The London School of Economics and Political Science

An Economic Evaluation of Telehealth and Telecare in England

Catherine Henderson

A thesis submitted to the Department of Social Policy of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, April 2018

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 85,792 words.

This thesis was copy edited for conventions of language, spelling and grammar by Lisa Findley.

Statement of conjoint work

Work presented in this thesis draws on three publications of which I was first author. Methods to calculate costs of telehealth and telecare interventions are presented in Chapter 4 and results reported in Chapter 5. These drew on work reported in the following:

Henderson, Catherine, Jennifer Beecham, and Martin Knapp (2013) "The costs of telecare and telehealth" in Lesley Curtis, eds. *Unit Costs of Health and Social Care 2013*. Canterbury: Personal Social Services Research Unit.

Henderson, C., M. Knapp, J. L. Fernandez, J. Beecham, S. P. Hirani, M. Cartwright, L. Rixon, M. Beynon, A. Rogers, P. Bower, H. Doll, R. Fitzpatrick, A. Steventon, M. Bardsley, J. Hendy, and S. P. Newman (2013) "Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial" *BMJ*, 346:f1035. doi: https://doi.org/10.1136/bmj.f1035

Henderson, C., M. Knapp, J. L. Fernandez, J. Beecham, S. P. Hirani, M. Beynon, M. Cartwright, L. Rixon, H. Doll, P. Bower, A. Steventon, A. Rogers, R. Fitzpatrick, J. Barlow, M. Bardsley, and S. P. Newman (2014) "Cost-effectiveness of telecare for people with social care needs: the Whole Systems Demonstrator cluster randomised trial" *Age and Ageing*, 43 (6):794-800. doi: https://doi.org/10.1093/ageing/afu067.

I have indicated the Unit Costs publication as the source of a boxed description of the methods of costing the intervention (in Box 4.1 Costing the telehealth and telecare interventions), although the text presented there is not a direct quotation.

Methods and results of the multilevel analyses (presented in Chapter 4 and Chapter 6) were not presented in the above publications.

The methods and results of the cost-effectiveness analyses (presented in Chapter 4 and Chapters 7 and 8) draw on work initially presented in the BMJ and Age Ageing publications listed above. However the multiple imputations and cost-effectiveness analyses presented in the thesis are new.

Abstract

In the English health and social care system, budgets are now constrained more than ever, while an increasing proportion of the population is expected to require care. There is an urgent need to find new ways to enable people with long-term illness and disability to live well, within the national budget. Policymakers have embraced new assistive technologies such as telecare and telehealth as a means to achieve this goal. Evidence that telehealth is cost-effective is emerging but remains limited; evidence on the impacts of telecare is still more limited. In this thesis I investigate the effectiveness of two advanced assistive technologies, telehealth and telecare, in improving or maintaining health-related quality of life and other psychological outcomes, given the costs of providing these interventions.

I deploy cost-effectiveness methods to analyse questionnaire data from two large-scale randomised controlled trials of telecare and telehealth in England. Drawing on provider data collected during the evaluation, I describe the inputs to production of the telehealth and telecare interventions and calculate their unit costs. I describe the health and social care costs of telehealth and telecare participants and explore participant characteristics associated with cost variations.

The results of cost-effectiveness analyses of telehealth and telecare indicate that these technological interventions did not produce the hoped-for improvements in self-reported quality of life and other psychosocial outcomes, nor reduce the overall estimated annual costs of health and social care. Policymakers and practitioners would benefit from better evidence on the mechanisms by which telecare and telehealth 'work', and for whom, to direct future investments of resources into these technologies.

Acknowledgements

First and foremost I thank my academic supervisors Professor Martin Knapp and Dr José-Luis Fernández for their guidance and encouragement throughout my registration.

I thank all the people who participated in the Whole Systems Demonstrator (WSD) Trials and Questionnaire Studies, the health and social care professionals and case study organisations assisting the evaluation, the WSD evaluation team and the Policy Research Programme of the Department of Health for England, which funded the evaluation. The views expressed here are not necessarily those of the Department.

I am grateful to my colleagues at PSSRU - Nicola Brimblecombe, Francesco D'Amico, Jackie Damant, Emily Freeman, Derek King, Juliette Malley, Madeleine Stevens and Lisa Trigg - for helpful discussions and moral support. I thank Sarah Helias for her help throughout the PhD process. My special thanks go to Tiziana Leone and Philipa Mladovsky for all their support and encouragement.

Lastly, I would not have finished this thesis without the love and support of my family. I thank my husband Mick Robson for the emotional support and wonderful cooking that kept me going through these years. I am indebted to my aunt and uncle, Bev and Ross Dowie, for all their help through some difficult times. I dedicate this work to my mother Diane Henderson.

Table of Contents

Table of Figures

Table of Tables

Table of Boxes

List of Abbreviations

A&E: Accident and emergency department

ADL: Activities of daily living

AAT: Advanced assistive technology

AT: Assistive technology

CI: Confidence interval

COPD: Chronic obstructive pulmonary disease

DD: Difference-in-difference

DDD: Difference-in-difference-in-difference

ED: Emergency department

EQ-5D: EuroQol 5 dimensions

HF: Heart failure

HRQoL: Health-related quality of life

ICECAP-O: ICEpop CAPability measure for Older people

ICER: Incremental cost-effectiveness ratio

ICT: Information and communications technology

IT: Information technology

LTC: Long-term conditions

NICE: National Institute for Health and Care Excellence

PA: Population-averaged

PCT: (NHS) Primary care trust

PERS: Personal emergency response systems

QALY: Quality-adjusted life year

SD: Standard deviation

SE: Standard error

SF-12: 12-Item Short-Form Health Survey

SS: Subject-specific

SUR: Seemingly unrelated regression

TC: Telecare

TH: Telehealth

TM: Telemonitoring

TS: Telephone support

WSD: Whole Systems Demonstrator

Chapter 1

Introduction

In this thesis I examine the costs and outcomes of two interventions. One intervention, "telehealth" falls largely within the ambit of the National Health Service; the other, "telecare" has been traditionally a service funded and/or provided by local authorities. Both interventions can be described as "advanced assistive technologies". Along with related innovations such as mobile health and telemedicine, telehealth and telecare have attracted substantial interest from governmental and non-governmental actors within the health and social care systems of the UK. Telehealth and telecare technologies promise new ways for health and social services to assist people with long-term and chronic conditions to take charge of their care and their lives. These technologies also promise to reduce unwanted contact with health services, such as hospital stays, and to reduce reliance on social services such as domiciliary and residential care. They hold the potential to address the twin governmental preoccupations of containing costs and meeting public expectations about health and care services. The relationships between the service use, costs and benefits associated with telehealth and telecare therefore merit close examination.

1.1 Background

Expenditure on health and social care has risen steadily over recent years.[1] Furthermore it is likely, if not inevitable, that expenditure on health and social care will continue to rise at a faster rate than today if it is to keep pace with growing need. In this chapter I begin by exploring the potential extent of future expenditure and drivers of this expenditure. I then move to a discussion of the important role that advanced assistive technologies may play in managing future pressures on health and social care.

The age profile of the UK population is changing. The population 85 years and over is set to double from 1.6 million in 2016 to 3.2 million in 2041 (Office for National Statistics

[1] Social care spending has experienced periods of decline but grown overall since 2010 in cash terms (Cromarty 2017).

2016, Barlow et al. 2012). The potential for compression of morbidity in western countries is debatable. The UK cannot count on an increase in disability-free life expectancy to counteract the impact of projected increases in life expectancy on demand for health and social care services (Howse 2006). Despite gains in life expectancy, years lived free of ill health and disability have not kept pace. Between 1990 and 2016, life expectancy in the UK increased from 78.47 to 82.86 years for women and 72.85 to 78.92 years for men (GBD 2016 DALYs and Hale Collaborators 2017). Healthy life expectancy in 2016, on the other hand, was estimated to be 70.97 years for women and 69.11 years for men. Jagger et al. (2016) suggest that there has been an expansion of morbidity in England (comparing 1991 and 2011 survey data) in terms of mild physical disability, a smaller proportion of life expectancy at age 65 being spent without a disability than previously. Guzman-Castillo et al. (2017) forecast that between 2015 and 2025, life expectancy will increase (by 1.7 years) but also a 65-year-old individual will live a quarter of his or her remaining life with a disability. Also, while the prevalence of disability in the older population is projected to decline, the numbers of older people are growing. The numbers of older people with a disability will increase by an extra 560,000 people in 2025 compared to a decade earlier, to a total of 2.81 million (Guzman-Castillo et al. 2017). These figures suggest that for many, it will be important to find ways to cope with ill health or disability in later life.

1.2 The Health and Long-Term Care System in the UK, Long-Term Conditions and Rising Future Expenditure

The National Health Service in the UK is funded by general taxation. The English NHS is overseen at a national level by the Secretary of State for Health; oversight in the other countries of the UK is devolved to ministers in the regional parliaments. Most NHS services in the UK are free at the point of access (Cylus et al. 2015). Long-term care consists of services such as home-based personal care, day- and other community-based care facilities and residential and nursing home care. Local authorities (‘councils’) in the UK are funded to provide care services from a mixture of general and local taxation. In England, councils are responsible under the Care Act 2014 (HM Government 2014) for ensuring that residents have access to information about care and to a range of care and support services; they are also responsible for enabling residents to receive care services to mitigate or delay the impacts of their needs (Department of Health 2016). This type of care and support is commonly known as ‘social care’. Local authority-funded care and support services in England are mostly

means-tested (but assistive devices and small-scale home adaptations are provided free of charge).

The terms ‘social care’ and ‘long-term care’ cover very similar ground; however, long-term care can encompass services such as long-stay hospital care, rehabilitation and intermediate care that are funded at least in part by the NHS (and sometimes jointly with local authorities). In this thesis I will use the term ‘social care’ as synonymous with paid care funded by local authorities directly or indirectly (for instance through grant-funding voluntary organisations to provide services). I will for the most part concentrate on the English health and care system.

Long-term or chronic conditions can be life-changing. These are incurable conditions that must be controlled over long periods of time (Department of Health 2012b). The management of those with long-term or chronic conditions is increasingly under the spotlight, given that, as the population ages, such conditions may be expected to rise. Department of Health estimates suggest that treatment costs of those with long-term conditions (LTC) make up 69 per cent of all health and social care spending in England, and that the numbers of those with at least one LTC will rise from three million to eighteen million by 2025 (Department of Health 2008b). Consequently the volume of treatment and care for those with chronic conditions is likely to rise in the future. Thus, purely from a cost-containment perspective, there is pressure on governments to seek ways to prevent or reduce the impact of chronic conditions on health and care services. In terms of the size of expenditure to be contained, a much higher proportion of GDP goes towards the NHS than to long-term social care services (7.3 per cent vs. 1 per cent in 2016/17 (Office for Budget Responsibility 2017)).

1.2.1 Health Care Expenditure

Projections of health expenditure typically incorporate assumptions about demographic, organisational and economic factors (health planning, financing, productivity and prices), consumer behaviour and income and also medical technological innovation (Astolfi, Lorenzoni, and Oderkirk 2012). While some technological innovations will decrease the costs of treatment, in many cases the technologies are high-cost. The diffusion of new medical technologies has been seen as a driver of dramatic rises in health expenditure within OECD countries over the past forty years (Cutler 2002, Appleby 2013). Demographic change accounts for a relatively small amount of rising health spending compared to technological progress and increases in national income levels (OECD 2010).

Health expenditure as a proportion of GDP in England has risen steadily over recent years. Spending rose by 36.6 per cent between 1997 and 2002 (The King's Fund 2005). There were annual real spending increases in the order of 5.1 per cent over the period from 1994 to 2011 (Appleby, Crawford, and Emmerson 2009). Estimates of future health spending vary: the Office for Budget Responsibility (OBR) projects UK-wide health expenditure as a proportion of GDP rising steeply from 6.9 per cent in 2021/2 to 12.6 per cent in 2066/67 (Office for Budget Responsibility 2017). The OECD has estimated a similar rise from 6.5 per cent in 2010 to 12.4 per cent in 2060; however with stronger government policies to contain costs, there could be a more modest rise of 8.5 per cent (OECD 2013). In the economic climate of recent years, it appears that NHS funding might be squeezed a little more than might have been expected in the previous decade. Projections carried out by the King's Fund and the Institute for Fiscal Studies (Appleby, Crawford, and Emmerson 2009) suggest that spending on the English NHS would need to increase by about 1.1 per cent a year to maintain service quality in light of demographic changes, yet the NHS faced the possibility of a gap of about £4 billion opening up between expenditure required to maintain quality in the NHS and the available budget[2] by 2017. Recent OBR figures suggest that UK-wide health spending as a proportion of GDP is expected to fall in the short-term, from 7.2 per cent of GDP in 2015/16 to 6.8 per cent in 2019/20: this equates to a rise of 0.5 per cent annually in real terms over this period (Office for Budget Responsibility 2016).

1.2.2 Social Care Expenditure

The system of care provision for disabled older people is, and has been under, a great deal of strain for a number of years. Recent reviews by the regulator of social care[3] have identified a number of problems in current patterns of provision, one being a lack of choice and flexibility in the way the state meets people's needs (Commission for Social Care Inspection 2009), another being inequities in access to publicly-funded care (Commission for Social Care Inspection 2008). Concerns have been raised over the past decade that the threat of unmet need in the population of older disabled people is increasing, due to the funding constraints and the progressive tightening of eligibility criteria used by local authorities when they assess

[2] The authors calculated the gap between projected expenditure under the their 'tepid' spending scenario and the 'fully engaged' Wanless projections (Wanless 2002), a scenario involving dramatic improvement in health outcomes, high levels of public engagement in health, high confidence in the NHS, high life expectancy, a high level of responsiveness in the NHS and more efficient use of resources.

[3] The current Care Quality Commission, which superseded the Commission for Social Care Inspection.

older people for care (Commission for Social Care Inspection 2008). Forder (2007) estimated that some 15 per cent of those who appear to have low level needs for help do not receive it, through formal or informal channels. An Age UK report found that the numbers in need but not receiving social care support in England had increased by 383,900 between 2010 and 2016, rising to as many as 1.2 million people (Age UK 2016, Cromarty 2017). Recent research has found that more than half of people aged 65 and over in England had unmet needs (drawing on the Care Act 2014 definition of qualifying need)[4].

In addition to the current pressures on the system, future demand for care is likely to increase, whereas it is uncertain that there will be a corresponding rise in the supply of informal care (Pickard et al. 2007). The majority of disabled older people receive help and support from relatives, such as spouses and children (Pickard et al. 2007). Projections suggest that as the population of those 65 years and over increases, the number of disabled older people likely to be in receipt of at least 20 hours of care a week from their (adult) children will almost double, from 665,000 in 2005 to 1,270,000 in 2041 (Pickard 2008). If the ratio of numbers of adult caregivers to parents receiving care remains constant, the gap between provision and receipt of intense care is likely to rise very substantially, supply exceeding demand in the near future. Any change in the level of supply of informal care is likely to have an impact on demand for formal care. In the UK, availability of informal care traditionally has been considered along with other criteria in local authority decision-making on providing funding for care (Comas-Herrera, Wittenberg, and Pickard 2010). The oldest-old (85 years and over) being the fastest growing segment of the older population (Office for National Statistics 2014, 2009) and the segment mostly likely to receive formal care (Pickard et al. 2007, Comas-Herrera, Wittenberg, and Pickard 2010) only adds to the demand pressures upon both formal care providers and adult children providing care over the coming years. Changes in living circumstances of older people, from communal to solo living may decrease the availability of assistance when there is no co-resident carer on hand (Falkingham et al. 2010).

Personal Social Services Research Unit (PSSRU) projections estimate that expenditure on social care and disability benefits for older people could increase from 1.2 per cent of GDP in 2005 to 2 per cent in 2041, assuming that current funding systems and

[4] Depending on the survey examined – English Longitudinal Study of Ageing or Health Survey for England - 58 per cent or 73 per cent had unmet needs respectively (Dunatchik et al. 2016, Ipsos MORI 2017).

patterns of care do not change (Wittenberg et al. 2008). The OBR on the other hand projects slower rises, such that 2 per cent of GDP will be spent on long-term care in 2066-67 (Office for Budget Responsibility 2015).

Savings targets imposed by central government in recent years have intensified budget pressures on English local authorities. The 2010 Spending Review (decreasing the settlement to local authorities by 27 per cent) was estimated to have resulted in a social care funding 'gap' in 2014/15[5] between available funding and anticipated expenditure of around £267 million (Appleby and Humphries 2010). In the first half of this decade, the National Audit Office found that local authorities had made deep cuts to their budgets, reducing their spending on adult social care by 8 per cent, or £1.4 billion from 2010 to 2013. English councils' expenditure on social care for people aged 65 and over fell by 9 per cent over this period, while numbers in receipt of council-funded care declined from 1.1 million to 0.85 million (The King's Fund 2016). The Local Government Association recently estimated that the funding 'gap' would reach £1.3 billion by 2020 (Local Government Association 2016). Several central government policies have added to cost pressures on local authorities' social care budgets. The introduction of the National Living Wage for workers aged 25 and over in 2016 increased the labour costs of social care, as many workers are on low wages. Changes in the scope of the Deprivation of Liberty Safeguards following a Supreme Court ruling and the winding up of the Independent Living Fund increased the number of assessments to be carried out by the local authority workforce (Local Government Association 2015, Cromarty 2017).

1.2.3 Policy Responses

The English Department of Health has long recognised that its share of the national budget cannot grow at the same pace as health and care costs will grow as a share of GDP. As long term conditions' "cost and prevalence continues to grow, doing more of the same is not an option if NHS and social services are to be sustainable in the future" (Department of Health 2012b, p.4). As well as the need for cost containment, there are compelling quality of life arguments for tackling long-term conditions (Department of Health 2005d). Evidence suggests that people do not consistently receive the information and advice they need to be able to manage their conditions (Blendon et al. 2003, Department of Health 2005b). The

[5] The projections assume that there will be a 4% real annual growth in the budget for social care needed due to demographic pressures and unit cost rises, and that existing eligibility criteria will continue to apply.

government has taken note of the personal and human cost of long-term health conditions over recent years and responded with initiatives such as community matrons, case-finding, expert patient groups and increased funding for the assessment and treatment of such conditions (Department of Health 2005d). At the same time, those in need of assistance with activities of daily living, for instance people with disabilities and long-term conditions, have not always been well served by the lack of choice and flexibility in the help provided by local authorities (Patmore and McNulty 2005, Williams et al. 2008, Commission for Social Care Inspection 2009). The last Labour government acknowledged calls for a system of assistance tailored to the individual's needs and wishes. It advocated a personalisation agenda to allow those with chronic long-term conditions and disabilities to direct their own care, through initiatives such as the individual budgets programme (Department of Health 2008a) and personal health budgets (Department of Health 2010a). Social care strategy under the Coalition government continued to promote this agenda. "A Vision for Adult Social Care: Capable Communities and Active Citizens" (Department of Health 2010b) emphasised prevention, allowing people to maintain their independence through initiatives such as re-ablement, the involvement of 'the Big Society', promoting plurality of provision and personalisation, along with partnership, high quality care and support services and the use of advanced assistive technologies. More recently, integrated personal commissioning has been introduced for people who have chronic needs for support (including adults with multiple long term physical or mental health conditions and people with frailty). This model was intended to provide people with access to integrated health and social care funding for personalised support packages (Bate 2017, NHS England 2017).

1.3 Advanced Assistive Technologies: The Route to a Sustainable Health and Care System?

The previous discussion has established that the landscape is one of rapid technological and demographic change creating pressures for higher expenditure over the long-term. Yet budget constraints threaten to undermine the quality of health and social care in the short-term as well as in the longer term. There are good reasons for central and local government, health and social care providers, interest groups and those in industry to seek new ways of managing long-term conditions and disability in order to mitigate these important drivers of future expenditure. In the first decade of this century, there was a growing interest in the potential of advanced assistive technologies (AAT) such as telemedicine, telehealth and telecare.

Proponents of AAT have given cost savings as one rationale for advocating their introduction and widespread implementation (Department of Health 2005a) and maintaining or improving quality as another (Department of Health 2005a, 2010b).

Hopes to improve the evidence base for making the best use of public funding for assistive technologies and for shaping the growing market for AAT underlay the decision to fund the £31 million Whole System Demonstrator (WSD) programme, the largest-scale trial of telehealth and telecare to be carried out in the UK (Department of Health 2007, 2010b). At the outset of the study, research evidence on telehealth was growing but had many gaps, particularly in terms of good-quality cost-effectiveness studies. The research evidence for telecare had a much weaker base than in the case of telehealth. This formed the context in which I began working on my thesis, with the aim of contributing to the economic evidence for the two technologies.

1.3.4 The Role of Economic Evaluation in Health and Social Care Decision-making

Policymakers in the UK, as elsewhere, must balance fiscal constraints and rising demand for health and social care services. Because public resources are scarce, decision makers face choices as to which services should be funded, at the expense of funding some other service. In other words, each service has an opportunity cost. Because there are often several credible alternatives that could be funded, some framework or organising principle is required for decision makers to allocate resources in an efficient and equitable way (Brazier 2007, Drummond et al. 2015). Economic evaluation provides such a framework, being a "comparative analysis of alternative courses of action in terms of both their costs and consequences" (Drummond et al. 2015, p.4). There are several approaches to economic evaluation. In cost-effectiveness analysis (CEA), alternative treatments with the same objective are compared in terms of outcomes measured in 'natural' units (for instance, changes in blood pressure readings) and the results presented as cost per unit. The objective of the treatment is unquestioningly presented as worthwhile. The approach has limitations: interventions differing in more than a single outcome cannot be compared simultaneously: so for instance. as a monetary value cannot easily be attributed to each outcome, the net benefit of the different interventions cannot be established (Brazier 2007). In cost utility analysis (CUA), a comparison of costs can be made across a broader range of alternative health programmes with the benefits measured using an indicator of "utility", usually in terms of a measure of quality adjusted life years (QALY) (Brazier 2007, Drummond et al. 2015). The

QALY combines a measure of length of life with a measure of health related quality of life (HRQoL) on one scale. This has the advantage of allowing the comparison of interventions with more than one outcome, allowing comparison of interventions with different outcomes and allowing comparison of interventions for different conditions with different outcomes (Brazier 2007).[6]

In England, the National Institute for Health and Care Excellence (NICE)[7] was founded to appraise new and existing health technologies in terms of clinical and cost-effectiveness, develop clinical guidelines and promote clinical audit (Rawlins 1999). NICE has adopted the 'reference case' as part of a process of the appraisal process. "A reference case goes beyond recommendations of good practice for economic evaluation, and attempts to standardize the scientific value judgements required in the conduct of economic evaluation, thereby improving quality and comparability of results" (Gray and Wilkinson 2016, p.112). Cost per QALY (an incremental cost-effectiveness ratio, or ICER) features prominently in the NICE decision framework (Gray and Wilkinson 2016); the ICER threshold, above which a technology could be considered not to represent value for money, has been set at between £20,000 to £30,000 per QALY since 2001 (McCabe, Claxton, and Culyer 2008, National Institute for Clinical Excellence 2004). Thus in England the consideration of cost-effectiveness is an important aspect of the technology appraisal process.

1.3.5 The production of welfare

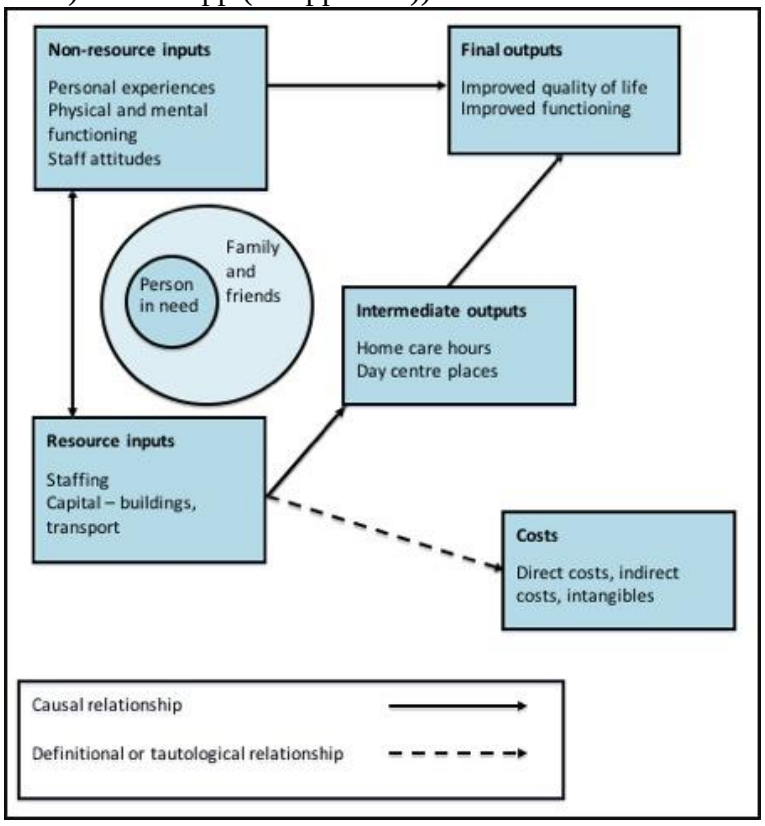
The production of welfare (PoW) approach offers a framework useful for evaluating the economic consequences of interventions taking place across health and social care (Davies and Knapp 1981, Knapp 1984). This approach has been refined in studies of a number of social care services and service innovations. Essentially, welfare is seen as a function of the relationship between outputs, in terms of quality of life and three kinds of inputs. Outputs "include all those consequences that so directly reflect aspects of welfare that they are valued in their own right" (Davies and Knapp 1981, p. 5). These can be further categorised as final

[6] There are other approaches to economic evaluation. In cost-benefit analysis, costs and consequences are valued in money terms so that if the benefits exceed the costs then the intervention is worth implementing (Brazier 2007); cost consequence analyses present the results of a number of relevant analyses rather than just the one summary CEA; and cost-minimisation seeks to establish the "least-cost alternative", if the outcomes of the treatment are the same (Brazier 2007, Drummond et al. 2015).

[7] NICE has had one acronym but several titles since its inception in 1999: National Institute for Clinical Excellence (1999 to 2005); National Institute for Health and Clinical Excellence (2005 to 2013); National Institute for Health and Care Excellence (from 2013).

and intermediate outputs, final outputs being for instance, improved mobility and functioning or more social engagement, intermediate outputs being the services given such as the number of home care hours or day centre places (Knapp 1984). Inputs encompass both 'resource inputs'– labour and capital – but also 'non-resource' inputs – intangible factors that are within the control of the producer, such as 'atmosphere' or 'friendly staff members' in a care home, and 'quasi-inputs', that are outside of the service producer's control, for instance, the personal characteristics of older people using home care services or entering a care home. Resource inputs influence the extent of achievement of the aims of care; but such influences are mediated by non-resource inputs (Knapp 1984).

Figure 1.1 The production of welfare (adapted from Davies & Knapp (Davies and Knapp 1981) and Knapp (Knapp 1984))



The model is illustrated in Figure 1.1. Thus "outputs are determined by the levels and modes of combination of the resource and non-resource inputs (which are mainly under the control of the administrator or policymaker[...], given the exogenously determined values of the quasi-inputs" (Davies and Knapp 1981, p. 8). A further component of the model is cost, this being a way of reflecting the resource inputs within the production relation (Knapp 1984).

I have adopted standard methods of cost-effectiveness and been guided by the PoW approach throughout the thesis.

1.4 Research Questions

The main question to be addressed within the thesis is:
What are the costs and benefits of introducing telehealth and telecare in England?
This leads to the following sub-questions:

1. What are the patterns of service use for people with and without telecare or telehealth support?
2. What are the total and component (service-specific) costs per person of the support/treatment received, and
3. What patient/user characteristics are associated with cost variations?
4. Are telecare and telehealth cost-effective compared to standard support/treatment?

Put another way, are these two instances of advanced assistive technologies efficient ways to improve or maintain the outcomes of health and social care, given the costs of that provision? I address these questions in the chapters that follow.

1.5 Overview of the Thesis

In chapter 2, I describe the context within which my research takes place, setting out some definitions of the technologies and introducing some important sub-classifications of the terms ‘telecare’ and ‘telehealth’. I discuss potential areas of overlap between these terms and closely associated technologies. I briefly review the conceptual literature to ask what the mechanisms are whereby the technologies ‘work’, focusing on service use and costs as outcomes.

In chapter 3, I begin the thesis by reviewing the evidence on telehealth and telecare, asking: what do we know about the effectiveness and cost consequences of implementing these technologies? I summarise the evidence base available on the effectiveness of telehealth and telecare, reviewing what is known about their impacts on preventing disease and disability and promoting independence. I concentrate on the literature in the years running up to the publication of results of the Whole Systems Demonstrator (WSD) Trials and Questionnaire Studies. I conclude the chapter with an in-depth discussion of the evidence base for costs and cost-effectiveness of these technologies, identifying gaps in the literature. I examine whether the technologies are reducing or increasing costs to the health and social

care system, and also whether on balance the evidence base on costs and benefits has been interpreted as suggesting cost-effectiveness.

In chapter 4, I describe the methodologies employed in the empirical chapters. I give an overview of the methodologies employed in the WSD Trials and Questionnaire Studies, the source of the data for the analyses in this thesis. I explain how I estimated costs of health and social care; and describe in detail how I estimated the costs of the telecare and telehealth interventions. The chapter covers the methods employed to investigate subgroup variations in costs in the study samples, including imputation of missing data. I finish by describing the methods used to carry out the economic evaluations.

In chapter 5, the first empirical chapter, I set the scene by describing the participants of the WSD Telehealth/Telecare Questionnaire Studies in terms of their socio-demographic characteristics, health and social care service use, and costs, split by experimental group. The unit costs of the interventions are also described. This provides necessary background information on the sample and addresses sub-question (1) and addresses sub-question (2).

In chapter 6, moving onto the more analytical investigations of the data, I address sub-question (3) in each of the telecare and telehealth questionnaire samples. These are subgroup analyses, examining cost variations of people in terms of socio-demographic and needs-related characteristics in the telecare and telehealth questionnaire study samples. The analysis of the telehealth sample data focuses on the role of three long-term conditions, diabetes, chronic obstructive pulmonary disease (COPD) and heart failure; the analysis of the telecare sample data concentrates on the impact of living arrangements (living with others and living alone).

In chapters 7 and 8, I present the results of the cost-effectiveness analyses of the telehealth and telecare interventions. In chapter 7, I address research question (3) by looking at the Telehealth Questionnaire study sample, discussing salient points from chapter 6, and presenting the results of the statistical models of the QALYs, and other quality of life and psychological outcomes, associated incremental cost-effectiveness ratios (ICERs), net-benefit lines and cost-effectiveness acceptability curves. I discuss the results in terms of the implications for policy and practice, and discuss limitations and future directions for research. In chapter 8, I address research question (3) by examining cost-effectiveness of telecare. I present the main findings. I also discuss a subgroup analysis of cost-effectiveness in terms of people living with others and living alone.

In the final chapter, I review the aims of the thesis and the research questions I sought to answer. I discuss the findings of the empirical chapters, setting them in the context of the

conceptual and empirical literature. I summarise the limitations of the work carried out. I draw out the implications of the findings and make recommendations for policy and for further research.

1.6 My Contribution to the Study Research Effort and Relationship with the PhD

I was involved, under the supervision of Martin Knapp, in planning the economic evaluation component of the WSD study since the inception of the project in 2008. I was responsible for designing the cost-effectiveness analyses. Because of the enormous scale of the study, it would not have been practical for me to collect the quantitative data personally. The evaluation questionnaires, including those needed for the economic evaluation, were administered by interviewers from a company specialising in research interviewing. For this reason administering the questionnaires did not form part of my planned fieldwork, although I carried out all the subsequent cleaning and processing of the data collected through the cost collection instruments. I worked within a team of researchers assembled across a number of other institutions, whose objective was to investigate both outcomes and costs. We worked together to discuss and interpret the emerging results. I was first author of three published papers (Henderson, Beecham, and Knapp 2013, Henderson et al. 2014, Henderson et al. 2013) on the costs and cost-effectiveness of the interventions, based on data from the questionnaire studies; I was co-author on five others (Steventon et al. 2012, Steventon et al. 2013, Hirani et al. 2013, Cartwright et al. 2013, Bower et al. 2011). All data were collected between 2008 and 2010. All the statistical analyses carried out for this thesis are my own work, with the guidance of my supervisors. While the cost-effectiveness analyses were presented in the published papers noted above, the analyses presented here feature new work. All errors are my own.

Chapter 2

Telehealth and Telecare: The Context

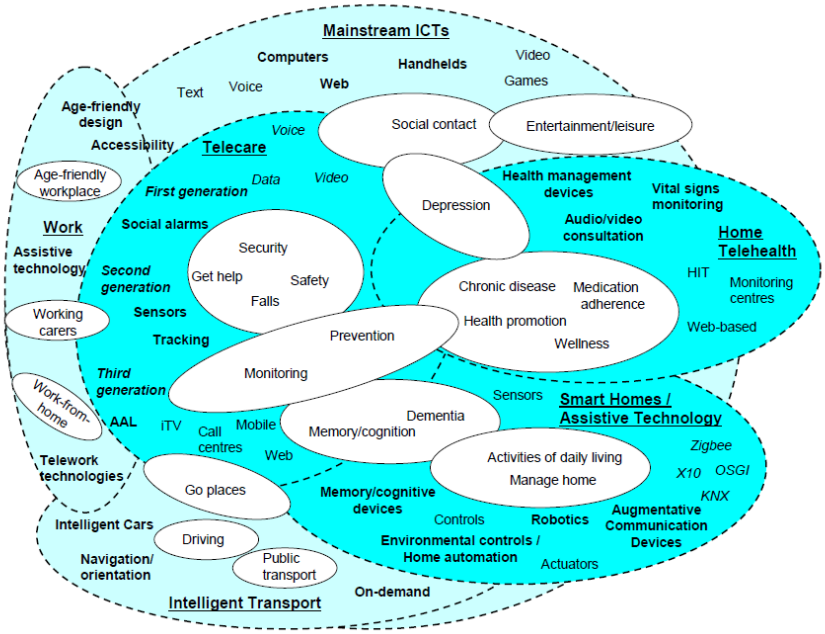
2.1 Introduction

The purpose of the thesis is to investigate whether telecare and telehealth, as defined within the Whole Systems Demonstrator evaluation, were cost-effective interventions. It is necessary to understand something of the nature of these interventions and where they lie in the assistive technology landscape, prior to considering cost-effectiveness. In this chapter, I discuss assistive technology broadly, setting out some definitions of the technologies as they are understood in the UK and internationally. I introduce some important sub-classifications of the terms 'telecare' and 'telehealth', in order to set the material characteristics of these technologies into context. I set out the systems-level context and address supply-side and demand-side perspectives on the purpose and function of the technologies. I briefly review the conceptual literature to ask what are the mechanisms whereby the technologies 'work'.

2.2 Telehealth and Telecare are Assistive Technologies

The term 'assistive technology' can cover a broad range of different technologies, from low-tech (ramps, rails, bath equipment) to advanced, or high-tech (telemedicine via video-link, health 'apps', robotics). Kubitschke and Cullen (2010) carried out a wide-ranging survey of information and communications technology (ICT) that could be used to address the needs and challenges created by ageing. Their map relating different kinds of ICT to these needs is reproduced in Figure 2.1. The diagram illustrates how a need may be addressed by more than one technological domain: for instance, aspects of depression-related needs might be addressed by both telehealth and telecare. These technologies may therefore cross service boundaries, both offering opportunities and creating tensions between services in the process.

Figure 2.1 The spectrum of needs and technology reproduced from Kubitschke and Cullen (2010)



2.2.1 An Array of Telehealth and Telecare Terminology

It should be noted that there is a variety of terminology used within the literature on telecare and telehealth, depending on the discipline and nationalities of the authors. Both technologies have been described variously as types of 'telemonitoring' and aspects of 'remote care' (Barlow et al. 2012). Researchers reviewing the literature have not always differentiated *telecare* 'telemonitoring' from *telehealth* 'telemonitoring', rather viewing the two as variants of technology with similar purposes.

Telecare has been characterised as a form of "health smart homes and home-based consumer health technologies that support aging in place" (Reeder et al. 2013) or "assisted living technologies (ALTs) that specifically enable older people to 'age in place'" (Graybill, McMeekin, and Wildman 2014), as well as "ADL telemonitoring" (Gokalp and Clarke 2013) and "gerontechnology" (Piau et al. 2014).

Telecare in the UK was once used to describe any preventative or supportive technology, but as medical applications for telecare were introduced, these have tended to be labelled telehealth (Doughty et al. 2008). This can be confusing, both for those seeking to understand the evidence base for AT, and for commissioners considering the purchase of AAT services (Doughty et al. 2008).

While telehealth and telecare can be viewed as elements of a common 'remote care' approach, the technologies are more often treated as separate domains of care. Section 2.2.2 covers definitions and classifications within telehealth and Section 2.3, within telecare.

2.2.2 Definitions of Telehealth

The term 'telehealth' as employed in the literature covers several distinct areas of clinical activity. There are 'telehealth' activities that could be characterised as telemedicine or distance medicine, where health care professionals use telecommunications to deliver health care through, for instance, joint teleconsultations (Currell et al. 2000). Then there are activities that are delivered by health care professionals to patients directly, which might involve 'telephone support' (or 'coaching') or might involve 'telemonitoring'. Telephone support involves the use of a simple or 'plain old' telephone system by health care providers to deliver support to patients or carers; and transmission of vital signs data does not occur over that system (Polisena, Tran, et al. 2009, Inglis et al. 2010). Telemonitoring can involve vital signs being monitored in real-time, for instance via video-link (also known as synchronous monitoring), or monitored using store-and-forward systems, where data such as video clips or sound files are submitted by the patient and transmitted to the health professional for later assessment (asynchronous monitoring) (Polisena, Tran, et al. 2009, Bergmo 2009). This non-invasive vital signs data are usually delivered by newer telecommunications technologies such as broadband or wireless data transfer (Inglis et al. 2010).

Telehealth as 'telemonitoring' can be classified into four generations (Anker, Koehler, and Abraham 2011, Cartwright et al. 2013):

First generation or 'non-reactive data collection and analysis systems': measurements are transferred to healthcare providers by store-and-forward systems asynchronously. Healthcare providers cannot respond in real time.

Second generation or 'non-immediate analytical or decision-making structure' systems: measurements are transferred in real time; the system processes and analyses the data provided by the patient. Healthcare providers are available to respond in real time, but their responses may be delayed if the systems are not running continuously out-of-hours.

Third generation or 'remote patient management systems': in addition to real-time processing and analysis of patient data, the monitoring centre is led by physicians and carried out by specialist nurses.

Fourth generation or ‘fully integrated remote management systems’: in addition to the features of third generation telehealth, patient data may be collected by non-invasive and also invasive (implanted) devices. The system is monitored by physicians.

2.3 Definitions of Telecare

The *Telecare Aware* website gives an expansive definition of Telecare:

> …from simple personal alarms (AKA pendant/panic/medical/social alarms, PERS, and so on) through to smart homes that focus on alerts for risk including, for example: falls; smoke; changes in daily activity patterns and 'wandering'. Telecare may also be used to confirm that someone is safe and to prompt them to take medication. The alert generates an appropriate response to the situation allowing someone to live more independently and confidently in their own home for longer. (Telecare Aware)

As illustrated in this quotation, telecare can be an umbrella term for rather different services such as “pendant alarms” and “smart homes”. Differences in the type and sophistication of telecare equipment and systems can also be classified in terms of ‘generations’ of telecare technology (Kubitschke and Cullen 2010). There appears to be some consensus in the literature that there are three generations.

First-generation: This form of telecare consists of a telephone unit and a pendant alarm with a button for summoning help; a monitoring centre receives the alert and identifies the user and can contact the user via the telephone unit; there is a protocol in place to alert the appropriate nominated responder (a paid or unpaid carer). The terms ‘social alarm’ and ‘community alarm’ and ‘personal emergency response system’ (PERS) are used synonymously to describe this generation of telecare (Kubitschke and Cullen 2010).

Second-generation: Here, automatic, passive alarm/sensor systems are added to the telephone unit; alarms/sensors can be triggered automatically and send an alert to the monitoring centre (Kubitschke and Cullen 2010).

Third-generation: These automatic, passive alarms/sensors in the home provide data that can be viewed by paid or unpaid carers to monitor the user’s well-being and evaluate the user’s care needs (Kubitschke and Cullen 2010). ‘Lifestyle monitoring’ is another term used to describe this generation of telecare. This involves the home installation of sensors “to monitor behaviour in order to gain an understanding of ‘normal’ activity so that any unusual changes over time can be recognised and responded to” (Brownsell et al. 2011, p.185).

2.3.1 Smart homes

Telecare also can be seen as fitting into a part of a broader concept, that of smart homes, "an innovative concept that integrates technology within residences in order to maintain and even enhance functional health, security, safety and quality of life of their residents" (Demiris and Hensel 2008, p. 35). Balta-Ozkan et al. (2013) give a broader definition: "A smart home is a residence equipped with a high-tech network, linking sensors and domestic devices, appliances, and features that can be remotely monitored, accessed or controlled, and provide services that respond to the needs of its inhabitants." Martin, Kelly et al. (2008) have created a useful hierarchy of smart homes, adapted from Aldrich (Aldrich 2003). At the most basic level, smart home environments can contain stand-alone "intelligent objects" for environmental control and monitoring; or at a greater level of sophistication, these objects can be networked within the home. In the more technologically complex "ubiquitous home", these networks extend beyond the home. The data are collected automatically by the technology without the resident having to initiate this process. The information thus accumulated can be used for care assessment and planning processes. Beyond this level, "learning homes" gather information on activity patterns, which can be compiled, so that the occupant's future needs can be anticipated and the technology adjusted in accordance; while "attentive homes" build on these systems to continuously record activity to the same end.

It might seem that 'telecare', 'smart homes' and 'lifestyle monitoring' descriptors are interchangeable; however there can be important differences between these technologies. Unobtrusive technology works behind the scenes to collect information for some proactive future use in 'learning' and 'attentive' smart homes, 'third-generation telecare' systems and 'lifestyle monitoring'. This continuous data-gathering aspect is absent from first- and second-generation telecare systems. In practice, however, the boundaries between the labels 'telecare' and 'smart homes' are sufficiently blurred that interventions that could be classified as second-generation telecare according to Kubitschke and Cullen (2010) are described elsewhere in the literature as examples of smart home technology (cf. Peek, Aarts, and Wouters 2015) (and see also Chapter 3, section 3.5).

2.3.2 Functional classifications of telecare and similar technologies

Several classifications of home-based remote care technologies have been proposed (Box 2.1). Demiris and Hensel (2008) suggest that health-related 'smart home' technologies could be categorised in terms of six functions: physiological monitoring; functional monitoring and emergency detection; safety monitoring and assistance; cognitive and sensory assistance; and social interaction monitoring and assistance. Categorisations by Doughty and Steele (2009) include similar functions; Brownsell, Blackburn, and Hawley (2008) list packages including specialist monitoring of people with specific conditions such as epilepsy, and lifestyle monitoring (see Section 2.3).

Box 2.1 Categorisations of smart home and telecare technologies

Six categories of smart-home technologies (Demiris and Hensel (2008), p. 34)

Technologies that collect and analyse data for:

- Physiological monitoring – measuring vital signs (e.g. blood pressure);
- Functional monitoring/emergency detection and response – measuring activity levels, motion, ADLs, and critical events such as falls;
- Safety monitoring and assistance – detecting environmental hazards and providing safety assistance e.g. automatic lighting, location technologies;
- Security monitoring and assistance – detecting human threats e.g. intruder alarms;
- Social interaction monitoring and assistance – tracking phone calls, visitors, e.g. video technologies to communicate with relatives, participate in groups online;
- Cognitive and sensory assistance – automated reminders; sensory devices for sight, touch e.g. water temperature sensors.

Categories of stand-alone telecare devices (Doughty and Steele (2008), p. 41)

- Safety – e.g. bath thermometers
- Security – e.g. timed lights
- Communication – picture phones
- Reminders – pill dispensers
- Safe walking – GPS trackers

Four types of telecare package (Brownsell, Blackburn, and Hawley (2008), p. 9)

- Security package – includes intruder alarm, flood and temperature detection, CCTV of entrance
- Falls package – fall detectors and lights on sensors
- Specialist devices – epilepsy bed monitors, vibrating pillow alerts, front door alarms
- Lifestyle reassurance (third generation system) – bed occupancy detectors, movement detectors, electrical usage

2.4 Markets for Telehealth and Telecare in the UK

Kubitschke and Cullen (2010) estimated the number of potential first-generation telecare users across the European (EU27) market to be in the vicinity of 2.6 to 12.8 million users in

2009; the number of potential users with a long-term condition of heart disease was estimated to be somewhere between 9.4 million and 13.9 million, and the number of potential users with diabetes between 3.8 and 5.4 million. The authors also reported that the UK and Ireland had the highest levels of first-generation telecare market penetration in Europe in 2010, estimated at 14-16% of people of 65 years and over; in contrast, only approximately 1% of older people had taken up second-generation telecare.

