsy or a family history of CHD, and 19 (37%) experienced a prenatal admission to hospital at a cost of £241 per bed-day.

Since the study, the fetal telecardiology service in Gillingham has continued to use the same videoconferencing system. During 2006, eight telemedicine sessions of 45–60 min duration were held and the videos of 84 women were assessed. Eleven of these women were followed up by London specialists and the hearts of four babies proved to be abnormal. Maternal obesity can impede visualization of the fetal heart during an anomaly scan. More than half the fetal cardiac images transmitted were from women with a body mass index above 30 but were of sufficient quality to distinguish abnormal cases. The ultrasound department remained committed to the telemedicine service because of its educational value for training staff in addition to its clinical effectiveness. Counselling is currently undertaken face to face in London as it is felt that this is better for achieving mutual understanding between the specialist and parents when major decisions, such as a termination of a pregnancy, are made, and specialist fetal cardiac liaison nurses are available to provide parental support. (It is also relevant that fetal images from specialist echocardiography machines are of a higher quality than images from obstetric ultrasound machines.) In the longer term, however, it may become acceptable to counsel parents during teleconferences, especially for more straightforward CHD lesions.

Only a small proportion of the 75 paediatric cardiology consultants in England are also specialists in fetal cardiology, so if their workloads are to be kept in check, a change of practice is needed to reduce the amount of time spent by them on the great majority of high-risk women whose fetuses are normal. The number of women with fetal cardiac risk factors in the UK is likely to increase as more and more survivors of CHD (both males as potential fathers, and females) reach adulthood, and the incidence of type 2 diabetes continues to rise.[25] As maternal obesity becomes more prevalent, a subpopulation with poor fetal heart imaging is likely to contribute to the workload.

New challenges to the existing workload of fetal cardiologists will arise as standards of antenatal screening in maternity services rise,[26] particularly as visualization of outflow tracts of the fetal heart is increasingly performed during second trimester anomaly scans.[3] Higher detection rates of equivocal and unequivocal cardiac anomalies by obstetric sonographers will, in turn, lead to appropriately higher rates of referral for perinatal echocardiography. Nuchal translucency screening in the first trimester is already available in one-quarter of English hospital trusts[2] in accordance with recently issued national guidelines.[27] This will substantially increase the number of women requiring an early and more complex fetal cardiac scan performed by a specialist, given the independent higher risk of CHD despite a normal karyotype.

In conclusion, the present study demonstrates the potential for fetal telecardiology to support the few perinatal cardiologists in the UK in providing timely specialist consultations for women with frank or suspected fetal abnormalities. Setting up networks between specialist units and referring district hospitals for the transmission of pre-recorded video or digitally-stored ultrasound images, where up to 12 patients can be discussed in hour-long videoconference sessions, is likely to be an efficient and economic way of providing support and training to local obstetric screening programmes and will avoid unnecessary referral to tertiary centres.

Acknowledgements: We are grateful to staff at the obstetric department of Medway Maritime Hospital in Gillingham for participating in this study. We thank Helen Kennett, Daphne Brown, Dr Mike Rigby and Dr Gwyn Weatherburn for assistance. Funding for the study was from the Policy Research Programme of the Department of Health and the Charitable Funds Committee of the Royal Brompton and Harefield NHS Trust. The views expressed in the publication are those of the authors and not necessarily those of the funders.

References

1 UK National Screening Committee. *Antenatal Ultrasound Screening. Ultrasound Survey of England: 2002.* See: http://www.screening.nhs.uk/fetalanomaly/ultrasound_survey.pdf (last checked 6 September 2008)
2 Healthcare Commission. *Key findings of the 2007 Maternity Service Review.* See: http://www.healthcarecommission.org.uk/_db/_documents/Maternity_FINAL_BRIEFING_NOTE.doc (last checked 6 September 2008)
3 Bricker L, Garcia J, Henderson J, *et al.* Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views. *Health Technol Assess* 2000;4:1–193
4 Acherman RJ, Evans WN, Luna CE *et al.* Prenatal detection of congenital heart disease in southern Nevada: the need for universal fetal cardiac evaluation. *J Ultrasound Med* 2007;12:1715–19
5 Bull C. Current and potential impact of fetal diagnosis on prevalence and spectrum of serious congenital heart disease at term in the UK. *Lancet* 1999;354:1242–7
6 Kilby M. *Fetal Medicine.* See http://www.bwhct.nhs.uk/print/index/about_clinicalreport_intro/about_clinical_report-200304_intro/about-clinicalreport_200304/about-clinicalreport-fetalmed.htm (last checked 25 September 2008)
7 Begg L, Chan FY, Edie G, Hockey R, Wootton R. Minimum acceptable standards for digital compression of a fetal ultrasound video-clip. *J Telemed Telecare* 2001;7 (Suppl. 2):88–90
8 Fisk NM, Sepulveda W, Drysdale K, *et al.* Fetal telemedicine: six month pilot of real-time ultrasound and video consultation between the Isle of Wight and London. *Br J Obstet Gynaecol* 1996;103:1092–5
9 Chan FY, Soong B, Watson D, Whitehall J. Realtime fetal ultrasound by telemedicine in Queensland. A successful venture? *J Telemed Telecare* 2001;7 (Suppl. 2):7–11
10 Sharma S, Parness IA, Kamenir SA, *et al.* Screening fetal echocardiography by telemedicine: efficacy and community acceptance. *J Am Soc Echocardiogr* 2003;16:202–8
11 Tsimiklis A, Maraka S, Tsekoura T, *et al.* Eighteen months' experience with remote diagnosis, management and education in congenital heart disease. *J Telemed Telecare* 2001;7:239–43
12 Dowie R, Mistry H, Young TA, *et al.* Telemedicine in pediatric and perinatal cardiology: economic evaluation of a service in English hospitals. *Int J Technol Assess Health Care* 2007;23:116–25
13 Mistry H, Dowie R, Young TA, *et al.* Costs of NHS maternity care for women with multiple pregnancy compared with high-risk and low-risk singleton pregnancy. *BJOG* 2007;114:1104–12
14 Curtis L, Netten A. *Pay and Prices Index. Unit Costs of Health and Social Care 2006.* Canterbury: Personal Social Services Research Unit, University of Kent at Canterbury, 2006
15 Jacklin PB, Roberts JA, Wallace P, *et al.* Virtual outreach: economic evaluation of joint teleconsultations for patients referred by their general practitioner for a specialist opinion. *BMJ* 2003;327:84–8
16 HM Treasury. *The Green Book: Appraisal and evaluation in central government.* See http://www.hm-treasury.gov.uk/media/9/C/Green_Book_03.pdf (last checked 6 September 2008)
17 Multimap. See http://www.multimap.com/ (last checked 6 September 2008)
18 Automobile Association. *Archive of Running Costs 1998–2007.* See http://www.theaa.com/motoring_advice/running_costs/archive.html (last checked 6 September 2008)
19 Manly BFJ. *Randomisation, Bootstrap and Monte Carlo Methods in Biology.* 2nd edn. London: Chapman and Hall, 1997
20 StataCorp. Stata Statistical Software: Release 10.0. College Station, TX: Stata Corporation, 2007
21 MathSoft, Inc. S-PLUS 2000 Professional Release 2. Seattle, WA: MathSoft, Inc., 1999
22 Rubin DB. *Multiple Imputation for Non-response in Surveys.* New York, NY: Wiley, 1987
23 Allan LD, Chita SK, Sharland GK, Fagg NL, Anderson RH, Crawford DC. The accuracy of fetal echocardiography in the diagnosis of congenital heart disease. *Int J Cardiol* 1989;25:279–88
24 Anon. Fifth report on the provision of services for patients with heart disease. *Heart* 2002;88 (Suppl. 3):1–56
25 Reig DS, Palda VA. Type 2 diabetes in pregnancy: a growing concern. *Lancet* 2002;359:1690–2
26 Haicombe J, Fairgrieve S. *National training needs analysis for antenatal screening services across England.* See http://www.perinatal.nhs.uk/screening/resources/Training_%20Needs_Analysis.pdf (last checked 6 September 2008)
27 UK National Screening Committee. *Fetal Anomaly Screening Programme.* See http://www.nsc.nhs.uk/hottopics/hottopics_fnd.htm (last checked 6 September 2008)

Appendix 11: Published paper – Mistry et al (2007)

DOI: 10.1111/j.1471-0528.2007.01458.x
www.blackwellpublishing.com/bjog

Costs of NHS maternity care for women with multiple pregnancy compared with high-risk and low-risk singleton pregnancy

H Mistry,[a] R Dowie,[a] TA Young,[b] HM Gardiner[c] on behalf of the TelePaed Project Team

[a] Health Economics Research Group, Brunel University, Uxbridge, Middlesex, UK [b] Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK [c] Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College, London, UK
Correspondence: H Mistry, Health Economics Research Group, Brunel University, Uxbridge, Middlesex UB8 3PH, UK.
Email hema.mistry@brunel.ac.uk

Accepted 11 June 2007. Published OnlineEarly 26 July 2007.

Objective To compare antenatal and obstetric costs for multiple pregnancy versus singleton pregnancy risk groups and to identify factors driving cost differentials.

Design Observational study over 15 months (2001–02).

Setting Four district hospitals in southeast England.

Population Consecutive women with multiple pregnancy and singleton women with risk factors for fetal congenital heart disease (CHD) (pregestational diabetes, epilepsy, or family history of CHD) or Down syndrome, and a sample of low-risk singleton women.

Methods Clinical care was audited from the second trimester anomaly scan until postnatal discharge, and the resource items were costed. Multiple regression analysis determined predictors of costs.

Main outcome measures NHS mean costs of antenatal and obstetric care for different types of pregnancy.

Results A total of 959 pregnancies were studied. Three percent of 243 women with multiple pregnancy reached 40 weeks of gestation compared with 54–55% of 163 low-risk and 322 Down syndrome risk women and 36% of 231 cardiac risk women. Antenatal costs for cardiac risk (£1,153) and multiple pregnancy (£1,048) were nearly double the costs for other two groups ($P < 0.001$). As 63% of multiple births were delivered by caesarean section, the obstetric cost for multiple pregnancy (£3,393) was £1,000 greater overall. Pregestational diabetes was the most influential factor driving singleton costs, resulting in similar total costs for multiple pregnancy women (£4,442) and for women with diabetes (£4,877).

Conclusions Our analyses confirm that multiple pregnancies are substantially more costly than most singleton pregnancies. Identifying women with diabetes as equally costly is pertinent because of the findings of the Confidential Enquiry into Maternal and Child Health that standards of maternal care for diabetics often are inadequate.

Keywords Costs and cost analysis, delivery, multiple, obstetric, pregnancy, pregnancy in diabetes, prenatal care.

Please cite this paper as: Mistry H, Dowie R, Young T, Gardiner H on behalf of the TelePaed Project Team. Costs of NHS maternity care for women with multiple pregnancy compared with high-risk and low-risk singleton pregnancy. BJOG 2007;114:1104–1112.