The number of telehealth users in Great Britain was about 350,000 in 2010 according to Barlow et al. (2012). Estimates of the number of telecare users (whether first- or second-generation) in England vary substantially, depending on the source. According to a report by Deloitte Centre for Health Solutions (2012), the actual number of users in England in 2010 was 1.6 million. On the other hand, based on data obtained via Freedom of Information Act 2000 requests from 121 (of 152) English councils, Corbett-Nolan and Bullivant (2012) reported that there were only 204,809 telecare users in 2009/10 and 241,582 in 2011/12. Over the 2011/12 financial year, councils spent a total of £50 million on telecare (an average of £500,529 per council). In terms of the proportion of the older population using some form of telecare, the figure appears to be low. Nyman and Victor's (2014) analysis of the English Longitudinal Study of Ageing (ELSA) dataset (Marmot et al. 2008) found that only 6% of adults 65 years and over reported the use of a personal call alarm (180 of a sample of 3091).[8] Using ELSA data (wave 4) and drawing on a more expansive definition of telecare, Lloyd (2012b) reports that 2% (375,000) of individuals 50 and over in England used mobile personal alarms[9] while 4% (720,000) had an alerting device.[10]

Market analyses and projections have predicted considerable scope for growth in uptake of tele-technologies. Barlow et al. (2012) suggested the existence of substantial untapped demand for "preventative remote care" in the UK: if the technology was expanded to cater for not only the most intensive users of health services but targeted at a broader pool of occasional health service users, this could open up a market of as much as 1.4 million people. In a similar vein, Lloyd (2012a, b) predicted a potential market of as many as 4,175,000 users in England if the service was broadened to include the younger-old (people

[8] 'Personal call alarms' were defined as alarms that are used to call for assistance after falls (Nyman and Victor 2014).
[9] 'Mobile personal alarms' were defined as alarms to call for assistance after falls, excluding other types e.g. attack alarms.
[10] These 'alerting devices fixed to the home' were defined as devices such as pendant alarms. The authors further define these devices as being fixed to the home, and also that they could be used in the event of a fall but the term could refer to a broader range of devices.

aged 60 and over) at 'low-risk' as well as the current population of telecare users, who tend to be older and frailer people at 'high-risk'.

Globally the market for telehealth and telecare was £1.7 billion in 2015 (Monitor Deloitte 2015). This market has grown significantly over the past decade. In the UK, the total market for telecare and telehealth brought in revenues of £141.7 million in 2010 (Deloitte Centre for Health Solutions 2012). By 2014, the market for telehealth alone in the UK was £90 million;[11] and the telecare market was worth £246 million (Monitor Deloitte 2015). The telehealth market in the UK was predicted to grow strongly between 2014 and 2018, at a compound annual growth rate (CAGR) of 13% to £148 million; but the telecare market was expected to grow more slowly at a CAGR of 4% to £292 million over the same period (Monitor Deloitte 2015).

Although the technologies are relatively new, they already face disruption from further smartphone-based developments such as mHealth (mobile applications that monitor vital signs) and 'connected homes' and 'internet of things' (IoT) applications (aimed at the general public to monitor aspects of the home environment such as temperature and security). A more privatised telecare market is emerging, manifested by increased co-payments, bundling of monitoring and response services and devices through a managed service, and increasing numbers of private payers (Monitor Deloitte 2015).

2.5 Supply and Supplier Perspectives

From the point of view of industry, there have been a number of barriers to the expansion of the market for telehealth and telecare. A number of these spring from the demand-side. Many health and social care markets in the UK feature public payers with strong purchasing power: the markets for telecare and telehealth are no exception to this general picture. The vast majority of UK telecare sales are to councils, only about a tenth being privately purchased; while telehealth services are typically commissioned by NHS organisations. Health and social care organisations' budget constraints can limit their flexibility and attitude to risk, hampering innovation (Barlow et al. 2012). Health providers' fondness for pilot projects has limited the growth of the telehealth market (Monitor Deloitte 2015, Barlow et al. 2012). For instance, parallel services have to be set up for a typically small number of pilot patients as a temporary arrangement, but this offers limited lessons for delivering the technologies at scale (Barlow et al. 2012). Lack of commissioning expertise with the technologies and short-term

[11] Their definition of telehealth was more inclusive as it included teleconsultation.

thinking can hamper successful procurement; for example, health service commissioners have tended to think of telehealth as merely a set of devices rather than a potentially transformative model of care requiring both expert software support and organisational learning (Barlow et al. 2012). The funding model does not encourage NHS commissioners to invest in new technologies as the funding is based on patient admissions, procedures or visits (Deloitte Centre for Health Solutions 2012). The telehealth and telecare manufacturing sector is relatively small in the UK. In 2015, around 20 companies were involved in manufacturing telehealth and/or telecare equipment (Monitor Deloitte 2015). About a quarter (23 per cent) had annual revenues of £1 million to £10 million (Monitor Deloitte 2015). Barlow et al. (2012) note that companies supplying telehealth and telecare technology have been relatively small-scale, typically employing about 40 people; according to Monitor Deloitte (2015), only 12 per cent of companies in this sector had more than 50 employees.

As it typically requires multiple stakeholders to produce remote care, the business model needs to allow for partnerships. As Barlow et al. (2012) say,

> The key features of a successful business model are an identified market, a value chain to create and distribute the offer, an understanding of the value chain's cost structure and profit potential, and understanding of the roles of different suppliers in the value chain, and finally – for private sector companies – a competitive strategy to gain and hold advantage over rivals. In remote care in the UK, these features are only partially developed. (p. 13)

In addition, there can be problems related to the technologies themselves. Some of the rapid development of devices was induced by suppliers looking to expand their share of the market, creating complexity for purchasers; also a lack of interoperability locked purchasers into ordering most of their equipment from one supplier (Barlow et al. 2012). At the beginning of the current decade, purchasers were hesitant to invest in telehealth because the devices were perceived to be expensive. Yet while costs of newer devices were falling, they could still "end up sitting on a shelf" (Deloitte Centre for Health Solutions 2012 p.15). There was a large number of purchasers (Primary Care Trusts (PCTs), and post-2013, Clinical Commissioning Groups (CCGs)) whose budgets for telehealth varied substantially. Clinician buy-in or resistance was a problem, such that a number of CCGs reported problems with clinical take-up and with supply chains, resulting in some CCGs deciding to cease commissioning telecare and telehealth (Monitor Deloitte 2015).

On the telecare side, incentives were not well aligned, in that local authority purchasers could question allocating their limited funds to telecare services, only to achieve reductions in hospital admissions and other benefits accruing to the health system. Because

telecare has been a feature of the social care system for many years, opportunities for market expansion have become rather limited (Monitor Deloitte 2015). Despite this apparent near-saturation of the telecare market, there has been significant variation in numbers of telecare users between local authorities (Corbett-Nolan and Bullivant 2012).

It is evident that demand-side factors such as public payers' willingness and ability to pay the costs of the technologies are of great importance in determining the future success of the market for telehealth and telecare. The Department of Health (DH) entered into a concordat with the telehealth and telecare industry to show the extent of governmental commitment to promoting the uptake of telehealth and telecare (Department of Health 2012a). The DH simultaneously launched a five-year campaign (called *3millionlives*) aimed at enabling 3 million people with social care need and long-term conditions to access these technologies.

2.6 Demand and Demand-side Perspectives

While the role of public payers is clearly important to the growth of the telehealth and telecare market and to the success of private-sector suppliers of the technologies, the role of other demand-side factors deserves consideration. These include consumers' preferences as to material aspects of the technologies; their attitudes towards the substitutability and complementarity of the technologies with other services; and their willingness and ability to pay for the technologies. The academic literature that touches on consumer preferences for telehealth and telecare encompasses disciplines such as health care research, sociology, gerontology, political science, engineering and computer science.

Research into stakeholders of telecare and telehealth by Greenhalgh et al. (2012) identified four discourses. In the modernist discourse, technology was a rational solution to a demographic problem. To age well was to use technologies proficiently; contact with paid carers was seen as a needless waste. By and large, the modernist discourse reflected the UK policymaking position. The change-management discourse portrayed the technologies as societally useful but challenging to implement and in need of project management; the humanist vision portrayed the lived reality, with technology having meanings that could be positive or could be stigmatising; the political economy discourse portrayed technology as potentially an agent of social control. Stakeholders' discourses and agendas were found not to be in alignment, with no single 'organising vision'. As a consequence, the authors predicted

that there could be a considerable gap between the cheerleaders for the technologies and actual take-up.

A complication in considering where demand-side issues begin and end lies in the co-productive aspect of telehealth and telecare services. These services are rarely produced by only one organisation or within one sector. Telehealth services require the involvement of a health care practitioner (usually an NHS employee) as well as telemonitoring software and vital-sign monitoring equipment supplied by private sector companies. Telecare services are delivered by monitoring call centres typically run by local authority or voluntary sector staff (although purely private sector call centres do exist), using telecare call-handing software and monitoring equipment from private sector suppliers. Thus, while not 'suppliers', telemonitoring clinicians and telecare call centre providers are producers of these services. In the case of telehealth, clinicians (whether or not directly involved in telemonitoring) have an important demand-side role, with the power to influence commissioners' purchasing decisions and patients' attitudes towards this new technology. In the following discussion, these actors are considered as consumers rather than suppliers.

2.6.1 Adoption and Acceptance of Telehealth and Telecare: Stakeholder Perspectives

A number of studies have investigated factors influencing the acceptance of and resistance to health and care technologies. Clinicians' resistance to telehealth has been remarked on within the policy and academic literature (Standing et al. 2016). Health professionals' predisposing attitudes – whether scepticism about the evidence base, or fears of duplication – can influence their engagement in the use of telehealth. Tensions are likely to arise between the goals of policymakers and clinicians because of the nature of the telehealth intervention: the introduction of new technologies, new systems of information management and new clinical protocols and processes could potentially increase the clinical workload. Clinicians may see telehealth as a threat to their traditional work roles. Also they may fear that patients will become over-reliant on and have unrealistic expectations of telemonitoring (Salisbury et al. 2015, Morton et al. 2017, Segar et al. 2013, Vassilev et al. 2015).

Some of the literature on adoption and acceptance of tele-technologies documents results of consulting stakeholders in the process of designing or market-testing new telemonitoring products; some of it reports stakeholder perspectives as part of research intended to inform system-level introduction of these technologies. The bulk of adoption research focuses on the prospective or pre-implementation stage (Peek et al. 2014).

Stakeholders in the adoption process include prospective users, carers and designers, as well as professional caregivers. In this literature, the concerns of potential users and carers about telemonitoring turn on the safety, reliability and costs of technology and threats to privacy and autonomy (Bentley et al. 2016, Cook et al. 2017, Milligan, Roberts, and Mort 2011, Peek, Aarts, and Wouters 2015, Percival and Hanson 2006, Powell et al. 2010, Rahimpour et al. 2008) . Nonetheless, there is also evidence that potential users can be receptive to the potential usefulness of new technologies (Williams, Victor, and McCrindle 2013), indeed they may be more receptive than clinicians (Standing et al. 2016).

In the case of telecare, there is an apparent disconnect between prospective users' receptiveness to these technologies and actual use. While older people may consider these technologies potentially useful, they may also think that they are for other, more disabled people than themselves (Bentley et al. 2016, Peek, Aarts, and Wouters 2015). Carers in Powell et al.'s study (2010) similarly saw new technologies as likely to be used increasingly in the future, and potentially helping them to be able to care at a distance, yet also envisaged them as being for 'other people' than themselves. Consumer acceptance and use of telecare is also influenced by the design of the devices; for instance, people do not want to wear devices that might mark them out as vulnerable (Williams, Victor, and McCrindle 2013). Prospective users' involvement at the design stage could be greater: older people do wish to be involved in the design of and choice of telecare devices and services (Peek, Aarts, and Wouters 2015). This is problematic, according to some assessments (Roberts, Mort, and Milligan 2012, Milligan, Roberts, and Mort 2011), because telecare is subject to a strong 'technological push' from industry, not responding to the actual needs of potential users, a situation that could lead users to experience disempowerment.

An individual's level of social care need may influence the choice to seek out and use telecare technologies. Nyman and Victor (2014) found from their analyses of ELSA data (wave 3) that having difficulty with activities of daily living and being in an older age category predicted personal call alarm use in people of 65 years and older.[12] Their findings also differed by household composition, in that people aged 75 to 84 years who were living with others were more likely to report using an alarm than those aged 65-74 years who were living with others; yet those living alone aged 85 years and older were more likely than those living alone aged 65-74 years to report using an alarm. This perhaps suggests that people living alone have to be much frailer before contemplating the use of these devices. Lloyd

[12] Their analyses controlled for other characteristics such as wealth, functional ability and quality of life scores.

(2012b) examined ELSA data (wave 4), finding that telecare users[13] were in poorer health than non-users, more likely to have difficulties with mobility and activities of daily living (ADL) than non-users. Users were also more likely to report particular reasons for difficulty in walking such as difficulty with balance, dizziness, fear of falling or fatigue than non-users. Users also had lower scores on memory and executive function measures than non-users. An Australian study by De San Miguel et al. (2015) compared people who had purchased a personal emergency alarm with people who had enquired about but decided not to purchase an alarm. Purchasers were older than non-purchasers and more functionally dependent in activities of daily living. These statistics appear to support Peek, Aarts, and Wouters (2015) surmise that there is a limit to the number of older people who see a need for telecare services, beyond those who are already disabled and in a position to consider that telecare will be a useful way to mitigate frailty and disability. Golant's (2017) model, discussed below, offers a more formal framework covering similar territory.

2.6.2 Models of Health Information Technology Adoption and Implementation

There is a substantial body of conceptual literature on consumer acceptance and adoption of new information technologies. These include the Technology Acceptance Model (Davis, Bagozzi, and Warshaw 1989, Davis 1989), the Unified Theory of Acceptance and Use of Technology models (Venkatesh, Morris, Davis, & Davis, 2003; Venkatesh, Thong, & Xu, 2012) and the Value-based Adoption Model (Kim, Chan, and Gupta 2007). It is beyond the scope of this chapter to present more than a brief explanation of these generic models.

Technology Acceptance Model (TAM): The model examines acceptance in the organisational context. The concepts of perceived usefulness (the belief that the technology will improve job performance) and perceived ease of use (the degree of effort believed to be needed to master the system of technology) are key to acceptance behaviours. Perceived ease of use influences perceived usefulness; perceptions of both ease of use and usefulness influence the attitude to using a technology, which in turn influences behaviour. Perceptions of usefulness also directly influence acceptance behaviours. External factors (e.g. quality of the system) can affect perceptions of usefulness and ease of use.

Unified Theory of Acceptance and Use of Technology (UTAUT): Venkatesh et al. (2003) built on models of technology acceptance in the organisational context to construct the UTAUT. The model has four constructs that determine acceptance: performance expectancy

[13] Users of mobile personal alarm and alerting devices.

(belief that the system will assist job performance, effort expectancy (ease of using the system), social influence (belief that others think the prospective user should use the system) and facilitating conditions (the belief that the system has the infrastructure - organisationally and technically – to support the system). Individual characteristics (age, gender, voluntariness of use and experience) are moderators of some relationships between these determinants and behavioural intentions. Behavioural intentions and facilitating conditions both directly influence use of the technology, moderated by age and experience.

Unified Theory of Acceptance and Use of Technology 2 (UTAUT2): This was adapted from UTAUT to suit a consumer context, with three additional constructs to that model: hedonic motivation (pleasure derived from the technology), price value (trade-off between potential benefits and costs) and habit (prior use). The model posits complex relationships between the consumer's intention to accept and use technology. For instance, individual characteristics (age, gender and experience of use) moderate the relationship between hedonic motivation and behavioural intention.

Value-based Adoption Model (VAM): The VAM seeks to explain adoption of more recent ICT such as mobile internet by consumers. In this model, building on consumer choice theories, there is a benefit component and a sacrifice component to the consumer's perception of the value of the product. Benefit consists of two concepts: usefulness and enjoyment; sacrifice consists of the concepts of technicality (quality of service) and perceived fees (the consumer's subjective assessment of the service price). These four concepts influence the perceived value of the technology which in turn influences the consumer's intention to adopt the technology.

The models discussed above are generically applicable to acceptance of ICT by workers and/or consumers and not specific to health information technologies (IT). I next discuss models of acceptance and adoption of health IT, beginning with a model of consumer acceptance with particular relevance to telehealth and telecare.

Golant (2017) proposes a model of smart technology adoption by older people seeking to age in place (here, 'smart' technologies include telecare and telehealth but also ICT). The model builds on previous models such as TAM, UTAUT and UTAUT2, and incorporates theories from other disciplines such as marketing and social psychology. Older people may have unmet needs related to ageing (including disability and chronic ill health). In this situation, an individual will examine a set of coping options that include 'assimilative' coping strategies such as: adoption of smart technologies; traditional solutions such as paid and unpaid assistance with ADL, 'low-tech' adaptive equipment and relocation to other

housing; or a mixture of the smart technology and traditional solutions. The person may alternatively adopt 'accommodative' coping strategies that involve taking no action (for instance, denying or accommodating mentally to the problem). In the model, the smart-technology adoption decision is positively influenced by the degree of perceived stress of having unmet needs, which influences the extent to which the person will take into account external information (e.g. from media, relatives, professionals) and past personal experiences. Both perceived stress and personal resilience impact upon adoption indirectly to influence an individual's overall appraisal of the 'efficaciousness' (usefulness), usability and 'collateral damages' (unintended consequences of technology such as loss of identity) of the technology. As a result, "when older people feel more stressed because of their unmet needs, they will be more motivated to attend to and evaluate information about their possible coping solutions." Older individuals facing the need to decide on a coping option will not adopt the smart technology option unless they appraise that option as better than 'traditional' options in terms of efficaciousness, usability and (lack of) collateral damages.

Golant's model chimes with themes identified in the smart home and telecare stakeholder research discussed in the previous section (2.6.1). Barriers to take-up may arise for complicated reasons, not least that older people may see telecare technologies as not all that useful to their own situations. "Many older adults have the desire to age in place, and many older adults also believe that smart home technology can contribute to independent living, yet these conditions often do not translate into a willingness to accept smart home technology" (Peek, Aarts, and Wouters 2015, p.4).

Health information technology must be not only accepted by consumers (users and front-line professionals) but implemented by a variety of stakeholders. Greenhalgh et al. (2017) propose a new framework for examining technology implementation in terms of "Nonadoption, abandonment, spread, scale-up, and sustainability of patient-facing health and care technologies" (NASSS). The framework was devised to appraise the likelihood of success of new health- or social care-related IT systems (e.g. ICT, smartphone apps, telecare and telehealth) at the individual level (non-adoption/abandonment) and the organisational level (failure to scale-up, spread or sustain the new system). NASSS consists of six domains: type of condition; technology; the value proposition; system adopting the technology; wider institutional context; adaptation between domains over time. The domains contain sub-questions, the answers to which are graded into three classifications (simple, complicated, complex). For instance, the key features of the technology (in the technology domain) could be 'simple' or off-the-shelf, 'complicated' because not properly developed, or 'complex'

because of serious dependability issues. Applying the technology implementation framework to real-world case studies, the authors found that adoption could fail because a technology might be thought to be complicated (having several components but predictable to implement) when it was in fact complex (constantly changing during implementation). Or a technology could be built for 'textbook' conditions that did not fit individuals well; it could be under-developed or unreliable; workers could find the data unhelpful and choose not to use it; its value could be unclear to users; the organisation was unable to find a commercial partner; there could be regulatory hurdles; it could be unadaptable for the local health and social care system. In particular, technology systems that were 'complex' across several NASSS domains were very difficult to bring into mainstream use.

2.6.3 How Do Telehealth and Telecare 'Work'?

The previous discussion has covered some definitions of telehealth and telecare, markets for these technologies and the supply-side and demand-side factors that have influenced their development. A further question relevant to setting the context of the economic evaluation of these technologies concerns their 'active ingredients'. How do these technologies work to produce outcomes of interests to patients, practitioners and policymakers? Health telemonitoring technologies purportedly play a role in the management of long-term conditions: I explore this first. I then examine the question of the role of telecare in fostering independence and improving quality of life outcomes.

2.6.4 Long-term Conditions, Self-management and Telehealth

To consider what role telehealth could play in the management of chronic conditions requires a little background on disease management. In the health policy literature, perhaps the best-known approach to disease management is the Chronic Care Model (CCM), an evidence-based framework for quality improvement and condition management involving system-level redesign (Wagner et al. 2001, Wagner, Austin, and Von Korff 1996). According to the model, effective self-management support linked to community resources are essential components of good chronic disease care, made possible by appropriate health care organization, delivery system design, *decision supports* and *clinical information systems*. In this model,

> High-quality chronic illness care is characterized by productive interactions between practice team and patients that consistently provide the assessments, support for self-management, optimization of therapy, and follow-up associated with good outcomes. (Wagner et al. 2001, p. 68)

The 2005 White Paper *Our Health, Our Care, Our Say* (Department of Health 2006) summarised the elements of current disease management in lay terms:

> At the moment, half the people with long-term conditions are not aware of support or treatment options and do not have a clear plan that lays out what they can do for themselves to manage their condition better. If people have a clear understanding of their condition and what they can do, they are more likely to take control themselves. (p. 8)

Self-management support – 'helping people to help themselves to manage their condition' – is strongly emphasised in Our Health, Our Care, Our Say and subsequent documents (Department of Health 2008b, 2012b). To unpack terms widely deployed in the disease-management literature, 'self-care', 'self-management' and 'self-management support' all refer to aspects of the practice of involving the patient in managing a chronic condition (Rijken et al. 2008). In self-care, the emphasis is on lay experiences of managing health problems in the context of everyday living (Rijken et al. 2008); in self-management, patients take on managing their conditions between their usual appointments with health care practitioners; in self-management support, patients and practitioners collaborate to treat the condition. Support includes joint goal setting and treatment planning, problem solving around barriers to self-management, information provision and efforts to enhance patient confidence (Wagner et al. 2001, Rijken et al. 2008). Patients who are active participants in their care can engage in more productive interactions with their health care providers (Wagner et al. 2001).

In the extensive conceptual literature on the self-management of chronic conditions, a number of theories (such as rational choice; self-regulation models (Petrie and Weinman 1997); and the stress coping model (Ridder and Schreurs 1996, Lazarus and Folkman 1984) describe how behaviours are related to psychological and social factors. These approaches form a starting point for health care practitioners seeking to effectively support patients in self-management (Rijken et al. 2008). Depending on the theoretical underpinning, the choice of support strategy could differ. For instance rational choice theory assumes that people make decisions on the basis of welfare-maximisation: the decision to modify one's behaviour will depend on whether the resultant benefits are assessed to be greater than the costs (Rijken et al. 2008). Strategies suggested by this theory include providing financial incentives to change patient behaviour and education about potential consequences of the illness and availability

of information about treatments. In the self-regulation model, behaviours are shaped by illness representations (which are beliefs about the illness including its aetiology and the possibility of influencing the outcome of the illness). Thus the behaviours of patients with chronic conditions may be changed by changing their illness representations (Weinman and Petrie 1997, Rijken et al. 2008).

Strategies to support people in managing their long-term conditions depend on several aspects of the person: knowledge of the illness, beliefs about the illness, the individual's attitudes towards healthy or unhealthy behaviours, levels of confidence, personal motivation and the characteristics of the individual's social networks (Rijken et al. 2008). UK health policymakers have asserted that telehealth technology has the potential to facilitate chronic disease management and also to facilitate self-management and self-management support (Department of Health 2006, 2008b). How might this work? As Salisbury et al. (2015) contend, theories to *explain* the workings of telehealth to achieve the goals of chronic conditions management have been largely lacking. The Chronic Care Model, for example, "does not in itself provide a model for the design of telehealth interventions" (p. 4). Salisbury et al. (2015) propose a conceptual model specific to telehealth (TElehealth in CHronic Disease, "TECH"). The model describes the means by which telehealth works to achieve beneficial outcomes for patients. This consists of 5 components: effective chronic disease management (promoting self-management, optimising treatment and care coordination); partnership between telehealth, primary care and other health care providers; contextual factors (characteristics of patients, wider health and social care system); and engagement of patients and primary care providers. The model posits that contextual factors and engagement moderate the relationship between managing the chronic condition and outcomes (health outcomes, costs, patient access to care and experience), while telehealth services delivered in partnership with other health providers will also produce beneficial patient outcomes. The model puts telehealth in a facilitating position between condition management and desired outcomes.

Gee et al. (2015) propose a role for eHealth (including telehealth but also mobile health and electronic health records) in the Chronic Care Model (CCM). They suggest that the "eCCM" would require the introduction of eHealth education (training patients to use eHealth tools, and health care practitioners to implement the tools and educate patients in their use); and the addition of a "complete feedback loop" (CFL). The CFL is a five-stage cycle whereby health data are transmitted, data are interpreted on the basis of clinical information (e.g. guidelines), addressing specific patient needs, providing timely and specific

feedback to patients and regular repeating of the loop. This feedback loop will contribute to the productive patient-physician interaction (Wagner et al. 2001) and will be needed to enable eHealth technologies to promote better outcomes in people with long-term conditions. For example, CFL is important to patient-provider communications, in that the time it takes for a clinician to respond to a patient-initiated eHealth communication is likely to have an impact on a patient's satisfaction with the eHealth system.

There is some qualitative evidence on the nature of the relationship between self-management and the kinds of outcomes that policymakers desire of telehealth (for instance, better disease control, better quality of life). Vassilev et al. (2015) conducted a realist synthesis to identify three mechanisms by which telehealth interventions 'work' to produce successful outcomes: these were relationships, fit and visibility. They refined these initial concepts by interrogating the qualitative evidence base on the use of telehealth in specific long-term conditions. Their findings reflect the mediating function that telehealth plays in self-management. *Relationships* provide telehealth users with support (professional/clinical/social) for behavioural change. For instance, telehealth can elicit practitioner feedback, thereby reinforcing positive changes in behaviour. The extent to which a telehealth intervention can be integrated within everyday life (its *'fit'*), determines patients' likelihood of continued use. The patient's ability to use the technology can facilitate his/her ability to benefit from telehealth. This suggests that simple technologies that fit in with existing technologies (e.g. messaging systems giving reminders or health behaviour prompts) will be easier to use and thus more effective than unfamiliar technologies. In particular, *visibility*, "how telehealth care makes an illness or condition apparent to the self and others" (p. 23) is a mediator between self-management tasks and the patient's motivation and understanding of the condition. How the technology reveals the condition to the patient is in some way related to the patient's capacity to manage that condition. Feedback on these visible signs by professionals and peers can improve engagement with self-management activities.

There is some evidence that telehealth can be effective even in the absence of clinical involvement, particularly if the intervention is visibly tailored to the person's needs (Vassilev et al. 2015). Morton et al. (2017) conducted a meta-ethnographic review of patient and practitioner experiences of "self-management digital interventions" (including telehealth) for chronic physical health conditions. They noted that patients reported experiencing increased *motivation to change their behaviour*, on the basis of vital signs readings alone when using "stand-alone telemonitoring systems" (monitoring systems where patients send vital signs to

practitioners for feedback or health advice but which have no explicit educational or behavioural change support elements).

The above discussion suggests that telehealth could have more than one 'active ingredient'. The intervention could effect positive outcomes by presenting vital signs in such a way as to make the condition manifest to the patient, even without the involvement of a clinician. Also telehealth may facilitate positive patient-clinician interactions through a continuous feedback loop related to the patient's specific needs. On the other hand, it is worth asking whether telehealth really needs to facilitate self-management in order to be effective. Morton et al. (2017) cite Schermer's (2009) assertion that telemonitoring systems usually facilitate a "compliant self-management" by patients – following the instructions of their health care practitioners. This would suggest that telehealth does not necessarily require much engagement on the part of the patient beyond following clinical advice; however, as Schermer (2009) suggests, a "concordant" approach is also possible, integrating clinical instructions with patients' own understandings of their illness to plan treatment. Morton et al. (2017) observe that practitioners may face problems helping patients with self-management goals, if these conflict with clinical guidelines.

There is little quantitative evidence on the nature of the links between telehealth, self-management and desired outcomes. Hanlon et al. (2017) drew on the PRISMS taxonomy of self-management support (Pearce et al. 2016) to identify 5 components[14] as deliverable through telehealth. These are patient education and information provision; telehealth-facilitated clinical review; adherence support; psychological support; and lifestyle interventions. The authors examined systematic reviews of the telehealth literature to identify the reported impact of these self-management components, finding that reporting was of variable quality and that there was "little explicit evidence of the mediating role for self-management in telehealth interventions". It was not possible to establish whether telehealth could be used to support self-management. They recommended that further telehealth research should be based more explicitly on self-management theory.

[14] The PRISMS taxonomy has four overarching dimensions (mode of delivery, personnel delivering support; targeting and intensity, intervention's frequency and duration) and 14 components delivered to individuals directly, including: information about the condition and resources, agreeing clinical action plans, training in communication with clinicians, training for self-management and psychological strategies (Pearce et al. 2016).

2.6.5 Independence, Risk and Telecare

Industry and policy literature has emphasised the 'preventive' role of telecare in delaying dependence and promoting 'independence' and a better quality of life (Telecare Aware) (Department of Health Change Agent Team 2005, Department of Health 2005a, 2008c, 2015). Yet the policy message on telecare appears to take a shortcut between introducing the instrument and the desired outcome without suggesting any mechanism through which change in outcomes is to be effected. Several papers have addressed the problems that ensue from taking an over-simplistic policy line. Glasby, Lynch, and Robinson (2018) conducted a case study of telecare delivery in one English local authority. The national-level policy narrative – that telecare fosters independence and improves users' quality of life while saving money – was espoused by local social services managers and front-line workers, becoming a familiar storyline about 'better outcomes'. However, telecare service users' and social services managers' perceptions of independence differed. Managers equated 'independence' with the reduction of dependence on the state for assistance. Service users painted a more complex picture, for instance some had feelings of isolation as contact with council workers was reduced. Others aspired to 'independence' in the sense of being free to make their own decisions, even if this led to taking risks.

Aceros, Pols, and Domènech (2015) examined conflicts that arise from twin governmental policies of 'ageing in place' and 'active ageing'. The former seeks to keep older people in their homes; the latter to help older people to remain fit and active participants in their social networks and wider communities. Telecare's allure is as a mechanism to implement both policies at once, yet the policies are not in alignment. 'Active' older people may want to go out for leisure activities and to socialise, potentially putting themselves at risk for falls; 'ageing in place' can imply a much more restrictive regime, where older people are encouraged to stay put and stay safe rather than venture out and put themselves at risk. The authors suggest that the technology shapes "a particular type of user: a place-bound subject who, due to his or her age, is in need of constant alertness at home" (Aceros, Pols, and Domènech 2015 p.109). In a similar vein, Peek, Aarts, and Wouters (2015) observe that policymakers tend to focus on independence as 'being able to look after yourself' but other modes of independence (Sixsmith 1986) are in play in the use of smart home technology: self-direction and 'not being obligated to someone else'. As an example of this last, an older person might decide not to activate a pendant alarm to avoid disrupting the lives of relatives. If the technology encroaches on a person's sense of self-direction, it may

not be acceptable. All three modes of independence may come into play in response to these technologies – perceptions of usefulness alongside fears of loss of self-direction and of being a burden. Technologies designed to assist older people to age in place appear to arouse feelings of ambivalence in this group. Telecare can signify frailty and encroaching dependence and at the same time be perceived as a potentially useful means of warding off dependence (Aceros, Pols, and Domènech 2015, Bentley et al. 2016).

There is a dearth of conceptual literature addressing how telecare might act to promote 'independence' and, importantly, what is the association between a 'sense of independence' and other outcomes of interest to individuals and policymakers alike, such as self-perceived quality of life. Theoretical models of the technologies aimed at supporting 'ageing in place' are needed to guide future studies examining the relationship of these technologies and potential outcomes (Reeder et al. 2013).

There is a modest literature on the way that individuals choose to use or not use PERS in managing their lifestyles (Agboola et al. 2017, De San Miguel et al. 2015, De San Miguel et al. 2017b, Fallis et al. 2008, Williams, Victor, and McCrindle 2013). This literature is exploratory or descriptive rather than model-building; much of it focuses on compliance and patterns of use. McKenna et al. (2015) point out that the role of risk, the unpredictability associated with risk, and the decision-making around how to manage that risk are key factors in the use of these alarm systems. Users of PERS typically fail to use their devices during high-risk activities (using bathrooms, moving around at night) (De San Miguel and Lewin 2008). In a study by De San Miguel et al. (2017a), only a third of personal emergency alarm users reported using the alarm during an emergency, a quarter choosing to use the telephone to call relatives or medical assistance.

From the point of view of enhancing autonomy, one question we could ask is whether first-generation telecare offers greater scope for active decision (cf. McKenna et al. 2015, Hawley-Hague et al. 2014), which could reinforce an individual's sense of autonomy, than do later-generation technologies. While PERS users may choose to activate the alarm or not, in second-generation telecare, sensors may go off automatically, with consequences out of the control of the user. As discussed earlier in section 2.6, older people can be deeply concerned about the technology encroaching on their privacy and autonomy. In particular, Milligan's study uncovered "significant reservations about telecare systems whose primary purpose is active monitoring or surveillance and which does not rely on the older person to activate them" (Milligan, Roberts, and Mort 2011 p. 353). The important difference between first and subsequent generations of telecare – the addition of automatic monitoring to the (first-

generation) alarm-activation function – poses a further problem for locating conceptual frameworks that would be relevant in linking second-generation telecare to outcomes such as quality of life. Any lessons from the PERS evidence may have limited applicability to second-generation telecare.

To summarise, there is a gap in the telecare research literature in terms of theoretical frameworks linking telecare to final outcomes.

2.6.6 How Similar and How Different Are Telehealth and Telecare Interventions?

This thesis concerns two technologies that are often mentioned together in policy documents. While the empirical chapters examine the costs and cost-effectiveness of the technology separately, it is worth considering how much these technologies have in common. Telehealth and telecare share certain features: both can offer "remote care" – technologically-enabled monitoring by agents outside the home. The two technologies nonetheless differ in important ways. Telehealth is by and large provided by the NHS, and the monitoring, software and equipment are therefore free at the point of access. Telecare is delivered by a combination of local authority and voluntary sector providers and users can be charged for the service. Telehealth appears to require a more active role on the part of the user than telecare; for instance, the telehealth user can examine information extracted from vital signs monitors (oximeters, glucometers etc.) to manage the long-term condition. In contrast, there is no requirement for day-to-day active input from the telecare user. In this sense, telehealth is a technology that could actively promote the autonomy and decision-making capacity of the user and there is some evidence that it does so (see section 2.6.4). In the case of telecare and in particular second-generation telecare, as discussed in section 2.6.5, the picture is more complicated.

On the other hand, differences in the intensity of 'use' of the technologies may in practice be quite minimal. For instance, it is possible for little interaction to occur between the monitor and the monitored in either case. For instance, if the long-term condition is stable, with little day-to-day change in vital signs, then depending on the model of telehealth monitoring, the patient might be largely unaware of being monitored. If telecare sensors remain untriggered, the user can also largely remain unaware of being monitored.

2.7 Summary

In this chapter, I set out to define telehealth and telecare and explain the terminology used to describe these technologies. Assessing the evidence base for the technologies is made more difficult by the variety of different labels used, depending on the country and discipline producing the research. Different kinds of 'telemonitoring' are not always well delineated in the literature, with telehealth and telecare being described in some reviews as two variants of 'ageing-in-place' technologies.

I described the markets for remote technologies such as telehealth and telecare. These have been growing, despite demand-side barriers (e.g. budget constraints to health and social care financing, lack of commissioning expertise, consumer resistance) and supply-side issues (e.g. poor interoperability, small-scale production methods). The market for telehealth in particular has experienced strong growth; however, m-Health may make inroads into the demand for telephone and internet-based telehealth products.

The addition of telehealth to self-management support is a recent development but some frameworks have been proposed to understand how telehealth relates to the outcomes of self-management. Conceptual frameworks linking telecare technology to final outcomes such as quality of life are lacking.

One purpose of this chapter was to describe literature on *how* telehealth and telecare 'work'. In the next chapter, I consider the evidence on *whether* the technologies work.

Chapter 3

Evidence for Effectiveness and Cost-effectiveness of Telehealth and Telecare

3.1 Introduction

In this chapter, I examine the evidence base on effectiveness and cost-effectiveness of telehealth and telecare, the technologies that were to be implemented in the Whole Systems Demonstrators trials. The literature reviewed here focuses in large part on the literature available prior to 2013, because after this point, the evidence base contained the findings of the WSD studies (as those publications began to emerge in mid-2012). Relevant publications from more recent years are discussed in Chapter 9.

The chapter begins with an overview of the methods employed to search the knowledge base and continues with an exploration of the pre-2013 evidence on the effectiveness of the technologies, reviewing what was known up to that point about the effectiveness of the technologies in terms of preventing disease and disability and promoting better quality of life and other psychosocial outcomes. The chapter finishes with an in-depth discussion of the evidence base for costs and cost-effectiveness of these technologies, asking whether they are reducing or increasing costs to the public purse.

3.2 Methods

3.2.1 Scope

The terminology used to describe telehealth and telecare varies considerably depending on research discipline and country setting. This poses challenges to identifying literature relevant to telehealth and telecare. In selecting studies to be considered in this chapter, I used the WSD Telehealth Trial intervention and population as a guide. As described in Chapter 4, section 4.1, the trial population consisted of individuals with the long-term conditions COPD, heart failure or diabetes. The intervention involved the remote exchange of data between a patient and health care professional to assist in diagnosing and managing a health care condition (Chapter 4, section 4.2), where the patient was transmitting vital signs data using (non-implanted) devices based in the home (Chapter 5, section 5.7). The definition of telehealth was operationalised for this chapter as telemonitoring (TM) and/or telephone support (TS) for long-term respiratory, cardiac and diabetic conditions. I defined the

following activities as not within the scope of this review: health-professional-to-health-professional communication (distance medicine), technologically-enabled programmes without interaction/data exchange between patients and health professionals (e.g. online health education, peer-to-peer support), monitoring of invasive/implanted devices or of electrode-mediated devices (e.g. cardiac telemetry), smartphone-mediated health applications (m-health) or telehealth for patients within hospital/clinic settings (e.g. teleradiology). The intervention and population examined in the WSD Telecare Trial served as a guide to defining telecare for the selection of studies. The intervention involved the remote, automatic and passive monitoring of lifestyle changes and emergencies in order to manage the risks of independent living; the population consisted of community-dwelling individuals with social care needs (Chapter 4, sections 4.1 and 4.2). I considered any first-, second- or third-generation telecare interventions to be within the scope of the review rather than limit the pool of evidence for consideration any further, given the anticipated sparsity of studies on telecare.

3.2.2 Search Strategy

Over the course of writing the thesis, I assembled evidence on the effectiveness and cost-effectiveness of telehealth and telecare from a combination of sources. Initial searches were conducted in April 2011 via EBSCOhost in CINAHL Plus with Full Text and in May 2011 in PubMed, using search concepts for telehealth and telecare, including keywords for telecare, remote monitoring, home telecare, telemedicine, teleconferencing, teleconsultation, telephone support and telephone monitoring. Searches in PubMed also included keywords for costs and effectiveness. I undertook further searches on these concepts in March 2014 via EBSCOhost simultaneously on several databases: CINAHL Plus with Full Text; EconLit; MEDLINE; PsycARTICLES; and PsycINFO. The keywords for telehealth and telecare that had been used in the April 2011 search were combined with the text string ‘systematic review’ in a search where results were limited to publication in 2008 (the first year of the trial) to 2014 and to peer-reviewed journals. Additional searches of the same databases were then carried out on combinations of keywords related to costs and telehealth trials and to trials of technologies related to telecare, including assistive technology and smart homes. Searches on combinations of keywords for costs, telehealth, telemonitoring and telecare were also carried out in Cochrane Library and Google Scholar with results limited to publication in 2008 and thereafter. Papers that were published after 2013 located in these searches were collected for

use in the discussion in Chapter 9. Results were compiled in EndNote and Mendeley databases. These compilations of references were frequently updated over the next three years through alerts from academic publishers based on the 'telehealth' and 'telecare' keywords, from intermittent searches on those keywords in Mendeley's literature catalogue, in PubMed and Web of Science from EndNote and in Google Scholar.

3.2.3 Selection Criteria

Having defined the scope of the telehealth and telecare activities to be considered, I also developed other criteria to choose whether to include papers for discussion in this chapter. In adopting further selection criteria, I necessarily balanced questions of relevance to the WSD trial populations with the availability of evidence. I included single studies where these were more recent than the last available systematic review. In the case of the telehealth literature, I focused on systematic reviews and single studies comparatively examining clinical outcomes (including service use and costs) in adult populations with one or more conditions related to the long-term conditions examined in the WSD Telehealth trial (COPD, diabetes and heart failure). I therefore excluded systematic reviews and studies that were focused exclusively on children, or on adult populations with other than diabetic, cardiac or respiratory conditions. I also excluded systematic reviews that only examined out-of-scope activities (e.g. web-based health education interventions, m-health, teleradiology, telemetry) and systematic reviews that included telemonitoring of invasive devices/implants. Single studies and systematic reviews of studies of experimental and quasi-experimental design were considered. Systematic reviews of economic evaluations of telehealth and telecare for any condition were considered.

3.2.4 Procedure

The quality of papers and systematic reviews was appraised in terms of the clarity of research aims and questions, the methods employed (e.g. selection criteria, study design and sample size; in the case of systematic reviews, evidence of having summarised and considered included studies, evidence of having considered the quality of studies in terms of study design and sample size) and the extent to which the conclusions reflected the evidence that had been presented. I summarised papers qualitatively, in more or less depth depending on my assessment of study quality, and grouped the results of the review into sections, separately presenting evidence on effectiveness, service use and costs, and cost-effectiveness.

3.3 Telehealth Effectiveness: Clinical and Health Related Outcomes

Reviews of the literature on telehealth have identified some weaknesses in the evidence base in relation to clinical and health-related quality of life (HRQoL) outcomes, but nevertheless there is a rapidly growing body of evidence on the subject, much of it based on studies carried out in the USA. Whereas a pre-2005 study of telemedicine publications judged good quality studies to be scarce (Hailey 2004), more recent systematic reviews have noted promising findings, particularly with regard to clinical, or surrogate outcomes, for some populations of telehealth users. Systematic reviews of telephone support (TS) and telemonitoring (TM) have been carried for a variety of populations: those with cardiac conditions, hypertension or congestive heart failure (Inglis et al. 2010), diabetes (Polisena, Tran, et al. 2009), respiratory disorders (Polisena et al. 2010a, Jaana, Paré, and Sicotte 2009) and depression (Garcia-Lizana and Munoz-Mayorga 2010) among others. Most have reviewed the literature on resource use and costs associated with the interventions. The following sub-sections cover the evidence base on effectiveness, in terms of clinical and health-related quality of life outcomes.

3.3.5 Cardiac Conditions

Evidence on the effectiveness of telephone-based technologies is perhaps most robust in the area of congestive heart failure (CHF). Polisena et al. (2010b) carried out a systematic review of telemonitoring studies involving adults and children with CHF, locating 11 randomised controlled trials (RCTs) and 10 observational studies (total N=3082). The review included RCTs and prospective observational studies and rated these for quality; meta-analyses were conducted on studies rated as of fair-to-good, good or high quality. A meta-analysis of mortality from all causes (data from 6 studies, N=1304) found that mortality was reduced in telemonitoring (TM) vs. usual care (UC) (the relative risk of death was 0.64, 95 per cent CI 0.48 to 0.85). 13 studies with quality ratings ranging from high to poor-to-fair reported various measures of HrQoL and satisfaction (condition specific and generic): seven found no differences between TM and UC and five reported better outcomes. The authors acknowledged the diversity of QOL and satisfaction measures used. They concluded that TM was "generally clinically effective" (Polisena et al. 2010b, p.75) but recommended more high-quality research into the clinical outcomes of TM in this population. The Inglis et al. (2010) systematic review of structured non-invasive home-based telephone support (TS) and

telemonitoring (TM) for CHF gave detailed descriptions of methods followed and of characteristics of included studies, presented risk of bias assessments and where possible carried out meta-analyses (fixed effects models). All included studies were RCTs. The meta-analyses indicated that the risk of all-cause mortality in patients with heart failure receiving either TS (pooled over 15 studies) or TM (pooled over 11 studies) was reduced, by 12 per cent (a non-significant positive effect) and 34 per cent (a significant positive effect) respectively, in comparison to those in usual care (UC). In terms of health-related quality of life (HRQoL) and satisfaction outcomes, six of nine TS studies reported improvements; of seven TM studies reporting these measures, three noted significant improvements. The reviewers concluded that both TS and TM interventions could improve quality of life.

There is also some evidence on other cardiac conditions, namely hypertension. A systematic review by Clark et al. (2010) examined nurse-led telephone monitoring (TM), community monitoring and clinics. The review included only RCTs, provided detailed summaries of all included studies, and assessed for risk of bias. The reviewers conducted a meta-analysis (applying a random-effects model) of three studies that they had rated as good-quality, finding that there were no significant differences between pooled TM treatment and usual care groups for either systolic blood pressure (BP) or diastolic BP. On the other hand, pooled data from three studies (one of which was rated by the reviewers as being of good quality) showed a significantly higher achievement of study BP targets with TM, with a relative risk of 1.24 (95 per cent CI 1.08 to 1.43).

Two trials reported in 2010 were not covered by these systematic reviews. Chaudhry et al. (2010) report a large-scale US-based RCT of patients with a recent hospital admission for heart failure randomised to "automated telemonitoring" or usual care (826 in TM and 827 in UC). The "automated telemonitoring" intervention involved daily calls into the system to complete automated questionnaires about general and HF-related health symptoms. TM did not reduce the risk of the combined 180-day endpoint of readmission (for any reason) or death (of any cause). In a US-based RCT (Datta et al. 2010), patients with a hypertension diagnosis using a Veterans Administration primary care clinic were randomised to a telephone-based nurse-administered patient education behavioural intervention to assist hypertensive patients to attain and maintain blood pressure control (N=294) or to usual care (N=294). Intervention patients received a telephone call from a nurse every 2 months for 24 months. Nurses provided educational and behavioural information, feedback on recent blood pressure values and medication and appointment reminders. Intervention patients experienced improvement in blood pressure control (from 40.1 per cent to 54.4 per cent); control patients

experienced a smaller improvement (from 38.2 per cent to 43.9 per cent), a non-significant difference between groups of 10.5 per cent (p=0.17).