Introduction

In England and Wales over the 5 years to 2004, the annual multiple pregnancy rate per 1000 maternities (women delivered) was relatively constant, ranging between 14.7 and 15.0.[1] Maternal care for twin pregnancies appears to be twice as costly as care for mothers of singletons, although little relevant information is available. Women experiencing multiple pregnancies are much more likely to give birth before reaching full term and to be delivered by caesarean section than those with singleton pregnancies—in a national audit in 2000–01, the twin pregnancy caesarean rate for England and Wales was 58.5% compared with 21.3% for women overall.[2]

Henderson *et al.*[3] estimated mean hospital maternity costs (in 1998 prices) in a London health region as £1,360 for mothers of singletons and £2,836 for mothers of twins. Assisted conception occurs in about 18% of all multiple pregnancies in the UK,[4] and Ledger *et al.* undertook a modelling study to determine direct costs to the NHS of pregnancies conceived through *in vitro* fertilisation (IVF). They estimated maternal

costs as £3,122 for a singleton birth and £6,058 for twin births (in 2001–02 prices).[5] In the Netherlands, Lukassen *et al.*[6] similarly observed that maternal costs for twin pregnancies after IVF treatment were at least double the costs for singleton IVF pregnancies.

No comparative information appears to have been published on maternal costs of multiple pregnancies and high-risk and low-risk singleton pregnancies in Britain. We had an opportunity to explore the relative costliness of care for these groups using a dataset that was assembled for a study evaluating the role of telemedicine in fetal cardiology in four maternity services in southeast England.[7] To establish the annual incidence of fetal cardiac anomalies, cohorts of women with fetal cardiac risk factors, including multiple pregnancies and pregestational diabetes or risk factors for Down syndrome, were identified and their clinical events were audited prospectively from midway through the second trimester. A sample of low-risk women was also monitored. Unit costs were applied to the clinical events. Thus, the dataset can be used to estimate costs of maternal care in the later months of pregnancy and to isolate patient attributes that may be driving cost differentials between risk groups.

Methods

Study design

The project took place in four district hospitals where between 3100 and 3800 women were delivered annually. Over 15 months, three risk groups of pregnant women were identified at the time of their anomaly scan (usually between 18 and 20 weeks of gestation). The groups are as follows:

1 All women with increased risk factors for fetal congenital heart disease (CHD).[8]
2 All women with an elevated risk for Down syndrome following a serum test or nuchal translucency ultrasound scan or chromosomal test.
3 A sample of women of low risk.

The risk factors for fetal CHD are described in Table 1. Women whose screening results were within normal limits formed the low-risk category. It was beyond the scope of the original telemedicine study to collect information on all low-risk women to provide baseline costs. Instead, women in groups (1) and (2) were matched in terms of the woman's year of birth, parity (primigravida or multigravida), month of gestation, and calendar month of anomaly scan, with low-risk women to obtain a representative sample of low-risk women. One-fifth of the women in groups (1) and (2) were matched.

Project facilitators in the district hospitals identified the eligible women and entered demographic and clinical details from their maternity records in an audit database. The women were followed prospectively from the date of the anomaly scan until they delivered and were discharged from the postnatal ward. Multicentre and local research ethics committees approved the project.

Table 1. Factors defining high-risk maternal populations for fetal CHD and Down syndrome

Maternal fetal CHD factors[8] were as follows:
Pregestational diabetes mellitus;
Drugs (e.g. antiepilepsy or lithium drug therapies);
Family history or maternal CHD;
Current multiple fetal pregnancy.
Down syndrome risk factors were as follows:
Woman's adjusted risk based on serum test result greater than the 'at risk' cutoff level adopted by the local screening programme;
Nuchal translucency measurement of greater than 3.0 mm;
Abnormal chromosome test result.

NHS resources and costs

Resource items for antenatal care from the time of the women's anomaly scan covered antenatal and specialist ultrasound scans, antenatal and outpatient clinic attendances inclusive of staff time, specialist consultations either face-to-face or through telemedicine, and antenatal inpatient admissions. Resource items covering obstetric care included the inpatient admission prior to transfer to a labour ward or obstetric theatre, labour/delivery bed day (recorded in hours), mode of delivery, blood products that were transfused, and the mother's stay on the postnatal ward. Home births were included.

Finance departments in the district hospitals and two fetal medicine centres in London provided unit costs, including overheads, for the resource items at 2001–02 prices. As there were variations in both the hospitals' caseloads of eligible women and the district costs, weighted unit costs were attributed.[9] The weights for the unit costs were derived according to the overall distribution of surveyed women in each hospital. For specialist care in London, average unit costs were attributed. Staff time for midwives and doctors was valued using NHS salary scales[10,11] and Netten and Curtis[12] and prorated according to the mean duration of antenatal consultations. Table 2 shows the weighted costs and the distribution of women who used the resources.

Statistical analysis

Comparative analyses between women with a completed multiple pregnancy and three risk groups of women with completed singleton pregnancies (cardiac risk, Down syndrome risk, and low risk) are presented. As the distributions of patient costs within the risk groups were skewed, bias adjusted nonparametric bootstrapping were performed to obtain a bootstrapped mean antenatal cost and a bootstrapped

Table 2. Weighted costs for key resource components of antenatal and obstetric care and the numbers of women who used the items

Resource components and number of women for whom the event occurred on one or more occasions	Weighted cost (£) (2001–02 prices)	Singleton pregnancy			Multiple pregnancy (n = 243)
		Low risk (n = 163)	Down syndrome risk (n = 322)	Cardiac risk (n = 231)	
Antenatal ultrasound scans at the district hospitals, i.e. anomaly scan or other scans	£22.71	163	317	228	239
Antenatal ultrasound scans at the specialist hospitals (specialist's mean time 20 minutes)	£55.36	1	50	50	59
Antenatal clinics: hospital or community antenatal or other (review)	£36.00–£50.51	163	322	231	243
Clinician consulted in clinic (mean time)					
Consultant (12.5 minutes)	£7.80	—	—	—	—
Midwife (10.7 minutes)	£2.33	—	—	—	—
Telemedicine consultation	£143.00	0	4	33	0
Prenatal maternity bed day					
During the antenatal period	£279.84	65	86	98	115
Prior to transfer to labour ward or obstetric theatre	£279.84	45	94	86	112
Mode of obstetric delivery*					
Normal birth	£941.16	100	207	116	69
Forceps birth	£1,043.81	5	3	7	8
Ventouse birth	£1,026.47	6	14	13	11
Caesarean birth (without complications)	£1,582.65	47	85	86	154
Home birth	£480.75	5	9	7	0
Blood transfusions (various products)	—	7	19	5	27
Labour/delivery bed day	£412.09	158	312	227	242
Postnatal maternity bed day	£281.90	129	264	191	227

Antenatal clinic attendances and hospital bed days covered nursing staff, clerical staff, consumables, equipment, overheads, and capital costs. Telemedicine consultations covered the costs of equipment, telephone line rental and calls, and the time of the clinicians.[7]
*Three women gave birth in specialist hospitals (one was a normal birth and the other two were delivered by caesarean section) and four opted for a water birth.

mean obstetric cost per woman, taking 2000 iterations of the data, to generate confidence intervals around the means.[13] The analyses were performed using Stata version 9[14] (College Station, TX, USA) and S-PLUS (Seattle, WA, USA);[15] t tests, F tests, and chi-square tests were used, and the tests were two-sided. A P value of ≤0.05 was considered to be statistically significant for the comparative analyses.

Modelling

To determine what risk factors were important predictors of the cost of maternity care for singleton and multiple pregnancies, a multiple regression model was fitted to the observed caseloads of women. A backward regression approach was adopted for the multiple regression analysis, whereby all potential risk factors were included in a multiple regression model and the least significant variables were removed, one at a time, until only those variables with a significance level of <0.1 remained. Models were fitted separately for antenatal costs and obstetric costs. Goodness-of-fit of the regression models was assessed by the regression error specification test (RESET).[14]

Results

A total of 1165 women with completed pregnancies met the selection criteria identified earlier. However, 206 women were excluded from the analyses because information on their care during the final weeks of pregnancy was incomplete. In total, our analysis concentrates on 959 women, of whom 243 (25%) had a multiple pregnancy (238 were expecting twins and 5 were expecting triplets). Of the 716 singleton women, 163 were low-risk women, 322 were at risk of Down syndrome, and 231 had cardiac risk factors. Within the cardiac risk group, 19 women had two risk factors, so overall, 57 (25%) had pregestational diabetes, 56 (24%) were treated for epilepsy, 126 (55%) had a maternal or family history of CHD, and 12 (5%) women had other factors (such as lithium therapy or vanishing twin syndrome). Over 90% of the women with Down syndrome risk had a serum test result above the cutoff level adopted by the local screening programme, although their baby was normal. Most of the 206 women excluded from the analyses became eligible during the last

3 months of recruitment, so the obstetric care for those who had a full-term pregnancy occurred outside the project's follow-up period. One hundred and ninety (92%) of the excluded women had a singleton pregnancy (77 low risk, 70 Down syndrome risk, and 43 cardiac risk) and 16 (8%) had multiple pregnancies.

Demographic characteristics and birth outcomes

There were significant differences in the attributes of the four groups of women. (Table 3)

- Maternal age: The low-risk and Down syndrome risk women were older than the cardiac risk and the multiple pregnancy women (*t* test, $P < 0.001$).
- Parity: Three-quarters of cardiac risk women were experiencing their second or subsequent pregnancy compared with two-thirds of the low-risk and Down syndrome risk women and over half of the multiple pregnancy women (chi-square test, $P = 0.003$).
- Timing of anomaly scan: Down syndrome risk women were scanned for anomalies earlier in their second trimester than the other groups of women (chi-square test, $P < 0.001$).
- Gestation at birth: Only 3% of women expecting twins reached 40 weeks of gestation compared with over half of the low-risk and Down syndrome risk women and more than one-third of cardiac risk women (chi-square test, $P < 0.001$).
- Birthweight: The mean birthweight for each singleton group was between 3.3 and 3.4 kg; the mean for multiple birth babies was significantly lower at 2.4 kg (*t* test, $P < 0.001$).

Key resource components

The resource items of greatest cost were inpatient bed days and mode of obstetric delivery (Table 2). Cardiac risk and multiple pregnancy women who were admitted to prenatal wards during their second or third trimester accumulated significantly more bed days than Down syndrome risk or low-risk women: the mean (SD) bed days for the first two groups of admitted women were 5.0 (6.5) and 4.0 (4.0), respectively, and 3.3 (3.8) and 2.5 (2.4) for the other groups (*t* test, $P = 0.003$). Women who experienced multiple births had significantly longer stays in hospital after delivery than

Table 3. Demographic characteristics and birthweights

	Singleton pregnancy			Multiple pregnancy ($n = 243$)
	Low risk ($n = 163$)	Down syndrome risk ($n = 322$)	Cardiac risk ($n = 231$)	
Maternal age				
Mean (SD)	31.1 (5.8)	32.7 (6.2)	28.6 (5.6)	29.7 (5.0)
Median	31	34	29	30
IQR	28–36	29–37	24–33	27–33
Parity, *n* (%)				
Primiparous	53 (32.5)	107 (33.2)	58 (25.2)	100 (41.2)
Multiparous	109 (66.9)	214 (66.5)	172 (74.8)	141 (58.0)
Gestation at anomaly scan, *n* (%)				
≤18 weeks	6 (3.7)	135 (41.9)	14 (6.1)	13 (5.3)
19–21 weeks	121 (74.2)	102 (31.7)	140 (60.6)	156 (64.2)
≥22 weeks	36 (22.1)	85 (26.4)	77 (33.3)	74 (30.5)
Gestation at birth, *n* (%)				
≤30 weeks	1 (0.6)	2 (0.6)	7 (3.0)	10 (4.1)
31–35 weeks	4 (2.5)	22 (6.8)	18 (7.8)	82 (33.7)
36–37 weeks	9 (5.5)	26 (8.1)	51 (22.1)	82 (33.7)
38–39 weeks	61 (37.4)	96 (29.8)	72 (31.2)	61 (25.1)
40 weeks	41 (25.2)	69 (21.4)	41 (17.7)	5 (2.1)
≥41 weeks	47 (28.8)	107 (33.2)	42 (18.2)	3 (1.2)
Birthweight (kg)				
Number of babies	158	312	223	489
Mean (SD)	3.4 (0.5)	3.4 (0.7)	3.3 (0.7)	2.4 (0.6)
Median	3.4	3.5	3.4	2.4
IQR	3.1–3.7	3.1–3.8	2.9–3.8	2.0–2.8