3.3.6 Diabetes

Polisena et al. (2009) carried out a systematic review of the impact of home telehealth, either TM or TS, for people with diabetes. The authors clearly described review objectives and methods and provided details of characteristics of the 26 included studies and rated these for quality. Both RCTs and prospective observational studies were included. Data on HbA1c were meta-analysed: pooling results of twelve RCTs, the reviewers found that TM improved control of diabetes as measured by HbAc1[15]: the HbA1c levels in the TM group were significantly lower than in the UC group, with a weighted mean difference of -0.22 (95 per cent CI -0.35 to -0.08). The evidence on HRQoL and patient satisfaction was more mixed. Eleven studies measured HRQoL or patient satisfaction using a number of instruments. In four studies, telehealth (TM/TS) was better than usual care, for instance in terms of reliability and ease of use. No differences between groups were found in four telehealth studies in terms of satisfaction or HRQoL. The authors considered that participants in the TM and TS intervention groups had HRQoL outcomes that were similar to, or less good than, the UC groups. Limitations were acknowledged in terms of the number and variable quality of available studies for some outcomes. The authors concluded that, notwithstanding these limitations, home telehealth interventions were clinically effective.

3.3.7 COPD

Polisena et al. (2010a) carried out a systematic review of literature on the use of home-based TH for patients with chronic obstructive pulmonary disease (COPD). The reviewers included nine studies (RCTs and prospective observational designs), quality-rated the included studies and carried out meta-analyses (random-effects models). The authors examined papers reporting studies of both TM and TS. There were fewer studies involving the COPD population than were available for the systematic reviews of diabetes or cardiac conditions described above. The quality of the studies as rated in the review ranged from high to low; sample sizes ranged from 18 to 240. A meta-analysis found a higher mortality rate in TS patients compared to usual care (relative risk 1.21 (95 per cent CI 0.84, 1.75)), based on three

[15] This is a measure of the control of diabetes.

studies. The review listed four studies reporting quality of life and satisfaction, finding improved outcomes in the telehealth relative to the usual care groups in two studies and no difference between groups in outcomes in the other two. Having acknowledged the small number of available studies and variations in their quality, they concluded that "home telehealth is generally clinically effective" (Polisena et al. 2010a, p.127).

A systematic review by Jaana, Paré, and Sicotte (2009) identified 23 studies on telemonitoring (TM) of people with respiratory conditions (including COPD and asthma) where monitoring consisted of various technologies such as short messaging services, internet monitoring and electronic diaries. The reviewers presented characteristics of included studies and rated these for strength of evidence (using a rating system based on the type of study design): 13 studies were rated as being of fair to poor quality; most relied on small samples (ranging from 5 to 300) and 13 had no control group. Most of the clinical effects reported involved disease markers and respiratory symptoms rather than HrQOL. The reviewers summarised the psychological effects of TM as positive in most of the included studies. The review noted improvements in some clinical indicators such as disease control, to be interpreted with caution given small sample sizes, unsystematic and uncontrolled designs in the reviewed studies.

3.3.8 Reviews across Chronic Conditions

One systematic review (Barlow et al. 2007) surveyed the evidence on a broader range of telephone or internet-based interventions, categorised as 'vital signs monitoring' (TM), 'information and support' (TS) and 'safety and security' (telecare, as defined in section 2.1). The reviewers included RCTs and observational studies as a means of ensuring quality and did not formally score studies on quality. The authors found that the evidence on TM was inconsistent: three studies involving populations with diabetes reported improvements in clinical measures but in five studies, there was no difference in blood glucose control between groups. The authors concluded from an examination of the systems' outcomes across fifteen studies that the automated transmission of clinical readings was as effective as usual care. Studies of telephone information and support (TS) found improved clinical outcomes in those with depression (six studies), heart disease (three studies), diabetes (seven studies), asthma (one study), COPD (one study) and frail older people (one study). Another two studies found no improvements for people with diabetes clinically; and one found no improvement in HRQoL. The reviewers concluded that the evidence on the clinical impact of

vital signs monitoring was equivocal. They noted that while most studies they had reviewed were of randomised design, in many cases these were based on small samples.

A study by Pare et al. (2010) examined peer-reviewed articles reporting the results of studies involving people with a diagnosis of diabetes, asthma, heart failure or hypertension receiving TH or TM. 45 percent (28/62) of the studies were US-based. The review used a strength of evidence-rating system based on the type of study design. The authors found a trend towards better glycaemic control in patients with diabetes. There were improvements in peak expiratory flows, reductions in related symptoms and improvements in self-reported QOL in asthma patients. For patients with hypertension, there were reductions in systolic and/or diastolic blood pressure. However, the findings of studies of TH or TM in heart failure patients were equivocal; no reductions in mortality or hospitalisations were reported. The authors observed that in diabetes, asthma and hypertension patients, TM allowed more frequent follow-ups, enabling early detection of deterioration in patient health.

A large-scale trial of care coordination through case management and disease management (Brown et al. 2007) included a number of projects using home telemonitoring but concluded that "few programs had statistically detectable effects on patients' behavior or use of Medicare services" (Brown et al. 2007, p. xviii).

3.4 Telecare Effectiveness: Clinical and Health Related Outcomes

Whereas there is demonstrably a mounting evidence base from controlled trials and observational studies for telehealth applications, the same cannot be said for telecare. It was difficult to find empirical literature on the relationship between the impacts (for users) and characteristics of telecare systems. Reeder et al. (2013) make a useful point about the evidence base for 'health smart homes' (HSH) and 'home-based consumer health technologies' (HCH) (terms which encompass second-generation telecare):

> HSH/HCH research has been conducted in both health services and technology disciplines and scientific findings have been published in different literature repositories that do not always overlap in their indexing […]This fragmentation of reported evidence represents a knowledge gap concerning what research has been done and communication barriers for knowledge translation to relevant stakeholders. (Reeder et al. 2013, p.566-7)

In the Barlow et al. (2007) review, the authors identified no RCTs or observational studies on the monitoring of safety and security reporting individual level outcomes that met their quality criteria; the authors concluded that the evidence base for the effectiveness of home

safety systems (telecare) was insufficient. One systematic review of smart home technologies, including telecare (Martin et al. 2008), found *no* studies of sufficient quality for inclusion. The reviewers considered studies of social alarms, electronic assistive devices, telecare platforms, environmental control systems, automated home environments and "ubiquitous" homes. Of 62 papers they excluded fourteen as actually being about telemedicine; most of the papers were only discussions or editorials.

One of the papers retrieved but excluded from the Martin et al. (2008) review (by Vincent et al. (2006)) does address the effectiveness of a home telesurveillance scheme in older people and is worth noting, given the dearth of relevant studies, although the study sample size was small (n=38), had an uncontrolled before-after design and measured outcomes over a short period, of six months. The intervention equipment consisted of a big-button telephone and call transmitter with emergency button worn as pendant/bracelet. The telephone had a programmable voice reminder feature. The monitoring was conducted by a telesurveillance call centre operated by nurses rather than unqualified call operators. The authors found that there was no significant improvement in HRQoL (SF-12) after six months of using the service.

A systematic review of health smart homes (HSH) and home-based consumer health (HCH) technologies to support ageing in place by Reeder et al. (2013) covered the literature from 1980 to 2011. Searches were conducted in databases of health care and also informatics publications. The search uncovered 31 publications from Europe, North America and Asia. The review covered technologies to support older people (60 years and older) in residential settings that supported, or prevented threats to, independence and collected data for monitoring health or communication. HSH were residential settings with embedded technologies for passive monitoring; HCH were health technologies used by older people in their homes. Strength of evidence was assessed by classification into four categories covering stage of technological readiness (from validity and feasibility testing through to larger-scale evaluation and implementation stages), sample size and study design. The review uncovered only 3 studies that were evaluations on a larger scale than 10 people, all of which had methodological weakness such as unequal dropout between comparator groups, non-randomised comparators or historical controls. The authors cited a paper by Tomita et al. (2007) reporting an RCT which found that a smart home technology intervention group (N=46) maintained their physical and cognitive status while controls (N=67) declined in these measures over 2 years. Almost half of intervention participants chose not to use some of the smart home technology, because of problems with usability. The reviewers identified a

study by Brownsell, Blackburn, and Hawley (2008), which compared participants receiving a mix of second and third generation telecare technologies (N=24) to non-equivalent controls in similar housing (N=28). Intervention participants spent more time outside the home than controls and felt safer during the day and night than controls. The reviewers cited a paper by Kelly (2005) (a historical controlled study, N=1700) which reported that a home safety (second-generation) telecare package reduced hospital admissions, hospital lengths of stay and nursing home lengths of stay.

A systematic review of telecare outcomes for carers (Davies, Rixon, and Newman 2013) identified very few papers reporting relevant quantitative analyses. Only seven studies fulfilled the criteria for inclusion, of which just two had been peer reviewed, the others being unpublished work or reports. Telecare was defined as "the continuous, automatic and remote monitoring of real-time emergencies and lifestyle changes over time in order to manage the risks associated with independent living"(Davies, Rixon, and Newman 2013 p.584, citing Telecare Aware website (accessed 14th May 2013)). The included evaluations were from US, UK and Norway. Four studies had sample sizes of less than 30, others had sample sizes of between 100 and 300. The equipment used in evaluations were mostly passive sensors to monitor activity, e.g. bed sensors; or response to emergencies, e.g. flood detectors; or assistive devices (stand-alone), e.g. calendar clocks. Davies, Rixon, and Newman (2013) find some evidence of positive impacts on carer well-being in terms of reductions in stress and strain but not QOL, burden or impact in terms of time; however their conclusions were tentative as all the studies reported in the review were methodologically weak.

Despite the prevalence of pendant alarms or PERS, there is very little hard evidence on the outcomes of these systems (De san Miguel 2017). de Miguel Diez et al. (2008) reported the results of a retrospective survey of 1476 users of a first-generation telecare service (PERS) in Western Australia. Respondents reported positive impacts such as greater sense of security, being less anxious about having a fall and more confidence in carrying out ADLs. Lee et al. (2008) randomised older people 70 years and over discharged from the emergency department after a fall to PERS (N=43) or standard discharge planning (N=43). Outcomes measured were fear of falling and anxiety, emergency department visits, hospital admissions and lengths of stay. The study found no differences between PERS and standard discharge planning groups after 60 days in terms of reductions in anxiety or fear of falling; nor in subsequent visits to the emergency department, numbers of admissions or lengths of stay.

There is some evidence on the effects of more advanced versions of telecare. A recent systematic review of "lifestyle monitoring" found just 4 papers reporting trials with over 20 participants, and 21 papers on trials involving fewer than 20 participants (Brownsell et al. 2011). The authors concluded that, given the small number of papers available on lifestyle monitoring, much remains to be understood on how such systems can be made effective. A systematic review of smart home projects (Demiris and Hensel 2008) (see also Chapter 2 on smart homes) located 114 publications reporting 21 projects from Europe, the US and Asia. The authors were not able to locate evidence on health outcomes or impacts on delaying admission to nursing homes. Most studies were at the stage of examining feasibility issues or had very limited sample sizes.

3.5 Telehealth- and Telecare-Related Use of Health and Social Services Resources, Costs and Cost-Effectiveness

It is important to consider whether AATs represent a useful and effective route for delivering health care. It is equally important to investigate whether AATs represent the best use of the available public funds. A relatively small proportion of telehealth evaluations have considered the relationship between the outcomes of the interventions and the costs associated with implementing the interventions (Whitten et al. 2002, Bensink, Hailey, and Wootton 2006). Bensink, Hailey, and Wootton (2006) found that less than a fifth of published studies on home telehealth gave economic data "judged sufficient for economic strength of evidence evaluation" (pp. 12-13). The information on the costs and cost-effectiveness of interventions featuring home telehealth is also scarcer than that on their effectiveness (Barlow et al. 2007). On the other hand, a number of recent systematic reviews were able to locate data on health service use, particularly in terms of either numbers of hospitalisations or numbers of bed-days. The evidence on the costs and cost-effectiveness of telehealth and telecare are reviewed in the following sections.

3.6 Telehealth: Resource Use and Costs

3.6.1 Cardiac Conditions

The Inglis et al. (2010) review of TS and TM for heart failure conditions identified a number of studies reporting hospitalisations and lengths of stay, but a smaller number of studies reporting the costs of the intervention and/direct service costs. The risk of all-cause hospitalisation in TS (pooled over 11 studies) significantly reduced risk of all-cause

hospitalisation by 10 per cent; the risk was non-significantly reduced by 6 per cent by TM (pooled over 8 studies). CHF-related hospitalisations were reduced significantly by 23 per cent in TS (pooled over 13 studies); and also in the TM studies, by 24 per cent (pooled over 4 studies). There was less evidence on length of stay (LOS): of six TS studies, only one study reported a significant reduction in LOS; in one TM study there was large decrease in the number of days but in another the trend towards a shorter stay did not reach significance. Twelve studies of TS or TM included information on health service costs. Nine reported cost reductions in hospital service use; three studies reported no reductions or an increase in health service costs. The reviewers reported that where studies identified decreases in resource use and costs, the range of savings was 35 per cent to 86 per cent. Five studies involving TS provided details on reductions in service costs. Four of the five were US-based studies. Inglis et al. (2010) report cost savings per patient expressed in a variety of ways in different studies: for instance as savings on inpatient care and also in terms of cost per QALY gain. Figures on reductions in service costs were given in fewer TM than TS studies (in terms of costs of readmissions, hospital care, medications).

Four studies involving TS provided details on the intervention costs according to Inglis et al (2010): these were also within a wide range and came from a variety of different health systems and countries. The review gave figures from US-based studies of between US $23.60/patient (n=34) (Barth 2001, cited by Inglis et al. (2010)) and US $2177 per patient (n=406) (Hebert, Sisk et al 2008, cited by Inglis et al. (2010)) Of European studies covered by the review, one industry-funded Dutch study (Balk et al. 2007, Balk et al. 2008) (n=214) noted that costs of TM (MOTIVA) for intervention patients increased the total costs for the intervention group, but did not give the actual cost of the intervention; and Giordano et al. (2009) gave a mean annual cost per patient for TM of 185 EUR +/-39 EUR.

In Polisena et al. (2010b), a meta-analysis of patients hospitalised (N=891 in 4 studies) found lower numbers hospitalised in TM than UC (a relative risk of 0.77 (95 per cent CI 0.65, 0.90)). Emergency department (ED) visits (in 8 studies) and all-cause bed days of care (in 4 studies) were decreased in the TM relative to the UC groups. While two studies found outpatient and primary care visits had increased in TM vs. UC, two others found outpatient visits reduced in TM vs. UC. The authors concluded that evidence on health care utilisation in TM was less than was available on clinical outcomes.

A study by Chaudhry et al. (2010) (RCT, n=1653) found that automated TM did not reduce the risk of hospitalisation for heart failure and there were no differences between groups in number of days in hospital. The authors noted that 14 per cent of the TM group

never used the automated system and only 55 per cent were still using the system at least three times weekly by the end of the study period.

Datta et al. (2010) (RCT, n=588) describe the service use and costs of a telephone-based nurse-administered patient education behavioural intervention for hypertensive patients. Over 2 years, 27.5 per cent of TS patients had 162 admissions vs. 25.2 per cent of control patients who had 150 admissions. The mean length of stay in TS was 9.57 days vs. 9.72 in UC. TS patients had fewer primary care clinic visits than control patients (4.2 vs. 7.5). The groups did not differ in terms of overall costs (including inpatient, outpatient and primary care services) over 2 years. The mean annual intervention cost was estimated at $112, constituting about 3 per cent of total costs. The same authors reported a cost-effectiveness analysis using a decision model: ICERs generated by this model were between $42,457 to $87,300 per life-year saved (for women and men of normal weight respectively). TS was described as "potentially cost-effective" (Datta et al. 2010, p.262).

3.6.2 Diabetes

The Polisena et al. (2009) systematic review of home telehealth for diabetes identified some studies reporting on health service use outcomes. There was evidence of significant reductions in proportions hospitalised in TM vs. UC groups in two studies. There was a significant reduction in hospitalisations in TM vs. UC groups in one study and significant reductions in bed days in three, although these results were limited in that no measures of variation were reported. In terms of visits to emergency departments, two observational studies reported contradictory results, one with significant reductions in the TM relative to the UC group and one with the opposite finding. There were no significant results related to proportions of hospitalisations or ED visits in the case of TS. The review's authors thus found some evidence of reductions in hospitalisations and bed days; evidence was limited to one or two studies for some types of clinic and primary care use. The authors concluded that, although impacts on the use of health services had a limited evidence base, home telehealth showed "great potential in some studies" (Polisena, Coyle, et al. 2009, p. 928).

A US-based RCT (Moreno et al. 2009) examined the costs to Medicare of introducing a computer-based monitoring system for patients with Type II diabetes, in tandem with nurse case management and guidelines-based recommendations to patients' primary care physicians. The study recruited Medicare beneficiaries in medically underserved areas of New York State (US). The study found that costs of TM participants (N=825 in a first cohort

and N=243 in a second) were 71 to 116 per cent greater than those of controls (N=800 in the first cohort and 248 in the second). The annual costs of the intervention itself were well over USD $8000 per patient ($8924 to $8437 depending on the enrolled cohort), which were described as “excessive”.

3.6.3 COPD

Polisena et al. (2010a) reported in their systematic review that home telehealth had reduced hospital admission rates (in one study) and numbers of hospitalisations (in eight studies) and visits to emergency departments (in four studies), while evidence on the impact of TH on the use of other health care services was more limited. The impact on use of hospital bed days was quite varied between the six studies reporting this outcome, particularly in regard to TM: higher than in UC in one study, and lower in two others. The reviewers concluded that overall there were limitations to the evidence on health service use in home telehealth. The systematic review by Jaana, Paré, and Sicotte (2009) found no consistent evidence that telemonitoring for respiratory conditions reduced health care utilisation (visits to primary care or emergency department, hospitalisations and lengths of stay or visits).

3.6.4 Reviews across Chronic Conditions

The Barlow et al. (2007) review suggested that automated vital signs monitoring could decrease utilisation of health services in the case of COPD and CHF, based on the evidence of 11 trials, but that the evidence was more mixed in the case of diabetes.

3.7 Telehealth: Economic Evaluations

Systematic reviews have been carried out on the economic evaluation evidence for telehealth (Bergmo 2009, Polisena, Coyle, et al. 2009, Vergara Rojas and Gagnon 2008). Polisena et al. (2009) identified 22 studies on economic evaluations of home telehealth on populations with CHF, COPD and diabetes. The review included comparative economic evaluation designs (cost-effectiveness, cost-utility, cost-minimisation and cost-benefit studies; costs-analyses where the intervention was assumed to be as effective as the alternative). Quality was assessed using a checklist for economic evaluation, adapted for examining telehealth studies; study interventions and comparators and study costs and consequences were summarised in some detail. The majority of these economic evaluations (14) were based on RCTs, four on case-control studies, four on pre-post study data; twenty-one were cost analyses, one, a cost

utility analysis (CUA). All but two of the studies found that home telehealth led to reduced costs of health care, from a system or an insurance provider perspective. The authors caution that although most studies found home telehealth to be cost-saving, conclusions drawn on cost-effectiveness "must be qualified as the quality of the studies in terms of economic evaluations was poor" (Polisena, Coyle, et al. 2009, p. 347). They found that the relatively few economic evaluations of telehealth for chronic disease management were mostly of poor quality; most failed to address perspective, use marginal analysis or carry out sensitivity analyses. The studies were characterised by small sample sizes, and lacked information on patient characteristics as well as clinical outcomes.

Vergara Rojas and Gagnon (2008), in a review of cost-effectiveness indicators used in telehomecare (telehealth), report similar findings: the great majority of the 23 economic analyses they had identified found telehomecare to be cost-effective. However they also caution that the studies were "far from providing the basis to make a good decision" (Vergara Rojas and Gagnon 2008, p. 902), finding flaws in the methods employed and the interpretation of results in many of the papers. The authors acknowledged a limitation of their review in that they did not assess the quality of the papers reviewed.

Bergmo (2009) set out to review telehealth-related economic evaluations in terms of their quality and validity, finding 33 economic evaluations covering a number of specialties, including six on diabetes, six on cardiology, and a range of others in areas such as dermatology and psychiatry. Studies comparing full economic evaluations (excluding costs analyses) were included; characteristics and results of all included studies were briefly summarised. Among the evaluations were thirteen RCTs, two case control studies, three before and after studies, two crossover trials and six decision modelling studies using secondary data. There were five cost utility analyses, the rest being cost-effectiveness analyses. Twenty-one evaluation papers did not report perspective, but where reported, the perspective was more often the health provider (eight studies) than societal (two studies), and two studies combined health provider and patient perspectives. A variety of effectiveness measures had been used, including process outcomes such as diagnostic accuracy, and surrogate measures of outcome, e.g. blood glucose. Seven studies employed the SF36 or EQ5D, and two used condition-specific HRQoL instruments. Four reported QALY gain. However, only two studies of TM were based on RCTs: one (Mason et al. 2006) (for diabetes) reported a cost-per-QALY of £43 400; another (for asthma) (Willems, Joore, Hendriks, Wouters, et al. 2007) reported a cost-per-QALY of €31 035 per QALY gained for

adults (the other studies being before-after or decision models) (Willems, Joore, Hendriks, van Duurling, et al. 2007).

In terms of the methods used to calculate costs, all studies had calculated direct health care costs, including investment, installation, call costs, personnel costs and other health costs. Fewer (eight studies) reported travel costs for personnel while eleven estimated patients' travel costs. However five studies gave little to no cost information, while the author found that in half the studies, the methods for costing were unclear. Less than half gave details of resources used in their physical units and reported the unit costs or prices they had used to value the resources. Three studies calculated marginal costs. In terms of estimation of uncertainty of the reported costs, 23 studies had calculated confidence intervals for point estimates, three employing non-parametric bootstrapping techniques. Only five reported incremental cost-effectiveness ratios (ICER). Less than half reported sensitivity analyses. This review prompted the author to conclude that the "evidence base for telemedicine decisions is alarmingly scarce" (Bergmo 2009, p.6) and that "few economic evaluations can be trusted to provide reliable information for decision making" (Bergmo 2009, p. 8).

In summary, there were several problems common to most published economic evaluations of telehealth: sample sizes were generally small; most evaluations failed to state perspective; which costs have been considered and included were not presented transparently; there was a lack of information on patient characteristics; and few conducted marginal analysis (Bergmo 2009, Polisena, Coyle, et al. 2009, Vergara Rojas and Gagnon 2008). Also, costs of older evaluations might not reflect current conditions in a fast-changing market with the potential for rapid price decreases. There was a need for those carrying out economic evaluations in this field to address issues of local variation, use more diverse populations to boost external validity, use a standardised approach, such as an explicit economic evaluation framework, to include all relevant costs and to be clear about inclusions and exclusions (Bergmo 2009, Polisena, Coyle, et al. 2009). Another issue for policymakers in the UK is how generalizable are the widely varying estimates of intervention costs and cost savings given in the literature to this health and social care system, particularly as many of these studies have been conducted in the US (Vergara Rojas and Gagnon 2008).

3.8 Telecare: Resource Use, Costs and Cost-effectiveness

If there were concerns about the quality of the economic evidence for telehealth, then they applied doubly to telecare, for which virtually no good quality studies of its impacts exist.

The more general question of telecare's efficacy was also not well supported by the evidence base. In contrast to the bourgeoning of evaluation literature on telehealth applications over recent years, where more randomised controlled trials and observational studies had been conducted, the same could not be said for telecare.

While the quantity of small-scale evaluations is growing, the quality was generally poor. In their 2007 review, Barlow et al. (2007) identified no RCTs or observational studies on the monitoring of safety and security reporting individual level outcomes that met their quality criteria. The review identified just two observational studies of 'safety and security', or telecare, interventions at 'systems level', one of which was case-controlled. That study (Woolham 2005) found that a home alert system might help people to stay at home and improved function in the 'intervention' group. The other observational study compared 170 people in care homes to the same number of people living in their own homes, finding that telecare was associated with fewer hospital admissions and improvement in discharge rates and cost savings (West Lothian Council 2004). Vincent et al. (2006) found that the provision of a telesurveillance call centre operated by nurses decreased hospitalisations (from twelve to ten admissions, in a three month period per client) and use of home care services (from eighteen to ten visits, on average in a three month period) at the end of the six month study.

There was a systematic review with some findings relevant to telecare by Graybill, McMeekin, and Wildman (2014), covering literature up to July 2012. The review covered economic analyses of assisted living technologies facilitating ageing in place (defined as home and environmental adaptations and/or telemedicine). The study populations were home-dwelling people 65 years and older with complex co-morbid conditions or functional limitations. Quality was assessed using an economic evaluation checklist; intervention characteristics and quality ratings were summarised. The reviewers searched two economic databases (NHS EED and HEED) for relevant studies and located eight. One concerned (low-tech) assistive devices (e.g. equipment for daily living, wheelchairs); seven others involved telemedicine or tele-rehabilitation interventions. The Vincent et al. (2006) study (also reviewed by Barlow et al. (2007)) was reviewed and classified as telemedicine. All eight studies were assessed to be of low methodological quality. Thus it seems that up to 2012, no economic evaluations of telecare (of any generation) were available within these economic databases. While most of the interventions reported in the studies covered by this review were not similar to telecare, the nature of the intervention described by Vincent et al. (2006) appeared as related to telecare as to telemedicine, in providing a safety surveillance system, even if delivered by health rather than social care workers. The authors cite Bowes and

McColgan (2006) (not included in the review) as an example of the gap in research evidence on costs of ALTs for ageing in place, because that evaluation did not consider the cost of the (telecare) intervention. Bowes and McColgan (2006) themselves presented costs and outcomes results from their observational study but conducted no formal cost-effectiveness analysis. Nonetheless they concluded that the evidence base for the cost-effectiveness of home safety systems was insufficient.

A curious feature of the telecare literature is that there appeared to be as much written about the potential for cost savings as on the effectiveness or cost-effectiveness of the intervention. In particular, a reduction in service utilisation was discussed in this literature as a measure of a satisfactory outcome. Yet there were few peer-reviewed publications that provided information on not just cost savings but such basic details as the typical composition of a telecare package, the cost of the equipment and of supporting the monitoring service and the range of support services available to respond to sensor activations. The cost of a "home safety and security package" in England, based on information from telecare pilots in England, was estimated in 2005 to be £360 (Department of Health 2005c),while the cost of monitoring was estimated to be about £5 per week. Uplifted to 2010 prices, the cost of equipment and monitoring might cost approximately £735 in the first year, or using an annual equivalent cost for the equipment, it might cost £403 per year, or about £7.70 per week. A report on the West Lothian telecare programme (Bowes and McColgan 2006) gives details of a cost study based on a sample of 57 older people receiving telecare in both home and new-build sheltered care settings. This estimated the costs of telecare for people in their own homes ("Opening Doors Dispersed") at £7 per week. These costs included the weekly cost of the technology, taking battery replacement and depreciation into account, and the support costs of the monitoring centre. More recently, the cost of a package of telecare in a Welsh local authority, providing a response as well as a monitoring service, was estimated at approximately £9 per week (including equipment), with revenue costs of £5.30 per client per week (Bayer and Barlow 2010). There was very little empirical information on the cost-effectiveness of lifestyle monitoring versus standard or usual care (Brownsell et al. 2011). Reeder et al. (2013) report that the costs of the smart home intervention in Tomita et al. (2007, cited in Reeder et al. (2013)) were less than €304; also that a small-scale study by Mahoney, Mahoney, and Liss (2009, cited in Reeder et al. (2013)) of nine family members of older people found that five were willing to pay €45 and four were willing to pay €23 per month for smart home technology.

3.9 Conclusion

In this chapter I gave an overview of the literature on outcomes, service use, costs and cost-effectiveness associated with telehealth (including telephone support and telemonitoring) and telecare (including smart homes, first-generation/PERS and third-generation telecare). I focused on the evidence available prior to the publication of the evidence emerging from the WSD research programme in mid-2012. The evidence for clinical outcomes of telehealth across the specific conditions of HF, COPD and diabetes, and across chronic conditions, was in general promising or at least suggested some equivalence between telemonitoring, telephone support and conventional medical management. On the other hand, telehealth economic evaluations were of variable quality and many did not adhere to standard economic evaluation checklists. It was evident that the quality and quantity of literature on telehealth differed from that of telecare. Much of the telecare literature reported small-scale investigations of weak methodological quality. The evidence base was somewhat fragmented between health indexing and informatics databases.

Chapter 4

Methods

In this chapter I first describe the context within which my research took place. I set the scene by giving an overview of the methodologies employed in the Whole Systems Demonstrator trials. The trial and accompanying evaluation required the endeavours of a large number of people, and it is not within the scope of the thesis to describe their work in great detail; however it would be impossible to understand the context of the economic evaluation without knowing something of the trial design. I describe in greater detail the methods used to carry out the economic evaluation.

4.1 The Whole Systems Demonstrator Evaluation: an Overview

The Whole Systems Demonstrator (WSD) evaluation examined two telemonitoring technologies, telehealth and telecare (Bower et al. 2011). The evaluation sought to assess the outcomes and costs of the technologies in the context of integrated health and social care and support. Evaluators used a range of quantitative and qualitative methods to investigate outcomes. There were two pragmatic cluster-randomised controlled trials investigating the impact of the interventions in two populations:

1. Telehealth: individuals with an index long-term condition (chronic obstructive pulmonary disease, heart failure or diabetes)
2. Telecare: individuals with social care needs

Alongside the quantitative research, qualitative studies examined the experiences of professionals involved in implementing the technologies (MacNeill et al. 2014) and of trial participants and their carers (Sanders et al. 2012), and an ethnographic study examined the organisational challenges to mainstreaming the technologies (Hendy et al. 2012).

The main trials drew on administrative data to address the question of effectiveness on utilisation and costs to the health and social care systems (Steventon et al. 2013, Steventon et al. 2012). Two questionnaire sub-studies, involving about half of the trial participant population, collected data from participants on self-reported outcomes and use of health and social care services (Cartwright et al. 2013, Henderson et al. 2013, Bower et al. 2011).

The WSD pilots operated in three English local authority areas, geographies covered by four Primary Care Trusts. The three sites were chosen to exemplify continuing

engagement in ‘whole-systems redesign’ between health and social care. The technologies were implemented with the objective of supporting integrated care services and processes (Bower et al. 2011).

The overall design of the WSD evaluation was devised by the principal investigator (Professor Stanton Newman) and other lead investigators of the study (Dr Jennifer Dixon, Professor Raymond Fitzpatrick, Professor Martin Knapp, Professor Anne Rogers and Professor James Barlow, Dr Peter Bower, Dr Helen Doll). I undertook my doctoral research in the context of this study. While the overall design of the evaluation was not part of my doctoral work, it is important background and its description will help to understand the origins, strengths and limitations of the data analysed in subsequent chapters.

4.2 Trial Interventions

The WSD trial employed the following definitions of telehealth and telecare:

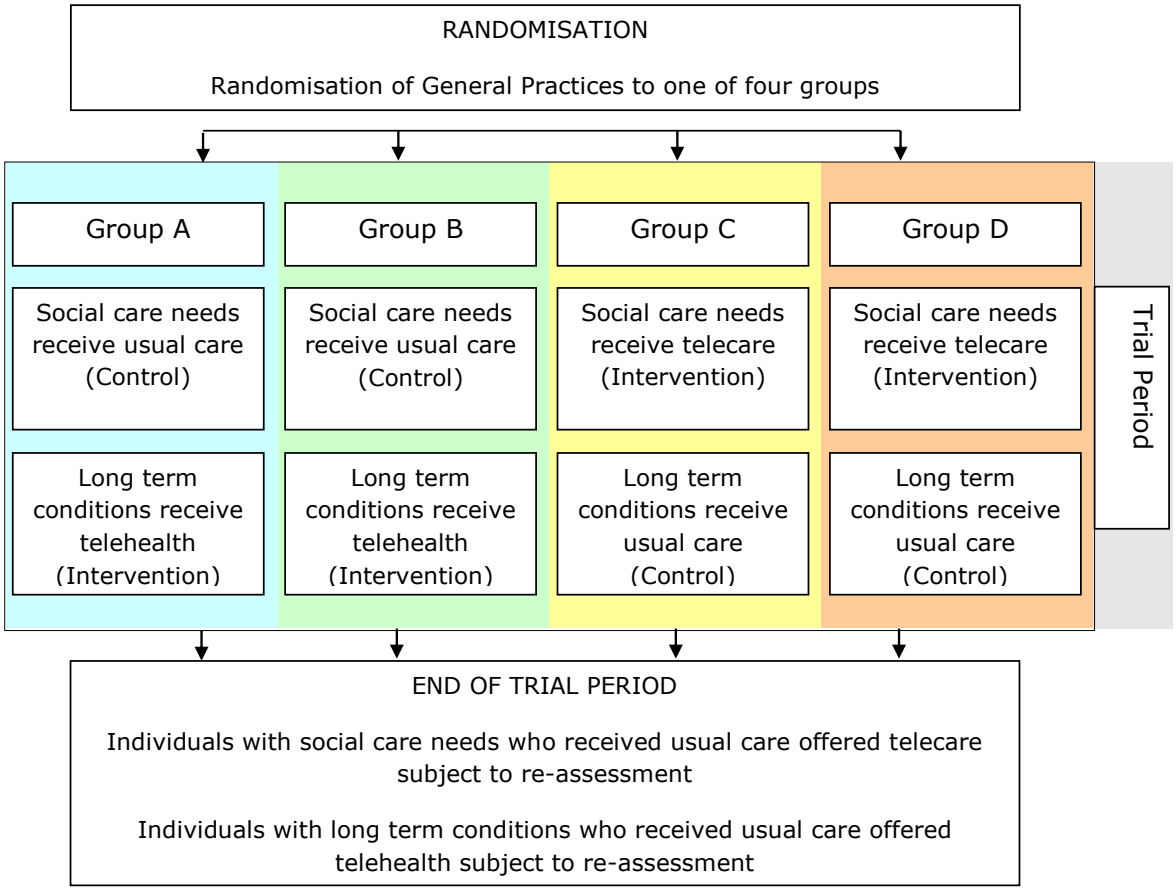
Telehealth (TH): "the remote exchange of data between a patient and health care professional to assist in the diagnosis and management of a health care condition. Examples include blood pressure and blood glucose monitoring” (Bower et al. 2011 p. 2).

Telecare (TC): “Telecare is the remote, automatic and passive monitoring of changes in an individual’s condition or lifestyle (including emergencies) in order to manage the risks of independent living. Examples include movement sensors, falls sensors, and bed/chair occupancy sensors. These technologies are generally provided to patients with social care needs” (Bower et al. 2011 p. 2).

4.3 Trial Sample Size, Randomisation Procedures

The quantitative study design was a pragmatic cluster-randomised controlled trial. The level of randomisation was at the general practice level (Figure 4.1). Practices were allocated to either the telecare or telehealth intervention, and acted as a control for the other intervention, so that those allocated to the telehealth control would be allocated to the telecare intervention and vice versa. In this way all participating practices had access to one of the two technologies.

Figure 4.1 Cluster randomisation design (Newman and Whole System Demonstrator Programme Evaluation Team 2014)



The telehealth trial sample size of 3000 was determined so as to power the study to detect a relative change of 17.5 per cent in the proportion of participants admitted to hospital at 12 months, from 25 per cent at baseline (80 percent (1-Type II error) power and two-sided p-value < 0.05 (type 1 error)). A minimisation procedure was put in place in order to allocate general practice clusters to the telehealth or telecare intervention or control, while balancing characteristics that might be associated with outcomes of the intervention across the trial arms. The characteristics considered were: size of practice (small/medium/large), prevalence of each of the index long-term conditions (low/medium/high), proportion of white/non-white patients (low/medium/high), Index of Multiple Deprivation 2007 (Noble et al. 2008) (low/medium/high) and site.

4.4 Trial Eligibility Criteria

All general practices within participating sites were eligible to join the trials.

Telehealth trial: Potential participants were deemed eligible for inclusion in the TH trial if they met any of the following criteria: they were included in a relevant QOF (Quality

Outcomes Framework) register; they had a confirmed diagnosis in either secondary or primary care health records (by ICD 10 or GP Read code); a local clinician (e.g. hospital or primary care medical practitioner or community matron) confirmed their disease status.

Telecare trial: Potential participants were eligible for inclusion in the TC trial if they met one or more of the following criteria: receiving or considered to need night sitting; receiving one or more days of day care or 10 or more hours of home care per week; having mobility difficulties; having falls or considered at high risk of falling; having a live-in or nearby carer facing difficulty providing support; having cognitive impairment, with a live-in or nearby carer (Bower et al. 2011). Potential participants were not excluded for having such basic forms of telecare as pendant or other community alarms (i.e. alarms that do not remotely collect and automatically send data to monitoring centres), or having items that were not part of a telecare package (e.g. smoke or carbon monoxide detectors).

4.5 Cluster and Participant Selection and Recruitment

General practices in each site were invited to join the trial. Once the practices had consented, a process of identification of potential TH and TC participants began. General practice registers were used to identify potential TH participants. For the TC trial, Social Services records were used to identify potential participants. Identified individuals were sent letters requesting their initial consent to share data with the WSD research team. Consenting individuals were invited to join the trial. Members of the sites' project teams then made a 'light-touch' visit to check eligibility for the trial. They also made initial checks on the home environment to assess suitability for TH/TC equipment, provided information on the trial and took informed consent to participate in the trial and also the nested questionnaire study (described below). People with cognitive impairments were eligible to take part in the TH trial, as long as family or friends were able to assist them with operating the TH equipment. Because the TH systems had an interactional element (for instance users had to respond to short questions about their health), the project teams also checked that potential users had the English language literacy required (Bower et al. 2011). People who wished to participate were visited again to receive a TH/TC needs assessment.

4.6 Trial Data Collection Procedures

The data collected and analysed for the WSD telecare and telehealth trials was extracted from routine data sources and included hospital and primary care service use, mortality, and social

services such as residential and nursing home and domiciliary care (Steventon et al. 2013, Steventon et al. 2012). A small set of participant characteristics at baseline were derived from administrative sources. Index of Multiple Deprivation 2007 scores were assembled by the trial team from participants' postcodes for use in analyses carried out within both the trial and questionnaire studies (see below). In addition, an index of the number of comorbidities was assembled based on a count of conditions diagnosed in the hospital episode statistics over three years prior to the trial. Reasons for withdrawal from the trial were also recorded by the trial team.

4.7 Questionnaire Studies

The questionnaire studies were nested within the parent TH and TC trials. The necessary total sample sizes for both questionnaire studies were powered to detect a small effect size of 0.3, given an intra-cluster correlation coefficient (ICC) of 0.05, with power of 80 per cent and $p<0.05$ (Newman and Whole System Demonstrator Programme Evaluation Team 2014). A required sample size of 550 telecare participants and 550 telehealth participants was estimated by inflating the planned numbers by 10 per cent to allow for possible increases related to varying-sized clusters. However, in order to reach the power needed for planned secondary LTC-specific subgroup analyses in each of the three long-term conditions, the Telehealth Questionnaire study aimed to attain a total sample size of 1650 telehealth participants (Newman and Whole System Demonstrator Programme Evaluation Team 2014). In addition the evaluation programme included a smaller-scale carers' questionnaire study, which will be referred to herein as the 'carers' study'. The following sections report the methods employed within the questionnaire studies.

4.7.1 Questionnaire Studies: Participant Eligibility, Selection and Recruitment

Individuals who had been identified as eligible to take part in the telehealth and telecare trials were invited to participate in the nested questionnaire studies. Those assenting were subsequently contacted by trained interviewers from a market research company, who visited to take written consent for the study and administer the study instruments (Newman and Whole System Demonstrator Programme Evaluation Team 2014, Bower et al. 2011). People who had been assessed as having cognitive impairments prohibiting them from completing the outcome measures on their own were ineligible for the questionnaire studies, but eligible for the parent trials (Cartwright et al. 2013, Hirani et al. 2013).

4.7.2 Questionnaire Studies: Data Collection

Questionnaire packs containing the Client Service Receipt Inventory (see 4.8.1) and other study instruments were administered by interview at baseline and posted to participants at 4- and 12-month follow-up. Participants who had not returned their questionnaire at 12 months were contacted to arrange an interview. 57 per cent of TH study questionnaires and 52 per cent of TC study questionnaires returned were completed by interview. Socio-demographic information about participants' characteristics was collected at baseline and covered age, sex, educational attainment and ethnicity. Housing tenure and household living arrangements were covered by the CSRI and therefore collected at all three time points.

4.8 Outcomes

The choice of outcomes and instruments to measure outcomes was made by evaluators of both the outcomes and cost-effectiveness research streams, in planning the questionnaire study. While I contributed to the choice of instruments used for the economic evaluation, the process of adopting the instruments for use in the study was made by the wider evaluation team.

Outcomes considered in the cost-effectiveness analyses reported in the thesis are presented in Section 4.16. The effectiveness analyses of the data from the Telehealth and Telecare questionnaire studies (Cartwright et al. 2013, Hirani et al. 2013) examined health-related quality of life, anxiety and depressive symptoms.

4.8.1 Outcome Measures – Instruments

Both generic and condition-specific health outcomes for participants were measured in the studies. Generic health-related quality of life (HRQoL) measures included the EQ-5D (Brooks 1996), the SF-12 (Jenkinson et al. 1997) and the ICECAP-O (Coast, Flynn, Natarajan, et al. 2008). Psychological outcomes (depression and anxiety), self-care behaviours, self-efficacy and social networks were measured, as well as long-term condition-specific quality of life measures (in the telehealth study only). Further information on all generic and condition-specific outcome measures used in the questionnaire studies can be found in other publications (Cartwright et al. 2013, Hirani et al. 2013, Newman and Whole System Demonstrator Programme Evaluation Team 2014). The following instruments were used to measure the outcomes examined in the cost-effectiveness analyses.

EQ-5D-3L (EuroQol Group 1990): The EQ-5D-3L is a generic preference-based measure of health (Brazier 2007). This 6-item instrument consists of the EQ-5D-3L descriptive system, which covers 5 dimensions of health-related quality of life (mobility, self-care, usual activities, pain/discomfort, anxiety), and a Visual Analogue Scale on which participants rate their health at the current time. Each item of the descriptive system in this original version of the EQ-5D-3L has three levels (no problem, moderate/some problems, severe/unable to perform). The system can be used to create a utility score, a single index value for health status. The index (the York A1 tariff) was derived using societal weights: to create the weights, 42 health states were valued by a representative sample of the UK population using the time-trade off technique and a statistical model created to estimate valuations of all 243 possible health states (Dolan et al. 1995, Brooks 1996, Dolan 1997). The instrument is suitable for use with older populations (see (Haywood, Garratt, and Fitzpatrick 2005, Hawton et al. 2011)). The NICE 'reference case' specifies the EQ-5D-3L as the preferred measure of health-related quality of life (National Institute for Health and Clinical Excellence 2008).

ICECAP-O (Coast, Flynn, Natarajan, et al. 2008, Grewal et al. 2006): The ICECAP-O (ICEpop CAPability measure for Older people) is a measure of capability in people aged 65 years and older. The descriptive system comprises five attributes of well-being: attachment, security, role, enjoyment and control. Population values (for people aged 65 years and over) for the attribute levels were estimated using best-worst scaling methods to construct a capability index. The index is "anchored" at 0, for no capability, and at 1, for full capability. The instrument was designed to be used for economic evaluations that span health and social care (Coast, Flynn, Sutton, et al. 2008).

SF-12 (MCS-12 and PCS-12) (Jenkinson et al. 1997): Summary mental health and physical functioning scores (Mental Component Summary (MCS-12) and Physical Component Summary (PCS-12)) were constructed from the 12-Item Short-Form Health Survey (SF-12). Differences of 2 to 2.5 points on the SF-36 summary scores have been suggested as clinically meaningful (Ware et al. 2007); larger values for the SF-12 summary scores of 2.5 and 10 points have been estimated (Parker et al. 2012).

SF-6D: 249 health states from the SF-6D descriptive system (derived from the SF-12), were valued using standard gamble technique by a sample representative of the UK population (the UK tariff) to produce a preference-based index (Brazier and Roberts 2004, Brazier 2007).

CESD-10 (Andresen et al. 1994): The short form Center for Epidemiologic Studies Depression scale (CESD-10) is a 10-item screening instrument for depression symptoms. The CESD-10 scale summary score ranges from 0 to 30 (where 0 is the lowest and 30 is the highest level of symptomatology). A difference of five points or more has been interpreted as clinically meaningful (i.e. showing depressed symptoms) (Steffens et al. 2002).

Spielberger State-Trait Anxiety Inventory (Brief STAI) (Marteau and Bekker 1992): The six-item short form of the instrument measures "state anxiety" (feelings of anxiety at the current time) and has been widely used, including for people with diabetes (Park et al. 2008). Inventory scores range from 6-24 (where 6 is lowest and 24 is highest).

CSRI (Beecham and Knapp 2001): The Client Services Receipt Inventory (CSRI) for this study collected comprehensive information from participants on their service use in the prior three months, living arrangements, employment status and welfare benefits. A carer module of the CSRI collected information (for use in the accompanying carers' study) on patterns of unpaid care and support provided by family and other carers.

4.9 Economic Evaluation: Choice of Evaluative Approach

Economic evaluators have, broadly speaking, a choice of two theoretical approaches to guide the methodology for assessing the costs and benefits of health care interventions. Cost-utility and cost-effectiveness analysis spring from the extra-welfarist or 'decision-maker's approach', aimed at maximising health outcomes from a given budget (Brouwer and Koopmanschap 2000). Cost-benefit analysis involves measuring health gains and the costs of achieving those gains in monetary terms (Pauly 1995): this approach has its roots in welfare economics and aims to assess whether a new technology is *worth* the expenditure from a given budget (Buchanan and Wordsworth 2015, Drummond et al. 2015).