IQR, inter-quartile range.
Not all birthweights were recorded. Parity was unknown for five women.

the singleton groups: the means (SDs) were 5.6 (4.0) bed days for mothers of twins and 2.7 (1.8) to 3.5 (2.7) bed days for the singleton mothers (F test, $P < 0.001$). The differential was attributable to the mode of delivery: 63% of the mothers of twins were delivered by caesarean section compared with 30% of all singleton mothers (37% for cardiac risk women). Mothers who underwent surgery required longer stays to recover, while those with twins, who had lower birthweights, may have stayed on until the babies were ready for discharge.

Bootstrapped mean antenatal and mean obstetric costs

Bootstrapped antenatal mean costs for cardiac risk women (£1,153) and multiple pregnancy women (£1,048) following their anomaly scan were nearly double that of the Down syndrome risk and low-risk women, and the cost differences across the four risk groups were significant ($P < 0.001$) (Table 4 and Figure 1). The bootstrapped obstetric mean cost of £3,393 for multiple birth maternities was approximately £1,000 greater than the mean for each singleton group, and the difference was significant ($P < 0.001$). Consequently, the bootstrapped mean total cost for multiple pregnancies was also significantly higher than the means for other groups: £4,442 for women expecting twins or triplets versus £3,625 for cardiac risk women, £2,751 for Down syndrome risk women, and £2,616 for low-risk women ($P < 0.001$). Figure 1 indicates the magnitude of the total mean costs for the different groups of women.

Sensitivity analysis for caesarean section costs

The unit cost for a caesarean section without complications for each of the four hospitals differed by £350: two hospitals had 'high' costs of £1,797 and £1,771, and the other two hospitals had 'low' costs of £1,425 and £1,420. For the sensitivity analysis, the bootstrapped analyses of the obstetric costs for all women were repeated with the baseline cost for a caesarean section of £1,583 (Table 2) being replaced first by a weighted high cost of £1782.88, and second, by a weighted low cost of £1,422.60. For all four groups of women, there was a significant difference in the mean obstetric costs when comparing the high versus low caesarean section analyses. For the multiple pregnancy women, in particular, the high and low mean obstetric costs were £3,520 and £3,292, respectively (Paired Wilcoxon test: $P < 0.001$).

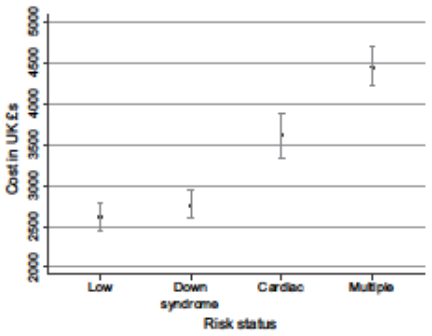


Figure 1. Bootstrapped total mean costs. ●, mean cost; ⊢⊣, 95% CI.

Modelling

The first step for the regression modelling was to identify demographic characteristics and risk factors that may generally influence antenatal and/or obstetric costs. These included:

- Age: As mothers' ages increase, obstetric costs increase since age is a confounder with respect to preterm delivery, induction of labour, and caesarean section.

Table 4. Bootstrapped costs for pregnant women

	Singleton pregnancy			Multiple pregnancy ($n = 243$)	F test* (P value)
	Low risk ($n = 163$)	Down syndrome risk ($n = 322$)	Cardiac risk ($n = 231$)		
Antenatal costs					
Mean (SD)	£598 (£643)	£594 (£769)	£1,153 (£1,397)	£1,048 (£989)	$P < 0.001$
95% CI	£512–£718	£523–£695	£1,001–£1,366	£938–£1,194	
Obstetric costs					
Mean (SD)	£2,018 (£839)	£2,157 (£1,210)	£2,472 (£1,399)	£3,393 (£1,567)	$P < 0.001$
95% CI	£1,887–£2,142	£2,040–£2,304	£2,304–£2,666	£3,225–£3,633	
Total costs					
Mean (SD)	£2,616 (£1,134)	£2,751 (£1,567)	£3,625 (£2,094)	£4,442 (£1,927)	$P < 0.001$
95% CI	£2,450–£2,794	£2,601–£2,940	£3,341–£3,873	£4,220–£4,709	

*Statistical tests conducted on non bootstrapped means.

- Parity: Primiparous mothers may receive extra antenatal care, leading to higher antenatal costs. They may also have higher obstetric costs because they are more likely to have a hospital birth than a home birth to minimise the risk of complications arising during delivery and also they may spend longer periods on postnatal wards after delivery than multiparous mothers.
- Type of pregnancy: Mothers pregnant with twins or triplets have higher antenatal and obstetric costs than mothers of singletons overall because of the increased risk of both maternal complications and fetal difficulties (such as twin to twin transfer syndrome).
- Pregestational diabetes: As the disease may precipitate maternal complications,[16] women with diabetes require extra care resulting in higher antenatal and obstetric costs (for women with diabetes, the mean antenatal cost was £1,604 and the mean obstetric cost was £3,273).
- Epilepsy: This disease may also precipitate maternal complications,[17] so women on antiepileptic therapy require additional care resulting in higher antenatal and obstetric costs (for women on antiepileptic therapy, the mean antenatal cost was £1,386 and the mean obstetric cost was £2,411).
- Family history: Maternal CHD may also lead to complications, so these women may require extra antenatal and obstetric care, leading to higher antenatal and obstetric costs.[18]
- Gestation at time of anomaly scan: If women have an anomaly scan before 19 weeks, they may have additional antenatal check-ups before delivery, leading to higher antenatal costs.
- Gestation at time of delivery: If a mother is delivered prematurely because of complications, she may require additional obstetric care, which results in higher obstetric costs. She may have also been hospitalised for antenatal care, with high antenatal costs as a consequence.

For the univariate (one variable at a time) and multiple regression modelling, we considered all 959 women individually. Separate models were fitted to identify which of the above factors influenced antenatal and obstetric costs (*t* tests were used to assess statistical significance, and the tables of results are available from the authors). Univariate analysis identified the following factors associated with antenatal costs: age ($P = 0.001$), type of pregnancy ($P < 0.001$), diabetes ($P < 0.001$), epilepsy ($P < 0.001$), and gestation at time of delivery ($P < 0.001$); and for obstetric costs: age ($P = 0.044$), parity ($P < 0.001$), type of pregnancy ($P < 0.001$), diabetes ($P < 0.001$), and gestation at time of delivery ($P < 0.001$).

Multiple regression analysis was applied to the antenatal and obstetric costs to establish all significant factors that affect costs (Table 5). For antenatal care, the multiple regression model showed four factors that were significant predictors of high costs: type of pregnancy ($P = 0.001$), diabetes ($P < 0.001$), epilepsy ($P < 0.001$), and gestation (preterm) at time of delivery ($P = 0.054$). According to the model, antenatal care for women with diabetes was on average £800 greater than for women overall, the antenatal costs for epileptic women were £660 greater, and for women with multiple pregnancies, the costs were £289 greater. The model for obstetric costs found five significant predictors: the risk factor of diabetes ($P < 0.001$); and for all women, age ($P = 0.065$), parity (primiparous pregnancy [$P < 0.001$]), type of pregnancy ($P < 0.001$), and gestation at time of delivery (earlier gestation being more costly [$P < 0.001$]). Obstetric costs for both multiple pregnancy women and women with diabetes were £750 greater than for women overall. The *P* value results from the RESET test for both the antenatal and obstetric costs were not significant, indicating that the models were well specified.

Table 5. Multiple regression reduced models for antenatal and obstetric costs

	Coefficient in £s	Standard error	*t* statistic	*P* value
Antenatal costs ($R^2 = 0.080$)				
Diabetes				
Yes	803.34	136.50	5.89	<0.001
No	Base	*	*	*
Antiepileptic therapy				
Yes	661.14	134.49	4.92	<0.001
No	Base	*	*	*
Multiple birth				
Yes	289.07	86.90	3.33	0.001
No	Base	*	*	*
Gestation at time of delivery**	−25.46	13.18	−1.93	0.054
Constant	711.82	50.16	14.19	<0.001
Obstetric costs ($R^2 = 0.242$)				
Parity				
Primiparous	Base	*	*	*
Multiparous	−504.84	87.27	−5.79	<0.001
Diabetes				
Yes	752.59	170.07	4.43	<0.001
No	Base	*	*	*
Multiple birth				
Yes	760.98	107.10	7.11	<0.001
No	Base	*	*	*
Gestation at time of delivery**	−131.18	16.42	−7.99	<0.001
Mother's age***	12.93	7.01	1.84	0.065
Constant	2381.52	217.07	10.97	<0.001

*Base case scenario.
**Standardised to most frequent gestation week for model cohort.
***Standardised to most frequent age (in years) for model cohort.

Discussion

The project was carried out in four district hospitals with a total caseload of 17 700 maternities over 15 months in

2001–02. A total of 262 multiple pregnancies were identified, giving a rate of 14.8/1000 maternities, which was similar to the national rate of 15.0.[1] We do not know the proportion of women whose conception was assisted. Among all the surveyed maternities were 61 women with pregestational diabetes managed with insulin, giving a rate of 3.4/1000 maternities, but this number may have included some women with type II diabetes who changed their treatment to insulin before or early in their pregnancy.[19] The comparative rate of type I diabetes cases for southeast England in the 2002–03 audit of maternity units for the Confidential Enquiry into Maternal and Child Health was 3.0/1000 all maternities.[16] Around 2500 women with epilepsy in the UK give birth annually[20] or 1.8/1000 maternities; in our cohort, the rate was substantially higher at 3.3 (59 women). Our rates for diabetes and epilepsy may have been attributable to the efforts of the telemedicine project facilitators in monitoring maternity records in the district hospitals for eligible high-risk women. But at the same time, the low-risk women in our sample may not be truly representative of the population of pregnant women because they were identified from matching demographic attributes of the high-risk women.

While the study took place in four district maternity services, the derived costs of maternity care may not be nationally representative because of variations in the delivery of care, including different policies regarding admission to prenatal wards during the antenatal period.[21] Moreover, our resource use data applied to events occurring from the time when the women underwent their anomaly scan midway through the second trimester, since structures of the fetal heart cannot be imaged satisfactorily through telemedicine until 18 weeks of gestation. Thus, we omitted two or more ultrasound scans performed earlier in the women's pregnancies, a blood serum screening test for Down syndrome that was offered to almost everyone and performed usually at 15–16 weeks and chromosomal diagnostic procedures (usually amniocentesis) performed soon afterwards on selected women in the Down syndrome risk group.