4.9.1 Welfare economics, Welfarism and Extra-welfarism

In welfare economics, individuals maximise their utility (thereby improving their welfare) by making choices that suit their own preferences (Drummond et al. 2015). The 'welfarist' approach dictates that social welfare is a function of individual utility alone (Culyer 1989). If there are gains for some individuals without losses for others (a Pareto improvement), welfare is improved and efficiency increased (Coast 2009, Brouwer et al. 2008). But there are obvious limitations to this approach since it offers no steer for policy makers allocating resources to public services, in deciding which group in society will benefit and which will

lose out (Drummond et al. 2015, Coast, Smith, and Lorgelly 2008). To allow comparisons between individuals, the compensation principle has been proposed: welfare improvement can still result if those who gain from the introduction of a new technology are able to compensate the losers and still be better off (a potential Pareto improvement) (Coast, Smith, and Lorgelly 2008, Drummond et al. 2015). The compensation test being met, resource transfers do not actually have to be paid for the innovation to be judged cost-beneficial (Pauly 1995). Benefit is measured in terms of a person's maximum willingness to pay for a new technology (whether a health or other technology); willingness to pay (in money terms) being the maximum amount of other goods/consumption opportunities the person would forgo to get that benefit (Drummond et al. 2015, Pauly 1995). If the total willingness to pay of all affected individuals outweighs the costs of supplying the technology, then it is welfare-maximising and therefore efficient (Coast, Smith, and Lorgelly 2008, Brouwer and Koopmanschap 2000). Thus those gaining (getting the benefit, or outputs, of the technology) could pay compensation to the suppliers of inputs to the technology and still be better off (Pauly 1995). Measuring improvement requires the measurement of benefits (gains) and costs (losses) in terms of their (monetary) value to individuals. However, in situations where markets do not exist or do not function well, as may arise in the case of health services, valuing relevant outcomes is not straightforward and must be established by other means, typically by establishing how much people are willing to pay for these types of outcomes (Drummond et al. 2015).

There are several possible objections to this approach. Welfare in the welfarist approach depends on a measure of individuals' willingness to pay, which may be influenced by ability to pay, so allocation on this basis could be biased towards the better-off (Coast, Smith, and Lorgelly 2008). Preferences other than those of the affected individuals could be important to society if people perceive inherent merit in providing certain services (Brouwer et al. 2008). Importantly, decision making on what constitutes the social good may depend upon more than individuals' welfare alone (Drummond et al. 2015).

The extra-welfarist (or decision maker's) approach proposes another way of determining the social welfare function. A decision maker delegated to make choices by society is provided with information on the valuation of individuals' preferences, and the implications of the decision to choose a particular option. In extra-welfarism, the social welfare function may maximise individual characteristics, typically health (Drummond et al. 2015). Brouwer et al. (2008) have summarised four key features of this approach. Evaluative outcome measures frequently involve health gain but others exist, such as well-being and

satisfaction with services. Valuation of costs and benefits can be undertaken by not only the individuals likely to be affected by the technology but also by representative samples of individuals, or societal decision makers. Weighting is often applied to outcomes in extra-welfarism, for instance to reflect need. Lastly, interpersonal comparisons of evaluative outcomes are explicitly of interest in the extra-welfarist framework and the use of measures of health can facilitate such comparisons.

The role of economic evaluation depends on the choice of methodological approach. From the welfarist perspective, the role is to determine whether a new technology is worth funding on the basis of social welfare maximisation. In the extra-welfarist framework, the role is "more modest, claiming to inform social decisions in health rather than prescribing social choice. […I]t exposes the policy implications of the social values implicit in existing policies and the resources allocated by those who claim some legitimacy to make such decisions." (Drummond et al. 2015, p.37)

Cost-effectiveness and cost-utility analytic approaches have increasingly dominated the practice of economic evaluation of health care programmes since the 1980s. This trend emerged as the demand for evaluations increased, the ability to measure health-related quality of life outcomes improved, and importantly, as numerous national health technology assessment (HTA) agencies (including NICE) adopted the extra-welfarist perspective (Buchanan and Wordsworth 2015, Coast, Smith, and Lorgelly 2008). Cost-benefit analytic approaches to evaluating health care technologies appear to have fallen out of favour for several reasons[16]. Drummond et al. (2015) suggest that governments have not been keen to distribute funding on the basis of preferences that might be skewed by existing inequalities in income distribution. Also, in the UK, monetary valuations of health are unpalatable to the general public (Coast, Smith, and Lorgelly 2008). The economic evaluation evidence base for telehealth described in Chapter 3 (cf. Polisena, Coyle, et al. 2009, Vergara Rojas and Gagnon 2008, Bergmo 2009) rests on extra-welfarist foundations (the evidence from any formal cost-effectiveness or cost-benefit analyses of telecare being scant). The choice to take a cost-effectiveness and extra-welfarist approach to the economic evaluation of telehealth and telecare was straightforward in the current HTA context.

[16] It has been argued, on the other hand, that other approaches than extra-welfarism (for instance examining capabilities) could be taken to evaluate certain kinds of health interventions. For example, interventions that cross sectors and government departments (for example such public health interventions as alcohol abuse or obesity prevention) may have non-health outcomes that could be measured within a different evaluative framework (Buchanan and Wordsworth 2015, Coast, Smith, and Lorgelly 2008).

4.10 Economic Evaluation of a Complex Intervention: Methodological Issues

Evaluators considering the impacts of telemonitoring upon the use of health and social services face a number of challenges. The populations using these technologies may be very diverse; the technologies to be compared may be offered to populations that differ in terms of needs and expectations; the interventions are sensitive to local conditions, constraining generalisability and reliability; and furthermore the technologies themselves differ in terms of the responses that are required from their users and from the health and social care agencies that provide them (Bergmo 2009, Polisena, Coyle, et al. 2009).

The types of challenges that Bergmo (2009) and Polisena (2009) identify will look familiar to evaluators of complex health and social care interventions. As Byford and Sefton (2003) suggest, studies of even quite standardised interventions may vary considerably in practice, given the skills and preferences of practitioners. A number of additional issues are likely to arise in evaluating complex interventions. Users of services may be a heterogeneous group. They may be highly involved in the production of care; the more active the user involvement, the more complicated the relationship between inputs and outputs. This is also a source of increased heterogeneity, with implications for the sample size required and number of user-related variables that should be controlled for. Some interventions, particularly social care and mental health services, may be deliberately flexible and complicated, tailored to the user, and covering several service areas, e.g. housing, day care and health. The goals of the intervention may be numerous and complicated. Finally, multiple agencies may be involved in delivering the intervention. The authors make an important point that measuring costs can be more difficult because the questionnaires may need to be long and broad in focus to try to capture a broad range of possible services; however this may limit the accuracy of self-report methods. While accuracy might be improved by electronic records, these may be limited to the use of services provided by the agency in question, so that the data from records of several agencies may be required. Valuation of the relevant costs is challenging as some services do not have national applicable unit costs readily available, and these must be directly calculated, which in turn requires more recording of more service components. The authors recommend that evaluators spend time at the design stage on understanding the components of the intervention in order to understand the mechanisms that influence outcomes and costs.

Recent guidance on complex interventions (Craig et al. 2008) makes a number of similar observations, defining these interventions as containing multiple interacting

components, which can be complex in terms of the numbers of interactions between components, the number of required behaviours (on the part of the intervener or the recipient), the number of levels of organisation or groups involved, the number of outcomes and the extent of flexibility allowed the intervention. The guidance addresses the role of economic evaluations, advising researchers to use information about the additional cost of the intervention to calculate how much more effective the intervention would need to be to be cost-effective (however this does require the existence of such information in the first place). They also suggest that the intervention be clearly defined so that relevant resource use can be identified; that resource use and outcomes are recorded consistently across time points, as cost and effects might differ at different points; that the perspective be identified and preferably a societal viewpoint be used; and that using the QALY as the outcome measure should be considered.

4.11 Economic Evaluation Methods

In order to address the question of cost-effectiveness in a trial context, and from a local authority or NHS commissioning perspective, it is necessary to gather information on costs to health and social services. Johnston, Buxton et al. (1999) suggest that in terms of health service costs, there are "direct" costs relating to the intervention itself; costs of illness more generally; costs of future use of health services; and trial-specific costs. There are also costs outside of the commissioner's perspective: these include costs to other public agencies immediately and in the future; to patients, service users and carers (travel, time lost in receiving treatment, productivity losses); and non-resource costs such as transfer payments (e.g. benefits). The following sections outline the data collection activities necessary to 'identify, measure, value and compare the costs and consequences of the alternatives being considered', which are the basic tasks of economic evaluation (Drummond et al. 2015, p. 4, Drummond et al. 1999, p. 9).

The cost-effectiveness analyses reported in this thesis took a NHS and local authority (LA) perspective. The analyses of costs and cost-effectiveness in the thesis are in the main focused on results at 12-month follow-up. The time horizon was limited to the trial period, and no discounting of costs or outcomes was undertaken.

The data from the "short-term", or 4-month follow-up assessments were of a poorer quality and in different sample sizes than those obtained at baseline and long-term follow-up assessment points. Whereas all of the baseline and more than half of the 12-month follow-up

questionnaires were administered by interviewers, all 4-month follow-up questionnaires were administered by post. The sample of participants returning data at this point did not entirely overlap with that returning data at 12-month follow-up (see 5.14). An example of the problems encountered with the short-term data collection involved a number of respondents opting to provide details of services they had used in the 'other' boxes on the CSRI, even when the option of reporting use of that particular service had been presented in a previous question. Such issues required making intensive scrutiny of individual forms and re-classifying such “other” services where relevant. I undertook extensive checks and all cleaning of the CSRI-generated data.

4.12 Health and Social Care Service Use and Costs

Health and social care utilisation data were collected from trial participants on the CSRI as described in section 4.7.2. Unit costs were then attached to the units of services used to calculate per-participant costs. In the next sections, I describe the methods used to locate and assign unit costs. I then describe in more detail the methods I followed to calculate the costs of the intervention.

4.13 Valuation Strategy for Self-reported Service Use

Unit costs (in 2009/10 prices) were applied to units of self-reported service use. Unit costs applied are summarised in Table 4.1; a table of unit costs with detailed descriptions of their sources and calculations or assumptions made is given in Appendix 1. The majority of unit costs for social care, primary and community health care were sourced from Personal Social Services Research Unit Costs compendium (PSSRU UC) (2010); if unavailable from this source, costs were taken from other published sources. For instance, some equipment costs were located in the national catalogue and tariff for Simple Aids to Daily Living (Department of Health Care Services Efficiency Delivery Programme 2010). Unit costs for hospital services (accident and emergency (A&E), inpatient overnight and day case bed-days, and outpatient attendances) were taken from the National Schedule of Reference Costs 2009-10 for NHS Trusts and PCTs Combined (2011).

Table 4.1 Unit costs summary

Cost category	Unit	Unit cost (£, 2009-10)
A&E	Attendance	103 - 133
Inpatient care	Attendance	116 - 1657
Day Hospital care	Bed-day	156 - 1496
Outpatient appointments and procedures	Attendance	23 - 306
Community health services/primary care	Visit	Range: 24-192
Community health services/primary care	Contact	Range: 20.26-86.85
Community health services/primary care	Minute	Range: 0.95-4
Community mental health care	Minute	Range: 0.83-4.72
Community care	Minute	Range: 0.42-0.92
Equipment	Item	Range: 1.5 - 455
Adaptations	Item	Range: 0.1-97.5
Care home respite	Days	Range: 63.72-70.57
Day care and other day services	Attendance	Range: 36-155.82
Medications	Standard Quantity Unit	Range: 0.01-419.62

Medications costs were obtained from the NHS Information Centre's Prescription cost analysis (PCA) (Health and Social Care Information Centre 2011). All medications at each time point were examined and assigned a medication code. The medication data in the dataset were then assigned the corresponding code using a string search method, matching the reported name to a stub of between four and eight characters. A look-up table was developed including the price per unit, the medication unit (such as milligrams or micrograms) and the dosage. Where all information on medication, medication unit and dosage was known, the specific unit price from the PCA was matched to the reported medication. Where the medication dosage was not known, the medication was matched to an average weighted price for the medication and medication unit from the PCA. Where only the name of the medication was known, the medication reported was matched to an average weighted price for that medication from the PCA.

In general the assumption was made that costs were borne by health and social care, including in cases where participants might make co-payments (e.g. use of services of dentists, chiropodists and opticians); however in the case of household adaptations and equipment for daily living, costs of those items reported to have been provided by participants or their relatives were excluded.

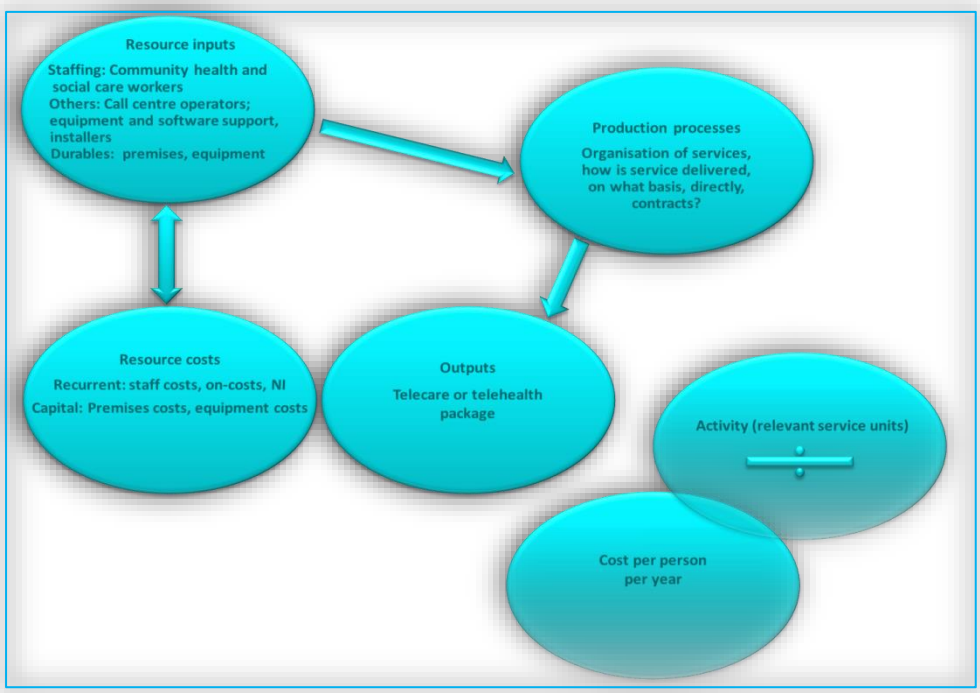
In the cost-effectiveness analyses, the costs of service use over the three-month periods prior to the baseline and 12-month follow-up were calculated and then multiplied by four, giving annual equivalent costs for the pre-baseline year and the year over which the intervention was delivered.

4.14 Intervention Costs

In order to estimate 'direct' costs associated with any intervention services consumed during the trial ('intervention costs'), I added a stream of work to the planned data collection to support the economic evaluation. This was to carry out a series of interviews with those responsible for implementing the interventions, in order to be able to describe the implementation of the interventions in all three sites (there were telecare and telehealth programmes in each site, requiring descriptions of six programmes altogether).

This work was to describe the interventions, agencies, staffing and other resource inputs and understand the activities, production processes and mechanisms that may influence costs. Estimating the intervention costs involves four stages (cf. Allen and Beecham 1993, Beecham 2000): (i) describing the interventions in terms of their typical resource inputs and associated routine activities; (ii) calculating relevant service units; (iii) collecting cost data; (iv) calculating a unit cost for the intervention. The elements involved in the estimation of a unit cost of a telecare/telehealth package are summarised in Figure 4.2. The arrows show the direction of influence: resource inputs influence resource costs and vice versa; the way that resource inputs are used in the production process influences not only the quantities of the outputs but the scope and quality of the outputs. Finding a relevant service unit is necessary to construct a useful and meaningful unit cost, in this case an annual per-person cost of a telecare/telehealth package (Box 4.1).

Figure 4.2 Elements involved in estimating a unit cost of telehealth/telecare



Box 4.1 Costing the telehealth and telecare interventions

Source: Henderson, Beecham, and Knapp (2013)

To develop an understanding of production inputs and processes, information was collected using a "bottom-up" approach, involving 19 interviews (by telephone or face-to-face) with key informants and drawing on correspondence with three on-site WSD project teams. A more "top-down" approach was taken to collecting cost and activity data on the delivery of the intervention, whereby a spreadsheet-based pro-forma was used to guide collection from the project teams. These bottom-up and top-down data were used to establish a unit cost, the direct annual per-person cost of a telecare or telehealth package. Unit costs were calculated based on 2009/10 service configurations (when most trial participants were recruited) in order to approximate the costs of running the services at the sites' planned capacity, rather than in the 2008/09 start-up phase.

The aim of the work was to establish the average costs of the interventions across the three sites. The ways in which telecare or telehealth services were delivered were determined locally and were not prescribed by the trial evaluation team.

The first step to building a detailed picture of the services in each site was to understand the inputs and processes involved in producing the interventions, examining

Box 4.1 (continued)

important features of the delivery systems put in place, in terms of equipment supply, systems and infrastructure enabling the appropriate equipment to operate (assessment, installation, servers, maintenance), and monitoring and response services interacting with participants through the telehealth/telecare technology.

As a condition of the WSD trials, participants were not to be charged for telehealth or telecare equipment or support services. Participants were however expected to have telephone lines and power supplies for telecare; in the case of telehealth, participants in one site were expected to have a television set. Data transmission by participants was also provided free of charge to them.

Telecare and telehealth equipment

The sites' project teams provided data on participants' telecare/telehealth equipment, and the prices that had been paid for the equipment, for the evaluation. Equipment costs were calculated for each participant. While most of the equipment was purchased for the trial, telehealth base units and most peripherals were rented in one pilot site. In either case, the purchased base units were annuitised over 5 years (Department of Health 2001), while costs of purchased "peripherals" (alarms, sensors or items attached to the base unit, e.g. blood pressure monitors) were annuitised over the same period or over the peripheral's lifetime if this information was available from sites or manufacturers' specifications. One site provided equipment rental charge information.

Telehealth

Telehealth users received a base unit, that could be either free-standing or a set-top box for a television, and 'peripherals' appropriate to their long-term condition. The latter consisted of cabled or bluetoothed pulse oximeters, blood-pressure cuffs, glucometers and weighing scales, which transmitted the observations data to the base unit. While a description of clinical processes and behavioural regimens associated with the telehealth intervention was not within the scope of this thesis, interested readers will find these details in Cartwright et al. (2013).

Telecare

Telecare users received equipment consisting of a telecare 'base unit' (Tunstall Lifeline Connect or Connect+), a pendant alarm and at least one other sensor or device.

Box 4.1 (continued)

Up to 27 types of device were available for use by trial participants, for instance 'key safes', bed sensors, temperature extremes sensors, and fall detectors; amongst those participating in the WSD questionnaire study, participants received between one and 11 items.

Costs of supporting the delivery of the interventions

Methods for calculating support costs were similar across both interventions. Support personnel were assumed to comprise individuals working to monitor and respond to alarms/sensor alerts and to triggers flagged by algorithms in the telehealth software programmes; supervisors of these workers; and on-site WSD team managers, trainers and back-office staff. The cost calculations excluded posts/parts of posts that involved trial evaluation or recruitment. On-costs, administrative, premises and capital overheads of directly-provided workers were calculated based on the WSD teams' information. Where sites could not provide details for calculation of administrative overheads, these were assumed to be 16 per cent of salary costs (Curtis 2010). Other relevant costs were: server maintenance, software licences, and costs of providing free-phone numbers and data transmission from base units to servers.

Installation and maintenance costs were partly variable and partly fixed. One site had maintained a detailed breakdown of spending on these activities in 2009/10; these proportions of expenditure were applied to costs in the other sites where less detailed information was available. Fixed costs were spread over five years, the assumed lifetime of the base units, while the variable costs were taken to be incurred within 2009/10. Costs of installers, their associated overheads and of storage and transport of equipment were all taken into account. For telecare, the split between fixed and variable costs was 65 per cent and 35 per cent respectively, and for telehealth 90 per cent and 10 per cent respectively.

Telecare monitoring services and dedicated response services were provided under contract and the assumption made that such contracts covered the providers' costs. The costs of contracts in 2009/10 were divided by the number of trial participants in order to obtain an annual per-participant average cost.

Telehealth monitoring services were calculated either top-down or bottom-up, depending on the components of the service. All sites had centralised monitoring call centre teams: the costs of these directly provided or contracted central teams were calculated in terms of annual expenditure on their staff in 2009/10 (included associated overheads).

Box 4.1 (continued)

However two sites provided some monitoring services through local nursing teams (community matrons or specialist nurses): their costs were estimated from the bottom-up, counting their time spent in telehealth training and in monitoring the telehealth screen. The annual total monitoring costs were calculated by applying the relevant unit costs (based on WSD project team information on NHS pay bands and local nursing team staffing complements, and including on-costs and capital, indirect and direct overheads) to the total estimated monitoring time. This latter was based in turn on the average daily screen-monitoring time (calculated using data provided by WSD project teams) of two minutes (Henderson et al. 2013). The costs of central and local monitoring were aggregated and divided by number of study participants monitored over the year, for an average annual per-participant cost of monitoring.

Mean annual telehealth and telecare support costs per participant (including monitoring, equipment infrastructure, installation and maintenance) were calculated and allocated to participants who had received the telehealth/telecare equipment. Because the support costs were estimated mostly top-down, these data did not vary between participants in the same site, although equipment cost data did vary between individual cases. In addition, variations of these total costs were calculated for use in future sensitivity analyses: annual costs for telehealth and telecare, excluding staff posts and contracts specifically related to WSD project management; and in the case of telecare only, the annual costs of support, excluding costs of the dedicated WSD telecare response services.

4.15 Economic Analyses

4.15.1 Intention-to-Treat and Per-protocol Populations

Cost-effectiveness analyses were carried out for all randomised participants, adhering to the principle of intention-to-treat (ITT). If data for 12 months were available for a participant, they were included in the cost-effectiveness analyses (modified intention to treat). The cost-effectiveness analyses were conducted using baseline and 12-month follow-up data (see section 4.11 for an explanation of the issues with the 4-month follow-up data). Participants' data were analysed within the group to which they were originally allocated, within each trial. In-depth analyses of cost variations by subgroup of interest in the telehealth and telecare samples were both analysed by ITT. The intention-to-treat analysis is important to understand what the benefits of changing to a new treatment are generally and not only within those

adhering precisely to the intervention (Hollis and Campbell 1999). The per-protocol population reflected participants' de facto allocations (for instance control participants who were provided with telehealth/telecare and intervention participants who declined their telehealth/telecare equipment but continued to participate in the questionnaire study).

4.15.2 Sample Characteristics at Baseline

Baseline characteristics of all samples are described for each experimental group in terms of means and standard deviations for continuous variables and in terms of percentages and number of observations for binary and categorical variables. Between-group raw differences are presented. In this thesis, conforming to CONSORT guidelines (Schulz, Altman, and Moher 2010, Moher et al. 2010), I have not presented tests for baseline differences in characteristics between experimental groups. Differences in the samples due to loss to follow-up have been compared within the group to which participants were allocated, and tested for differences using clustered t-tests for continuous variables and clustered chi-squared tests for dichotomous variables (Donner and Klar 2000, Herrin 2012) where appropriate, or a z-test of proportions where the variable is not grouped by general practice cluster.

4.16 Economic Evaluation Outcome Measures

4.16.1 Primary Outcomes

Cost per quality-adjusted life year was adopted as the primary outcome in both telehealth and telecare cost-effectiveness analyses.

Quality-adjusted life years (QALY) were constructed by calculating the utility scores derived from the EQ-5D-3L using societal weights (Dolan et al. 1995, Brooks 1996). QALYs were calculated taking the 'area under the curve' with linear interpolation between the baseline and 12-month follow-up assessment scores.

4.16.2 Secondary Outcomes

Secondary outcomes examined were: psychological well-being (ICECAP-O), state-trait anxiety (Brief STAI), depression symptoms (CESD-10), and summary mental health and physical functioning scores (MCS-12) and (PCS-12) (examined in the telecare cost-effectiveness analyses). The baseline standard deviations of the PCS-12 and MCS-12

measures (PCS-12 SD=9.0168 and MCS-12 SD=11.965) were multiplied by the effect size (Samsa et al. 1999), which had been set at 0.3 for the study HRQoL instruments as the smallest size of effect that was meaningful (Bower et al. 2011), to give differences (rounded to integer) of 3 and 4 points respectively. The Brief STAI scores were rescaled to between 0 and 1, to indicate effectiveness in terms of, respectively, lowest and highest levels of anxiety. QALY were also constructed from the SF-6D index and QALYs calculated in the same manner as described in Section 4.16.1.

In the case of telecare, psychological well-being, state-trait anxiety and summary mental health and physical status were examined in the cost-effectiveness analyses. For telehealth, cost-effectiveness analyses included psychological well-being, state-trait anxiety and depression symptoms.

4.17 Descriptive Analyses and Cost Categories

All descriptive analyses of the Telecare and Telehealth questionnaire study samples are presented in Chapter 5. Raw data on service use over the prior three months have been summarised, by trial and by experimental group in terms of the numbers and proportions of participants that used each service, and in terms of numbers of service units (reporting the means and standard errors). Intervention and control group imputed costs over the prior three months have been summarised in terms of their means and cluster-adjusted standard errors. Raw differences between groups in mean utilisation and mean costs are also presented.

Descriptive statistics in Chapter 5 have been organised by cost-reporting categories, the agency to which costs are assumed to fall and resource use sub-categories, as given in Table 4.2.

Table 4.2 Cost reporting categories and sub-categories

Cost category	Agency	Resource use
Hospital services		
	NHS	A&E attendance
	NHS	Inpatient bed days
	NHS	Day Hospital attendances
	NHS	Outpatient attendances
Community health		
	NHS	Paramedic
	NHS	Community matron
	NHS	Community or district nurse
	NHS	Practice nurse
	NHS	Night nurse
	NHS	Specialist nurse
	NHS	Physiotherapist or occupational
	NHS	GP (home)
	NHS	Dentist
	NHS	Chiropodist
	NHS	Optician
Community mental health		
	NHS	Psychiatrist visit
	NHS	Mental health nurse visit
	NHS	Medications
Community social care		
	Local Authority	Social worker visit
	Local Authority	Council home help visit
	Local Authority	Private/independent home care/home
	Local Authority	Paid night carer visit
	Local Authority	Meals on Wheels meal
	Local Authority	Laundry (incontinence) service
	Local Authority	Community alarm*
	Local Authority	Long-term care/respite stays

Cost category	Agency	Resource use
Equipment		
	Local Authority	Equipment items
	NHS	Equipment items
Adaptations		
	Local Authority	Adaptations items
	NHS	Adaptations items
Day care		
	Local Authority	Day care and other day services
	NHS	Day care and other day services

*For the participants in receipt of a telecare package (per-protocol allocation), community alarm costs were excluded from all follow-up cost calculations to avoid double-counting with the cost of the intervention. For examination of variations in social service costs (excluding intervention costs) in the telecare sample (Chapter 6), receipt of community alarms was excluded from cost calculations. This exclusion was necessary to avoid confounding receipt of social services as an indirect result of the intervention with the direct result of allocation to the intervention.

4.18 Costs and Cost-effectiveness Analyses

I used analytical methods appropriate to the cluster-randomised nature of the trial, to avoid biasing the standard errors of the regression coefficients (Bartholomew et al. 2008); hence preventing errors in inference and inefficient parameter estimates (Manca et al. 2005). My approach varied depending on the objectives of the analysis and necessarily balanced the requirements of the analysis against the time available for analysis, the capabilities of the software packages to which I had access and my ability to use them appropriately.

4.19 Multivariate Analyses of Service Use and Costs Data

To address research sub-question 3, "what patient/user characteristics are associated with cost variations?" I undertook in-depth explorations of variations in the costs of the study participants.

4.19.1 Variations in Costs of Telehealth and Telecare: Subgroup Analyses

In other parts of the thesis, I compare the costs of intervention and control participants. However in chapter 6, I drill down beyond the pooled analyses to explore whether:

1) the three-month costs of participants allocated to telehealth or usual care differed between baseline and long-term follow-up time points, depending on their index long-term condition (diabetes, COPD or heart failure); and
2) the three-month costs of participants allocated to telecare or usual care differed between baseline and long-term follow-up time points, depending on their living arrangements (living alone or with others).

Costs were examined at different levels of aggregation: total health and social care costs (including and excluding the costs of the intervention), and agency-specific costs. Thus all NHS costs were considered (including all items in the following cost categories in Table 4.2: hospital, community health and mental health, NHS day services, equipment and adaptations). Hospital costs were also examined separately. Social care costs included all items in the following cost categories in Table 4.2: community social care (except community alarms in the case of the Telecare costs analyses), local authority day services, equipment and adaptations.

In the following sections I first describe the motivation for my choice of modelling approach; I then describe the multilevel models and underlying assumptions, specifying equations generically to cover both telehealth and telecare analysis strategies (in Chapter 6, I set out equations covering the telehealth- and telecare-specific models); lastly I describe the dependent and independent variables of interest.

4.19.2 Econometric Modelling Approach

I adopted a multilevel framework for modelling health and social care costs. This approach met the requirements of the cluster-randomised data structure, with the flexibility to reflect within-subject differences in three-month costs between the baseline and 12-month follow-up. In this framework, at the third level, data were clustered by general practice; at the second level (subject-level), data were clustered by participant; the first level consisted of cost observations at each time point (two observations per participant).

There are several approaches to the issue of clustering effects, depending on whether the clustering unit is of interest to the analyst, or merely a nuisance factor. From either perspective, the analysis must take account of clustering in order to avoid downwardly biasing the standard errors of model estimates (Bartholomew et al. 2008). Exploring the influence of clustering can yield valuable information about the interaction between and within clusters simultaneously. The relationship between an outcome of interest and the

characteristics of cluster members is not necessarily the same at the cluster level. For instance, wards in a hospital may have different outcomes depending on their characteristics, while hospital-level outcomes and hospital-level characteristics may also be related, but not in the same way. Perhaps some hospitals are more likely to have lower infection rates because of other hospital-level characteristics (e.g. leadership, staff relations, geographical location), even if individual wards have high infection rates because of certain characteristics (staff relations, infection-control procedures, medical specialty) at the ward-level. Aggregating data to the hospital level could sacrifice a great deal of information on the degree of variation in infection rates within the hospitals and lead to inappropriate inferences. A two-level model could be used in this instance to examine hospital-level (level-two or cluster-level) and ward-level (level-one) infection rates.

Turning to a more general discussion, we might want to know (i) the degree to which the clusters' means vary around the overall (or population) mean of the response variable; and also (ii) to what extent the units within clusters vary around the cluster means. I have drawn on the work of Bartholomew, Steele et al. (2008) and Rabe-Hesketh and Skrondal (2012) in the following illustration. A simple linear regression model with no covariates, ignoring clustering, might be:

$$y_i = \beta_0 + e_i \,, \tag{4.1}$$

where y_i is the response of unit *i*, β_0 is the population mean, and e_i is the distance between the population mean and the value of *y* for unit *i* (otherwise known as the 'residual'or 'error'). Extending this to consider cluster membership, we could instead write

$$y_{ij} = \beta + u_j + e_{ij} \,, \tag{4.2}$$

where y_{ij} is the response of unit *i* in cluster *j* and the overall mean is β. We have modelled the two-level structure by partitioning the residual, or variance, into two components. The departure of cluster *j* from β is known as its 'random effect', u_j, which has between-cluster variance σ_u^2, uncorrelated across clusters; the departure of unit *i* from the cluster mean, e_{ij} , has a constant within-cluster variance σ_e^2. Residuals follow normal distributions:

$$u_j \sim N(0, \sigma_u^2)$$
$$e_j \sim N(0, \sigma_e^2)$$

This 'mixed effects' model has a fixed part containing the covariates (β in this case) and a random part (u_j and e_{ij}here). The fixed part (or mean structure) describes the population-averaged relationship between the response and predictor variables (Rabe-Hesketh and Skrondal 2012).

Extending the constant-only model above, the response of a normally distributed outcome variable to a single covariate x_{ij} is

$$y_{ij} = \beta_1 + \beta_2 x_{ij} + u_j + e_{ij}. \quad (4.3)$$

In this case, the relationship between the response variable and the covariate is described by a straight line with constant β_1 and slope β_2. The cluster *j* intercept $\beta_1 + u_j$ represents the deviation from the overall mean by u_j (Bartholomew et al. 2008). In this linear case, the mean of the response variable, y_{ij} , is a linear function of the covariate. Equations (4.2) and (4.3) describe linear multilevel models.

Adopting now a multilevel generalised linear modelling approach, there are three model components to be described (Rodrıguez 2008, Hox 2010):

1. an outcome variable *y* with a specific error distribution and mean μ and variance σ^2
2. a linear equation producing a latent predictor η for the expected value μ of outcome variable *y*
3. a link function g describing the relationship between this expectation, μ, and predicted values or transformed mean η, where $\eta = g(\mu)$.

Equation (4.3) can thus be re-written as a multilevel generalised linear model:

$$g(\mu_{ij}) = \eta_{ij} = \beta_1 + \beta_2 x_{ij} + u_j,$$

where the conditional mean μ depends on the random effect $u_j$and a covariate x_{ij} so that the outcome is independent and follows the linear equation (Rodrıguez 2008).

In this multilevel generalised linear framework, the conditional outcome distribution is now assumed to be exponential, rather than linear (Rodrıguez 2008), as would be the case in the general linear multilevel model. If the errors are normally distributed and the link function is identity, we have $\mu_{ij} = \eta_{ij}$, as in the general linear case (Hox 2010). In this case, the mean of the response is modelled on the same scale as the covariates (Baldwin,

Fellingham, and Baldwin 2016). Where the outcome variable is not expected to be normally distributed, non-linear approaches are required to appropriately model the relationship between the conditional mean of response variable and covariates. Non-linear models are discussed in the following section.

The multilevel models so far described explicitly partition the residual variation of the response variable from the population mean according to the source (level 1 or level 2). The approach can be useful in examining variability in the costs of a particular participant receiving telehealth, given that person's characteristics. These hierarchical models have come into more common use within health economics in recent years in studies examining costs across regions, countries and centres (Thompson, Nixon, and Grieve 2006, Vazquez-Polo et al. 2005, Manca, Hawkins, and Sculpher 2005). On the other hand, if we are only interested in the response at the 'population' level – for instance to examine how participants differ in response to particular factors, such as allocation to a telehealth intervention, we might choose a 'marginal' or 'population-averaged' approach. While population-averaged and subject-specific (or conditional) coefficients will be similar in the linear model, even if the interpretation is different (Baldwin, Fellingham, and Baldwin 2016), they will not be the same in the non-linear case, as discussed in section 4.19.5.

4.19.3 Three-level Model Specification

Three-level 'null' linear model with Gaussian distribution and identity link: In the simplest case, we consider a linear, multilevel random-intercept cost model: y_{ijk} denotes cost at occasion i $(i = 1,2)$ for person j $(j = 1, \dots, n)$ in general practice cluster $(k = 1, \dots, n)$. A random intercept model without covariates can be written as:

$$y_{ijk} = \beta + u_{jk}^{(2)} + u_{k}^{(3)} + e_{ijk} \tag{4.4}$$

where β represents the overall mean, $\mu_{jk}^{(2)}$ is random intercept for subject j and cluster k and $\mu_{k}^{(3)}$ is the random intercept for cluster k. The GP-level and participant-level random intercepts are assumed to have means of zero and variances of ${\sigma_{\mu}^{2}}^{(3)}$ and ${\sigma_{\mu}^{2}}^{(2)}$ respectively, and the time-level error term, a mean of zero and variance ε_{ijk} (Rabe-Hesketh and Skrondal 2012). In this model, participants in each allocation group (practice-level clusters allocated to either intervention or control) have the same trajectory of costs over the two time points.

Three-level 'null' linear model with gamma distribution and log-link: Health and social expenditure data might be expected to follow a non-normal distribution. Some people receiving services may have very low (or no) costs but a few will have very considerable costs. To accommodate the zero-truncated and right skewed data, we can modify the model to fit these to a gamma distribution with mean μ and variance σ^2, and log-link (Thompson, Nixon, and Grieve 2006, Rabe-Hesketh, Toulopoulou, and Murray 2001, Manning, Basu, and Mullahy 2005, Liu et al. 2010) and write the model as:

$$y_{ijk}\left|\, \mathrm{x}_{ijk,}\, u_{jk}^{(2)}, u_{k}^{(3)} \sim Gamma\left(\mu_{ijk,}\sigma^2\right),\ \log(\mu_{ijk}\right) = \beta + u_{jk}^{(2)} + u_{k}^{(3)}, \tag{4.5}$$

where $\mu_{ijk,} = E\left(y_{ijk}|\, \mathrm{x}_{ijk,}\, u_{jk}^{(2)}, u_{k}^{(3)}\right)$.

Three-level linear model with covariates: The aim of the analyses was to examine the difference in baseline and 12-month follow-up costs between treatment groups in terms of the subgroups (e.g. living arrangements; long-term conditions). This requires an extension to the model to allow the calculation of these differences, taking a difference-in-difference-in-difference (DDD) approach (Das and Smith 2012). The approach is useful for examining effects of changes over time between groups, removing the influence of unobserved differences between the groups, provided that these do not vary with time (Cameron and Trivedi 2005). The approach has been taken in numerous observational and randomised controlled studies (Ikenwilo 2013, Bardsley, Steventon, and Doll 2013, Jacobs and Barrenho 2011). The model is extended to include interaction terms for the subgroup of interest and treatment allocation, for time point and subgroup, and for the triple interaction of time point, subgroup and treatment allocation, as follows:

$$\begin{aligned} y_{ijk}\left|\, \mathrm{x}_{ijk,}\, u_{jk}^{(2)}, u_{k}^{(3)} \sim Gamma\left(\mu_{ijk,}\sigma^2\right),\ \log(\mu_{ijk}\right) \\ = \beta_1 + \beta_2 Treat_k + \beta_3 Time_{ijk} + \beta_4 Treat_k Time_{ijk} \\ + \beta_5 Subgroup_{jk} + \beta_6 Treat_k Subgroup_{jk} \\ + \beta_7 Time_k Subgroup_{jk} + \beta_8 Treat_k Time_{ijk} Subgroup_{jk} \\ + u_{jk}^{(2)} + u_{k}^{(3)}, \end{aligned} \tag{4.6}$$

where *Subgroup* stands for the subgroup of interest, *Treat* for treatment allocation and *Time* for time point. Also,

$$u_k^{(3)} \sim N\left(0, \sigma_u^{2(3)}\right), u_k^{(2)} \sim N\left(0, \sigma_u^{2(2)}\right)$$

$$\text{cov}\left(u_{jk}^{(2)}, u_k^{(3)}\right) = 0$$

In other words, GP and participant-level random intercepts are assumed to have means of zero and variances of $\sigma_u^{2(3)}$ and $\sigma_u^{2(2)}$ respectively, given the model covariates. The level 2 and 3 random effects are assumed to be independent.

In this multiplicative model, the exponentiated coefficient on time (exp(β_3)) must be interpreted as a ratio of costs between the baseline and follow-up points within participants (Thompson, Nixon, and Grieve 2006). The coefficient of interest is on the triple interaction term β_8 which represents the effect on costs in the intervention period of participants who were randomised to the intervention and who were in one of the subgroups. In the case where there are two subgroups, this can be understood as (Wooldridge 2008):

$$\beta_8 = \left[\left(\overline{y}_{T,A,2} - \overline{y}_{T,A,1}\right) - \left(\overline{y}_{T,B,2} - \overline{y}_{T,B,1}\right)\right] - \left[\left(\overline{y}_{C,A,2} - \overline{y}_{C,A,1}\right) - \left(\overline{y}_{C,B,2} - \overline{y}_{C,B,1}\right)\right]$$

where $\overline{y}$ denotes conditional mean costs, T denotes intervention group, C control group, A is subgroup A, B is subgroup B, 1 denotes time 1 and 2 denotes time 2. The response of the outcome is conditional on both the covariate and the random effects for participants and general practice cluster.

Finally, *h* covariates z_{ijk} can be added to the model.

$$\begin{aligned} y_{ijk} \Big| \, \mathrm{x}_{ijk,}\, u_{jk}^{(2)}, u_k^{(3)} &\sim Gamma\left(\mu_{ijk,}\sigma^2\right),\ \log\left(\mu_{ijk}\right) \\ &= \beta_1 + \beta_2 Treat_k + \beta_3 Time_{ijk} + \beta_4 Treat_k Time_{ijk} \\ &+ \beta_5 Subgroup_{jk} + \beta_6 Treat_k Subgroup_{jk} \\ &+ \beta_7 Time_k Subgroup_{jk} \\ &+ \beta_8 Treat_k Time_{ijk} Subgroup_{jk} + \sum_h \beta_h\, z_{ijk} + u_{jk}^{(2)} \\ &+ u_k^{(3)}, \end{aligned} \tag{4.7}$$

4.19.4 Two-level Model Specification

If the general practice unit has little or no effect on the outcome, there will be little or no deviation from the overall mean in response to each model covariate (an ICC for general practice cluster of near-zero or zero). If so it would be reasonable to ignore GP-level

clustering in modelling the cost variations data. In this case, the treatment is effectively considered to be allocated at the participant level. A two-level random-intercept model with covariates is written:

$$
\begin{aligned}
y_{ij} \big|\, \mathrm{x}_{ij,}\, u_j \;\sim & Gamma\big(\mu_{ij,}\sigma^2\big),\ \log(\mu_{ij}) \\
& = \beta_1 + \beta_2 Treat_j + \beta_3 Time_{ij} + \beta_4 Treat_j Time_{ij} \\
& + \beta_5 Subgroup_j + \beta_6 Treat_j Subgroup_j \\
& + \beta_7 Time_{ij} Subgroup_j \; + \beta_8 Treat_j Time_{ij} Subgroup_j \\
& + \sum_h \beta_h \, z_{ijk} + \; u_j \;\;,
\end{aligned}
\tag{4.8}
$$

4.19.5 Telecare and Telehealth: Population-averaged Model Specification

So far I have set out costs models that are subject-specific (also known as unit-specific and conditional models). However, as discussed in 4.19.2, another approach to clustering can be taken. In the 'marginal' or 'population-averaged' approach, we are interested in the marginal expectation of the response variable (Zeger, Liang, and Albert 1988): $\mu_{ij,} = E(y_{ij})$. The link function describes the relationship between μ_{ij} and covariates z_{ij}; the variance of y_{ij} is a variance function g multiplied by a scale parameter ϕ, or formally

$$\mathrm{var}(y_{ij}) = g\big(\mu_{ij}\big) \cdot \phi \,.$$

Marginal approaches, for instance Generalised Estimating Equations (GEE), account for dependency between responses within a cluster (intra-cluster correlation) on covariates by specifying an appropriate working correlation matrix (Zorn 2001). This working correlation "is assumed to be the same for all subjects, reflecting *average dependence* among the repeated observations over subjects" (Hu et al. 1998, p.695).

A population-averaged model with covariates can be written as (the coefficients have been subscripted as β_{PA} to denote that these are parameters of a population averaged model):

$$
\begin{aligned}
y_{ij} \mid \mathrm{x}_{ij} \sim Gamma(\mu_{ij,}\rho_j),\ \log(\mu_{ij}) \qquad (4.9)\\
= \beta_{\mathrm{PA1}} + \beta_{\mathrm{PA2}} Treat_j + \beta_{\mathrm{PA3}} Time_{ij} + \beta_{\mathrm{PA4}} Treat_j Time_{ij}\\
+ \beta_{\mathrm{PA5}} Subgroup_j + \beta_{\mathrm{PA6}} Treat_j Subgroup_j\\
+ \beta_{\mathrm{PA7}} Time_j Subgroup_j + \beta_{\mathrm{PA8}} Treat_j Time_{ij} Subgroup_j\\
+ \sum_h \beta_{PAh}\, z_{ij}
\end{aligned}
$$

In the unit-specific model described in equation (4.3), β_2 can be interpreted as the response in outcome y_{ij} to a change in covariate x_{ij} for a particular individual i, being the effect of x_{ij} for an observation having the same random effect u_j . In the population-averaged model, the interpretation of $\beta_{2\mathrm{PA}_j}$ is very different, being the average effect of a change in covariate x_{ij} across the population on the response in outcome y_{ij} (Zorn 2001, p. 474).

Thus the subject-specific approach has the benefit of explicitly modelling how clustered observations such as repeated measurements are correlated. However the population-average approach is useful for examining how groups differ in response to a change in the covariate(s). In linear models, estimates from either approach would be similar; however in models where the linear model is estimated through a non-linear link function (such as the natural log) this is no longer the case. The random effects and coefficients in equation (4.5) are on the log scale. When re-transforming logged expectations of the mean response conditional not only on the model covariates but also the random effects, through the application of the exponent, then a random effect will have a mean of 1 rather than 0 (see assumptions for the distribution of random effects above). Thus it is non-ignorable when interpreting the estimate coefficients (Baldwin, Fellingham, and Baldwin 2016, Heagerty and Kurland 2001, Rabe-Hesketh and Skrondal 2012).

4.19.6 Marginal Effects

I explored the marginal effects of treatment allocation on expenditure at baseline and follow-up by index condition (telehealth) or living arrangement (telecare) in the models (which in each case included the DDD interaction term for index condition or living arrangement, time and allocation). Measuring marginal effects is to measure the change in the conditional mean

of a dependent variable when an independent variable changes by a unit (Cameron & Trivedi, 2005). Marginal effects were calculated at the average response of all cases (average marginal effects, or AME). As implemented in Stata software, standard errors of the AME conditional on covariates were calculated using the delta method (StataCorp 2015b). In the two-part population-averaged models (estimated by generalised estimating equations) (see below), marginal effects of gamma and logistic regressions were estimated by bootstrapping together their recycled predictions, clustering on subject.

4.19.7 Two-part Models

A two-part approach was employed to substantiate inferences about social and hospital care costs, given substantial zero costs for these service categories. Such data can be viewed as being semi-continuous or having a mixture of distributions (Liu et al. 2010). Leaving aside considerations of clustering for now, the first part of the two-part approach (Cameron and Trivedi 2005) consists of a model of binary probability of 'participation', so that the outcome (*d*) is observed only for participants, the outcome for non-participants being zero. In the second part, the conditional density of the dependent variable for participants (y>0), is $f(y|d=1)$ for a given density $f(\cdot)$. This mixture model can be stated as

$$f(y|\mathbf{x}) = \begin{cases} \Pr[d=0|\mathbf{x}] & \text{if } y=0, \\ \Pr[d=1|\mathbf{x}]\, f(y|d=1,\mathbf{x}) & \text{if } y>0. \end{cases} \qquad (4.10)$$

The two-part model allows the mechanisms driving the zero (first) and non-zero (second) parts to differ (Cameron and Trivedi 2005 p.544): in the present case, the reason for receipt or non-receipt of a service does not have to be the mechanism driving the quantity of service consumed.