A prenatal bed day was the costliest antenatal resource item, with a weighted bed day costing £280. Our analysis found that women with multiple pregnancies and singleton women with cardiac risk factors who were admitted to these beds stayed for a total of 4 or 5 days on average, significantly more than 2–3 days for Down syndrome risk and low-risk singleton women. Henderson and Petrou[21] likewise observed in North West Thames in 1996 and 1997 that the duration of antenatal bed days for mothers of twins was nearly twice the duration for singleton mothers (5.24 versus 2.68 in 1997), although in 1998 there was no difference between the two cohorts. The ultrasound unit costs of £23 for scans performed in district hospitals and £55 for specialist scans were comparatively similar to the costs (in 1998–1999 prices) of £14–£16 for routine antenatal scans and £52 for specialist detailed scans derived by Bricker *et al.* from a detailed costing exercise in a Liverpool teaching hospital.[22]

The most costly obstetric resource item was mode of delivery, with the weighted costs for a spontaneous vaginal delivery and a caesarean section without complications being £941 and £1,583, respectively. Two district hospitals provided separate unit costs for elective and nonelective caesarean sections. As our survey dataset did not distinguish between women who had an emergency or an elective caesarean section, a single cost for each hospital was derived by weighting the elective and nonelective values by 37 and 63%, respectively, according to published statistics.[23] Our dataset also did not distinguish between deliveries with and without complications. However, blood products (such as units of whole blood or standard red cells) transfused to women were identified and costed (Table 2), and for this reason, we used a weighted caesarean section without complications in the analyses.

Our costs for vaginal delivery and caesarean section (in 2001–02 prices) are comparatively similar to those identified for singleton women in the Grampian Region of Scotland by Petrou and Glazener[24]—£712 and £1,733 in 1999–2000 prices. Ledger *et al.* used NHS reference costs for 2002 to derive estimates for delivering women after IVF infertility treatment. Unlike our study and the Scottish study, these researchers differentiated between elective and emergency caesarean sections, with the reference cost for an emergency caesarean section exceeding an elective section by £640 (£2,581 versus £1,942).[5]

In southeast England, in 2001, the caesarean section rate overall was 22.6%,[2] whereas our rates for the four hospitals in Essex and Kent were 26–30% for the Down syndrome risk women, the low-risk women, and the cardiac risk women excluding those with diabetes. However, the caesarean section rate for multiple pregnancies of 63% was similar to the rate for southeast England of 61.7%,[2] while the rate of 61% for the women with pregestational diabetes was also similar to the national rate of 67%.[16] The sensitivity analysis comparing high and low unit costs for caesarean section without complications aptly demonstrated the effect of this mode of delivery upon mean obstetric costs for cohorts of women.

The type of delivery naturally impacted on the length of time women remained on the postnatal wards before discharge. For mothers of twins or triplets (in five cases) in our study, the mean duration was 5.6 (SD 4.0) days compared with 3.5 (SD 2.7) for cardiac risk women. We did not, however, assess the cost of high dependency care for women with complications. Thus, the mean cost of obstetric care for the multiple pregnancy mothers was significantly higher, even though our reported costs from four district hospitals may be lower than national averages in 2001/2002 prices.

The identification of pregestational diabetes in the multiple regression model as a significant predictor of costlier

antenatal and obstetric care is entirely consistent with epidemiological evidence from national audits of diabetic pregnancies. The Confidential Enquiry into Maternal and Child Health in England, Wales, and Northern Ireland found a three-fold increase in perinatal mortality and a two-fold increase in the fetal congenital anomaly rate, particularly of the heart, in women with diabetes compared with the general maternity population.[19] Blood glucose levels require careful monitoring, so women attend diabetic clinics as well as antenatal clinics. The events around childbirth are also more acute: in the audit of diabetic pregnancies for the Confidential Enquiry, delivery before 37 weeks of gestation was five times that in the general population.[16] In our study, only two of the 57 singleton pregnancies in women with pregestational diabetes extended beyond 38 weeks. Consequently, the total mean (SD) cost of £4,877 (£2,192) for the women with diabetes exceeded the total for multiple pregnancies by £435.

Conclusion

Our analyses of a dataset covering cohorts of women receiving maternity care in four district hospitals have confirmed that the cost of care from midway through the second trimester until the mother's discharge from hospital is substantially higher for multiple pregnancies than that for most singleton pregnancies. We have also shown that diabetic pregnancies are particularly costly. This finding is pertinent because the Confidential Enquiry found in its case–control study of pregnancies in women with type I or type II diabetes that maternity care could have been improved for almost half the women, and suboptimal clinical care was associated with poor pregnancy outcome.[25] If additional resources are allocated in NHS maternity services for managing women with diabetes, then mean costs for diabetic pregnancies will rise accordingly, and at a time when the prevalence of type II diabetes in women in older childbearing ages is increasing.[26] Our results should help inform healthcare decision makers in the UK about key drivers of maternity costs for high-risk singleton and multiple pregnancies.

Acknowledgements

We thank the staff for their kind support and help in the obstetric departments of Basildon Hospital, Colchester General Hospital, Southend General Hospital, and Medway Maritime Hospital in Gillingham. The project facilitators in these hospital departments were: L Evans, V Asker, S Anstis, and D Brown. The TelePaed project team members are: R Dowie, project director, and H Mistry (at Brunel University), T Young (University of Sheffield), G Weatherburn (Buckinghamshire Chilterns University College), R Franklin, H Gardiner, M Rigby, and G Rowlinson (at the Royal Brompton Hospital, London).

Funding for this study was from the Department of Health and the Charitable Funds Committee of the Royal Brompton and Harefield NHS Trust. The Health Economics Research Group at Brunel University receives funding from the Department of Health Policy Research Programme. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health. ■

References

1 Office of National Statistics. Birth Statistics. Review of the Registrar General on births and patterns of family building in England and Wales, 2004. London: ONS, 2005. Series FM1 no.33 [http://www.statistics.gov.uk/downloads/theme_population/FM1_33/FM1_33.pdf]. Accessed 27 February 2007.