The model is estimated by maximum likelihood. The first part may be modelled as a probit or logit regression; the second part fits the data to a positive values-only distribution such as the log-normal (Duan et al. 1983) or gamma. Modelling costs with the gamma distribution has several useful properties. The response variable does not require transformation, as in the log-normal model. Retransformation of the logged response variable in the presence of heteroscedastically distributed errors can lead to biased estimates (Mullahy 1998). Exponentiating the coefficients of a log-normal model also produces geometric mean estimates rather than the arithmetic mean estimates of the log-gamma model (Baldwin, Fellingham, and Baldwin 2016).

While it is possible that the reason for receipt and the quantity consumed if received may not be linked, in the context of a within-cluster and within-subject framework, the assumption appears strong (Liu et al. 2010). Multilevel two-part models have been proposed in more recent years that make use of the generalised gamma distribution in the second part that allow for correlation between the random effects of the first and second part, or "cross-part" correlation (Liu et al. 2010, Lee et al. 2010, Baldwin, Fellingham, and Baldwin 2016). If the outcomes of the first and second parts are correlated but their covariance is assumed to be zero, this can bias the second part estimates (Lee et al. 2010). While the equation for the second-part models has already been described (equation (4.1), the first-part subject-specific model of receiving care services can be written as a generalised linear multilevel model (GLMM) (in the three-level case) (cf. Rabe-Hesketh and Skrondal 2012):

$$y_{ijk}|\pi_{ijk} \sim \text{Binomial}(1, \pi_{ijk}),\ \text{logit}(\pi_{ijk}) \tag{4.11}$$
$$= \beta_1 + \beta_2 Treat_k + \beta_3 Time_{ijk} + \beta_4 Treat_k Time_{ijk}$$
$$+ \beta_5 Subgroup_{jk} + \beta_6 Treat_k Subgroup_{jk}$$
$$+ \beta_7 Time_k Subgroup_{jk} + \beta_8 Treat_k Time_{ijk} Subgroup_{jk}$$
$$+ u_{jk}^{(2)} + u_k^{(3)}$$

where $\pi_{ijk} \equiv \Pr\left(y_{ijk} = 1 \middle| \mathbf{x}_{ijk}, u_{jk}^{(2)}, u_k^{(3)}\right)$.

And in the two-level case:

$$y_{ij}|\pi_{ij} \sim \text{Binomial}(1, \pi_{ijk}),\ \text{logit}(\pi_{ij}) = \beta_1 + \beta_2 Treat_j + \beta_3 Time_{ij} + \tag{4.12}$$
$$\beta_4 Treat_j Time_{ij} + \beta_5 Subgroup_j + \beta_6 Treat_j Subgroup_j +$$
$$\beta_7 Time_{ij} Subgroup_j + \beta_8 Treat_j Time_{ij} Subgroup_j + u_j\ ,$$

where $\pi_{ij} \equiv \Pr(y_{ij} = 1 | \mathbf{x}_{ij}, u_j\ )$.

The random effects are assumed to be independent across clusters, and the responses have a Bernoulli distribution.

The population-averaged case is as equation (4.12) but $\pi_{ij} \equiv \Pr\left(y_{ij} = 1 | \mathbf{x}_{ij}\right)$.

$$y_{ij}|\pi_{ij} \sim \text{Binomial}(1, \pi_{ijk}),\ \text{logit}(\pi_{ij}) = \beta_1 + \beta_2 Treat_j + \beta_3 Time_{ij} + \tag{4.13}$$
$$\beta_4 Treat_j Time_{ij} + \beta_5 Subgroup_j + \beta_6 Treat_j Subgroup_j +$$
$$\beta_7 Time_{ij} Subgroup_j + \beta_8 Treat_j Time_{ij} Subgroup_j,$$

As previously explained, random-effects logistic models will have a subject-specific interpretation, as the log-odds of receipt will be conditional on not only the covariates but also the random effects in the model. However for binary response variables, marginal probabilities estimated from subject-specific models (by integrating over the random intercept distribution) will be similar to those estimated from population-averaged models (Hu et al. 1998, Fieberg et al. 2009). In the case of logistic-normal models, it is also possible to approximate the population averaged estimate of a coefficient of a conditional model by applying the formula (Hu et al. 1998, Zeger, Liang, and Albert 1988):

$$\beta_{PA} \simeq \beta_{SS}/\sqrt{1 + 0.346\sigma_u^2} \, .$$

An attractive feature of the subject-specific multilevel two-part model is the ability to estimate the correlation between the first and second parts of the model. Thus it is possible to examine the extent to which the likelihood of receipt is related to the cost of the services consumed. More formally, in the three-level case, the random effects of the first (u_{1j}) and second parts (u_{2j}) are assumed to have a multivariate normal distribution with a vector of zero means and variance/covariance matrix (Baldwin, Fellingham, and Baldwin 2016, Lee et al. 2010):

$$\begin{bmatrix} u_1^{(3)} \\ u_2^{(3)} \end{bmatrix} \sim \text{MVN}(0, \Sigma_3) \sum\nolimits_3 = \begin{bmatrix} \sigma^2_{u_1^{(3)}} & \\ \sigma_{u_1^{(3)}}\sigma_{u_2^{(2)}} & \sigma^2_{u_2^{(3)}} \end{bmatrix}$$

$$\begin{bmatrix} u_1^{(2)} \\ u_2^{(2)} \end{bmatrix} \sim \text{MVN}(0, \Sigma_2) \sum\nolimits_2 = \begin{bmatrix} \sigma^2_{u_1^{(2)}} & \\ \sigma_{u_1^{(2)}}\sigma_{u_2^{(2)}} & \sigma^2_{u_2^{(2)}} \end{bmatrix}$$

The vector of random effects of each part $\begin{pmatrix} u_1^{(3)} \\ u_1^{(2)} \end{pmatrix}$ and $\begin{pmatrix} u_2^{(3)} \\ u_2^{(2)} \end{pmatrix}$ are assumed to be independent.

In the two-level case, the random effects are also assumed normally distributed, with variance-covariance matrix:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim \text{MVN}(0, \Sigma) \sum = \begin{bmatrix} \sigma^2_{u_1} & \\ \sigma_{u_1}\sigma_{u_2} & \sigma^2_{u_2} \end{bmatrix}$$

The correlation of the random effects, ρ_{12}, is the covariance $\sigma_{u_1}\sigma_{u_2}$, divided by the product of the standard deviations of the random effects σ_{u_1} and σ_{u_2}.

4.19.8 Methods of Model Testing and Comparison

It is not possible to employ standard tests of fit such as the log-likelihood ratio test and AIC and BIC statistics within the framework of analysing multiply-imputed datasets. Instead, tests to evaluate whether the random intercepts for each level were jointly equal to zero were carried out using the conditional test of Li et al. (1991), cited by StataCorp (2015b). Whether other joint effects (e.g. of living together and intervention allocation, separately and together with their interaction term) were equal to zero were similarly tested.

4.19.9 Analyses

All analyses for this section were carried out in Stata 14 (StataCorp 2015b), using the <<*meglm*>>, <<*gsem*>> and <<*margins*>> commands. The <<*meglm*>> command fits generalised linear mixed models by maximum-likelihood, implemented by mean-variance adaptive Gauss-Hermite quadrature (StataCorp 2015a). The <<*gsem*>> command, which allows the estimation of systems of equations within a multilevel framework, approximates maximum likelihood in the same way. This estimation technique is used to evaluate marginal likelihood in GLMM models, necessitated for integrating out random effects in the absence of a closed-form solution (Rabe-Hesketh and Skrondal 2012). All results are reported for imputed data (see section 4.22).

4.19.10 Dependent Variables Included in the Analyses

Telehealth and telecare study data were examined separately. In terms of dependent variables, I examined first all health and social care costs, with and without the direct costs of the intervention (see Table 4.2 for the units contributing costs to the total costs). Second, I considered costs to the NHS of primary, secondary and community health care; to secondary care only; and social services. Table 4.2 lists the units contributing costs to each care category. For the purposes of examining the social care costs (excluding the costs of the intervention) in the telecare sample, the costs of community alarms were excluded from the calculations: telecare recipients would naturally report having this item as part of their telecare package (see Sections 4.2 and 5.16.3), and the receipt of care would be confounded with allocation to the intervention. All service costs were assumed to have been incurred by public agencies. The analyses focused on data collected at baseline and 12-month follow-up (see section 4.11).

4.19.11 Telehealth Dataset: Variables Used as Covariates

It is important, given the multilevel structure of the model, to characterise model covariates as belonging to a model level. For instance, in this time-varying model, those variables that change over time are characterised as being 'first level'. As well as the baseline/follow-up indicator, there were two first level variables, for self-care and tenure.

Self-care domain of the EQ5-D: This is an ordinal variable, where 1 is "I have no problems with self-care"; 2 is "I have some problems washing or dressing myself" and 3 is "I am unable to wash or dress myself". The variable was used as proxy measure of need for social care.

Tenure category: A categorical variable derived from the original tenure variable. Respondents were asked to indicate whether their accommodation was council-rented, housing association-rented, privately rented, owner-occupied or they could specify 'other' accommodation situations. Free-text answers were recoded into one of five categories: (1) council rented; (2) rented from housing association, registered social landlord or charitable trust; (3) privately rented; (4) owner occupied, shared ownership or equity release; (5) a category for all other housing types that could not be classified into the standard categories (e.g. family member's home, mobile home or temporary accommodation). This variable was recoded into two categories: (1) all types of rental and other types of accommodation (including living with relatives, temporary accommodation); and (2) owner-occupiers (including shared ownership and equity release).

Other demographic covariates are second-level variables, varying only at the person-level.

Age: This variable was categorised into 4 age-bands: under 65 or 'young' (reference category); 65-74 ('young old'); 75-84 ('old old'); and 85 years and over ('oldest old'). Another formulation, a continuous variable and its quadratic, was considered but initial explorations of model fit indicated that this performed less well. These categorisations have been long-established in the sociological literature on ageing (Suzman, Willis, and Manton 1992) and are useful in terms of formulating policy recommendations for different age cohorts.

Comorbidities: This is a count of chronic conditions sourced from acute hospital records (Steventon et al. 2012), treated as a continuous variable, and grand-mean-centred.

IMD: A continuous variable for Index of Multiple Deprivation 2007 (Noble et al. 2008), grand-mean-centred.

Educational attainment: A three-category variable for education was constructed from the existing variable for educational attainment in the dataset. The categories were no formal education (reference category); having O levels/GCSE or A levels; degree (undergraduate or graduate level).

Female: A dichotomous variable for female (1) and male (0).

Ethnicity: A dichotomous variable was constructed, coding ethnicity as White-British/non-White British.

Index conditions: The three index long-term conditions – COPD (reference category); heart failure; diabetes.

Two variables could be considered to vary at a higher level of organisation than the person-level.

Allocation: Allocation to intervention or control. As the general practices were the unit of randomisation, the allocation variable could be considered a level 3 predictor. However the variable is not an attribute of general practices as such and could also be considered a level 2 predictor.

Site: identifies the participating local authority (1, 2 or 3). While technically general practices are nested within local authority areas, sites have been treated here as a fixed effect.

4.19.12 Telecare Analyses: Variables Used as Covariates

These analyses focused on the costs for two groups: people living alone and people living with others. I outline the reasons for choosing to explore these subgroups' costs in chapter 6. Covariates included in the telecare models were first-level, time-varying variables: baseline/follow-up indicator; self-care domain of the EQ-5D-3L, and tenure, as described in Section 4.21.11. One additional covariate was also included:

Living arrangement: A dichotomous variable for living alone or with others was derived from two variables: number of adults in the household and number of children of 16 years or younger in the household. All participants reporting one adult and no children in the household were assumed to be living in a one-person household and therefore living alone; all other living arrangements were classified as living with others. The participant questionnaires did not contain an explicit question about the relationship between participants and other residents of their household, nor on marital status. There was no way to create a variable that expressed both numbers and relationships between residents in the household from the available data. Living alone was coded as 0; living together was coded as 1.

Several demographic covariates were second-level variables, varying only at the person-level. These included age bands, comorbidities, IMD, educational attainment, sex and ethnicity (as described in Section 4.21.11).

Comorbidities: This is a count of chronic conditions sourced from acute hospital records (Steventon et al. 2012), treated as a continuous variable, and grand-mean-centred.

IMD: A continuous variable for Index of Multiple Deprivation 2007 (Noble et al. 2008), grand-mean-centred.

Educational attainment: A three-category variable for education: no formal education (coded as zero); below degree (having O levels/GCSE or A levels); degree (undergraduate or graduate level; the reference category).

Sex: A dichotomous variable for female/male.

Ethnicity: A dichotomous variable (White-British/non-White British).

As in 4.19.11, two variables could be considered to vary at a higher level of organisation than the person-level.

Allocation: Allocation to intervention or control. The allocation variable could be considered a level 3 or a level 2 predictor.

Site: identifies the participating local authority, treated as a fixed effect.

4.20 Cost-effectiveness Analyses

4.20.1 Decision Rules for Cost-effectiveness

The relationships between the costs and outcomes of telehealth and telecare were examined in terms of the incremental cost-effectiveness and net monetary benefit of the respective interventions.

The incremental cost-effectiveness ratio (ICER) is the difference in the mean costs of the intervention and control groups ($\triangle C$) divided by the difference in the mean outcomes between the groups ($\triangle E$). The intervention will be seen as cost-effective if the ICER is less than some maximum amount (λ) that a societal decision maker/purchaser is willing to pay for a gain in outcome. This decision rule can be expressed as

$$\triangle C/\triangle E < \lambda$$

where λ represents willingness to pay (WTP) for the gain in outcome. This equation can be re-arranged to give the net monetary benefit (NMB), representing the pecuniary value of the additional gain in benefits associated with the intervention, for a given λ, net of the

additional cost of providing the intervention (Drummond et al. 2005). If the intervention is to be cost-effective, then the NMB must exceed zero:

$$\lambda \times \triangle E - \triangle C > 0$$

The National Institute for Health and Care Excellence (NICE) has generally considered that for adoption by the NHS, technologies should cost in the region of the £20,000 to £30,000 per QALY range (National Institute for Health and Clinical Excellence 2008, Cerri, Knapp, and Fernandez 2013).

A range of values, from £0 to £90,000, of willingness to pay for additional benefit were considered, including the £20,000 to £30,000 per-QALY range; these results were plotted to produce cost-effectiveness acceptability curves (CEACs). CEACs are useful as they provide decision makers with a visual representation of the probability that the intervention is effective and the error probability associated with that level of WTP (Drummond et al. 2005, Drummond et al. 2015).

4.20.2 Cost-effectiveness Analyses

The regression-based approach allows sampling uncertainty to be taken into account, and adjustment for any between-group baseline differences in individual characteristics (e.g. socio-demographic differences) (Drummond et al. 2005, Hoch, Rockx, and Krahn 2006).

It was agreed across the three quantitative themes of the WSD study that a core of baseline characteristics would be included in the analyses of the Telecare and Telehealth data, namely: age, gender, deprivation as measured by the Index of Multiple Deprivation (IMD), ethnicity, index condition (Telehealth participants only), site and presence of comorbidities. All covariates included in the cost-effectiveness models for Telecare and Telehealth are listed in the following sections.

Any concerns about the variability of telehealth and telecare interventions are likely to be magnified when comparisons are made across localities or centres (Bergmo 2009). Manca, Rice et al. (2005) suggest an analytical approach to the variations in costs related to the locality, or cluster. In their paper on the subject of generalising cost-effectiveness results from multi-centre, cluster-randomised trials, they note that a correlation of costs and outcomes may occur because of variations between locations, or clusters. They advocate the use of multilevel models (MLM), which can tackle the problem of variability between sites simultaneously in terms of resource use and costs and outcomes, and which could allow analysts to quantify how generalisable by location the cost-effectiveness results are. There are

alternative approaches. Seemingly unrelated regression models were fitted to both telecare and telehealth data. This is a system of equations that takes account of the correlation between error terms of the cost and outcome equations (Gomes, Ng, et al. 2012).

4.20.3 Telehealth

Telehealth data were analysed in Stata 14, using *gsem*. The models adjusted for a number of socio-demographic characteristics: site, age, sex, ethnicity, IMD score (Noble et al. 2008) and two indicators of health need: a comorbidity count constructed from a range of chronic conditions, sourced from acute hospital records (Steventon et al. 2012) and the index long-term condition (Bower et al. 2011). Baseline costs were also included in cost equations and baseline utility or baseline secondary outcome measures (Manca, Hawkins, and Sculpher 2005) were included in the outcome equations.

4.20.4 Telehealth Sensitivity Analyses

Decreases in the costs of equipment: Equipment costs might have a considerable effect on the overall costs of telehealth, and conclusions about cost-effectiveness could depend on the unit cost of equipment use. Equipment prices may fall over time as technology evolves. I explored the effect of falling input prices, using data obtained from the Department of Health on equipment prices in North American markets in 2010. I applied general price decreases of 50 per cent and 80 per cent to equipment costs calculated for the trial. Because North American equipment prices were 10-50 per cent of the price for equipment purchased in England before the trial, these assumptions were relatively conservative.

'At capacity' scenario: Telehealth teams may have been able to work at a higher capacity. The sites had originally expected to have about 1000 users each, at least for a few months over the trial period (as those allocated to the telehealth intervention were gradually joined by people who had been allocated to the control group). Instead, the monitoring teams had somewhere between half and three-fifths of the target number in 2009-10. The sensitivity analyses explored costs if each site monitored 1000 people, assuming that the central monitoring teams did not increase the number of staff available to handle additional demand, and that both service structures and participant outcomes did not change at the larger scale of the service.

The equipment costs and telehealth support costs parameters were varied as described in the two scenarios above for use in the cost-effectiveness analyses; models and covariates remained the same as in the main analyses.

4.20.5 Telecare

For the telecare data, the SUR models were fitted in Stata 14 (StataCorp 2015b) using the *gsem* command. Covariates used were pre-baseline costs and baseline utility (following Manca, Hawkins, and Sculpher 2005)/outcome measure, site, age, sex, ethnicity, Index of Multiple Deprivation (IMD) 2007 quintiles (Noble et al. 2008), a one-person household indicator, a count of chronic conditions (Steventon et al. 2013), level of dependency (baseline EQ5-D self-care domain score) and whether the participant had a 'personal/community alarm' at baseline. The self-care domain score was included as a covariate as much of the variation in receipt of social care is linked to the degree of difficulty individuals experience in tasks such as washing and dressing; self-care is highly correlated with the need for support in activities of daily living (ADL) (Forder and Caiels 2011). The presence of a personal/community alarm was included as an indicator that the local authority, the family or the individual considered the individual to be at risk. Cluster-robust standard errors were used in estimating regression coefficients (observations were clustered by general practice). Estimates from the results were used to calculate the ICER and the net monetary benefit (NMB) using Fieller's rules (Fieller 1954, O'Brien and Briggs 2002, Glick 2007). The formula for calculating the ICER point estimate and confidence intervals and the NMB point estimates and confidence intervals at different levels of willingness to pay employed the model estimates of the covariance between costs and outcomes, and the coefficients on the intervention variable in the cost and outcomes equations and their variances.

4.20.6 Telecare Sensitivity Analyses

The assumptions of the analyses of the primary outcome were subjected to further testing. One analysis explored the robustness of the base case to variations in the costs of the intervention by using a lower cost of telecare monitoring support, if telecare could be delivered for the cost of a service operating at scale (a 'mainstream' service). A study by Bayer and Barlow (2010) provided a figure of £5 per week. A second analysis examined the input prices for telecare equipment if that equipment had been purchased at half the price paid within the trial, given that with advances in technology, telecare equipment prices might

fall substantially. A third analysis examined the impact of assuming that costs and outcomes were normally distributed in the base case analysis. This involved performing the SUR analyses on 3000 bootstrapped replications generated using a two-stage bootstrap procedure in R, as described in Gomes, Ng et al (2012). SUR was carried out in R using the `systemfit` package (Henningsen and Hamann 2007). The procedure and analyses were performed separately on each complete dataset generated by the multiple imputation process (see 4.24) and estimates combined in NORM (NORM: Multiple imputation of incomplete multivariate data under a normal model (Version 2) [Software] 1999).

4.21 Missing Data

There were inevitably data missing at all assessment points. CSRIs that were not returned because the participant did not complete the assessment were considered to be missing and were not included in any imputations or analyses. When the CSRI section of the questionnaire pack was not filled out at all or just one or two answers given, data for the whole case was also considered as entirely missing. Apart from these cases, service costs were calculated for each participant. If the use of a particular service was indicated on the CSRI, but not the frequency of use of that service, then a per-contact unit cost was used to value that item (rather than a duration-based unit cost). Where no such unit cost was available, the mean duration of users who reported frequency of use was calculated and the mean duration applied to those missing information on duration of use; a mean cost of users was calculated, attaching duration-based unit costs to frequency, and this mean cost was allocated to cases where only use of the service was indicated. Service use costs were aggregated up to cost categories (see Table 4.2) which were in turn added together to produce the total cost. Where all individual costs making up a cost category were missing, the total for the cost category was calculated as missing whereas if only some of the costs were missing, then these were assumed to be zero costs and the total cost represented as the sum of the available costs.[17] If the case was not entirely missing, but category-level costs were missing, these costs were derived through multiple imputation.

[17] In Stata, the command used was: *egen newvar=rowtotal (var1, var2,…varx), m*

4.22 Multiple Imputation

The imputation models used for the costs and cost-effectiveness analyses drew on a number of predictors. In the case of telehealth, predictors included socio-demographics (age, sex, education, ethnicity, number of comorbidities, site, Index of Multiple Deprivation score, long-term condition indicator, general practice, household size, housing tenure), outcome and process measures (EQ-5D-3L, ICECAP-O, SF-12, depression symptoms, state-trait anxiety, self-efficacy) and category-level and total costs at all assessment points, as well as variables related to the trial itself (allocation, reasons for withdrawal). For telecare, predictors included: socio-demographics (as for telehealth), outcome and process measures (EQ-5D-3L, SF-12, ICECAP-O, state-trait anxiety, self-efficacy) and category-level and total costs at all assessment points, and variables related to the trial (allocation, reasons for withdrawal, function/classification of telecare equipment).

For both telehealth and telecare datasets, missing data were imputed using the data available from cases that had completed an assessment at that time point. Thus no data was imputed where the case was fully missing because the assessment had not been completed (for any reason) at that time point.

Data were assumed to be missing at random, which means that missingness is dependent on the observed values of the data and not on the unobserved values of the data (Little and Rubin 2002). Multiple imputations of costs and outcomes data for the telecare analyses in chapters 6 and 8 was carried out via predictive mean matching (PMM) using the MCMC procedure in SPSS v.21 (SPSS Inc., Chicago, Illinois).

Imputation of data for the telehealth analyses (work carried out in 2017) presented in chapters 6 and 7 was conducted within a multilevel framework. Multiple imputation by PMM (with 5 nearest neighbours) was implemented using the `mice` (Van Buuren and Groothuis-Oudshoorn, 2011) v. 2.25 and `miceadds` (v2.7-19) packages in R {R Core Team, 2016 #1195}. The process generated 10 completed datasets with a maximum of 15 iterations for the convergence of the MICE (Multivariate Imputation by Chained Equations) algorithm. In this approach, a linear mixed model (implemented using the mice.impute.2l.lmer function, itself dependent on the `lme4` R package) (Bates et al. 2015) serves as the basis of the imputation model; values predicted by the model are matched to the nearest observed value in the data. The multilevel approach explicitly accounts for the relationship between a predictor and an outcome at both the individual and at the cluster level (Ludtke, Robitzsch, and Grund 2017). The approach also recognises the possibility that missingness in an outcome may be

linked to the clustering unit, for instance because cluster members share common characteristics or because of specific attributes of that cluster (Gomes et al. 2013). Missing data were also imputed separately by experimental group, following Diaz-Ordaz, Kenward, and Grieve (2014) and Gomes et al. (2013). In each case, the procedure was followed to create ten datasets (Schafer 1999), to be analysed and then combined, taking the stochastic nature of the imputations into account (Carpenter and Kenward 2007, Rubin 1987).

4.23 Ethics

The study was given approval by the Liverpool NHS Research Ethics Committee (ref: 08/H1005/4).

Chapter 5

Telehealth and Telecare Samples: Characteristics, Service Use Patterns and Costs

5.1 Introduction

In this chapter I present results in terms of the availability of data from each trial at the three study assessment points. I give an overview of the baseline characteristics of the two trial populations and compare the samples completing the final assessment point with those who did not complete. I also present a comparison of the resource use and costs of the experimental groups at the three trial assessment points. Costs of self-reported service use over the last three months before each assessment point are reported by category; mean values summarise the (imputed) costs of the available cases, while the standard errors are adjusted for cluster. The discussions throughout this chapter focus primarily on the baseline and 12-month follow-up results. Issues with 4-month data were discussed in section 4.11; baseline characteristics of participants completing/not completing at 4 months are discussed in the chapter, and 4-months service use and costs results from the 4-months data are presented in Appendix 2.

5.2 Telehealth

There were 3230 participants from 238 general practices in the WSD telehealth trial. Of these individuals, 1573 (patients of 154 general practices) participated in the WSD telehealth questionnaire study: 845 (81 practices) were allocated to the intervention and 728 (73 practices) to usual care. Seventeen participants allocated to usual care received the intervention, and six allocated to the intervention did not receive any equipment.

Baseline data from the questionnaire study on services used were available for 841 telehealth and 728 usual care participants. At the 4-month follow-up, service use data were available for 972 participants (547 intervention and 425 control participants). At the 12-month follow-up, 969 (538 telehealth, 431 control) had service use data available. By 12-month follow-up, 599 (38 per cent) participants had dropped out of the questionnaire study, leaving outcomes data from 974 participants. Service use data from both baseline and 12-month follow-up were available for 965 participants (534 intervention; 431 control). Service

use data was available at all three assessment points for 743 participants (418 intervention; 325 control).

5.3 Socio-demographic Characteristics

Participants' characteristics at baseline assessment are presented in terms of completion of the baseline assessment and completion and non-completion of study instruments at the 12-month assessment (Table 5.1). The majority of the sample was male (59 per cent). While the mean age was 70.4, just under a third of participants were under 65 years of age. A large proportion (68 per cent) of the sample had at least one comorbidity. The sample was quite evenly distributed between the three participating local authorities, although a larger proportion (40 per cent) resided in site 2.

The experimental groups were broadly similar at the outset of the study, although a larger percentage of participants with heart failure were in the usual care group (38 per cent) than in the telehealth group (31 per cent). A larger percentage of patients with chronic obstructive pulmonary disease (40 per cent) were in the telehealth group than in usual care (34 per cent).

There were differences between telehealth and usual care participants in relation to the groups in the first and second IMD quintiles (the least deprived groups), although mean scores did not differ greatly. Within each experimental group, the baseline and 12-month follow-up samples were broadly similar in age, number of comorbidities, the proportion of females and proportions of participants with chronic obstructive pulmonary disease and heart failure. At baseline and follow-up, telehealth participants with an index condition of chronic obstructive pulmonary disease formed the largest group; in the usual care group, participants with an index condition of heart failure constituted the largest group.

Table 5.1 Baseline characteristics of participants with economic data available at baseline and 12-month follow-up across Telehealth sample

	Total baseline sample			Participants completing 12-month follow-up study instruments*			Participants not completing 12-month follow-up study instruments†		
	UC (n=728)	**TH (n=841)**	**Raw**	**UC (n=431)**	**TH (n=534)**	**Raw**	**UC (n=297)**	**TH (n=302)**	**Raw**
Mean years of age (SD)	70.6 (20.7)	70.1 (21.6)	-0.5	70.1 (16.1)	70.0 (17.0)	-0.1	71.3 (16.9)	70.5 (16.8)	-0.8
Under 65 (young)	215 (30%)	242 (29%)	-1%	131 (30%)	150 (28%)	-2%	84 (28%)	90 (30%)	2%
65-74 (young old)‡	214 (29%)	288 (34%)	5%	137 (32%)	199 (37%)	5%	77 (26%)	86 (28%)	3%
75-84 (old old)	239 (33%)	243 (29%)	-4%	130 (30%)	156 (29%)	-1%	109 (37%)	87 (29%)	-8%
85+ (oldest old)§	60 (8%)	68 (8%)	0%	33 (8%)	29 (5%)	2%	727 (9%)	39 (13%)	4%
Women	290 (40%)	347 (41%)	1%	162 (38%)	222 (42%)	4%	128 (43%)	124 (41%)	-2%
Mean IMD score (SD) ‖	28.6 (52.2)	27.7 (55.3)	–0.9	27.7 (40.1)	26.0 (43.8)	-1.7	29.8 (36.3)	30.6 (36.2)	0.7
1st quintile‖	130 (18%)	215 (26%)	8%	81 (19%)	151 (28%)	9%	49 (16%)	64 (21%)	5%
2nd quintile‖	164 (23%)	140 (17%)	-6%	105 (24%)	93 (17%)	-7%	59 (20%)	46 (15%)	-5%
3rd quintile‖	124 (17%)	155 (18%)	1%	79 (18%)	101 (19%)	1%	45 (15%)	53 (18%)	2%
4th quintile‖	168 (23%)	165 (20%)	-3%	87 (20%)	110 (21%)	0%	81 (27%)	54 (18%)	-9%
5th quintile‖	142 (20%)	166 (20%)	0%	79 (18%)	79 (15%)	-4%	63 (21%)	85 (28%)	7%

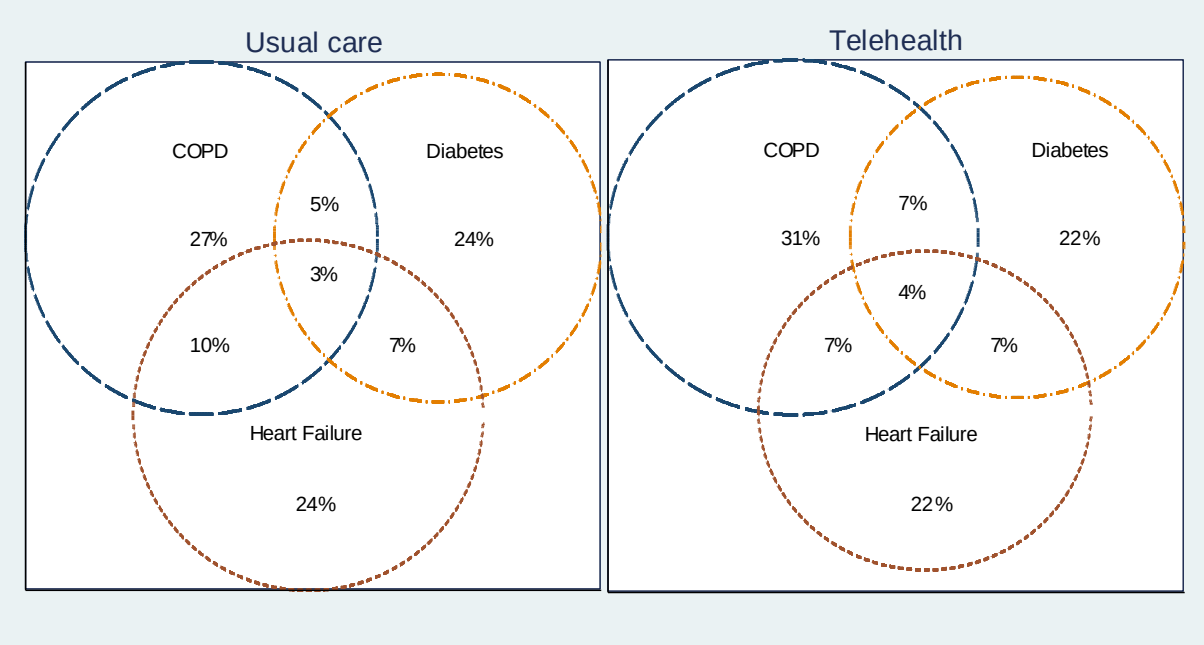
	Total baseline sample			Participants completing 12-month follow-up study instruments*			Participants not completing 12-month follow-up study instruments†		
	UC (n=728)	**TH (n=841)**	**Raw**	**UC (n=431)**	**TH (n=534)**	**Raw**	**UC (n=297)**	**TH (n=302)**	**Raw**
Index condition									
COPD	244 (34%)	334 (40%)	6%	140 (33%)	232 (43%)	11%	104 (35%)	99 (33%)	-2%
Heart failure	275 (38%)	263 (31%)	–7%	175 (41%)	177 (33%)	–8%	100 (34%)	86 (29%)	–5%
Diabetes	209 (29%)	244 (29%)	–0%	116 (27%)	125 (23%)	–3%	93 (31%)	117 (39%)	7%
1+ comorbidities	511 (70%)	560 (67%)	-4%	297 (69%)	345 (65%)	-4%	214 (72%)	211 (70%)	-2%
Mean no. comorbidities (SD)	2 (2.8)	1.8 (2.9)	–0.2	2 (2.7)	1.8 (2.9)	–0.2	2.1 (2.1)	2 (2.1)	–0.1
WSD site									
Site 1	234 (32%)	256 (30%)	–2%	132 (31%)	174 (33%)	2%	102 (34%)	81 (27%)	–8%
Site 2 ¶	283 (39%)	342 (41%)	2%	184 (43%)	236 (44%)	1%	99 (33%)	105 (35%)	1%
Site 3**	211 (29%)	243 (29%)	–0%	115 (27%)	124 (23%)	–3%	96 (32%)	116 (38%)	6%
White British ethnicity‖	630 (87%)	735 (87%)	1%	377 (87%)	478 (90%)	2%	253 (85%)	255 (84%)	1%
Living alone‖††	195 (27%)	229 (27%)	0%	119 (28%)	132 (25%)	-3%	76 (26%)	95 (31%)	6%

	Total baseline sample			Participants completing 12-month follow-up study instruments*			Participants not completing 12-month follow-up study instruments†		
	UC (n=728)	**TH (n=841)**	**Raw**	**UC (n=431)**	**TH (n=534)**	**Raw**	**UC (n=297)**	**TH (n=302)**	**Raw**
Owns‖	497 (68%)	569 (68%)	-1%	299 (69%)	373 (70%)	0%	198 (67%)	193 (64%)	-3%
Education									
No formal education‖	423 (58%)	501 (60%)	2%	235 (55%)	301 (56%)	2%	187 (63%)	198 (66%)	3%
GCSE/O/ A- level ‖‡‡	222 (31%)	247 (29%)	-1%	144 (33%)	169 (32%)	-2%	78 (26%)	75 (25%)	-1%
Degree‖	83 (11%)	93 (11%)	0%	51 (12%)	64 (12%)	0%	32 (11%)	29 (10%)	-1%

UC=usual care; TH=telehealth; COPD=chronic obstructive pulmonary disease; SD=standard deviation.
*cases where costs and outcomes data were available
† Outcomes instruments not completed and/or CSRI not completed
‡ Difference within TH: differences between completion/non-completion clustered χ^2= 4.591 and p<0.05
§ Difference within TH: differences between completion/non-completion clustered χ^2=14.456 and p<0.001
‖ Imputed data
¶ Difference within TH: differences between completion/non-completion *z*=-2.6641, p<0.01; difference within UC: differences between completion/non-completion: *z*=-2.5456, p<0.05
**Difference within TH: differences between completion/non-completion *z*=4.6633, p<0.001
†† Difference within TH: differences between completion/non-completion clustered χ^2= 4.464, p<0.05
‡ ‡ Difference within UC: differences between completion/non-completion clustered χ^2= 4.1405, p<0.05

While participants were recruited into the trial on the basis of having one of three 'index' conditions (COPD, diabetes or heart failure), some had more than one of these conditions. The proportion of the total questionnaire study sample at baseline with each long-term condition and the overlap between those conditions is illustrated for each experimental group in a proportional Venn diagram (Figure 5.1). It can be seen that in both groups about three-quarters of the sample had just one long-term condition. A slightly smaller proportion of participants with diabetes or with heart failure had only the one condition relative to participants with COPD. Very few had all three conditions.

Figure 5.1 Proportions of the Telehealth sample with a long-term condition, by experimental group



5.3.1 Characteristics of the Sample Completing and Not Completing the 12-month Follow-up

Comparing the characteristics of cases with data available at baseline and at 12-month follow-up within their ITT allocation-groups (Table 5.1), and taking clustering into account, there were no significant differences between most demographic characteristics at baseline and at follow-up within the control or intervention groups. There were, however, differences within the groups in the proportions from sites 2 and 3. Within each experimental group, the participants completing at 12 months had a significantly higher proportion of people living in site 2 than those not completing. Within the intervention group, the non-completing sample had a significantly higher proportion of participants living in site 3 than the completing

sample. Of participants allocated to the intervention, in the completing sample there was a lower proportion of participants living alone, a higher proportion in the young-old age band, a lower proportion in the oldest-old age band and a higher proportion of people with secondary-school qualifications (GCSE/O/A-levels) than in the non-completing sample. The characteristics of the whole sample completing at both time points were mostly similar to those not completing. A greater proportion of participants who completed the questionnaire study were in site 2 than was the case in the non-completers (43 per cent vs. 33 per cent; z=-3.716, P<0.001); conversely a smaller proportion of participants who completed the study were in site 3 than was the case in the non-completers (25 per cent vs. 35 per cent; z=4.509, P<0.001).

5.4 Cluster Numbers and Sizes

The clusters diminished in size between assessments and were slightly less well-balanced at the 12-month follow-up relative to baseline. Cluster size decreased from approximately 10 participants on average in both experimental groups at baseline, to 6 and 7 participants in the usual care and telehealth groups respectively (Table 5.2 and Table 5.3).

Table 5.2 Number and size of clusters, participants with economic data available in Telehealth sample

	Total baseline sample		Participants completing 12-month follow-up study instruments*	
	UC (N=73)	TH (N=81)	UC (N=69)	TH (N=76)
Cluster size	10 [1-44]	10.4 [1-48]	6.2 [1-26]	7 [1-32]

Note: Data are mean [min – max]
UC=Usual care; TH=Telehealth; COPD=chronic obstructive pulmonary disease.
*where costs and outcomes data were available

5.5 Characteristics of the Sample Completing and Not Completing the 4-month Follow-up

The characteristics of cases with data available at baseline and at 4-month follow-up are compared within their ITT allocation-groups in Appendix 2, Table A2.1. As with the 12-month follow-up, in each experimental group, the participants completing at 4 months had a significantly higher proportion living in site 2 than those not completing; also the participants in the usual care group completing at 4-months had a significantly lower proportion living in

site 3 than those not completing. In addition, telehealth group participants completing at 4 months had a significantly higher proportion living in site 1 than those not completing. The completing samples of telehealth participants differed significantly from the non-completing telehealth sample in terms of having lower proportions in the young-old, higher proportions in the oldest-old age bands (the completing telehealth sample was non-significantly older than the non-completing telehealth group). A smaller proportion of telehealth participants that completed had diabetes than telehealth participants not completing; the same pattern was observable in the usual care completing/non-completing samples. The completing/non-completing samples of telehealth participants also differed in terms of ethnicity, numbers in IMD and mean IMD score. Across the sample, a greater proportion of participants completing the questionnaire study were in site 2 than in the case of the non-completers (44 per cent vs. 33 per cent; z=4.412, p<0.001); and a smaller proportion of participants who completed the study were in site 3 than in the case of the non-completers (21 per cent vs. 41 per cent; z=8.357, p<0.001).

5.6 Telehealth Service Use: Descriptive statistics

Examining the data from 1569 cases that were available at the baseline (Table 5.3) and 965 cases at the 12-month follow-up time points (Table 5.4), mean reported service use was broadly similar between the telehealth and usual care groups. Due to the potentially large number of comparisons, between-group differences were not subjected to tests of significance. Proportions of the sample using services were generally greater in the follow-up sample. There was a broad pattern of slightly fewer reported mean contacts with hospital services for the telehealth group than for the usual care group at both assessment points, and these differences were generally larger at follow-up. In each group, the proportions using and the mean use of community matron and specialist nurse contacts was somewhat higher at follow-up, as were proportions and mean use of primary care (visiting the GP surgery and practice nurse contacts). The use of community care services was not negligible, particularly in terms of home care and home help, 15 and 14 percent using these services in control and intervention respectively; and 17 and 12 percent in control and intervention respectively at follow-up. 14 (16) per cent of the control and 13 (15) per cent of the intervention group used a community alarm at baseline (12-month follow-up).

Table 5.3 Number and percentage of groups using services and mean numbers of units (standard errors) used over previous 3 months across Telehealth sample, available cases at baseline

Resource item	Control (n=728)			Telehealth (n=841)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Hospital use							
A&E attendance	13	93	0.21 (0.03)	13	111	0.17 (0.02)	-0.04
Inpatient services	11	81	0.66 (0.11)	10	83	0.57 (0.10)	-0.09
Day Hospital attendances	6	42	0.24 (0.15)	7	61	0.22 (0.07)	-0.03
Outpatient attendances	50	359	1.56 (0.13)	49	410	1.25 (0.09)	-0.31
Community health services/primary care							
Paramedic contact	11	82	0.27 (0.06)	9	77	0.19 (0.04)	-0.08
Community matron visit	11	77	0.52 (0.08)	8	68	0.36 (0.06)	-0.16
Community matron telephone contact	3	22	0.11 (0.03)	2	13	0.06 (0.02)	-0.05
Community or district nurse visit	14	99	0.94 (0.22)	10	86	0.88 (0.28)	-0.05
Community or district nurse telephone contact	2	17	0.14 (0.09)	1	12	0.02 (0.01)	-0.11
Practice nurse visit	26	190	0.87 (0.11)	29	242	0.82 (0.08)	-0.05
Night nurse visit	0	2	0.00 (0.00)	0	2	0.00 (0.00)	0
Specialist nurse contacts	18	132	0.56 (0.26)	17	142	0.32 (0.05)	-0.24
Physiotherapist or occupational therapist visit	9	62	0.39 (0.08)	9	73	0.28 (0.04)	-0.11
GP (home) visit	13	95	0.27 (0.03)	11	95	0.26 (0.04)	-0.01
GP (surgery) visit	58	419	1.43 (0.08)	57	480	1.37 (0.07)	-0.06
GP (telephone) visit	11	83	0.25 (0.03)	12	101	0.25 (0.03)	0
Dentist visit	23	168	0.33 (0.03)	23	196	0.35 (0.03)	0.02
Chiropodist visit	23	166	0.43 (0.09)	23	196	0.35 (0.03)	-0.08
Optician visit	19	139	0.22 (0.02)	21	175	0.24 (0.02)	0.01

Resource item	Control (n=728)			Telehealth (n=841)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Community mental health							
Psychiatrist visit	1	6	0.01 (0.00)	1	8	0.04 (0.02)	0.03
Mental health nurse visit	1	6	0.02 (0.01)	1	5	0.01 (0.01)	-0.01
Community care services							
Social worker visit	3	19	0.07 (0.02)	4	31	0.07 (0.02)	0
All daytime home care/home help visit	15	112	7.42 (1.18)	14	115	5.17 (0.78)	-2.25
Council home help visit	6	45	5.38 (1.06)	5	42	3.47 (0.73)	-1.91
Private/independent home care/home help visit	11	77	2.05 (0.32)	10	80	1.71 (0.27)	-0.34
Paid night carer visit	1	4	0.39 (0.28)	1	8	0.46 (0.21)	0.07
Meals on Wheels meal	1	6	0.44 (0.23)	1	11	0.79 (0.28)	0.35
Laundry (incontinence) service	0	2	0.02 (0.02)	0	4	0.04 (0.02)	0.01
Community alarm	14	104		13	112		
Equipment inc. mobility aids, ADL NHS	1	8	0.02 (0.01)	0	3	0.01 (0.00)	-0.01
Major and minor adaptations NHS	1	6	0.01 (0.00)	0	3	0.00 (0.00)	-0.01
Equipment inc. mobility aids, ADL LA	2	15	0.03 (0.01)	1	12	0.02 (0.01)	-0.01
Major and minor adaptations LA	7	46	0.09 (0.01)	4	34	0.06 (0.01)	-0.03
Care home respite							
Days	0	1	0.03 (0.03)	0	0	0.00 (0.00)	-0.03
Day services	4	728	0.47 (0.12)	4	841	0.69 (0.18)	0.22
Day care and other day attendances - LA	3	25	0.44 (0.11)	3	29	0.59 (0.16)	0.16
Day care and other day attendances - NHS	0	3	0.03 (0.02)	1	6	0.10 (0.07)	0.07
Medications							
Number of medications	99	726	7.86 (0.13)	99	839	7.94 (0.13)	0.08

*Proportion of non-missing cases who reported using a service

Table 5.4 Number and percentage of groups using services and mean numbers of units (standard errors) used over previous 3 months across Telehealth sample, available cases at 12-month follow-up

Resource item	Control (n=431)			Telehealth (n=534)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Hospital use							
A&E attendance	20	85	0.38 (0.07)	14	73	0.23 (0.04)	-0.15
Inpatient services	15	63	1.23 (0.24)	11	56	0.98 (0.22)	-0.25
Day Hospital attendances	21	84	0.51 (0.12)	16	79	0.39 (0.10)	-0.13
Outpatient attendances	48	196	1.31 (0.13)	50	260	1.11 (0.08)	-0.19
Community health services/primary care							
Paramedic contact	12	46	0.18 (0.04)	10	47	0.13 (0.02)	-0.05
Community matron visit	13	51	0.76 (0.15)	14	67	0.70 (0.14)	-0.06
Community matron telephone contact	6	22	0.20 (0.04)	10	45	0.38 (0.10)	0.18
Community or district nurse visit	14	56	0.73 (0.25)	10	49	1.26 (0.74)	0.53
Community or district nurse telephone contact	5	19	0.14 (0.06)	7	34	0.24 (0.07)	0.1
Practice nurse visit	57	200	1.50 (0.15)	53	227	1.26 (0.11)	-0.24
Night nurse visit	0	1	0.00 (0.00)	0	1	0.01 (0.01)	0.01
Specialist nurse contacts	27	118	0.69 (0.10)	27	143	0.64 (0.08)	-0.05
Physiotherapist or occupational therapist visit	10	32	0.70 (0.30)	8	36	0.29 (0.08)	-0.41
GP (home) visit	16	48	0.37 (0.07)	10	35	0.23 (0.07)	-0.13
GP (surgery) visit	73	292	1.69 (0.08)	71	349	1.70 (0.10)	0.01
GP (telephone) visit	28	83	0.52 (0.06)	27	102	0.42 (0.04)	-0.1
Dentist visit	27	104	0.42 (0.06)	29	141	0.45 (0.04)	0.03
Chiropodist visit	28	106	0.61 (0.13)	28	133	0.60 (0.11)	-0.01
Optician visit	31	115	0.48 (0.09)	26	121	0.37 (0.04)	-0.11

Resource item	Control (n=431)			Telehealth (n=534)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Community mental health							
Psychiatrist visit	2	6	0.02 (0.01)	1	7	0.02 (0.01)	0
Mental health nurse visit	1	4	0.02 (0.01)	1	7	0.03 (0.02)	0.01
Community care services							
Social worker visit	6	22	0.35 (0.23)	5	26	0.16 (0.05)	-0.19
All daytime home care/home help visit	17	74	6.36 (1.40)	12	65	4.98 (1.50)	-1.38
Council home help visit	6	23	4.26 (1.29)	4	18	3.63 (1.57)	-0.62
Private/independent home care/home help visit	15	59	2.90 (0.81)	11	53	1.77 (0.36)	-1.13
Paid night carer visit	1	3	0.19 (0.11)	1	6	0.40 (0.24)	0.21
Meals on Wheels meal	2	6	0.45 (0.26)	1	5	0.65 (0.46)	0.2
Laundry (incontinence) service	0	0	0.00 (0.00)	1	5	0.05 (0.03)	0.05
Community alarm	16	65		15	79		
Equipment inc. mobility aids, ADL NHS	4	17	0.09 (0.03)	5	24	0.11 (0.03)	0.02
Major and minor adaptations NHS	1	3	0.01 (0.00)	1	4	0.01 (0.00)	0
Equipment inc. mobility aids, ADL LA	3	12	0.05 (0.01)	3	15	0.04 (0.01)	-0.01
Major and minor adaptations LA	3	10	0.04 (0.01)	4	21	0.06 (0.01)	0.02
Care home respite							
Days	0	1	0.02 (0.02)	0	1	0.03 (0.03)	0
Day services							
Day care and other day attendances - LA	5	23	0.58 (0.18)	2	13	0.41 (0.16)	-0.17
Day care and other day attendances - NHS	0	1	0.00 (0.00)	0	2	0.03 (0.02)	0.03
Medications							
Number of medications	100	315	8.57 (0.23)	100	411	8.64 (0.20)	0.07

*Proportion of non-missing cases who reported using a service

5.7 Telehealth Intervention

In each site, participants received telehealth equipment that consisted of a freestanding base unit or a television set-top box and 'peripherals' (monitors such as pulse oximeters, blood pressure cuff, glucometers and weigh scales). Sites provided monitors based on condition, their local protocols and clinical assessment. Three of the peripherals were considered 'critical' to a particular index condition (glucometer to diabetes; pulse oximeter to COPD; weighing scales to heart failure) and thus would be routinely provided for that condition. However clinicians could over-ride provision of the peripheral if they judged it to be clinically inappropriate for a particular patient (Cartwright et al. 2013). The peripherals sent vital signs data to the base unit/set-top box wirelessly or by cable. In each site, patient data thus collected was uploaded to a server. Computer algorithms then compared patients' readings to their baseline clinical parameters and classed them according to a risk-rating system. Each weekday, the readings would be reviewed by nursing staff who could then respond in various ways: for instance by further monitoring with no immediate action, contacting the patient for further discussion, contacting the patient's GP, or contacting the emergency services. The configurations of services in place to manage and respond to the received and processed vital signs data (or 'alerts') varied considerably between sites. The interventions, although varying between sites, could be classified as second-generation telehealth (Cartwright et al. 2013).

5.7.1 Description of Telehealth Support Services

Telehealth service configurations have been described in Table 5.5 in terms of equipment supply; installation, server and equipment maintenance, asset management and training; and monitoring and responses to alarm and sensor alerts. All sites had equipment supplied by private companies but only site 2 followed a rental model. The structures and processes in place for project management varied across sites. The local authority in site 2 provided both installation technicians and back-office support staff for non-clinical problems (e.g. troubleshooting for equipment malfunction), whereas these aspects were entirely managed within the NHS in site 1. In all cases there was a call centre staffed by nurses and support workers; in two sites, other community-based nurses had access to telehealth data. While the trial objective was to examine the impact of 'telehealth' per se rather than specific models, the variations in service configurations were reflected in the site-specific service costs.

Table 5.5 Features of delivery systems in the WSD Telehealth pilot sites, 2009/10

Producers / **Roles**	**Site 1**	**Site 2**	**Site 3**
Equipment supply	Equipment was procured from several suppliers.	Most telehealth equipment was procured from one supplier.	The base telehealth units were rented along with combinations of peripherals as a monthly package.
Installation, server and equipment maintenance, asset management, training	Installations, maintenance and asset management were the responsibility of local engineers employed by the primary care organisation; during installation engineers were accompanied by support staff from the central team who demonstrated how to use the equipment.	Technicians from the local authority carried out installations, maintenance and asset management, and provided users with an initial tutorial on use of the equipment.	Equipment was supplied, installed and configured by a private company that also trained patients in its use.
Monitoring and responses to alarms/sensors alerts	Telehealth data were reviewed by a central clinical team of nursing and support staff, and also by community matrons or specialist nurses based in community health settings. The central team also followed up non-clinical issues, e.g. where no data had been transferred by the user for some days, or in case of equipment or software problems.	Arrangements for clinical monitoring differed across the two Primary Care Trust areas within the participating local authority. Telehealth data in both areas could be monitored by call centre teams of qualified and unqualified nursing personnel. Data was also be monitored by specialist community nurses (specialist community matrons, community matrons or specialist nurses in one Trust; community matrons, but not specialist nurses, in the other). The telehealth core team (run by the local authority social services department) followed up non-clinical issues with users.	Telehealth data were reviewed by a central clinical team with nursing and support staff; also the team followed up non-clinical issues.

5.8 Telehealth Equipment

Telehealth participants received 2.8 (SD 0.6) items of telehealth 'peripheral' equipment on average (Table 5.6). Combinations of peripherals were common: for instance, 43 per cent received a combination of blood pressure monitor, pulse oximeter and weighing scales. Thus most participants (87 per cent (745/856)) received more than one item. There were striking variations in the distribution of the monitors between sites (Figure 5.2). The proportion of patients receiving equipment in Site 1 varied more distinctly depending on the index condition than in the other two sites: for instance, very small proportions of patients with COPD and heart failure received a glucometer. In all sites and across the conditions, almost all those in receipt of telehealth devices had blood pressure monitors; in Sites 2 and 3 and across the conditions, substantial majorities of telehealth device recipients received weighing scales.

Table 5.6 Telehealth equipment used by Telehealth study sample (N=856)

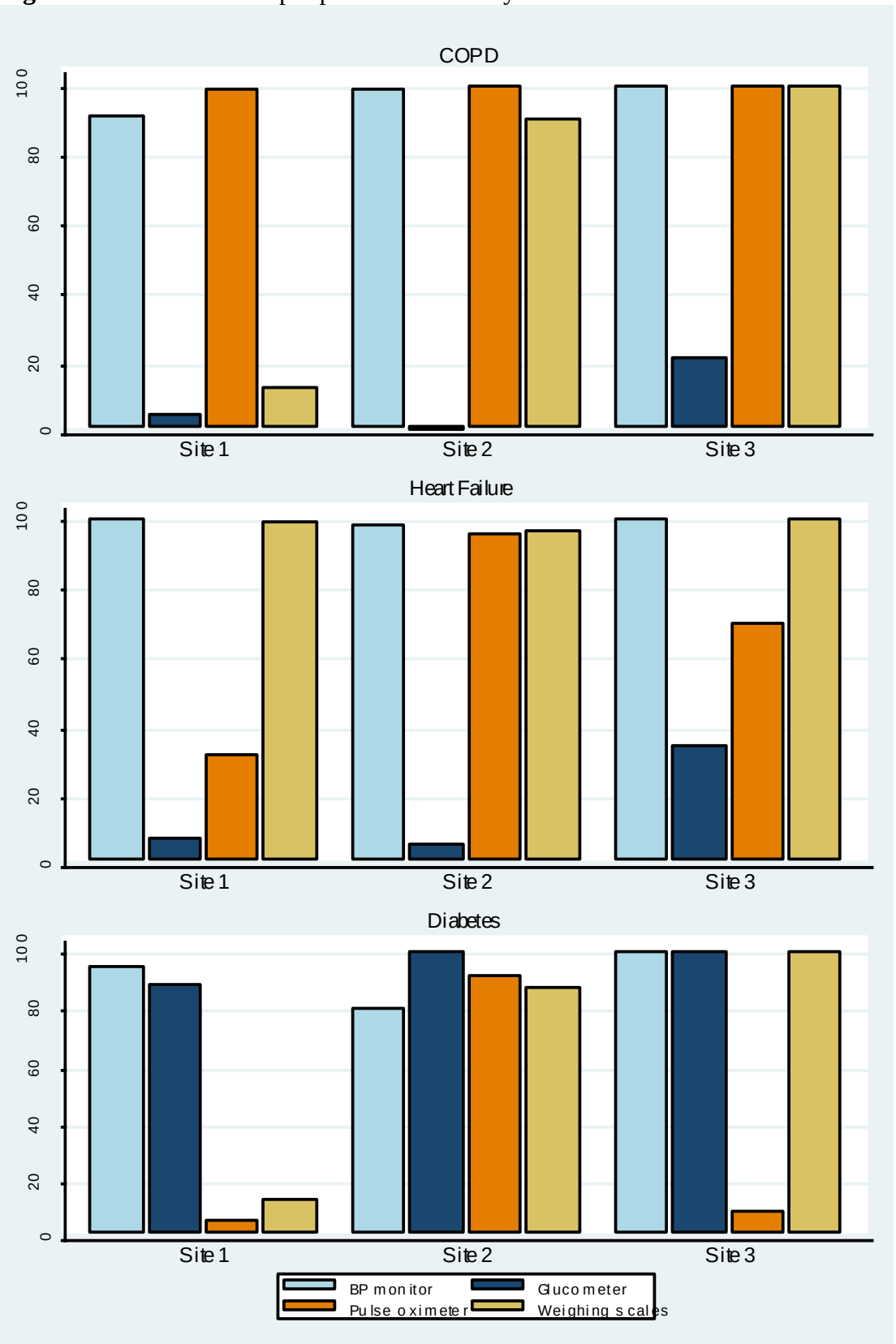
Items of Equipment	*N**	*% using*
BP monitor	831	97%
Glucometer	300	35%
Pulse oximeter	581	68%
Weighing scales	681	80%
Combinations of items†		
BP monitor + weighing scales	68	8%
BP monitor + glucometer	44	5%
BP monitor + oximeter	99	12%
BP monitor + pulse oximeter + weighing scales	364	43%
BP monitor + glucometer+ weighing scales	149	17%

Notes: BP monitor=Blood pressure monitor

*Number of questionnaire participants at baseline with equipment data available, including those not completing CSRIs.

† Combinations of equipment used by more than 1% of the questionnaire study sample.

Figure 5.2 Distribution of peripheral monitors by index condition and site



5.9 Unit Costs of the Telehealth Services

Annual telehealth intervention delivery and equipment unit costs, and intervention unit costs, are presented in terms of the range of costs across the three sites (Table 5.7). Some additional unit costs are presented here: the costs excluding project management costs, costs excluding equipment costs and costs had the sites been able to recruit their maximum planned number of participants and run at a higher capacity during the trial.

Table 5.7 Telehealth intervention costs in the three WSD sites

Cost category	**Range (£ per year, 2009-10)**
In-house staff*	338,598 – 540,381
Computer hardware and peripherals	188,249 – 490,748
Computer software	86,064 – 39,678
Installation	17,914 – 69,185
Contract costs/fees to other organisations	8,623 – 261,588
TOTAL DIRECT COST	840,464 – 1,168,671
DIRECT SUPPORT COST PER PARTICIPANT	1,487 – 2,042
Less total equipment cost†	1,134 – 1,241
Less posts/contracts specific to project management	804 – 1,199
Assuming 1000 participants recruited per site‡	580 – 733
Equipment costs per participant†	334 – 852

Costs rounded to the nearest £1.
*Excludes costs of installation staff, which were reported separately.
†Total equipment costs=costs of base units and peripherals-specific costs.
‡The monitoring costs of the service, assuming that it was functioning "at capacity" (for sensitivity analyses).

Per-site level unit costs of support, excluding equipment costs, were allocated to each participant on the basis of their receipt of telehealth equipment (the per-protocol rather than intention-to treat-allocation). The mean annual cost for telehealth equipment and support was £1847 (standard error £11.3) for participants in receipt of telehealth equipment with 12-month follow-up costs data.

5.10 Costs of Health and Social Care

Categories of costs of self-reported service use (imputed data) over the last three months before the baseline and the 12-month follow-ups are reported in Table 5.8 and .Table 5.9 Costs in the three months prior to baseline were similar in the intervention and control groups. Examining costs in the three months prior to the 12-month follow-up (excluding the direct costs of the intervention), hospital costs made up about half the total costs (47 per cent) for all participants, followed by primary care costs (18 per cent); medications (18 per cent);

and combined costs of social care (including community care, local authority-provided day care and equipment) (16 per cent). Costs in the telehealth group, excluding intervention specific costs, were not significantly lower than those in the usual care group, with a difference in costs of £243 (95% CI (-£565, £79)) between groups. The costs, including intervention-specific costs, in the telehealth group were higher than in the usual care group. For the telehealth group, three-month costs for equipment averaged £168 per person, about a tenth of the total. Total costs for health and social care, for the three months prior to the 12-month interview, were £1150 and £1394 for the telehealth and usual care groups, respectively, excluding the direct costs of the intervention; if direct costs were included, these costs were £1608 and £1403, respectively.

In terms of missing data, baseline costs at the category level were generally near-complete (less than 2 per cent missing) (LA and NHS equipment having the highest proportion of missingness at 2 per cent and 5 per cent of cases in both groups respectively). At follow-up, there were more missing in certain categories: care home costs (5 per cent of telehealth and 6 per cent of usual care cases), NHS day care costs (7 per cent of telehealth and 11 per cent of usual care cases), NHS adaptations and equipment (4 per cent of telehealth and 8 per cent of usual care cases), LA adaptations and equipment (4 per cent of telehealth and 9 per cent of usual care cases), and medications costs (23 per cent of telehealth and 27 per cent of usual care cases).

Table 5.8 Mean service costs (standard errors) over previous 3 months across Telehealth sample, available cases at baseline

Resource item	Control (n=728)	Telehealth (n=841)	Difference (units)
	Mean (SE)	Mean (SE)	Mean (95% CI)
Hospital use	461 (46)	432 (43)	-29 (-153, 95)
Community health services/primary care	235 (16)	195 (16)	-40 (-84, 5)
Community mental health	3 (5)	14 (5)	11 (-3, 24)
Community care services	131 (43)	138 (42)	7 (-111, 126)
Care home respite	2 (1)	0 (1)	-2 (-5, 2)
Day services LA	11 (5)	16 (4)	5 (-8, 17)
Day services NHS	3 (9)	12 (8)	9 (-15, 33)
Medications	440 (20)	464 (19)	24 (-30, 78)
Equipment/Adaptations LA	3 (1)	2 (0)	-1 (-3, 0)
Equipment LA/Adaptations NHS	1 (0)	0 (0)	-0 (-1, 0)
Total costs excl. TH delivery & equipment	1289 (71)	1273 (66)	-16 (-206, 174)

Note: Imputed data (10 completed datasets).

Table 5.9 Mean service costs (standard errors) over previous 3 months across Telehealth sample, available cases at 12-month follow-up

Resource item	Control (n=431)	Telehealth (n=538)	Difference (units)
	Mean (SE)	**Mean (SE)**	**Mean (95% CI)**
Hospital use	670 (90)	520 (83)	-149 (-389, 90)
Community health services/primary care	244 (20)	211 (20)	-33 (-86, 20)
Community mental health	8 (9)	6 (9)	-3 (-12, 7)
Community care services	197 (28)	149 (28)	-48 (-144, 48)
Care home respite	1 (8)	2 (9)	0 (-9, 9)
Day services LA	43 (14)	21 (13)	-22 (-59, 16)
Day services NHS	2 (5)	8 (4)	6 (-7, 20)
Medications	227 (8)	232 (7)	5 (-16, 25)
Equipment/Adaptations LA	1 (1)	1 (1)	1 (-2, 4)
Equipment LA/Adaptations NHS	0 (0)	0 (0)	0 (0, 1)
Total costs exc. telehealth delivery & equipment	1 394 (119)	1 150 (110)	-243 (-562, 75)
Telehealth intervention	6 (4)	289 (4)	284 (272, 296)**
Telehealth equipment	4 (8)	168 (8)	165 (141, 188)**
Total costs inc. telehealth delivery & equipment	1403 (120)	1608 (110)	205 (-114, 524)

Note: Includes cases where baseline cost data are missing. Imputed data (10 completed datasets).
*p<0.001 on clustered t-test

5.11 Clustering Effects

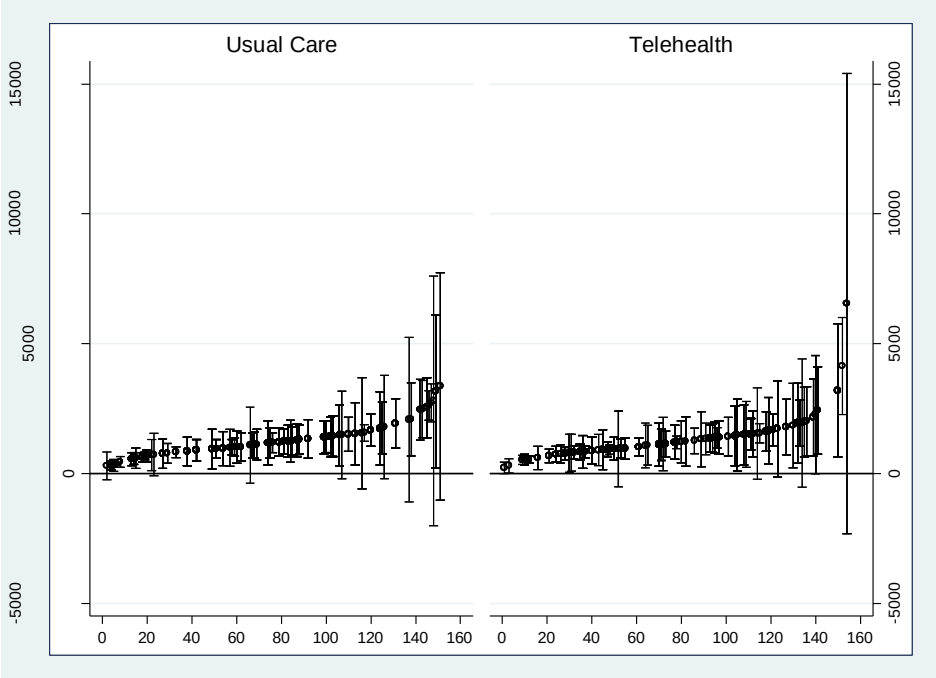
The clustering of costs (excluding the intervention) for telehealth participants is examined in Table 5.10. The ICCs presented for the general practice level are examined separately by time point. The estimated ICC values at each time point are higher in the intervention group than in controls; however the confidence intervals of the estimates overlap, suggesting that practice-level clustering is similar within the allocation groups at each time point. The variation in total costs (excluding intervention costs) between clusters at baseline and at 12-month follow-up is illustrated in Figure 5.3 and Figure 5.4.

Table 5.10 Health and social care service costs, Telehealth sample, prior three months: intra-cluster correlation coefficients (ICC) for general practice, per time point, ITT allocation

	Baseline ICC[a]	No. Practices	N	Follow-up ICC[a]	No. Practices	N
Control	-0.002 (-0.037, 0.032)	73	728	0.007 (-0.054, 0.069)	69	431
TH	0.022 (-0.016, 0.060)	81	841	0.061 (-0.005, 0.127)	76	538

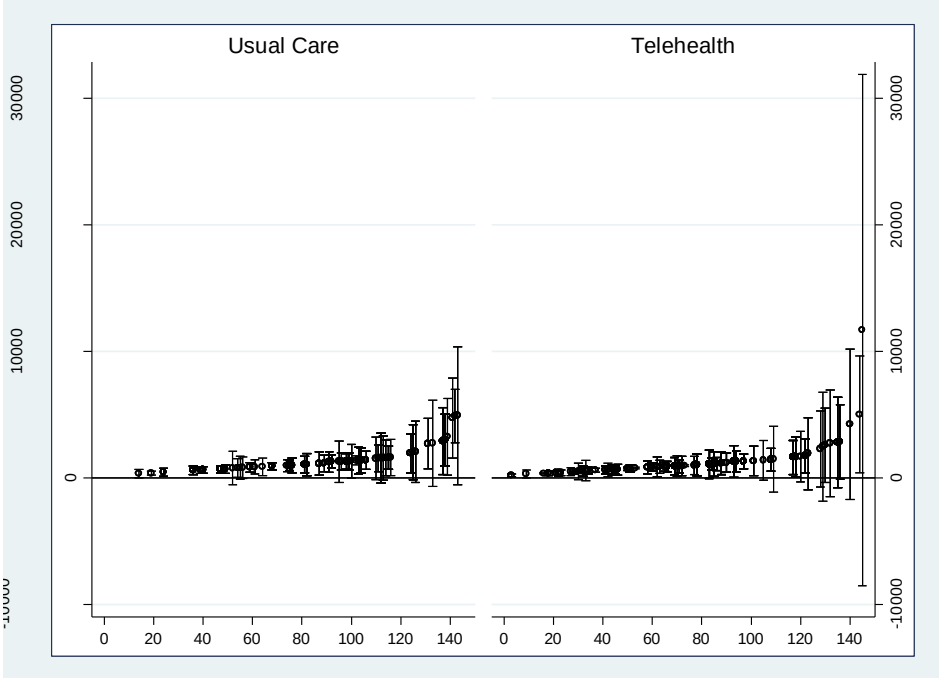
Note: costs exclude costs of the intervention. Imputed data (10 completed datasets).
a from one-way analysis of variance, Searle's Confidence intervals (Ukoumunne 2002).

Figure 5.3 Caterpillar plot of health and social care costs per cluster, Telehealth sample, three months prior to baseline, by experimental group



Note: graph of data from the first complete dataset generated by the multiple imputation process (graphs from other complete datasets are similar). The error bars represent the standard errors of the cluster means (including clusters of one).

Figure 5.4 Caterpillar plot of health and social care costs per cluster, Telehealth sample, three months prior to 12-month follow-up, by experimental group



Note: graph of data from the first complete dataset generated by the multiple imputation process (graphs from other complete datasets are similar). The error bars represent the standard errors of the cluster means (including clusters of one).

5.11.1 Telecare

There were 2,600 participants from 217 general practices in the WSD Telecare trial. 1,189 people from 204 general practices participated in the WSD Telecare Questionnaire study: 101 practices representing 550 participants were allocated to telecare and 103 practices with 639 participants, to usual care. At the baseline assessment point, service use data were available for 1,182 participants; these were available for 531 (269 telecare and 262 usual care) participants from 174 practices at the 4-month assessment, and for 757 participants from 191 practices at the 12-month follow-up assessment point (381 telecare and 376 control). Both baseline and 12-month follow-up costs were available for 378 telecare and 376 control participants.

5.12 Socio-demographic Characteristics

Baseline characteristics of the telecare sample with baseline service use data available are summarised in Table 5.11. Almost two-thirds of participants were women. Almost half of the sample resided in site 2. The average age of participants was 74.2. While the sample can generally be characterised as composed of older adults, a substantial proportion (23 per cent)

were participants under 65 years of age. Almost half of the people participating had at least one comorbidity. The samples with data available at baseline and 12-month follow-up were generally well-balanced in terms of age, sex and mean IMD score. At both time points there was a larger percentage of the telecare group in the second IMD quintile. There was little difference between groups in the proportions living alone.

5.13 Characteristics of the Samples Completing and Not Completing the 12-month Follow-up

A comparison of baseline characteristics of participants who did and did not complete study instruments at 12-month follow-up is presented in Table 5.11. Of participants allocated to the usual care intervention, participants completing were older (75.9 years) than participants who did not complete (73.2 years). Compared to the sample of usual care participants not completing study instruments, the usual care sample completing study instruments had a higher proportion of cases in the secondary-school education level group (GCSE/A/O-level), lower proportions of cases in site 2 and higher proportions of cases in site 3.

5.13.1 Cluster Numbers and Sizes

The clusters slightly diminished in size but became more balanced at the 12-month follow-up, from approximately 6 cases on average in the usual care group and 5 cases in the telecare group at baseline, to approximately 4 in both treatment groups.

5.14 Characteristics of the Sample Completing and Not Completing the 4-month Follow-up

At the 4-month follow-up, the majority of Telecare study participants did not complete questionnaires. However, there were few differences between the samples completing and not completing, within experimental groups (Appendix 2, Table A2.). The proportion of intervention participants from site 2 was greater in the completers group than in the non-completers (56 per cent vs. 44 per cent, z=-2.7860, p=.0053). Conversely the proportions of usual care participants from site 3 was lower in the completers than in the non-completers group (22 per cent vs. 33 per cent, z=0.6740, p=0.0075). These differences were apparent across experimental groups, with higher proportions of completers than non-completers residing in site 2 and lower proportions of completers than non-completers residing in site 3.

Table 5.11 Baseline characteristics of participants with economic data available at baseline and 12-month follow-up across Telecare sample

	Total baseline sample			Participants completing 12-month follow-up study instruments*			Participants not completing 12-month follow-up study instruments†		
	UC (n=634)	**TC (n=548)**	**Raw**	**UC (n=378)**	**TC (n=375)**	**Raw**	**UC (n=253)**	**TC (n=170)**	**Raw**
Mean years of age (SD)‡	74.3 (17.5)	74 (17.1)	-0.3	73.1 (17.6)	73.2 (17.4)	0.1	75.9 (14.5)	75.8 (14.4)	-0.1
Under 65 (young)	138 (22%)	129 (24%)	2%	91 (24%)	91 (24%)	0%	47 (19%)	36 (21%)	3%
65-74 (young old)	139 (22%)	116 (21%)	-1%	94 (25%)	89 (24%)	-1%	45 (18%)	27 (16%)	-2%
75-84 (old old)	208 (33%)	168 (31%)	-2%	118 (31%)	112 (30%)	-1%	88 (35%)	55 (32%)	-2%
85+ (oldest old)	149 (24%)	135 (25%)	-1%	75 (20%)	83 (22%)	2%	73 (29%)	52 (31%)	2%
Female	415 (65%)	344 (63%)	2%	250 (66%)	241 (64%)	-2%	163 (64%)	102 (60%)	-4%
1+ comorbidities	304 (48%)	252 (46%)	-2%	176 (47%)	167 (45%)	-2%	125 (49%)	84 (49%)	0%
Mean comorbidities (SD)§	1.1 (1.6)	1.1 (1.6)	-0.0	1.1 (1.5)	1 (1.5)	-0.1	1 (1.6)	1.2 (1.6)	0.2
White-British‡§	561 (89%)	482 (88%)	-1%	332 (88%)	328 (87%)	0%	228 (90%)	151 (89%)	1%

	Total baseline sample			Participants completing 12-month follow-up study instruments*			Participants not completing 12-month follow-up study instruments†		
	UC (n=634)	**TC (n=548)**	**Raw**	**UC (n=378)**	**TC (n=375)**	**Raw**	**UC (n=253)**	**TC (n=170)**	**Raw**
WSD site									
Site 1	137 (22%)	125 (23%)	1%	84 (22%)	83 (22%)	0%	52 (21%)	42 (25%)	4%
Site 2‖	309 (49%)	273 (50%)	1%	169 (45%)	187 (50%)	5%	138 (55%)	83 (49%)	-6%
Site 3¶	188 (30%)	150 (27%)	-2%	125 (33%)	105 (28%)	-5%	63 (25%)	45 (27%)	2%
IMD	28.8 (40.4)	27.8 (38.2)	-0.7	29.6 (32.8)	27.5 (32.1)	-2.1	27.8 (28.4)	28.6 (25.8)	0.8
1st quintile§	152 (24%)	127 (23%)	-1%	87 (23%)	89 (24%)	1%	67 (26%)	35 (21%)	-6%
2nd quintile§	82 (13%)	109 (20%)	7%	45 (12%)	71 (19%)	7%	36 (14%)	37 (22%)	8%
3rd quintile§	133 (21%)	100 (18%)	-3%	79 (21%)	71 (19%)	-2%	50 (20%)	28 (16%)	-3%
4th quintile§	120 (19%)	102 (19%)	0%	72 (19%)	70 (19%)	0%	50 (20%)	35 (21%)	1%
5th quintile§	146 (23%)	110 (20%)	-3%	95 (25%)	74 (20%)	-5%	50 (20%)	35 (21%)	1%
Living alone§	340 (54%)	285 (52%)	-2%	202 (53%)	188 (50%)	-3%	136 (54%)	95 (56%)	2%

	Total baseline sample			Participants completing 12-month follow-up study instruments*			Participants not completing 12-month follow-up study instruments†		
	UC (n=634)	**TC (n=548)**	**Raw**	**UC (n=378)**	**TC (n=375)**	**Raw**	**UC (n=253)**	**TC (n=170)**	**Raw**
Tenure									
Rents/Other§	242 (38%)	192 (35%)	-3%	152 (40%)	129 (34%)	-6%	89 (35%)	62 (36%)	1%
Owns§	392 (62%)	356 (65%)	3%	226 (60%)	246 (66%)	6%	164 (65%)	108 (64%)	-1%
Education									
No formal education§	406 (64%)	349 (64%)	0%	230 (61%)	235 (63%)	2%	173 (68%)	112 (66%)	-2%
GCSE/O/A-level§**	185 (29%)	129 (24%)	-6%	126 (33%)	89 (24%)	-10%	59 (23%)	39 (23%)	0%
Degree	43 (7%)	70 (13%)	6%	22 (6%)	51 (14%)	8%	21 (8%)	19 (11%)	3%

UC=usual care; TC=telecare; SD=standard deviation

*costs and outcomes data available

† Outcomes instruments not completed and/or CSRI not completed

‡ Age: within UC: differences between completion/completion sample t=1.972, p<0.05 on clustered t-test

§ Imputed data

|| Within UC: differences between completion/completion *z*=2.423, p<0.05 (z-test of proportions)

¶ Within UC: differences between completion/completion *z*=-2.198, p<0.05 (z-test of proportions)

**Within UC: difference between completion/completion clustered χ^2= 7.0546, p<0.05

Table 5.12 Number and size of clusters corresponding to participants with economic data available at baseline and 12-month follow-up across Telecare sample

	Total baseline sample		**Participants completing 12-month follow-up study instruments***	
	UC (N=103)	TC (N=101)	UC (N=95)	TC (N=96)
Cluster size	6.2 [1-26]	5.4 [1-21]	4 [1-14]	3.9 [1-14]

Note: Data are given as mean [min – max]
UC=Usual care; TC=Telecare
*where costs and outcomes data were available

5.15 Telecare service Use: Descriptive Statistics

Differences between experimental groups were small for most categories in the 3 months prior to baseline and 12-month follow-up (Table 5.13 and Table 5.14). The use of social services such as home care and social work was greater amongst the telecare participants. Telecare participants had 9.6 more daytime home care visits than control participants (42 (SE 4.3) vs. 33 (SE 3.7)), while telecare participants reported 1.6 more community nursing visits at the 12-month follow-up. As in the telehealth sample, there were large differences in the proportions reporting seeing a practice nurse at baseline (approximately 21 and 22 per cent in control and intervention respectively) and at 12-month follow-up (approximately 42 per cent and 41 per cent control and intervention respectively).

5.15.1 Community Alarm Usage

The percentages of participants who reported having a 'personal/community alarm' (a community alarm or a pull-cord) were very similar at baseline (51.5 per cent intervention, 50.5 per cent usual care). Given the nature of the intervention, we would expect a growth in the reporting of some form of personal or community alarm in the intervention group.

Table 5.13 Number and percentage of groups using services and mean numbers of units (standard errors) used over previous 3 months across Telecare sample, available cases at baseline

Resource item	Usual care (n=634)			Telecare (n=548)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Hospital use							
A&E attendance	15	95	0.18 (0.02)	17	91	0.23 (0.03)	0.04
Inpatient services	13	83	1.44 (0.26)	12	66	1.65 (0.34)	0.21
Day Hospital attendances	7	45	0.26 (0.07)	8	42	0.30 (0.10)	0.04
Outpatient attendances	43	268	1.05 (0.08)	44	240	1.20 (0.12)	0.14
Community health services/primary care							
Paramedic contact	16	104	0.31 (0.05)	20	111	0.36 (0.05)	0.04
Community matron visit	4	25	0.41 (0.29)	4	23	0.26 (0.07)	-0.15
Community matron telephone contact	1	7	0.07 (0.03)	1	6	0.06 (0.03)	-0.02
Community or district nurse visit	23	147	2.40 (0.46)	25	136	3.09 (0.62)	0.69
Community or district nurse telephone contact	5	29	0.11 (0.03)	5	26	0.14 (0.04)	0.03
Practice nurse visit	21	134	0.75 (0.11)	22	121	0.76 (0.13)	0.01
Night nurse visit	1	8	0.05 (0.02)	1	5	0.05 (0.04)	0
Specialist nurse contacts	8	53	0.32 (0.14)	9	48	0.54 (0.29)	0.22
Physiotherapist or occupational therapist visit	18	117	1.10 (0.17)	23	125	1.54 (0.27)	0.43
GP (home) visit	22	142	0.46 (0.05)	22	123	0.51 (0.06)	0.05
GP (surgery) visit	45	287	1.16 (0.08)	45	244	1.08 (0.09)	-0.09
GP (telephone) visit	17	109	0.44 (0.05)	15	82	0.53 (0.11)	0.08
Dentist visit	21	132	0.35 (0.04)	24	130	0.41 (0.04)	0.06
Chiropodist visit	34	213	0.56 (0.04)	34	186	0.57 (0.04)	0
Optician visit	23	145	0.30 (0.03)	20	108	0.26 (0.03)	-0.05
Community mental health							

Resource item	Usual care (n=634)			Telecare (n=548)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Psychiatrist visit	5	29	0.10 (0.03)	4	22	0.20 (0.12)	0.1
Mental health nurse visit	1	8	0.03 (0.01)	3	14	0.10 (0.04)	0.07
Community care services							
Social worker visit	16	99	0.49 (0.21)	20	109	0.62 (0.23)	0.13
All daytime home care/home help visit	50	315	49.67 (3.44)	50	276	47.23 (3.65)	-2.44
Council home help visit	39	245	44.72 (3.29)	35	194	42.17 (3.57)	-2.55
Private/independent home care/home help visit	19	121	5.10 (0.86)	22	121	5.30 (0.83)	0.21
Paid night carer visit	4	23	1.95 (0.54)	5	28	3.36 (0.92)	1.4
Meals on Wheels meal	5	34	3.49 (0.68)	5	30	4.20 (1.33)	0.71
Laundry (incontinence) service	1	7	0.12 (0.05)	1	5	0.07 (0.04)	-0.05
Community alarm	50	320		51	282		
Equipment inc. mobility aids, ADL NHS	4	20	0.06 (0.01)	6	28	0.09 (0.02)	0.04
Major and minor adaptations NHS	1	3	0.01 (0.01)	2	11	0.03 (0.01)	0.02
Equipment inc. mobility aids, ADL LA	9	53	0.16 (0.02)	10	50	0.17 (0.03)	0
Major and minor adaptations LA	14	82	0.20 (0.02)	17	80	0.24 (0.03)	0.04
Care home respite							
Days	1	8	0.24 (0.11)	2	9	0.31 (0.12)	0.07
Day services							
Day care and other day attendances - LA	16	103	3.41 (0.42)	17	91	2.77 (0.36)	-0.64
Day care and other day attendances - NHS	2	10	0.21 (0.09)	2	10	0.32 (0.13)	0.11
Medications							
Number of medications	98	620	6.83 (0.19)	98	534	6.80 (0.21)	-0.03

* Proportion of non-missing cases who reported using a service

† Reported having a personal/community alarm – means represent number of ‘yes’ responses

Table 5.14 Number and percentage of groups using services and mean numbers of units (standard errors) used over previous 3 months across Telecare sample, available cases at 12-month follow-up

Resource item	Usual care (n=381)			Telecare (n=376)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Hospital use							
A&E attendance	20	71	0.24 (0.03)	18	64	0.24 (0.04)	0
Inpatient services	14	51	0.82 (0.17)	16	57	1.05 (0.25)	0.23
Day Hospital attendances	16	54	0.40 (0.13)	16	58	0.39 (0.10)	-0.01
Outpatient attendances	49	176	1.32 (0.12)	46	162	1.17 (0.11)	-0.15
Community health services/primary care							
Paramedic contact	16	54	0.30 (0.08)	17	60	0.29 (0.05)	-0.01
Community matron visit	5	18	0.32 (0.11)	7	22	0.26 (0.08)	-0.06
Community matron telephone contact	3	9	0.06 (0.02)	4	13	0.18 (0.13)	0.13
Community or district nurse visit	21	74	1.29 (0.28)	26	93	2.90 (0.66)	1.61
Community or district nurse telephone contact	8	27	0.15 (0.04)	8	29	0.28 (0.09)	0.12
Practice nurse visit	42	132	1.43 (0.24)	41	126	1.19 (0.20)	-0.24
Night nurse visit	1	3	0.01 (0.01)	1	3	0.72 (0.56)	0.71
Specialist nurse contacts	16	59	0.35 (0.06)	14	52	0.49 (0.24)	0.14
Physiotherapist or occupational therapist visit	14	43	0.50 (0.14)	14	44	0.69 (0.15)	0.19
GP (home) visit	26	79	0.53 (0.08)	30	92	0.69 (0.08)	0.16
GP (surgery) visit	64	224	1.66 (0.12)	57	192	1.25 (0.09)	-0.41
GP (telephone) visit	34	100	0.72 (0.11)	32	99	0.89 (0.14)	0.17
Dentist visit	30	101	0.44 (0.05)	24	81	0.34 (0.04)	-0.1
Chiropodist visit	42	144	0.70 (0.07)	39	136	0.95 (0.18)	0.25
Optician visit	34	116	0.41 (0.04)	26	85	0.33 (0.03)	-0.09
Community mental health							

Resource item	Usual care (n=381)			Telecare (n=376)			Difference
	% using*	N using	Mean (SE)	% using*	N using	Mean (SE)	Mean
Psychiatrist visit	4	14	0.10 (0.05)	3	10	0.05 (0.03)	-0.05
Mental health nurse visit	1	4	0.02 (0.02)	3	10	0.11 (0.06)	0.09
Community care services							
Social worker visit	13	43	0.22 (0.05)	16	56	1.19 (0.87)	0.97
All daytime home care/home help visit	45	170	32.50	54	202	42.23	9.74
Council home help visit	29	101	29.29	33	114	36.33	7.05
Private/independent home care/home help visit	29	99	7.94 (1.38)	36	125	11.86	3.92
Paid night carer visit	2	8	0.42 (0.29)	3	10	0.79 (0.41)	0.37
Meals on Wheels meal	6	19	2.34 (0.69)	6	20	3.68 (0.91)	1.34
Laundry (incontinence) service	2	8	0.21 (0.08)	2	6	0.35 (0.29)	0.14
Community alarm	64	232		89	320		
Equipment inc. mobility aids, ADL NHS	11	39	0.30 (0.06)	14	47	0.34 (0.06)	0.04
Major and minor adaptations NHS	2	8	0.04 (0.01)	4	14	0.07 (0.02)	0.03
Equipment inc. mobility aids, ADL LA	7	24	0.11 (0.03)	8	28	0.19 (0.04)	0.08
Major and minor adaptations LA	12	40	0.22 (0.04)	10	32	0.13 (0.02)	-0.09
Care home respite							
Days	1	4	0.12 (0.08)	2	4	0.27 (0.18)	0.15
Day services							
Day care and other day attendances -LA	13	50	2.32 (0.42)	14	53	2.35 (0.40)	0.03
Day care and other day attendances - NHS	1	4	0.10 (0.06)	1	5	0.09 (0.05)	-0.01
Medications							
Number of medications	99	303	7.42 (0.23)	99	277	7.10 (0.24)	-0.32

*Proportion of non-missing cases who reported using a service

By the 12-month follow-up, 64 per cent of the usual care group reported using a community alarm, a relative increase of 26 per cent from baseline and an absolute difference in baseline and follow-up proportions of 13 (clustered χ^2= 105.8396, p=.0000). The proportion of the usual care group using a community alarm at the 4-month follow-up (61 per cent) showed a similar trend from baseline (clustered χ^2= 107.3696, p=.0000).

It is possible that community alarms being installed in the control group participants' homes during the trial period could have attenuated differences in outcomes related to the intervention between the experimental groups, if the telecare provided as part of the intervention did not vary substantially from the telecare provided through other routes. Participants were not asked to give details on the types of community alarms/telecare devices that were in their homes at the baseline assessment or (for non-WSD devices) during the follow-up period. No further information was available on the type of telecare packages (obtained through other routes than WSD) used by the control participants.

5.16 Intervention Costs

5.16.1 Telecare intervention

The WSD telecare intervention most closely resembled a second-generation telecare system, whereby the participant received a base unit and pendant/bracelet that could be used to alert call centre monitoring operators, and sensors were monitored remotely and automatically (see 5.16.3 for further details on the equipment).

5.16.2 Description of the Telecare Services

The configuration of telecare services varied considerably. Not surprisingly given the mixed market in social care, some aspects of WSD Telecare services were contracted out, to differing extents, to other public and private organisations (Table 5.15). All sites had call centre contracts with long-standing providers of telecare in their local areas, either district councils or arm's-length management organisations. One private company supplied the majority of telecare equipment across the sites. Installation and response arrangements in particular varied between and also within the sites. For instance one site did not organise a dedicated response to any telecare sensor alerts, whereas the other sites had the option to organise a visit to the service user in certain circumstances.

Table 5.15 Features of delivery systems in the WSD Telecare pilot sites, 2009/10

Producers / **Roles**	**Site 1**	**Site 2**	**Site 3**
Equipment supply	Equipment purchased by the local authority for the project from private provider.	Equipment purchased for the project through private provider.	Equipment purchased by the local authority for the project from private providers.
Installation, server and equipment maintenance, asset management, training	Most services provided by local authority-based project team: installations, responsible for asset management, routine maintenance and callouts for equipment-related problems such as battery replacements. Installations also provided 1 day/week by one of the 2 monitoring services. District council call centres hosted the servers.	Services provided by combination of local authority-based units. Local authority business unit carried out installations, handled warehousing, inventory control and equipment configuration; provided routine maintenance, maintenance callouts. Local authority project team provided project/contract management; assessment for equipment; point of contact for service users with equipment-related/technical problems. District council call centre hosted the server.	Services provided by combination of local authority-based unit and private providers. Private company responsible for transporting the equipment to the person's home, installing cabling, installing and configuring the equipment; also for asset management, scheduled maintenance visits and maintenance callouts. Local authority managed server. Arm's-length local authority call centre assessed for equipment, follow-up equipment testing visit.

Producers / **Roles**	**Site 1**	**Site 2**	**Site 3**
Monitoring and responses to alarms/sensors alerts	Monitoring: provided by call centres in 2 districts, later merged into single service, 24/7 service. Responses: via call centre staff, either to contact nominated carer to request response, or contact emergency services. Two types of responses to sensor alerts/activations: operator contacts the service user, if the user reports a problem, operator contacts a nominated carer to visit the service user; if no carer is available to respond but a response of some kind is required, operator contacts the emergency services. No dedicated WSD response service.	Monitoring: provided by a district local authority call centre, providing 24/7 service. Responses: initial response via call centre operators. Three types of responses to sensor alerts/activations: Operator contacts nominated carer to request response; if no carer available to respond, operator calls dedicated WSD response team; or contacts the emergency services. Dedicated WSD telecare response teams were organised on a district basis: in one PCT area, intermediate care and out-of-hours nursing teams; in the other, private providers of out-of-hours primary care services and a provider of health and domiciliary care.	Monitoring: provided by arm's-length local authority call centre, providing 24/7 service. Responses: initial response via call centre telecare officers. Three types of responses to sensor alerts/activations: telecare officer contacts the service user to investigate the nature of the problem. If necessary, telecare officer contacts the emergency services; the officer contacts named carers (relatives or friends) to ask carer to visit the service user; if nominated carers are not available, officer visits the service user. The call centre offers a dedicated response visiting service to people using telecare from the local authority.

5.16.3 Telecare Equipment

The number of items of telecare equipment provided to participants ranged between 1 and 11 items (Table 5.16). On average, participants had a mean of 4.7 items (a mode of 4). Of the available sensors, smoke detectors, carbon monoxide monitors, fall detectors, flood detectors and temperature extremes sensors were most frequently provided.

The WSD evaluation team developed a categorisation of telecare devices by assessing the function of devices and mapping these against groupings from the telecare literature (Bower et al. 2011, Brownsell, Blackburn, and Hawley 2008, Demiris and Hensel 2008, Doughty et al. 2008). For the purposes of the trial, the evaluation team also added a category for 'standalone' devices. These are devices that are not monitored but that can facilitate telecare (e.g. key safes, which allow authorised staff responding to a telecare alert to enter the person's home) (Hirani et al. 2013). The telecare equipment was mapped to one of the four functions: monitoring functional status; monitoring home security; monitoring the home environment; and standalone devices. The following list shows which devices were mapped to each function.