2 Thomas J, Paranjothy S, Royal College of Obstetricians and Gynaecologists Clinical Effectiveness Support Unit. National Sentinel Caesarean Section Audit Report. London: RCOG Press, 2001.

3 Henderson J, Davidson LL, Chapple J, Garcia J, Petrou S. Pregnancy and Childbirth. In: Stevens A, Raftery J, Mant J, Simpson S, editors. *Health Care Needs Assessment, Third Series*. Oxford, UK: Radcliffe Publishing; 2007. p. 735–89.

4 Office of National Statistics. Key Population and Vital Statistics. Local and Health Authority Areas. *Population and Vital Statistics by Area of Usual Residence in the United Kingdom, 2004*. London: ONS, 2006. VS no. 31, PP1 no. 27 [http://www.statistics.gov.uk/downloads/theme_population/KPVS31_2004/KPVS2004.pdf]. Accessed 27 February 2007.

5 Ledger WL, Anumba D, Marlow N, Thomas CM, Wilson ECF, the Cost of Multiple Births Study Group. The costs to the NHS of multiple births after IVF treatment in the UK. *BJOG* 2006;113:21–5.

6 Lukassen HGM, Schönbeck Y, Adang EMM, Braat DDM, Zielhuis GA, Kremer JAM. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. *Fertil Steril* 2004;81:1240–6.

7 Dowie R, Mistry H, Young TA, Weatherburn G, Gardiner HM, Rigby M, *et al*. Telemedicine in pediatric and perinatal cardiology: economic evaluation of a service in English hospitals. *Int J Technol Assess Health Care* 2007;23:116–25.

8 Gardiner HM. Access to perinatal cardiology in the United Kingdom. *Postgrad Med J* 2001;77:1–3.

9 Longworth L, Young T, Buxton MJ, Ratcliffe J, Neuberger J, Burroughs A, *et al*. Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. *Liver Transpl* 2003;9:1295–307.

10 Department of Health. *Advanced Letter (MD) 1/02: Pay and Conditions of Service of Hospital Medical and Dental Staff and Doctors in Public Health Medicine and the Community Health Service*. London: DoH, 2002a. [http://www.info.doh.gov.uk/doh/coin4.nsf/Circulars?FReadForm]. Accessed 27 February 2007.

11 Department of Health. *Advanced Letter (NM) 1/2002: Nurses' and Midwives' Pay for 2002/03*. London: DoH, 2002b. [http://www.info.doh.gov.uk/doh/coin4.nsf/Circulars?ReadForm]. Accessed 27 February 2007.

12 Netten A, Curtis L. *Unit Costs of Health and Social Care 2002*. Canterbury, UK: Personal Social Services Research Unit, University of Kent at Canterbury, 2002.

13 Manly BFJ. *Randomisation, Bootstrap and Monte Carlo Methods in Biology* (Texts in Statistical Science), 2nd edn. London: Chapman & Hall, 1997.

14 StataCorp. *Stata Statistical Software: Release 9.2*. College Station, TX: Stata Corporation, 2003.

15 S-PLUS. *Guide to Statistics, Volume 1 & 2*. Seattle, WA: Mathsoft, 2000.

16 Confidential Enquiry into Maternal and Child Health. *Pregnancy in Women with Type 1 and Type 2 Diabetes 2002-2003, England, Wales and Northern Ireland*. London: CEMACH, 2005. [http://www.cemach.org.uk/publications/CEMACHDiabetesOctober2005.pdf]. Accessed 27 February 2007.
17 Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, *et al*. Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. *J Neurol Neurosurg Psychiatry* 2006;77:193–8.
18 Uebing A, Steer PJ, Yentis SM, Gatzoulis MA. Pregnancy and congenital heart disease. *BMJ* 2006;332:401–6.
19 Macintosh MC, Fleming KM, Bailey JA, Doyle P, Modder J, Acolet D, *et al*. Perinatal mortality and congenital anomalies in babies of women with type 1 or type 2 diabetes in England, Wales, and Northern Ireland: population based study. *BMJ* 2006;333:177–80.
20 UK Epilepsy and Pregnancy Register. *UK Epilepsy and Pregnancy Register*. 2006. [http://www.epilepsyandpregnancy.co.uk/]. Accessed 27 February 2007.
21 Henderson J, Petrou S. Economic implications of multiple births. In: Keith LG, Blickstein I, Oleszczuk JJ, Keith DM, editors. *Triplet Pregnancies and their Consequences*. London: The Parthenon Publishing Group; 2002. pp. 429–35.
22 Henderson J, Bricker L, Roberts T, Mugford M, Garcia J, Neilson J. British National Health Service's and women's costs of antenatal ultrasound screening and follow-up tests. *Ultrasound Obstet Gynecol* 2002;20:154–62.
23 RCOG Clinical Effectiveness Support Unit. *The National Sentinel Caesarean Section Audit Report*. London: RCOG Press, 2001. [http://www.rcog.org.uk/resources/public/pdf/nscs_audit.pdf]. Accessed 27 February 2007.
24 Petrou S, Glazener C. The economic costs of alternative modes of delivery during the first two months postpartum: results from a Scottish observational study. *BJOG* 2002;109:214–17.
25 Confidential Enquiry into Maternal and Child Health. *Diabetes in Pregnancy: Are We Providing the Best Care? Findings of a National Enquiry: England, Wales and Northern Ireland*. London: CEMACH, 2007. [http://www.cemach.org.uk/Final%20CEMACH%20report%20070207.pdf]. Accessed 27 February 2007.
26 Lusignan S, Sismanidis C, Carey IM, DeWilde S, Richards N, Cook DG. Trends in the prevalence and management of diagnosed type 2 diabetes 1994-2001 in England and Wales. *BMC Fam Pract* 2005; 22:13. [http://www.biomedcentral.com/content/pdf/1471-2296-6-13.pdf]. Accessed 27 February 2007.