Functional monitoring: Lifeline base unit + pendant, Minuet watch, Pull Cord, Bed Occupancy, Chair Occupancy, Enuresis Sensor, Epilepsy sensor, Fall Detector, Medication dispenser

Security monitoring: Bogus caller button, PIR Movement Sensor, Property Exit Sensor

Environmental monitoring: Natural Gas Detector, Carbon Monoxide Detector, Smoke Detector, Heat Sensor, Temperature Extremes Sensor, Flood Detector

Standalone: Motion Light, Picture Phone, Timex USB Watch, Memo Minder, Dummy Bell Box, Key Safe, Magiplug, DDA Pager, Big Button Phone

Relatively few participants received safety and security monitoring sensors; in contrast, all participants had at least one "functional monitoring" sensor. More than half of the telecare participants had stand-alone devices. The most frequently used telecare packages combined functional, environmental and stand-alone devices.

Table 5.16 Telecare equipment used by Telecare study sample by function

	*N**	*mean (SD)*	*[Range]*	*% using*
All items of equipment	553	4.7 (1.77)	[1 - 11]	100%
Functional monitoring	553	1.8 (0.83)	[1 - 5]	100%
Environmental monitoring	522	2.1 (1.15)	[0 - 5]	94%
Stand-alone devices	302	0.6 (0.63)	[0 - 3]	55%
Security monitoring	79	0.2 (0.42)	[0 - 3]	14%
Participants with items from a single function category	17	1.7 (0.59)	[1 - 3]	3.1%
Combinations of function†				
Functional, environmental and stand-alone	240	5.2 (0.09)	[3 - 10]	43.4%
Functional and environmental	207	3.7 (0.09)	[2 - 7]	37.4%
Functional, environmental, safety/security and stand-alone	50	7.1 (1.64)	[4 - 11]	9.0%
Functional, environmental and safety/security	25	5.4 (0.25)	[4 - 8]	4.5%
Functional and stand-alone	10	3.4 (0.27)	[2 - 5]	1.8%

* Number of questionnaire participants at baseline, including those not completing CSRIs
†combinations of equipment used by more than 1% of the questionnaire study sample

5.16.4 Unit costs of the Telecare Services

The ranges of unit costs of telecare support across the sites are given in Table 5.17, along with the costs of the service excluding project-related posts and contracts from calculations, and excluding costs of dedicated WSD telecare responders. As I have noted, sites had quite different project management structures and local arrangements for monitoring and responding to telecare sensor activations; as might be expected, the unit costs of telecare support also varied substantially between the sites. The mean annual per-person cost of telecare equipment was £81 (SE £1.9) for participants who had completed baseline assessments and £82 (SE £2.3) for those who had also completed the 12-month follow-up.

Table 5.17 Unit costs, Telecare intervention in the three WSD sites

Cost category	Range (£ per year, 2009-10)
In-house staff	26,999 – 213,465
Monitoring base unit	6,951 – 12,228
Sensors and other peripherals	17,019 – 28,148
Maintenance	24,891 – 34,217
Installation	13,694 – 17,224
Contract costs/fees to other organisations	52,000 – 191,112
DIRECT NON-EQUIPMENT COST OF SUPPORT	170,432 – 456,019
DIRECT SUPPORT COST PER PARTICIPANT*	437 – 1004
Less project management-specific posts and contracts	423 - 870
Less response-related contract costs	408 - 908
"Mainstream" telecare support package of £5 per week †	261
Equipment costs	
Unit costs‡	1.05 - 93.20
Equipment costs per participant	73 - 93

* Excludes cost of equipment
† For sensitivity analysis
‡ Annual equivalent

5.16.5 Costs of Health and Social Care

At baseline, hospital services constituted about a quarter of health and social care costs (imputed and excluding intervention costs), with community primary and mental health care costs contributing 14 per cent and community social care 37 per cent (Table 5.18). At 12-month follow-up (Table 5.19), the composition of costs was largely unchanged, although medications accounted for less of the total (8 per cent vs. 15 per cent). The mean annual cost of a telecare support and equipment package (for those in receipt of equipment) was £791. Intervention group costs were somewhat higher than control group costs at the end of the trial.

In terms of missing data, most costs at the category level were near-complete (<2% missing) at baseline (however NHS equipment and adaptations were missing for 14 per cent of telecare and 12 per cent of usual care cases; LA equipment and adaptations missing for 6 per cent telecare and 3 per cent usual care cases). At follow-up, there were more missing data in certain categories: care home costs (6 per cent of telecare, 3 per cent of usual care cases), NHS day care costs (19 per cent of telecare and 18 per cent of usual care cases), NHS adaptations and equipment (7 per cent of telecare, 6 per cent of usual care cases), LA adaptations and equipment (7 per cent of telecare, 5 per cent of usual care cases), and medications costs (25 per cent of telecare cases, 20 per cent of usual care cases).

Table 5.18 Mean service costs (standard errors) over previous 3 months across Telecare sample, available cases at baseline

Resource item	Usual care (n=634)	Telecare (n=548)	Difference (units)
	Mean (SE)	Mean (SE)	Mean (95% CI)
Hospital use	662 (84)	660 (89)	-3 (-243,237)
Community health services/primary care	299 (36)	354 (37)	55 (-47,156)
Community mental health	21 (7)	25 (7)	4 (-16,24)
Community care services	880 (113)	913 (117)	32 (-286, 351)
Care home respite	16 (14)	22 (15)	6 (-35,46)
Day services LA	112 (13)	86 (14)	-26 (-63,12)
Day services NHS	33 (23)	51 (24)	17 (-49,83)
Medications	379 (24)	364 (25)	-15 (-83, 53)
Equipment/Adaptations LA	7 (1)	8 (1)	0 (-4, 4)
Equipment LA/Adaptations NHS	1 (0)	2 (0)	1 (0, 2)
Total costs exc. telecare delivery and equipment	2411 (166)	2484 (174)	73 (-398, 544)

Table 5.19 Mean service costs (standard errors) over previous 3 months across Telecare sample, available cases at 12-month follow-up

Resource item	Usual care (n=381)	Telecare (n=376)	Difference (units)
	Mean (SE)	Mean (SE)	Mean (95% CI)
Hospital use	466 (69)	512 (69)	45 (-147, 237)
Community health services/primary care	230 (21)	269 (21)	38 (-19, 96)
Community mental health	12 (9)	18 (9)	5 (-19, 30)
Community care services	724 (109)	855 (109)	131 (-172, 434)
Care home respite	8 (9)	19 (9)	11 (-13, 36)
Day services LA	175 (30)	184 (30)	8 (-75, 92)
Day services NHS	12 (7)	13 (7)	0 (-18, 19)
Medications	164 (10)	147 (9)	-17 (-44, 9)
Equipment/Adaptations LA	7 (2)	3 (2)	-4 (-9, 2)
Equipment LA/Adaptations NHS	2 (1)	3 (1)	1 (0, 3)
Total costs exc. telecare delivery and equipment	1801 (167)	2021 (166)	220 (-242, 681)
Telecare intervention	9 (7)	177 (6)	169 (151, 187)**
Telecare equipment	1 (1)	20 (1)	19 (18, 21)**
Total costs inc. TC delivery & equipment	1811 (169)	2218 (168)	408 (-59, 875)

Note: Includes cases where baseline cost data are missing.
*p<0.001 on clustered t-test

5.16.6 Clustering Effects

The clustering of costs (excluding the intervention) is examined in Table 5.20. The ICCs presented for the general practice level are examined separately by time point. The estimated ICC values at each time point are higher in the intervention group than in controls (in particular, the baseline ICC is much higher in the intervention group than in the control group and the confidence intervals of the ICC do not overlap). The confidence intervals of the 12-month follow-up ICC estimates overlap, suggesting that practice-level clustering is similar within the allocation groups at that time point.

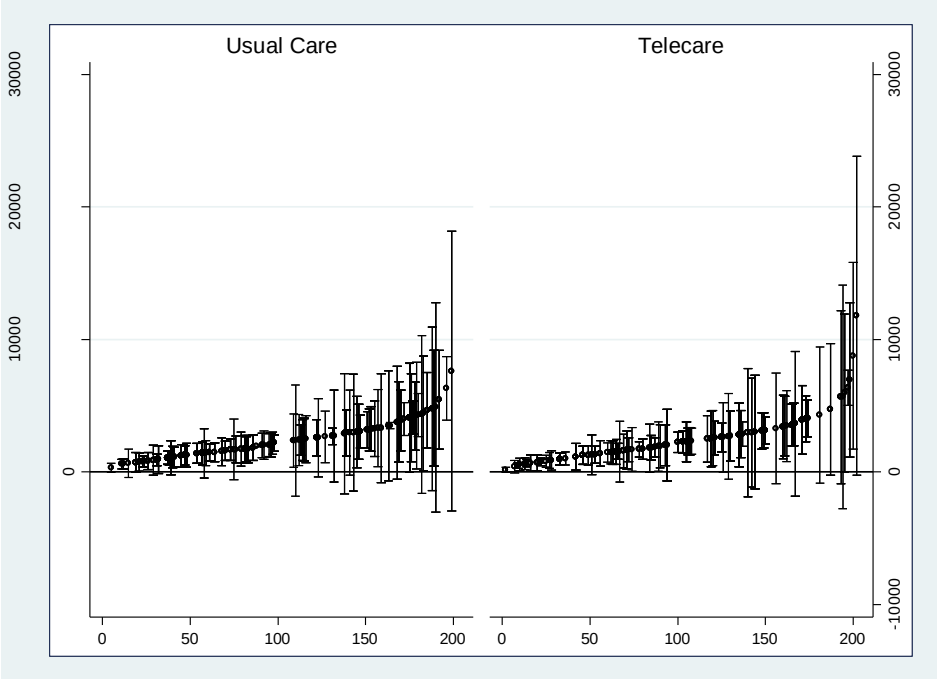
A caterpillar plot of these costs ranked in ascending order illustrates this point, as well as highlighting the variability in costs between general practice clusters.

Table 5.20 Health and social care service costs, Telecare sample, prior three months: intra-cluster correlation coefficients (ICC) for general practice, per time point, ITT allocation

	Baseline ICC[a]		No. Practices	N	Follow-up ICC[a]	No. Practices	N
Usual care	-0.008	(-0.056,0.039)	103	634	0.087(-0.010,0.183)	95	381
Telecare	0.196	(0.104,0.289)	101	548	0.094 (-0.005,0.193)	96	376

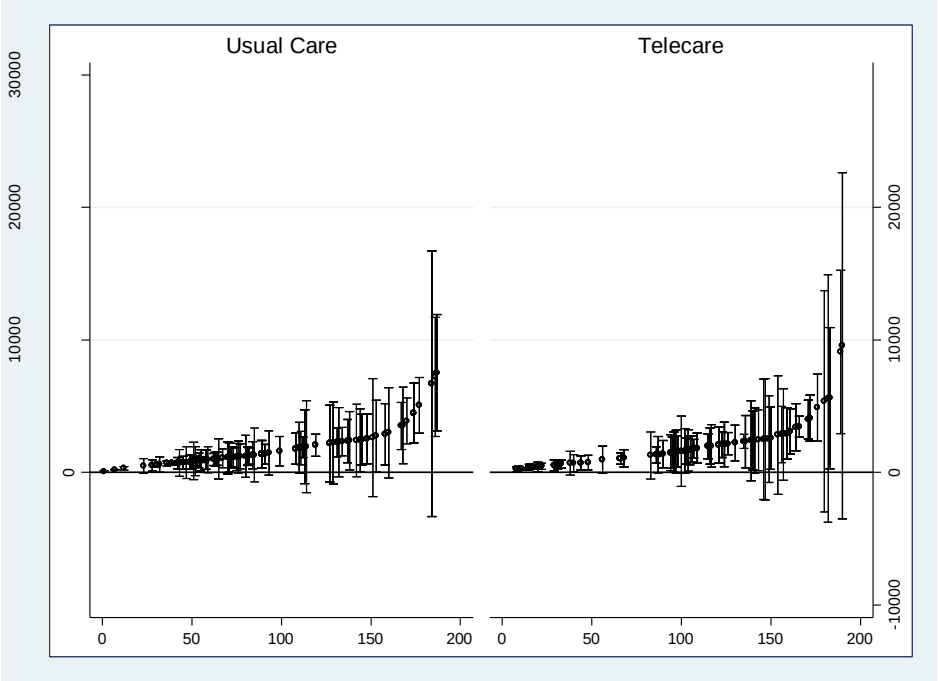
Note: imputed data; costs exclude costs of the intervention
a. From one-way analysis of variance, Searle's Confidence intervals (Ukoumunne 2002)

Figure 5.5 Caterpillar plot of costs per cluster, Telecare sample, three months prior to baseline, by experimental group



Note: graph of data from the first complete dataset generated by the multiple imputation process (graphs from other complete datasets are similar). The error bars represent the standard errors of the cluster means (including clusters of one).

Figure 5.6 Caterpillar plot of costs per cluster, Telecare sample, three months prior to 12-month follow-up, by experimental group



Note: graph of data from the first complete dataset generated by the multiple imputation process (graphs from other complete datasets are similar). The error bars represent the standard errors of the cluster means (including clusters of one).

5.17 Discussion

The trial participant samples were distinctly different, as would be expected given the different trial eligibility criteria. There was a preponderance of women in the telecare study sample, whereas there were more men than women in the telehealth study sample. Compared to the telehealth sample, more of the telecare sample resided in site 2. Compared to the resource-use profile of telehealth trial participants at baseline, the telecare trial participants were heavier users of hospital, mental health, social care and GP services (especially home visits and rehabilitation). Fewer telecare than telehealth study participants used community matron services.

Demographic and costs data from both studies exhibited considerable heterogeneity, attributable to a number of factors. Sites contributed participants in unequal proportions in both studies. Site 2 provided about 50 per cent of participants at baseline in telecare, with Site 1 providing about 23 per cent of the sample; Site 2 provided about 40 per cent in telehealth, with the other two sites contributing roughly equal numbers to the sample at baseline. These imbalances were somewhat exacerbated by attrition over the 12-month study period. There

was considerable between-site variation in the production processes of both the telehealth and telecare interventions.

In both of the questionnaire studies, there was a substantial loss to follow-up at both 4-month and 12-month assessment points. In the following chapters, analyses focus on the data from the 12-month follow-up. In neither case was there much evidence that, within experimental groups, the participants that completed the 12-month follow-up differed from those that did not, nor between groups apart from variations in the proportions of people with COPD in the telehealth sample.

Chapter 6

Cost Variations in the Telehealth and Telecare Samples

In this chapter I address research question 3 in each of the telehealth and telecare samples. In each case I examine key participant characteristics associated with variations in health and social care costs. I first present a subgroup analysis, examining whether the three-month costs of participants allocated to telehealth or usual care differ between baseline and long-term follow-up time points, depending on their index long-term condition (diabetes, COPD or heart failure). I next examine whether the costs of participants allocated to telecare or usual care differ between baseline and follow-up depending on their living arrangements (living alone or with others).

6.1 Telehealth

There is a small body of evidence on drivers of health and social care costs of recipients of telehealth with long-term conditions. The bulk of this literature is based on Medicare data on heart failure patients in the United States. A study of 168 Medicare heart failure patients using telehealth examined the association of key medical events (where action is taken by the monitoring practitioner in response to the telehealth alert) and use of emergency department (ED) and hospitalisations, finding that cancer comorbidity, anxiety comorbidity and the number of weekly alerts were significant predictors of all-cause emergency department visits and hospital admissions (Radhakrishnan et al. 2013b). A study of 403 Medicare heart failure patients discharged from hospital to the care of a home-health care agency found that 29 per cent of the sample had a re-hospitalisation within 60 days, and that the number of prescribed medications and non-use of certain cardiac medications prescribed (Angiotensin Converting Enzyme Inhibitor/ Angiotensin II Receptor Blockers) were predictors of increased risk of these events (Radhakrishnan et al. 2013a). A study of a similar population of telehomecare users with heart failure found that a higher risk of re-hospitalisation over 60 days was related to living arrangements (living with others), overall health status, severe pain and skin problems, and lower rates with independence in dressing the lower body (Kang et al. 2016). One small UK study (Biddiss, Brownsell, and Hawley 2009) of 45 community-dwelling heart

failure patients receiving telemonitoring found associations between key medical events and the number of alerts per week, self-rated mobility, self-rated health and self-rated anxiety.

Two studies noted possible problems with false alarms and many "non-events". Radhakrishnan, Jacelon et al. (2013b) remark on the low numbers of "meaningful" alerts recorded (less than 5 per cent of all alerts resulted in a key medical event; false alarms were frequent). Biddiss, Brownsell et al. (2009) likewise report that only 6 per cent of alerts resulted in a key medical event and also suggests that self-reported health and symptom alerts were better predictors of key events than physiological indicators (vital signs measurements such as blood pressure), 86 per cent of which did not result in a key medical event.

In terms of predictors of health care use and costs in the wider population with these chronic conditions, a number of studies have examined these using survey or administrative data. A variety of personal characteristics, area-level and organisational factors have been examined. Several studies have identified drivers of higher health care use in the COPD population. Personal characteristics can predict health care utilisation: older age (Hutchinson et al. 2010), being female (Menn et al. 2012, Hetlevik, Melbye, and Gjesdal 2016), having any comorbidities (Hetlevik, Melbye, and Gjesdal 2016), having particular comorbidities such as arthritis, cancer, diabetes, CVD and stroke (Menn et al. 2012), heart failure (Hutchinson et al. 2010), having more advanced stages of the disease (Menn et al. 2012), ADL difficulties (Garcia-Polo et al. 2012) and lower educational attainment (Hetlevik, Melbye, and Gjesdal 2016). Clinical measures constitute another set of health care utilisation predictors: FEV1 and BMI (Garcia-Polo et al. 2012, Darnell et al. 2013), higher peripheral blood leukocytes and fibrinogen and lower SPO2 (blood oxygen saturation) (Garcia-Polo et al. 2012). In addition, increased health care usage is associated with management by particular medical specialties (Darnell et al. 2013) and with characteristics of general practitioners and practices (Hetlevik, Melbye, and Gjesdal 2016). Lastly, use of pharmaceuticals such as the number of prescriptions and use of prescriptions such as inhaled corticosteroid and short acting anticholinergics (Darnell et al. 2013) and use of home oxygen (Hutchinson et al. 2010) are associated with increased utilisation in COPD. A study of patients with COPD, diabetes or at cardiovascular risk in the care of Dutch disease management providers found that having cardiovascular disease as a comorbidity, higher comorbidity on the Charlson index and lower EQ-5D-3L scores were associated with higher health care costs; also in patients with COPD, being in employment was associated with decreased costs (Tsiachristas and Rutten-van Mölken 2014).

I examined the impact of telehealth on subgroup costs by exploring the effect of the intervention at baseline and 12-month follow-up on the costs of participants with the three index conditions (COPD, diabetes and heart failure), as well as that of participants' socio-demographic and needs-related characteristics. Given that the approach taken was to examine the impact of the index conditions across the sample rather than separately examine costs of each condition, covariates in the models were necessarily generic to the whole telehealth sample, rather than condition-specific (e.g. severity of disease). The questionnaire dataset did not include any clinical measures or physician-level personal characteristics.

The analysis addressed the question: does the impact of telehealth on costs differ between the different index conditions, COPD, HF and diabetes? And does the answer change depending on the sector for which costs are measured – secondary (hospital) NHS care, primary and community NHS care, or social care?

6.2 Methods Used in the Telehealth Cost Subgroup Analyses

In this section I provide an overview of models employed to examine subgroup difference in costs related to receipt of telehealth. I constructed models of total health and social care costs (including and excluding intervention costs) and agency-specific costs (NHS and social care, hospital care). The composition of cost categories are given in Chapter 4, Table 4.2.

6.2.1 Models

Multilevel models were fitted to the costs data. Multilevel and population-averaged models of the relationship between a continuous response variable and a set of covariates were presented in the methods chapter (4.19.2). A three-level difference-in-difference-in-difference (DDD) approach was described in equation (4.6). The costs of telehealth participants can be described as a function that includes a set of covariates:

$$COSTS = f[TH, LTC, Followup, Age, Education, Female, Ethn, Comorb, Owns, Site, IMD, Selfcare],$$

where f(.) is any function described in the models given in Chapter 4, 4.19.2. Here, TH is the treatment allocation. *LTC* is a categorical variable for index long-term conditions. Dummy variables were created from the LTC variable to indicate which long-term conditions the participant had (these are labelled *COPD*, *HF* and *Diab*). *Followup* is an indicator for the 12-

month follow-up vs. the baseline time point. *Age* is a categorical variable[18], Education is a three-category variable (no formal, GCSE/O/A-level or degree-level qualifications), Female identifies women and men in the sample, *Comorb* is a count of chronic conditions sourced from acute hospital records (Steventon et al. 2012), *Ethn* is a binary indicator of white-British/non-white British ethnicity, *IMD* is a continuous measure of deprivation based on the Index of Multiple Deprivation 2007 (Noble et al. 2008), *Site* identifies the participating local authority, *Selfcare* is an indicator of ADL need based on the self-care domain of the EQ-5D-3L(no problems, some problems, unable to wash or dress) and *Owns* is an indicator of owner-occupation vs. renting and other forms of tenure.

As discussed in Chapter 4, Section 4.19.5, subject-specific and population-averaged models are underpinned by different assumptions and therefore have different interpretations. While the former are invaluable for exploring how a participant's costs changed in response to having telehealth, they are less useful in answering policy questions such as 'how did the costs in the intervention group differ from the costs of the controls?' An important, if self-evident point should be made, that marginal models can only estimate the average impact of one random effect, not the impact over multiple levels of nesting. Predicted probabilities generated by subject-specific regressions of dichotomous variables should match those produced by population-averaged models (Heagerty and Kurland 2001, Rabe-Hesketh and Skrondal 2012); however where non-linear transformations of continuous response variables are concerned this does not apply (see section 4.19.4).

Returning to the models described by equations (4.6) to (4.9), the coefficients on each triple-interaction term in the equations (the difference-in-difference-in-difference estimator), here called HF and $Diab$ (the reference category being COPD) can be understood as (see Section 994.19.3):

$$HF = \left[\left(\overline{y}_{T,COPD,2} - \overline{y}_{T,COPD,1}\right) - \left(\overline{y}_{T,HF,2} - \overline{y}_{T,HF,1}\right)\right] - \left[\left(\overline{y}_{C,COPD,2} - \overline{y}_{C,COPD,1}\right) - \left(\overline{y}_{C,HF,2} - \overline{y}_{C,HF,1}\right)\right]$$

$$Diab = \left[\left(\overline{y}_{T,COPD,2} - \overline{y}_{T,COPD,1}\right) - \left(\overline{y}_{T,DIAB,2} - \overline{y}_{T,DIAB,1}\right)\right] - \left[\left(\overline{y}_{C,COPD,2} - \overline{y}_{C,COPD,1}\right) - \left(\overline{y}_{C,DIAB,2} - \overline{y}_{C,DIAB,1}\right)\right]$$

In the marginal model, HF can be interpreted as the ratio of follow-up to baseline costs in intervention participants with COPD (the reference category) vs. this ratio in HF participants; $Diab$ can be interpreted as the ratio of follow-up to baseline costs in intervention participants with COPD vs. this ratio in diabetes participants. It is less straightforward to interpret the

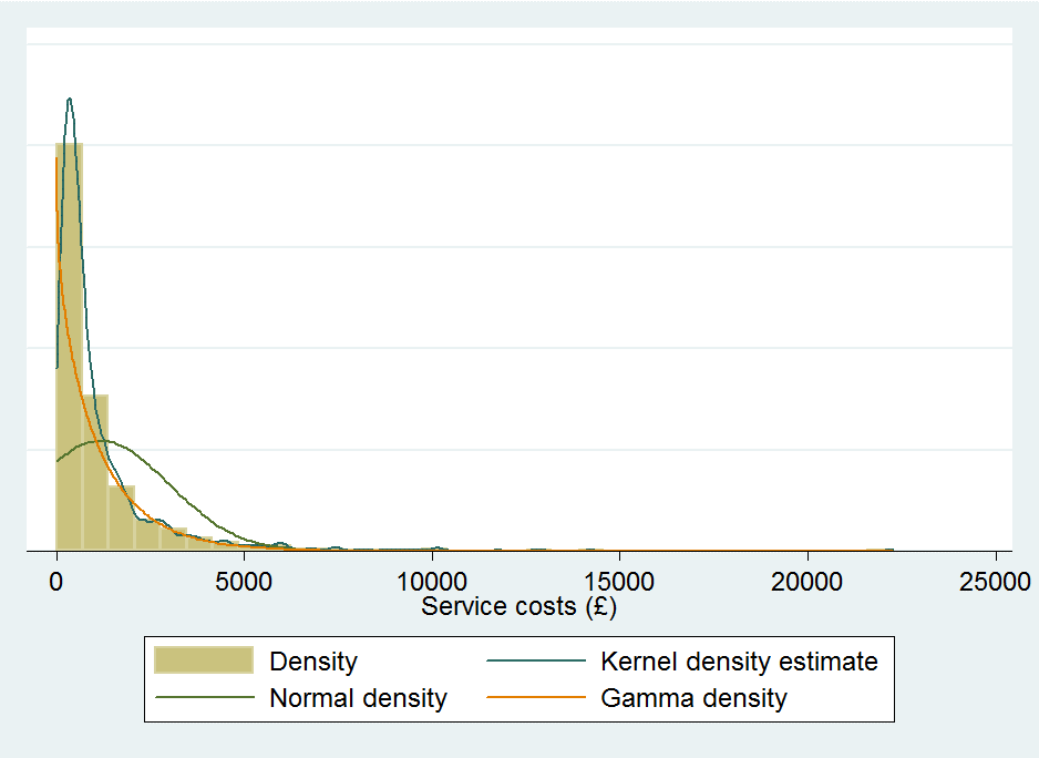
[18]Ages were categorised into 4 bands: 1 "under 65 (young)"; 2 "65-74 (young old)"; 3 "75-84 (old old)"; and 4 "85+ (oldest old)".

subject-specific model: the costs ratio for a *given* participant with COPD in the intervention group can be compared to the costs ratio of a participant with HF in the intervention group with the same values of the random intercept (Rabe-Hesketh and Skrondal 2012).

6.3 Distributions of Telehealth Costs and Clustering Effects

The costs of health and social care (excluding the intervention) for study participants were right-skewed, as can be seen in Figure 6.1. The figure depicts their frequency distribution and also the corresponding density functions of the gamma and normal distributions over both time points. Examining the density probability plots for these distributions, it is evident that these data fit the gamma better than the normal distribution.

Figure 6.1 Frequency and theoretical distributions of service costs



The clustering of costs for the 965 participants included in this analysis is examined in Table 6.3 at both participant and general practice level. The cost data (pooled across participants) were clustered within 145 general practices, with practices allocated to the intervention (76) outnumbering those allocated to control (69), where cluster sizes were smaller in the control (12.5) than in the intervention group (14.5). Costs pooled across baseline and follow-up points varied somewhat more within than between GP clusters in the intervention; the difference in the amount of variation within and between clusters was more

pronounced in the control group. At level 1, the pattern is reversed, with more variation between participants than within participants in both groups, not surprising given the limited time points under observation. This pattern can also be summarised by the intra-cluster correlation (ICC) (calculated by one-way analysis of variance (Ukoumunne 2002)). Pooling observations over time, the ICC for general practice is lower in the control than in the intervention group. The ICC at level two is slightly higher in the control than in the intervention group, but the ICCs are broadly similar between groups. These statistics suggest that costs vary somewhat by the cluster-randomisation unit (and more so in the intervention clusters), and to a greater extent within-person over time; however they do not take into account the influence of important confounders such as age and sex. The ICC of costs for general practice, examined by time point, is higher in the intervention than in the control clusters, particularly at follow-up; the ICC is higher and negative in the controls at baseline compared to the smaller and positive ICC at follow-up. This suggests that there is more variability within the control than the intervention GP clusters (one reason being that the average number of participants per control cluster is smaller).

6.4 Costs of Participant Subgroups with COPD, Heart Failure and Diabetes

The population of the Telehealth questionnaire study (see Section 5.2) was predominantly male; while many participants were in their older years, just under a third were less than 65 years of age. Around a quarter of the sample lived alone. Most people (67 per cent) had one or more comorbid conditions. The raw costs of participants with index conditions of COPD and heart failure (Table 6.2) were in general similar in both experimental groups at baseline; however costs were somewhat higher for the people with diabetes allocated to telehealth.

At follow-up, between-group differences in total costs (excluding those of the intervention) were slightly greater than at baseline for participants with COPD and heart failure but less for those with diabetes (6 per cent less in TH than in UC). With the addition of the cost of the TH intervention, the differences between experimental groups were uniformly somewhat higher.

Table 6.1 Health and social care service costs (£): cluster means, counts and intra-cluster correlation coefficients (ICC)

Time point (level 1)				
Usual care	Mean	SD	Count	ICC[a]
	1,251	1,798	862	0.204 (0.113,0.295)
		between subject	n	
		1,395	431	
		within subject	Mean n[b]	
		1,135	2	
TH	Mean	SD	Count	ICC[a]
	1,167	1,861	1,068	0. 204 (0.123, 0.285)
		between subject	n	
		1,444	534	
		within subject	Mean n[b]	
		1,175	2	
Participant unit (level 2)				
	Mean	SD	Count	ICC[a]
Usual care	1,251[c]	1,394	862	0.021 (-1.014,0.057)
		between practice	n	
		750	69	
		within practice	Mean n[b]	
		1,275	12.49	
TH	Mean	SD	Count	ICC[a]
	1167[c]	1443	1068	0.062 (0.017, 0.107)
		between practice	n	
		1001	76	
		within practice	Mean n[b]	
		1282	14.05	

Note: costs at baseline and follow-up, excludes costs of intervention. Imputed data.
a. Intra-cluster correlation, calculated by one-way analysis of variance; Searle's Confidence intervals report arithmetic mean cluster size for unbalanced data (Ukoumunne 2002)
b. Average number of units under observation
c. Participant costs pooled across time points

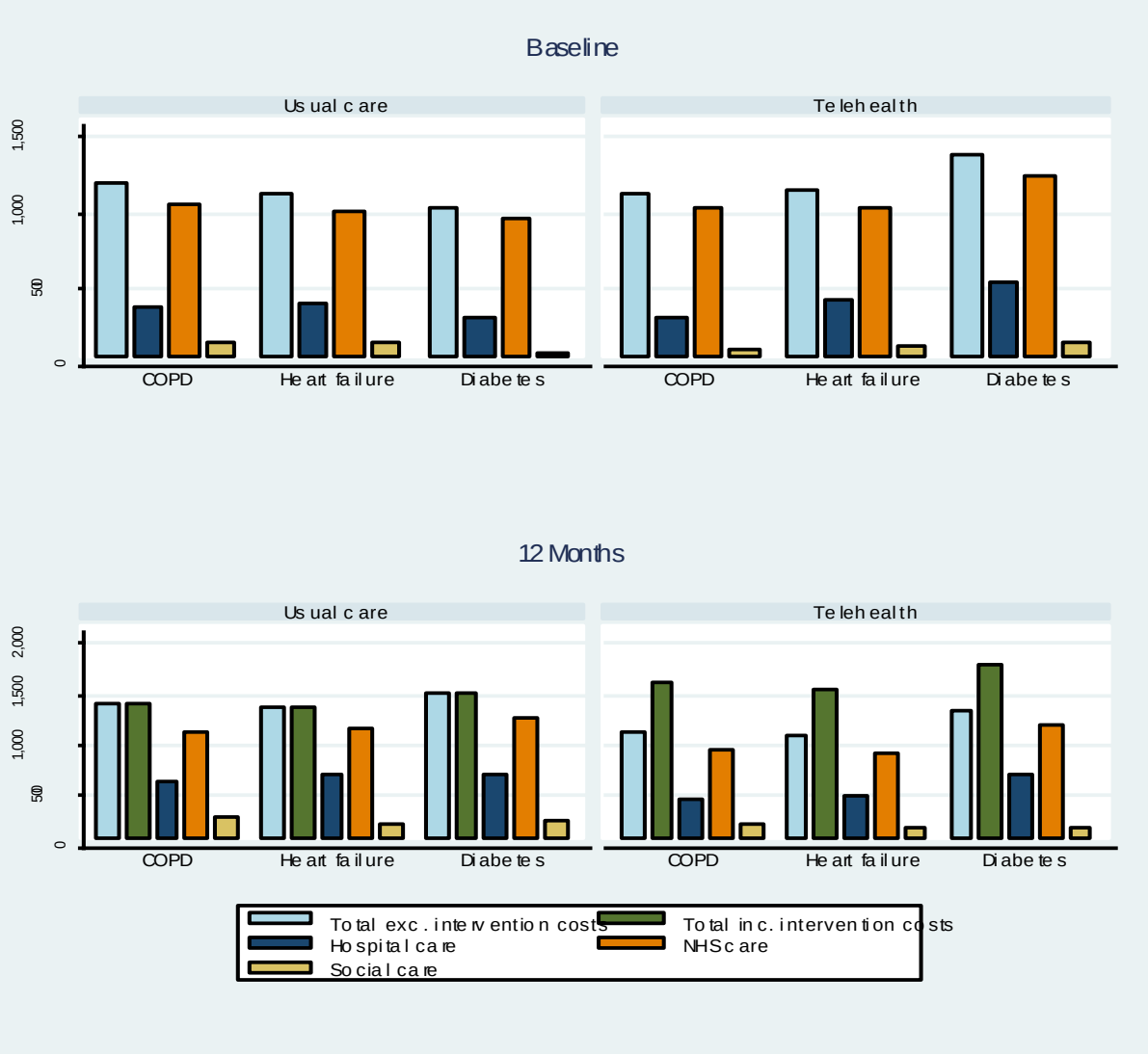
Table 6.2 Mean costs (clustered standard errors) at baseline and 12 months, by ITT allocation and index condition, participants with complete data

	COPD			**Heart Failure**			**Diabetes**		
	Usual care (SE) (n=140)	**Telehealth (SE) (n=232)**	**Diff.**	**Usual care (SE) (n=175)**	**Telehealth (SE) (n=177)**	**Diff.**	**Usual care (SE) (n=116)**	**Telehealth (SE) (n=125)**	**Diff.**
Total costs exc. delivery and equipment									
Baseline	1179 (117)	1119 (92)	-60	1112 (117)	1134 (117)	22	1015 (214)	1371 (218)	356
Follow-up	1387 (191)	1125 (148)	-262	1348 (137)	1065 (134)	-283	1471 (298)	1320 (302)	-152
Total costs inc. delivery and equipment									
Baseline	-	-	-	-	-	-	-	-	-
Follow-up	1391 (191)	1581 (148)	192	1357 (137)	1516 (134)	160	1489 (301)	1785 (306)	296
NHS[a]									
Baseline	1045 (116)	1030 (93)	-15	988 (110)	1016 (110)	28	948 (184)	1228 (184)	281
Follow-up	1105 (164)	931 (131)	-173	1142 (123)	903 (119)	-239	1224 (253)	1166 (255)	-58
Hospital only									
Baseline	367 (118)	308 (100)	-60	381 (79)	422 (79)	41	304 (158)	536 (157)	232
Follow-up	620 (133)	455 (107)	-164	687 (113)	483 (112)	-204	705 (246)	694 (253)	-10
Community care (LA)									
Baseline	134 (40)	89 (31)	-45	124 (51)	118 (52)	-6	67 (64)	142 (66)	75
Follow-up	282 (89)	193 (69)	-89	206 (54)	163 (54)	-43	248 (66)	154 (66)	-93

Note: Imputed data (10 completed datasets).

a. Hospital, primary and community mental health services.

Figure 6.2 Total, hospital, NHS and social care costs by index condition at baseline and follow-up



6.5 Model Results of Telehealth Subgroup Analyses: Total Costs

Several subject-specific models were fitted to the data in order to examine the impact of clustering at general practice and person levels. To begin with, three-level constant-only models were fitted to total costs and sub-cost categories (Appendix 3, A3.1, Table A3.1). For total costs, the level 3 variance ($\sigma_{\mu}^{2(3)}$, representing the ratio of the mean of the practice to the overall mean) was 1.041 and less than the level 2 variance ($\sigma_{\mu}^{2(2)}$, representing the ratio of the mean of person to the GP cluster mean), which was 1.366 and thus contributing less (33% vs. 43%) to the total variation in costs (A3.1, Table A3.1, model (1)). Approximately 25 per cent of the total cost variation in this model (coefficient of 0.781, representing the ratio of costs at each time point to the person's overall mean) is attributable to level 1. The inclusion of a

random intercept for general practice and DDD model with no other covariates was significantly different from zero (p=0.037) (Appendix 3, A3.1, Table A3.1, model (2)). The addition of the DDD variables reduced the amount of variation in costs due to practice-level factors and increased the amount of variation in costs due to participant characteristics. The inclusion of further covariates for needs-related and other personal characteristics (A3.1, Table A3.1, model (3)) greatly reduced the level 3 variance, suggesting that the full set of independent variables accounted for much of the variation at that level. A two-level constant-only model (Appendix 3, Table A3.1, model (4)) unsurprisingly had a higher level 2 variance, of 1.414 (95 per cent CI 1.319, 1.516), but a very similar scale parameter of 0.782 (95 per cent CI 0.716, 0.854). With the addition of other covariates, variation between participants was little different than in the three-level model (A3.1, Table A3.1, model (6)). In the two-level case, the estimated random intercept was 1.237, and the confidence limits of the estimate indicated that 95 per cent of participants can be expected to have an intercept between 1.169 and 1.309. In the three-level case, the random intercept estimate was slightly higher (1.234) but as 95 per cent of participants can be expected to have an intercept between 1.158 and 1.315, there is no evidence that the three-level model fits better than the two-level version.

Results of models fitted on the costs of health and social care (excluding intervention costs), NHS, hospital and social care were similar. In each case, F-tests of the level three random intercept indicated that the general practice level variance was not significantly different from zero at the 5 per cent level; although in some cases the level 3 intercept could not be estimated and models could not converge with the inclusion of the full set of covariates.

In summary, the inclusion of the third (general practice) level did not improve the fit of the cost models. I focus on the two-level model results hereon.

6.5.1 Two-level Subject-specific and Population-averaged Analyses: Total Costs

The interaction effects of allocation, time point and condition were jointly significantly different from zero, whether including or excluding intervention costs (Table 6.3). Looking at the time-by-allocation interaction term, total costs were greater in the last three months of the trial (by 53 per cent, p=0.001) for a given participant in the TH group relative to controls at the baseline. In the model excluding intervention costs, the time-by-allocation interaction term was not significantly different from zero (p=0.3).

The education dummy variables were jointly associated with costs (including or excluding intervention costs, p=0.003 and p=0.002 respectively), as were the age and ADL function dummies (p=0.000, including or excluding intervention costs). Having difficulty with self-care increased costs by 56 per cent and 64 per cent, including or excluding intervention costs, respectively. Being female was associated with increased total costs excluding those of the intervention (19 per cent higher). Each additional comorbid condition was associated with an increase in costs (including or excluding those of the intervention) in the order of 14 per cent to 16 per cent.

The population-averaged model results for health and social care costs differed somewhat from those of the subject-specific model. The interaction effects of allocation, time point and condition were not jointly significantly different from zero at the 5 per cent level, including intervention costs or excluding intervention costs (Table 6.3). Being female did not increase costs (excluding intervention costs) in this model.

In other respects, results were similar across subject-specific and population-averaged models in terms of association of costs with a number of characteristics. Tenure did not appear to influence the total costs of care and small-area level deprivation score was also not associated with costs at the 5 per cent level in any model. Costs of participants having difficulty with ADLs were estimated to be more than one and a half times as great as the costs of participants without ADL difficulty; being unable to self-care increased costs more than two-fold. Education, age and ADL function dummies were jointly significant as in the subject-specific models (at the 5 per cent, 1 per cent and 0.1 per cent level respectively).

6.6 Model Results of Telehealth Subgroup Analyses - Marginal Effects

To explore the implications of the models, the partial effect of treatment allocation on total costs (excluding/including intervention costs) was decomposed by time and index long-term condition (Table 6.4 to Table 6.7). In the subject-specific model results (Table 6.4), at the 12-month follow-up, the costs (including intervention costs) of intervention group participants were higher than in controls across the three conditions. For COPD participants, the difference in the differences between intervention and control costs at follow-up and baseline (DD) was significant (p=0.000). The pattern of results in the heart failure (HF) group were somewhat similar but the estimate was not significant at the 5 per cent level (p=0.059). There was little indication of difference in the cost differences between time points in the diabetic group.

Table 6.3 Parameter estimates, subject specific (random intercept) and population-averaged (GEE) models of costs (£) in 3 months prior to baseline and 12-month follow-up

	Subject-specific		*Population average*	
	Excluding intervention costs	*Including intervention costs*	*Excluding intervention costs*	*Including intervention costs*
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
TH	1.028	1.027	1.003	1.004
	(0.106)	(0.106)	(0.134)	(0.131)
Follow-up	0.955	0.965	0.995	1.017
	(0.105)	(0.107)	(0.133)	(0.136)
TH*Follow-up	0.865	1.534**	0.899	1.353+
	(0.121)	(0.207)	(0.159)	(0.221)
HF	0.881	0.887	0.859	0.868
	(0.098)	(0.099)	(0.130)	(0.129)
Diab	0.874	0.863	0.830	0.823
	(0.106)	(0.103)	(0.132)	(0.126)
TH*HF	1.057	1.048	1.097	1.080
	(0.158)	(0.156)	(0.210)	(0.203)
TH*Diab	1.249	1.247	1.351	1.353
	(0.202)	(0.202)	(0.281)	(0.275)
HF*Follow-up	1.255	1.252	1.356	1.319
	(0.191)	(0.193)	(0.292)	(0.281)
Diab*Follow-up	1.355*	1.375*	1.338	1.341+
	(0.206)	(0.210)	(0.240)	(0.238)
TH*Follow-up*HF	0.816	0.794	0.698	0.728
	(0.164)	(0.151)	(0.192)	(0.185)
TH*Follow-up*Diab	0.787	0.718	0.755	0.712
	(0.169)	(0.145)	(0.209)	(0.179)
Young old	1.161*	1.108+	1.131	1.085
	(0.078)	(0.067)	(0.088)	(0.077)
Old old	1.176*	1.115	1.205*	1.146
	(0.087)	(0.076)	(0.111)	(0.097)
Oldest old	1.823***	1.653***	1.616***	1.515***
	(0.173)	(0.149)	(0.171)	(0.154)
GCSE/O/A-level	1.230***	1.192***	1.211**	1.180*
	(0.072)	(0.062)	(0.087)	(0.076)
Degree-level	1.136+	1.121+	1.067	1.066
	(0.087)	(0.077)	(0.095)	(0.085)
Female	1.119*	1.099*	1.068	1.056
	(0.058)	(0.051)	(0.067)	(0.059)
White-British	1.212+	1.186+	1.198	1.172
	(0.132)	(0.119)	(0.170)	(0.152)
Comorbidities	1.162***	1.144***	1.160***	1.145***
	(0.017)	(0.015)	(0.020)	(0.018)
Owns	0.876*	0.888*	0.838*	0.853*
	(0.057)	(0.053)	(0.073)	(0.068)

	Subject-specific		*Population average*	
	Excluding intervention costs	*Including intervention costs*	*Excluding intervention costs*	*Including intervention costs*
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
Site 2	1.034 (0.065)	1.074 (0.061)	1.039 (0.082)	1.073 (0.075)
Site 3	1.006 (0.099)	1.078 (0.098)	1.065 (0.134)	1.124 (0.133)
IMD	1.005^{+} (0.003)	1.003 (0.003)	1.004 (0.003)	1.002 (0.003)
Some problems	1.636*** (0.090)	1.556*** (0.078)	1.657*** (0.106)	1.578*** (0.091)
Unable wash/dress	2.722*** (0.462)	2.460*** (0.381)	3.022*** (0.549)	2.675*** (0.432)
Level 1 constant	549.929*** (86.246)	577.705*** (84.578)	675.255*** (126.057)	705.170*** (120.005)
σ	0.798*** (0.044)	0.746*** (0.035)		
$\sigma^2[u]$	1.321*** (0.049)	1.237*** (0.036)		
N_i	1930	1930		1930
Interaction effects	F(7.000, 7951772)=2.701 p=0.008	F(7.000, 14186249)=2.055 p=0.045	F(7.000, 43309927)=1.913 p=0.063	F(7.000, 51257609)=0.877 p=0.523

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Table 6.4 Two-level subject-specific model: Partial effect/discrete difference in costs (£) between TH and UC, (ITT allocation) between baseline and follow-up, by index condition

	Excluding intervention costs						*Including intervention costs*					
	COPD		HF		Diab		COPD		HF		Diab	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	32 (122)	0.790	96 (124)	0.437	304 (156)	0.052	29 (115)	0.798	79 (115)	0.493	281 (146)	0.055
Follow-up	-126 (131)	0.335	-325* (148)	0.028	-182 (178)	0.308	623*** (135)	0.000	409** (149)	0.006	566 ** (184)	0.002
Follow-up-Baseline	-159 (158)	0.315	-421 (177)	0.017	-486 (213)	0.023	593 (160)	0.000	330 (174)	0.059	285 (213)	0.18

Table 6.5 Two-level subject-specific model: difference-in-difference-in-difference (DDD)

	Excluding intervention costs		*Including intervention costs*	
	Exp (β) (95% CI)	*p*	Exp (β) (95% CI)	*p*
DDD HF- COPD	-262 (-725,200)	0.267	-264 (-720, 193)	0.258
DDD Diab - COPD	-327 (-847,193)	0.218	-308 (-826, 210)	0.244
DDD Diab - HF	-64.377 (-606, 477)	0.816	-44 (-579, 490)	0.871
Total costs (sum of DD)	-1065 (-1691,-440)	0.001	1208 (573,1843)	0
N_i	1930		1930	

Table 6.6 Population-averaged model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by index condition

	Excluding intervention costs						*Including intervention costs*					
	COPD		HF		Diabetes		COPD		HF		Diabetes	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	4 (158)	0.980	109 (150)	0.466	370 (207)	0.074	4 (150)	0.978	90 (144)	0.534	361 (197)	0.067
Follow-up	-117 (164)	0.476	-476* (237)	0.044	-116 (233)	0.619	426* (166)	0.011	102 (229)	0.655	440 (230)	0.055
Follow-up-Baseline	-121 (204)	0.553	-585 (274)	0.032	-486 (286)	0.09	422 (202)	0.037	12 (265)	0.963	79 (278)	0.776

Table 6.7 Population-averaged model: difference-in-difference-in-difference (DDD) costs estimates (£)

	Excluding intervention costs		*Including intervention costs*	
	Exp (β) (95% CI)	*p*	Exp (β) (95% CI)	*p*
DDD HF- COPD	-464 (-1132, 204)	0.173	-409 (-1059, 240)	0.217
DDD Diab - COPD	-364.817 (-1056, 326)	0.302	-342 (-1020, 335)	0.322
DDD Diab - HF	99.485 (-677, 876)	0.802	67 (-684, 818)	0.862
Total costs	-1192.29 (-2065, 319)	0.007	513 (-339, 1366)	0.238
N_i	1930		1930	

The marginal effects (the DDD), whether comparing the cost differences over time between COPD and HF, HF and diabetes, or COPD and diabetes, were not significant at the 5 per cent level. Adding together the effects of the intervention over the last three months of the intervention period on all three condition groups gave an adjusted estimate of £1208 ($p=0.000$) greater than in the control group. Table 6.4 also presents the marginal effects results for health and social care costs excluding direct intervention costs. The cost differences between intervention and control were significantly lower ($p<0.05$) at 12-month follow-up (that is, the DD) in the case of HF and diabetic participants. However the results do not suggest that there were substantial differences between subgroups. Adjusting for demographic and needs-related characteristics, total estimated costs excluding those of the intervention of participants allocated to telehealth were significantly lower at the 1 per cent level, by £1065.

The population-averaged approach (Table 6.6) produced somewhat different marginal effects estimates of total cost differences. The DD for the COPD group was about two-thirds of the size predicted by the subject-specific model. The DDs for heart failure and diabetic groups were small and not significantly different from zero. The DDD estimates, as with the subject-specific estimates, were not significantly different from zero at the 5 per cent level. The impact of the intervention on costs across the conditions was not significant. The marginal effects estimates for health and social care costs excluding the intervention were on the other hand broadly similar to those produced by the subject-specific model. There were no differences in the DD between conditions; however the overall impact across conditions was to decrease three-month costs significantly at the 1 per cent level, by £1192.

6.7 Telehealth Subgroup Costs by Sector

6.7.1 NHS Costs

Use of NHS services was near universal, between 99 per cent and 100 per cent of participants in the subgroups having NHS costs at either time point. The results of the random intercept model of NHS costs are given in Table 6.8. The interaction terms for time, allocation and condition were significantly different from zero. Having some difficulty with self-care increased costs by 51 per cent; being unable to self-care increased costs by 88 per cent. Costs were 25 per cent higher for a given participant attaining secondary school qualifications (GCSE/O- or A-levels) compared to a participant with no formal education. The education dummy variables were jointly significantly different from zero ($F(2.000,151080)=6.937$,

p=0.001). The coefficients for age categories were significantly different from zero (F(3.000,550208)=3.912, p=0.008). White British ethnicity was associated with 25 per cent higher costs than for other ethnicities. Each additional comorbidity was associated with increased costs, as in the total costs models. The coefficient on the interaction of time point and diabetes was significantly greater than zero (p=0.030), suggesting that costs at follow-up diverged for the participants with similar values of random intercept who had COPD and diabetic conditions.

In the population-averaged model (Table 6.8), the triple interaction terms were not jointly significant at the 5 per cent level. The coefficient on the interaction of time point and diabetes was not significantly greater than zero (p=0.091). Overall, higher levels of education was significantly associated with increased costs (F(2.000,864443)=4.179, p=0.015). For other covariates, coefficients were similar to the subject-specific model, but standard errors were larger.

Marginal effects: Results of the subject-specific model (Table 6.9 and Table 6.10) were similar to those for health and social care costs excluding intervention costs. The end-of-trial cost difference, net of the baseline cost difference, in heart failure participants indicates that intervention participants' NHS costs were significantly lower than controls. A similar result is found for the diabetic participants. None of the differences in the difference-in-differences between subgroups (DDD) was significant (Table 6.10). The estimated total cost savings across conditions, adjusted for subgroup differences and other covariates, suggests a substantial decrease in NHS costs (not including direct costs of the telehealth intervention) of £804 between time points.

The results of the population averaged model (Table 6.11) suggested larger DD estimates in the heart failure and diabetes groups than predicted by the subject-specific model; the DD estimate in the heart failure group was significant. As with the subject-specific model, there were no significant differences in the difference-in-differences between subgroups (DDD) (Table 6.12). The estimated three-month savings across conditions on the other hand remained large (£990, p=0.014).

Table 6.8 Parameter estimates, subject specific (random intercept) and population-averaged (GEE) models of NHS costs (£) in the 3 months prior to baseline and 12-month follow-up

	Subject-specific	*Population average*
Parameter	Exp (β) (SE)	Exp (β) (SE)
TH	1.045 (0.108)	1.029 (0.135)
Follow-up	0.854 (0.094)	0.916 (0.123)
TH*Follow-up	0.906 (0.126)	0.912 (0.164)
HF	0.868 (0.098)	0.864 (0.131)
Diab	0.894 (0.108)	0.852 (0.131)
TH*HF	1.057 (0.159)	1.065 (0.206)
TH*Diab	1.191 (0.194)	1.303 (0.275)
HF*Follow-up	1.333^{+} (0.207)	1.430 (0.324)
Diab*Follow-up	1.392* (0.212)	1.363^{+} (0.250)
TH*Follow-up*HF	0.789 (0.161)	0.688 (0.198)
TH*Follow-up*Diab	0.795 (0.172)	0.764 (0.219)
Young old	1.082 (0.074)	1.058 (0.083)
Old old	1.063 (0.080)	1.102 (0.105)
Oldest old	1.391*** (0.137)	1.255* (0.135)
GCSE/O/A-level	1.246*** (0.074)	1.231** (0.091)
Degree-level	1.107 (0.085)	1.042 (0.096)
Female	1.054 (0.055)	1.013 (0.065)
White-British	1.253* (0.138)	1.260 (0.178)
Comorb	1.164*** (0.017)	1.162*** (0.021)
Owns	0.872* (0.058)	0.831* (0.075)

	Subject-specific	*Population average*
Parameter	Exp (β) (SE)	Exp (β) (SE)
Site 2	1.009 (0.064)	1.024 (0.083)
Site 3	1.006 (0.101)	1.079 (0.137)
IMD	1.002 (0.003)	1.001 (0.003)
Some problems	1.512*** (0.084)	1.547*** (0.103)
Unable wash/dress	1.876*** (0.279)	1.976*** (0.356)
Level 1 constant	562.191*** (90.052)	666.723*** (125.926)
σ	0.797*** (0.044)	
$\sigma^2[u]$	1.343*** (0.048)	
N_i	1930	1930
Interaction effects	F(7.000,3742241)=2.414 p=0.018	F(7.000,27278471)=1.727 p=0.098

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Table 6.9 Two-level subject-specific model of NHS costs: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by condition

	COPD		HF		Diabetes	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	47 (110)	0.669	101 (110)	0.359	241 (142)	0.090
Follow-up	-48 (108)	0.654	-240 (127)	0.058	-125 (154)	0.416
Follow-up-Baseline	-95 (134)	0.478	-342 (155)	0.028	-367 (186)	0.034

Table 6.10 Two-level subject-specific model of NHS costs: difference-in-difference-in-difference (DDD) costs (£) estimates and total costs

	Exp (β) (95% CI)	p
DDD HF- COPD	-246 (-648,155)	0.229
DDD Diab - COPD	-271.11 (-723,181)	0.24
DDD Diab - HF	-24.691 (-500,451)	0.919
Total costs (sum of DD)	-804 (-1347,-261)	0.004
N_i	1930	

Table 6.11 Population-averaged model of NHS costs: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by condition

	COPD		HF		Diabetes	
Intervention vs. control	Exp (β) (SE)	p	Exp (β) (SE)	p	Exp (β) (SE)	p
Baseline	30	0.827	92	0.501	322	0.098
	(138)		(137)		(195)	
Follow-up	-60	0.674	-406	0.058	-81	0.690
	(142)		(214)		(202)	
Follow-up-Baseline	-90 (178)	0.614	-498 (250)	0.047	-403 (260)	0.056

Table 6.12 Population-averaged model of NHS costs: difference-in-difference-in-difference (DDD) costs (£) estimates and total costs

	Exp (β) (95% CI)	p
DDD HF - COPD	-407.665 (-1009,193)	0.184
DDD Diab - COPD	-313 (-932,306)	0.322
DDD Diab - HF	95 (-614, 803)	0.793
Sum of DD	-990 (-1778, -203)	0.014
N_i	1930	

6.8 Numbers and Proportions Using Hospital and Social Care Services

55 per cent of participants (n=528) at baseline and 56 percent (n=545) at 12-month follow-up had used some hospital service. There were some variations in proportions having used these services in the prior three months at baseline and 12 months (Table 6.15). In particular, fewer participants with COPD in the TH group had any social care than in the UC group at 12-

months. Use of hospital services by people with diabetes was particularly high across time points and experimental groups. Social care was less used than hospital care across the subgroups. The imputed datasets differed slightly in the number using any social care service, given that missing category costs were estimated. The numbers using social care were 243 at baseline; complete datasets at follow-up varied between 335 and 336.

6.9 Hospital Costs: Two-part Models

Results are given in Table 6.14 and Table 6.15. In the first (logistic) part of the subject-specific two-part model of receipt of any hospital care and all hospital costs, the interaction effects of time, condition and time point were not jointly significant. The interaction of time point and diabetes was significantly different from zero. Educational level was strongly associated with an increase in receipt of any hospital care (Table 6.14). The odds of receipt were about 1.6 times greater for those with a secondary school education than no formal education. The education dummies were jointly significant at 1 per cent (F(2.000,931645)=5.872, p=0.003). White British ethnic background nearly doubled the odds of receipt of care compared to other ethnicities. Each additional comorbidity increased the likelihood of receipt by 27 per cent. The ADL need dummies were jointly significant (F(3.000, 2.836e+10)=5.05, p=0.002). Neither age nor tenure was significantly associated with receipt of hospital care. The exponentiated covariance between the random effects for receipt and costs was significantly greater than zero (1.365, 95 per cent CI 1.07, 1.743, p=0.012), indicating that receipt of and costs of hospital care were positively correlated. In the second (gamma regression) part (Table 6.15), each extra comorbidity increased the costs of hospital care by 19 per cent. ADL needs dummies were jointly significant (F(2.000,3.075e+11)=5.969, p=0.003). The interaction effects were not jointly significant. Site was associated with neither receipt nor costs of hospital care.

In the population-averaged logistic regression of receipt, the interaction terms of the DDD estimator were jointly significantly different from zero. Estimates for most covariates were similar to those of the subject-specific model, although the size of the estimates and their standard errors were somewhat smaller. As with the subject-specific model, the interaction of time point and diabetes was significantly different from zero at the 5 per cent level; also ADL need and education were significant at the 1 per cent level (F(2.000,3.983e+09)=6.958, p=0.001; and F(2.000,830650)=6.040, p=0.002 respectively).

Table 6.13 Use of hospital and social care in the prior 3 months at baseline and 12-month follow-up: percentage (number) using service

	COPD			**Heart Failure**			**Diabetes**		
	Usual care (SE) (n=140)	**Telehealth (SE) (n=232)**	**Raw**	**Usual care (SE) (n=175)**	**Telehealth (SE) (n=177)**	**Raw**	**Usual care (SE) (n=116)**	**Telehealth (SE) (n=125)**	**Raw**
Hospital only									
Baseline	48% (67)	47% (109)	-1%	56% (98)	62% (110)	6%	53% (61)	66% (83)	14%
Follow-up	48% (67)	46% (107)	-2%	62% (108)	56% (99)	-16%	71% (82)	66% (82)	-5%
Community care (LA)*									
Baseline	29% (41)	24% (56)	-5%	31% (54)	27% (47)	-7%	18% (21)	19% (24)	1%
Follow-up	42% (59)	30% (69)	-12%**	41% (71)	34% (61)	-6%	30% (35)	32% (40)	2%

Note: the number of cases of social care use vary between multiply imputed datasets 1 to 10 (numbers per dataset: 335 in 8 and 336 in 2 complete datasets)

*Imputed data.

**Difference between usual care and telehealth χ^2=4.048 and p=0.044

The population-averaged gamma regression of costs for those in receipt of any hospital services yielded estimates that were somewhat larger and standard errors were in general substantially larger than the two-part subject-specific model; the constant was almost twice as large as the subject-specific estimate. Each extra comorbidity increased the costs of hospital care by 15 per cent. ADL needs dummies were jointly significant (F(2.000,3.075e+11)=5.969, p=0.003). Age was not significantly associated with costs (F(3.000,1.484e+11)=1.266, p=0.284).

Marginal effects: Examining the predictions generated by the subject-specific model (Table 6.16), there were no significant differences in costs of hospital care between intervention and control at baseline or at follow-up in any of the condition subgroups. Likewise there was little indication of difference in the DD between conditions. While the predictions generated by the two-part population-averaged model (Table 6.18 and Table 6.19) were derived from separate models, the standard errors are drawn from a bootstrapped distribution of estimates from these models. These predictions differed quite markedly from the subject-specific-derived results (in which standard errors are adjusted for the correlation between random effects of the receipt and costs models), particularly for the difference-in-difference estimates for groups with heart failure and diabetes. For the most part, the overall interpretation of the effects does not change: the differences between intervention and control at each time point are not significantly different nor are the DDD between conditions. The total savings across conditions due to telehealth estimated from either model are not significant.

6.10 Social Service Costs: Two-part Models

In the logistic part of the subject-specific model (Table 6.14), the interaction and main effects of time, condition and time point were not jointly significant. The odds of receipt over the three months prior to the 12-month follow-up point were 3.2 times higher than at baseline; being female increased the odds of receipt almost four-fold. Being unable to wash or dress increased the odds of social care receipt enormously over having no difficulties with ADLs (by 28 times). Each additional comorbid condition increased the odds of receipt by 18. Being an owner-occupier decreased the odds of receipt by 72 per cent.

The exponentiated covariance between receipt and cost models' random effects was greater than one, denoting a (non-significantly) positive relationship between these (ρ =2.007 (95 per cent CI 0.953, 4.225, p=0.067). In terms of costs of participants in receipt of some

form of social care (Table 6.15), cost in the three months prior to follow-up (Table 6.15) were 88 per cent higher than at in the three months prior to baseline. Higher costs were significantly associated with age (F(5.000,27061166)=14.999, p=0.000). The education dummies were not jointly significant (F(2.000,11441499)=0.891, p=0.410). Costs increased slightly (1.8 per cent) with higher (more deprived) IMD scores. Having moderate ADL difficulties increased the costs of care by 67 per cent and being unable to wash or dress increased costs by 279 per cent. Site was not a significant predictor of receipt or costs of social care in the Telehealth sample.

In the population-averaged logistic regression of receipt (Table 6.14), the estimates and standard errors were quite similar to the subject-specific model. Being at the follow-up time point and being female doubled the odds of receipt. Comorbidities were associated with higher odds of receipt. ADL need increased the odds of receipt so that the most severely ADL impaired were 7.6 times more likely to receive than those without problems washing and dressing. Home ownership decreased the odds by 54 per cent. Age dummies were jointly significant (F(3.000,2.278e+08)=35.058, p=0.000). The population-averaged gamma regression results (Table 6.15) did not suggest a significant difference in costs between baseline and follow-up, or between owners and other forms of tenure. The interaction effects of the DDD estimator were not significantly different from zero. The standard errors of the estimates were in general slightly smaller. Levels of ADL need (F(2.000,8.870e+09)=19.886, p=0.000) and mean IMD score remained significant at the 5 per cent level as did the age dummies (jointly) (F(3.000,7.416e+08)=4.247, p=0.005). The intercept was more than four times larger than the subject-specific estimate.

Table 6.14 Two-level model estimates of receipt from two-part subject-specific and population averaged models in 3 months prior to baseline and 12-month follow-up

	Subject-specific				*Population average*			
	Social care Random intercept		*Hospital care Random intercept*		*Social care*		*Hospital care*	
Parameter	β	Exp (β) (SE)	β	Exp (β) (SE)	β	Exp (β) (SE)	β	Exp (β) (SE)
TH	-0.047	0.954	0.039	1.040	-0.055	0.947	0.036	1.037
	(0.421)	(0.402)	(0.254)	(0.264)	(0.269)	(0.255)	(0.218)	(0.227)
Follow-up	1.158**	3.183**	0.024	1.024	0.722**	2.058**	0.020	1.021
	(0.367)	(1.167)	(0.262)	(0.269)	(0.228)	(0.469)	(0.226)	(0.230)
TH*Follow-up	-0.782+	0.457+	-0.085	0.918	-0.472	0.624	-0.074	0.929
	(0.467)	(0.214)	(0.337)	(0.309)	(0.293)	(0.183)	(0.289)	(0.269)
HF	0.228	1.256	0.304	1.356	0.155	1.168	0.265	1.304
	(0.440)	(0.552)	(0.269)	(0.365)	(0.283)	(0.330)	(0.233)	(0.303)
Diab	-0.501	0.606	0.198	1.219	-0.352	0.703	0.174	1.190
	(0.526)	(0.319)	(0.305)	(0.372)	(0.337)	(0.237)	(0.263)	(0.313)
TH*HF	-0.255	0.775	0.371	1.449	-0.185	0.831	0.318	1.375
	(0.587)	(0.455)	(0.359)	(0.520)	(0.380)	(0.316)	(0.310)	(0.426)
TH*Diab	0.093	1.098	0.671+	1.955+	0.100	1.106	0.575	1.778
	(0.679)	(0.746)	(0.406)	(0.795)	(0.440)	(0.486)	(0.350)	(0.623)
HF*Follow-up	-0.307	0.736	0.371	1.297	-0.186	0.831	0.226	1.254
	(0.476)	(0.350)	(0.359)	(0.454)	(0.300)	(0.249)	(0.302)	(0.379)
Diab*Follow-up	0.023	1.024	0.671+	2.422*	0.035	1.036	0.762*	2.142*
	(0.567)	(0.581)	(0.406)	(0.918)	(0.359)	(0.372)	(0.327)	(0.699)
TH*Follow-up*HF	0.519	1.681	-0.543	0.581	0.308	1.361	-0.471	0.625
	(0.653)	(1.098)	(0.479)	(0.278)	(0.415)	(0.565)	(0.414)	(0.258)
TH*Follow-up*Diab	0.832	2.298	-0.862	0.422	0.516	1.675	-0.741	0.476
	(0.767)	(1.763)	(0.535)	(0.226)	(0.485)	(0.812)	(0.462)	(0.220)

	Subject-specific				Population average			
	Social care Random intercept		Hospital care Random intercept		Social care		Hospital care	
Parameter	β	Exp (β) (SE)	β	Exp (β) (SE)	β	Exp (β) (SE)	β	Exp (β) (SE)
Young old	1.097***	2.994***	0.116	1.123	-2.104***	2.059***	0.102	1.107
	(0.290)	(0.867)	(0.158)	(0.177)	(0.400)	(0.393)	(0.136)	(0.151)
Old old	2.104***	8.200***	-0.017	0.983	0.722***	3.918***	-0.008	0.992
	(0.311)	(2.547)	(0.165)	(0.162)	(0.191)	(0.761)	(0.143)	(0.142)
Oldest old	4.209***	67.322***	0.313	1.368	1.366***	14.399***	0.270	1.310
	(0.496)	(33.395)	(0.261)	(0.357)	(0.194)	(4.056)	(0.227)	(0.298)
GCSE/O/A-level	-0.175	0.839	0.449***	1.567***	2.667***	0.908	0.389***	1.475***
	(0.229)	(0.192)	(0.134)	(0.209)	(0.282)	(0.135)	(0.114)	(0.168)
Degree-level	0.272	1.313	0.317	1.373	-0.097	1.194	0.278	1.321
	(0.340)	(0.446)	(0.200)	(0.274)	(0.148)	(0.262)	(0.174)	(0.229)
Female	1.382***	3.982***	0.071	1.073	0.177	2.487***	0.064	1.066
	(0.221)	(0.880)	(0.122)	(0.131)	(0.220)	(0.332)	(0.106)	(0.113)
White-British	-0.025	0.975	0.665*	1.944*	0.911***	0.967	0.569*	1.766*
	(0.416)	(0.405)	(0.264)	(0.514)	(0.134)	(0.265)	(0.228)	(0.403)
Comorb	0.167**	1.182**	0.240***	1.272***	-0.033	1.114**	0.207***	1.230***
	(0.059)	(0.070)	(0.035)	(0.045)	(0.274)	(0.041)	(0.030)	(0.037)
Owns	-1.258***	0.284***	0.176	1.192	0.108**	0.463***	0.152	1.164
	(0.236)	(0.067)	(0.137)	(0.163)	(0.037)	(0.068)	(0.119)	(0.138)
Site 2	-0.007	0.993	0.075	1.078	-0.770***	1.016	0.065	1.067
	(0.247)	(0.245)	(0.143)	(0.154)	(0.147)	(0.162)	(0.123)	(0.132)
Site 3	-0.241	0.786	0.275	1.317	0.016	0.863	0.229	1.257
	(0.371)	(0.291)	(0.229)	(0.302)	(0.160)	(0.212)	(0.198)	(0.249)

	Subject-specific				Population average			
	Social care Random intercept		Hospital care Random intercept		Social care		Hospital care	
Parameter	β	Exp (β) (SE)	β	Exp (β) (SE)	β	Exp (β) (SE)	β	Exp (β) (SE)
IMD	0.006 (0.010)	1.006 (0.010)	0.007 (0.006)	1.007 (0.006)	-0.147 (0.245)	1.005 (0.007)	0.006 (0.006)	1.006 (0.006)
Some problems	1.220*** (0.200)	3.387*** (0.678)	0.452*** (0.130)	1.572*** (0.204)	0.005 (0.007)	2.123*** (0.269)	0.406*** (0.112)	1.501*** (0.168)
Unable wash/dress	3.243*** (0.603)	25.600*** (15.441)	0.487 (0.345)	1.627 (0.561)	0.753*** (0.127)	7.626*** (2.732)	0.450 (0.301)	1.569 (0.472)
Level 1 constant	-3.280*** (0.640)	0.038*** (0.024)	-1.326*** (0.386)	0.265*** (0.102)	2.032*** (0.358)	0.122*** (0.049)	-1.149*** (0.332)	0.317*** (0.105)
$\sigma^2[u]$	1.125+ (0.646)	46.509*** (35.693)	0.858*** (0.070)	2.140** (0.504)		- -		- -
N_i	1930		1930		1930		1930	
Interaction effects	F(7.000,3.339e+08)=0.659 p=0.707		F(7.000,7.489e+11)=2.004 p=0.051		F(7.000,57602794.181)= 1.018 p=0.416		F(7.000,6.947e+11)=2. 019 p=0.049	

Note: the number of cases of social care use vary between multiply imputed datasets 1 to 10 (numbers per dataset: 335 in 8 and 336 in 2 complete datasets)

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Table 6.15 Two-level model estimates of costs from two-part subject-specific and population averaged models in 3 months prior to baseline and 12-month follow-up

	Subject-specific		*Population average*	
	Social care Random intercept	*Hospital care Random intercept*	*Social care*	*Hospital care*
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
TH	0.851	0.840	0.981	0.869
	(0.357)	(0.166)	(0.433)	(0.228)
Follow-up	1.859*	1.371	1.267	1.475+
	(0.573)	(0.287)	(0.409)	(0.333)
TH*Follow-up	0.855	0.909	0.920	0.988
	(0.361)	(0.248)	(0.461)	(0.323)
HF	0.888	0.865	0.669	0.790
	(0.350)	(0.175)	(0.247)	(0.220)
Diab	0.730	0.733	0.652	0.701
	(0.379)	(0.165)	(0.339)	(0.192)
TH*HF	1.207	1.105	1.030	1.177
	(0.648)	(0.297)	(0.550)	(0.416)
TH*Diab	2.507	1.108	1.743	1.569
	(1.687)	(0.347)	(1.173)	(0.649)
HF*Follow-up	0.752	1.014	0.935	1.365
	(0.295)	(0.304)	(0.373)	(0.490)
Diab*Follow-up	1.590	1.139	1.574	1.107
	(0.897)	(0.342)	(0.928)	(0.356)
TH*Follow-up*HF	1.052	0.939	1.058	0.613
	(0.606)	(0.362)	(0.656)	(0.298)
TH*Follow-up*Diab	0.395	1.009	0.496	0.731
	(0.291)	(0.420)	(0.383)	(0.392)
Young old	1.333	1.121	1.373	1.098
	(0.364)	(0.133)	(0.294)	(0.154)
Old old	1.314	0.997	1.182	1.117
	(0.373)	(0.131)	(0.253)	(0.190)
Oldest old	3.195**	1.608*	2.245**	1.464+
	(1.181)	(0.314)	(0.612)	(0.289)
GCSE/O/A-level	1.171	1.245*	1.202	1.155
	(0.221)	(0.129)	(0.183)	(0.141)
Degree-level	1.399	1.105	1.204	0.861
	(0.334)	(0.144)	(0.219)	(0.118)
Female	1.055	1.001	0.895	0.927
	(0.203)	(0.095)	(0.120)	(0.105)
White-British	0.722	1.003	0.620	0.935
	(0.261)	(0.182)	(0.180)	(0.214)
Comorb	1.064	1.194***	1.080+	1.152***
	(0.048)	(0.030)	(0.050)	(0.034)

	Subject-specific		*Population average*	
Parameter	*Social care Random intercept*	*Hospital care Random intercept*	*Social care*	*Hospital care*
Owns	1.629*	0.853	1.264	0.807
	(0.318)	(0.099)	(0.212)	(0.124)
Site 2	1.316	1.074	1.193	1.099
	(0.256)	(0.128)	(0.209)	(0.154)
Site 3	0.857	0.971	0.839	1.070
	(0.265)	(0.166)	(0.247)	(0.236)
IMD	1.019*	1.005	1.019*	1.004
	(0.008)	(0.006)	(0.008)	(0.006)
Some problems	1.672**	1.389**	1.616***	1.424***
	(0.266)	(0.140)	(0.217)	(0.152)
Unable wash/dress	3.794**	1.444	5.685***	1.667
	(1.823)	(0.367)	(1.718)	(0.572)
Level 1 constant	67.258***	361.030***	289.364***	657.763***
	(50.311)	(103.730)	(134.110)	(198.042)
σ	1.208*	0.879***	-	-
	(0.090)	(0.022)	-	-
$\sigma^2[u]$	3.080+	2.359***	-	-
	(1.989)	(0.165)	-	-
ρ_{12}	2.212+	1.365*	-	-
	(1.049)	(0.170)	-	-
N_i	335-336	1073	335-336	1073
Interaction effects	F(7.000,3.339e +08)=0.659 p=0.707	F(7.000,2.138e +12)=0.216 p=0.982	F(7.000,1.197e +09)=0.238 p=0.976	F(7.000,1.242e +12)=0.558 p=0.791

Note: the number of cases of social care use vary between multiply imputed datasets 1 to 10 (numbers per dataset: 335 in 8 and 336 in 2 complete datasets)
+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Marginal effects: As with hospital costs, there were no significant differences in the three-month costs of social care between intervention and control at baseline or 12-month follow-up in any condition subgroup (Table 6.16). The results of the population-averaged model show similarly that there were no differences between intervention and control between conditions. The estimate of total savings due to telehealth was not significantly different from zero.

Table 6.16 Two-level subject-specific model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by living arrangement

Intervention vs. control	*Social care*						*Hospital care*					
	COPD		HF		Diab		COPD		HF		Diab	
	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	-14	0.692	-6	0.871	62	0.208	-52	0.489	18	0.819	46	0.620
	(36)		(37)		(49)		(74)		(80)		(93)	
Follow-up	-102	0.159	-38	0.484	-66	0.497	-123	0.274	-158	0.237	-128	0.387
	(72)		(54)		(96)		(112)		(134)		(148)	
Follow-up-Baseline	-87.71 (67)	0.192	-32 (55)	0.562	-127 (101)	0.206	-71 (121)	0.555	-177 (148)	0.233	-174 (159)	0.272

Table 6.17 Two-level subject-specific model: difference-in-difference-in-difference (DDD) costs (£) estimates and total costs

	Social care		*Hospital care*	
	Exp (β) (95% CI)	*p*	Exp (β) (95% CI)	*p*
DDD HF - COPD	-32 (-140, 76)	0.562	-105 (-479, 268)	0.581
DDD DIAB - COPD	-40 (-273, 193)	0.739	-103 (-495, 290)	0.608
DDD DIAB - HF	-95 (-320, 129)	0.404	2 (-422, 427)	0.991
Total costs (sum of DD)	-247 (-511, 17)	0.067	-423 (-911, 66)	0.09
N_i	1930		1930	

Table 6.18 Population-averaged model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by living arrangement

	Social care						*Hospital care*					
	COPD		HF		Diab		COPD		HF		Diab	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	-7 (68)	0.920	-11 (39)	0.787	64 (68)	0.345	-47 (99)	0.635	56 (93)	0.545	207 (159)	0.194
Follow-up	-73 (77)	0.342	-27 (39)	0.476	-46 (83)	0.581	-93 (149)	0.534	-327 (187)	0.080	-44 (177)	0.804
Follow-up-Baseline	-67 (87)	0.446	-17 (46)	0.716	-110 (104)	0.289	-46 (161)	0.774	-384 (212)	0.07	-251 (230)	0.275

Table 6.19 Population-averaged model: difference-in-difference-in-difference (DDD) costs (£) estimates and total costs

	Social care		*Hospital care*	
	Exp (β) (95% CI)	*p*	Exp (β) (95% CI)	*p*
DDD HF - COPD	50 (-148, 248)	0.622	-337 (-854, 179)	0.200
DDD DIAB - COPD	-43 (-312, 225)	0.752	-205 (-747, 337)	0.459
DDD DIAB - HF	-93 (-320, 133)	0.421	133 (-458, 723)	0.660
Total costs (sum of DD)	-193 (-464, 77)	0.162	-680 (-1397, 36)	0.063
N_i	1930		1930	

6.11 Discussion of Telehealth Subgroup Analyses Results

It appears that, controlling for socio-demographic and needs-related variables, there is little clear evidence of a difference in the impact of telehealth on the total, NHS, hospital or social care costs of Telehealth questionnaire study participants on the basis of their index long-term condition. The evidence from the subject-specific and population-averaged models of total costs (not including intervention costs) and the marginal effects (of treatment allocation by time point and index long-term condition group) suggests that people with diabetes and with heart failure in the intervention group had lower total costs than in the control in the last three months of their participation in the trial, taking the difference in baseline costs into account. However the between-group differences were not significantly different between conditions. A similar conclusion could be reached about the between-group cost differences for participants with diabetes and heart failure based on the subject-specific model of NHS costs (but these differences were not found in the population-averaged estimates). Across the conditions, the total costs (including intervention costs) at the end of the study were greater in the intervention than in the control group. On the other hand, across conditions, those in the intervention group had on average lower 3-month NHS, and overall costs at 12-month follow-up, if excluding the costs of the intervention. This is clearly important. While the high costs of the intervention at the time meant that there could be no overall cost savings if these were taken into account, there is substantial evidence here that telehealth did decrease participants' other health costs across the board. Interestingly, social care receipt was higher at the follow-up, across the sample. Several covariates were consistently associated with higher costs across sectors: ADL need, older age, number of comorbid conditions. These are all characteristics related to chronic disability and ill-health. Being female increased total costs in the subject specific model of total costs, and increased social care costs in both modelling approaches. Higher level of education was associated with increased total costs and costs of NHS and hospital care and odds of receipt of hospital care. Owner-occupation was associated with lower total, NHS and social care costs.

6.12 Limitations

The observations available for each participant were limited to three-month cost snapshots at baseline and follow-up. This limited the amount of information available on within-participant variation in costs over time. Making inferences about change over time must be considered in this light. Differences in ADL need may not have been fully reflected by the

EQ-5D-3L self-care subscale that was used as a proxy for this characteristic. Details of comorbid conditions other than COPD, diabetes and heart failure were not available in the dataset.

6.13 Telecare Subgroup Analyses

In chapter 5, I set out an overview of the costs of people with social care needs who had been allocated to either telecare or usual care. This descriptive summary suggested that there were no statistically significant cost differences between experimental groups at the 12-month follow-up, whether or not the direct costs of the intervention were included. While the focus of the WSD Telecare study was to examine the costs (and outcomes) of all participants allocated to the telecare or usual care programmes, there may well be subgroups of participants that have quite distinct costs that vary in response to the introduction of telecare. Policymakers and commissioners may want to know where to target telecare resources to make savings, nationally or locally, by implementing this intervention. In this section, I examine whether telecare has an impact on expenditure depending on the living arrangements of the user.

As I have set out in Chapter 2, there are a few reasons given in the policy and industry literature for advocating telecare as a 'good thing'. There is a strong emphasis on the promotion of 'independence', suggesting that telecare is useful in cases where people are at risk of becoming dependent. Thus telecare could be particularly useful for those 'at risk' because of age-related disability, and for people living alone (Lloyd 2012a) who might otherwise need more hands-on care or a move to more supportive accommodation. Particularly in the case of people living alone, a motivation for providing this group with telecare might be to ensure their safety and protect them from serious injuries or even prevent falls (Department of Health 2005a). People living alone may be at risk of more serious injuries after a fall than those living with others (Elliott, Painter, and Hudson 2009). This could be because an individual who cannot summon help after a fall may have a 'long lie' on the floor, with adverse consequences; equipped with a remotely monitored falls-detector, that person could benefit from a rapid response to the activation of the detector. Indeed, in England the proportions using a pendant alarm is much higher in people who live alone than in people who live with others (12 per cent versus 2 per cent) (Nyman and Victor 2014). For such reasons, it is worth asking whether expenditure on people living alone differs from expenditure on those living with others, because of the presence of telecare. The question is

of considerable policy relevance, given the intersection of old age, living alone and morbidity, all trends that will accelerate in future years. The majority of people aged 85 years and over, or 'oldest old' in England and Wales live alone (59 per cent) (compared to 38 per cent of the 'old old'); their number has grown rapidly in recent years (Tomassini 2006). Finding ways to support and protect this group, a sizeable proportion of which may be expected to experience morbidity, will be of increasing importance as the population ages (Office for National Statistics 2011).

In the next sections, I explore relationships between individual characteristics and patterns of health and social care expenditure and ask whether living arrangement (living alone or with others) has a differential impact on these patterns.

6.14 Methods Used in the Telecare Cost Subgroup Analyses

I presented generic models of participants' costs in Chapter 4, Section 4.19.2. The health and social care costs of telecare participants can be described as a function including covariates:

$$COSTS = f[TC, Age, Education, Female, Comorb, Ethn, IMD, Site, Selfcare, Owns, Livewith, Followup\ ],$$

where *f*(.) is a function as described in the models laid out in Section 4.19.2 . *TC* is the treatment allocation, *Age* is a categorical variable[19], *Female* identifies women and men in the sample, *Comorb* is a count of chronic conditions sourced from acute hospital records (Steventon et al. 2012), *Ethn* is a binary indicator of white-British/non-white British ethnicity, *IMD* is a continuous measure of deprivation based on the Index of Multiple Deprivation 2007 (Noble et al. 2008), *Site* identifies the participating local authority, *Selfcare* is an indicator of ADL need based on the self-care domain of the EQ5-D, *Owns* is an indicator of owner-occupation vs. renting and other forms of tenure and *Livewith* is an indicator for multi-person households (living with others) and Followup is an indicator for the 12-month follow-up vs. the baseline time point..

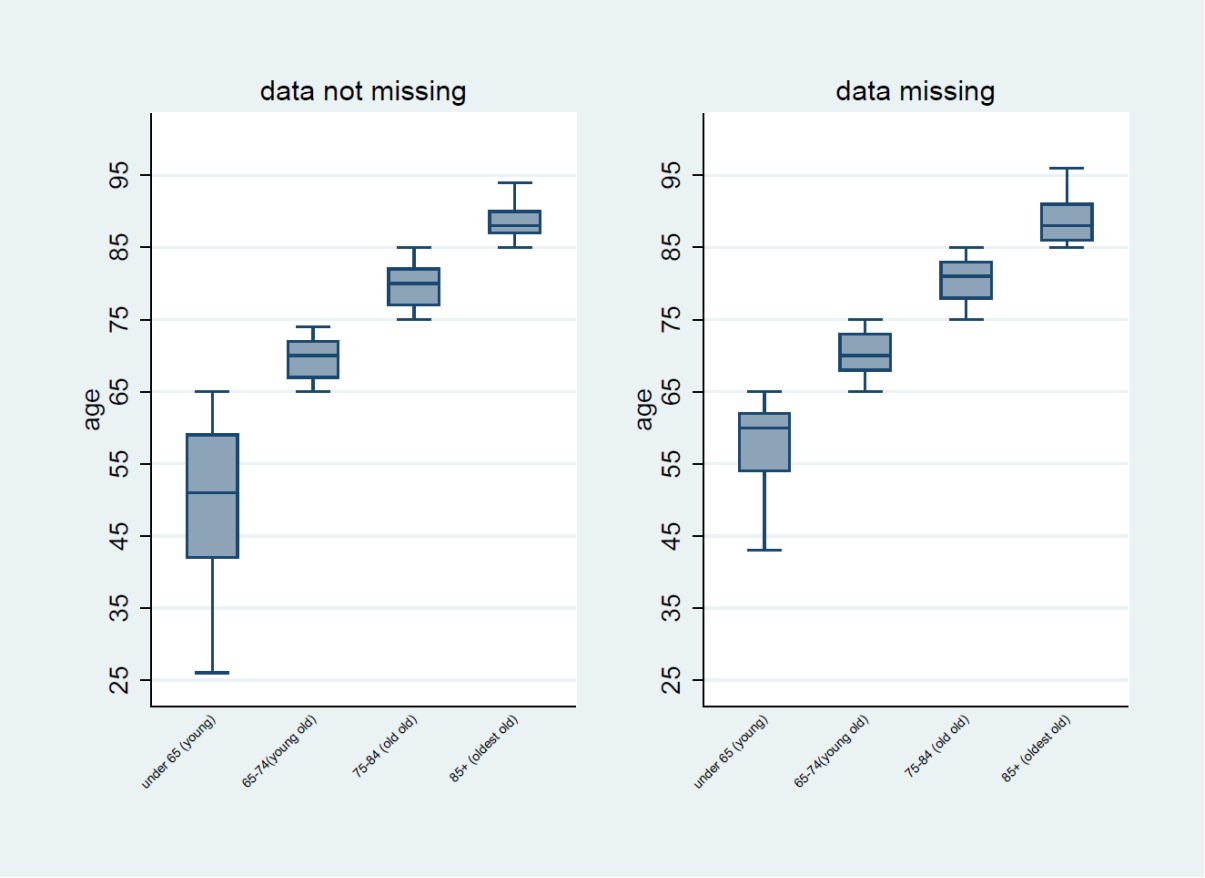
The discussion of population and subject-specific models in section 6.2 holds equally for the telecare costs analysis. The general approach to modelling the costs data was the same as taken in analysing the costs of telehealth participants. The investigation in this case centred on the impact of living arrangement on the costs of care for participants. I used a difference-in-difference-in-difference method to explore differences between experimental groups over

[19] Ages were categorised into 4 bands : 1 "under 65 (young)"; 2 "65-74 (young old)"; 3 "75-84 (old old)"; and 4 "85+ (oldest old)".

the two time points by living arrangement. I first considered total costs, with and without the direct costs of the intervention, then costs by sector (to the NHS, to secondary health/hospital services and to Social Services). Hospital and social care services were used by some but not all of the sample and therefore models were initially fitted only to the users of services. I examined the probability of use of hospital and of social care also through multilevel logistic models and applied these predicted probabilities to estimate the costs in the full sample. The average marginal effects of allocation, time and living arrangement were estimated in order to compare the differences between experimental groups over the two time points in those living with others and those living alone. All analyses reported in this chapter are based on imputed data (see Chapter 4, Section 4.22) and include all participants who had costs data available at both baseline and 12-month follow-up, 753 cases (375 intervention and 378 control). Participants' data were analysed by their ITT allocation.

As these analyses focus on the participants' living arrangements, the availability of data on living arrangements is important. The variable for living arrangement was derived from two variables: adults living in the household (including the respondent) and children under 16 years of age living permanently in the household. At baseline, there was one missing observation for numbers of adults in the household, but for children in the household, there were missing data in both groups (25 (7 per cent) controls; 33 (8 per cent) intervention). At follow-up, almost half of the responses to the number of children in the household were missing (212 (44 per cent) controls; 197 (47 per cent) intervention) (in contrast to 3 per cent and 4 per cent missing data for adults in the household). Perhaps this question seemed irrelevant to many respondents, which in combination with large proportion of the sample completing the questionnaire by post resulted in many missing observations (see for box-and-whisker plots displaying the missingness of this variable plotted against age, grouped by age band). These household composition variables were included in the multiple imputation model (see Chapter 4, Section 4.22); the variables for living alone and with others were then derived from the imputed data.

Figure 6.3 Box-and-whisker plots: missing data in the children 16 years of age in the household, by age of respondent grouped by age band



6.15 Characteristics of the Telecare Participants: Participants' Living Arrangements

As I have described in Chapter 5, the population completing the Telecare study was predominantly female. While approximately half of the participants were aged 75 years and over (the 'old-old' and 'oldest-old'), just under a quarter were under 65 years of age. Half of the sample lived alone.

The proportion of the sample living with others at baseline was roughly equal in both groups although slightly greater in the intervention group (47 per cent control vs. 50 per cent intervention); proportions at 12-month follow-up were similar (46 per cent control vs. 49 per cent intervention). Transitions from living with others at baseline to living alone at follow-up were rare in both experimental and control groups (3 per cent (n=12) in control vs. 4 per cent (n=14) in the intervention, respectively). Transitions from living alone at baseline to living with others at follow-up were equally rare (2 per cent (n=8) in control vs. 3 per cent (n=10) in the intervention, respectively).

6.16 Costs

The total costs of the telecare sample were calculated (see Chapter 4, section 4.17) including and excluding the cost of the intervention. In this chapter, costs have been further disaggregated by agency (NHS, local authority social care) and by health sector (secondary care; primary, community and mental health care).

Total unadjusted health and social care expenditure on people living alone and living with others in the three months prior to baseline and 12-month follow-up is summarised in Table 6.20 (by ITT allocation). At follow-up, total spend on those living alone who received the intervention was significantly greater than on those receiving usual care (£634, cluster-adjusted t=-2.757, p=0.006); corresponding spend on those living with others in receipt of telecare was somewhat (£111) higher than controls. The pre-baseline costs, in contrast, were somewhat greater in the intervention group regardless of living arrangement. People living with others and allocated to the intervention had much higher pre-baseline costs than controls, a difference rather more marked than in the corresponding groups who lived alone.

At follow-up, people living alone allocated to the intervention group had somewhat higher secondary care costs on average than those in the control group (difference of £237, t=-1.803, p=0.071), while those living with others in the intervention group had somewhat lower costs than those in the control group (£170, t=0.903, p=0.367). At follow-up and looking at health and social care expenditure separately, the difference between experimental groups was somewhat greater in those living alone than in those living with others; however only NHS expenditure on those living alone differed significantly at the 5 per cent level between experimental groups.

Looking at the sample as a whole (not shown in the table), people living alone at 12-month follow-up had lower unadjusted total costs than those living with others (£1702 vs. £2154), a difference of £452 (t=-1.9, p=0.057) and lower unadjusted social care costs than those living with others (£799 vs. £1,198), a difference of £399 (t=-2.244, p=0.026). In contrast, there was little difference in hospital costs between those living alone (£476) and with others (£500) across the sample at the 12-month follow-up.

Table 6.20 Costs at baseline and 12 months, by ITT allocation and living arrangement, participants with complete data

	Living alone[a]			**Living with others[b]**		
	Usual care (SE)	**Telecare (SE)**	**Difference**	**Usual care (SE)**	**Telecare (SE)**	**Difference**
Total costs exc. delivery and equipment						
Baseline	1892 (220)	2135 (223)	243	2206 (247)	2693 (245)	487
Follow-up	1480 (159)	1939 (163)	460*	2201 (314)	2109 (316)	-92
Total costs inc. delivery and equipment						
Baseline	-	-	-	-	-	-
Follow-up	1492 (161)	2126 (165)	634**	2207 (314)	2319 (316)	111
Hospital, primary and community mental health (NHS)						
Baseline	1191 (167)	1234 (170)	43	1333 (148)	1640 (144)	307
Follow-up	764 (101)	1041 (104)	276*	1036 (186)	875 (189)	-161
Hospital only						
Baseline	541 (156)	611 (158)	70	541 (118)	689 (115)	147
Follow-up	362 (91)	599 (94)	237	588 (132)	417 (134)	-170
Community care (LA)						
Baseline	692 (112)	892 (114)	199	868 (162)	1047 (161)	179
Follow-up	706 (115)	899 (118)	192	1160 (207)	1234 (205)	74

a. Numbers living alone at baseline: UC N=202, TC N=188. Numbers living alone at follow-up: UC N=206, TC N=192.

b. Numbers living with others at baseline: UC N=176,TC=187. Numbers living with others at follow-up: UC N=172, TC N=183.

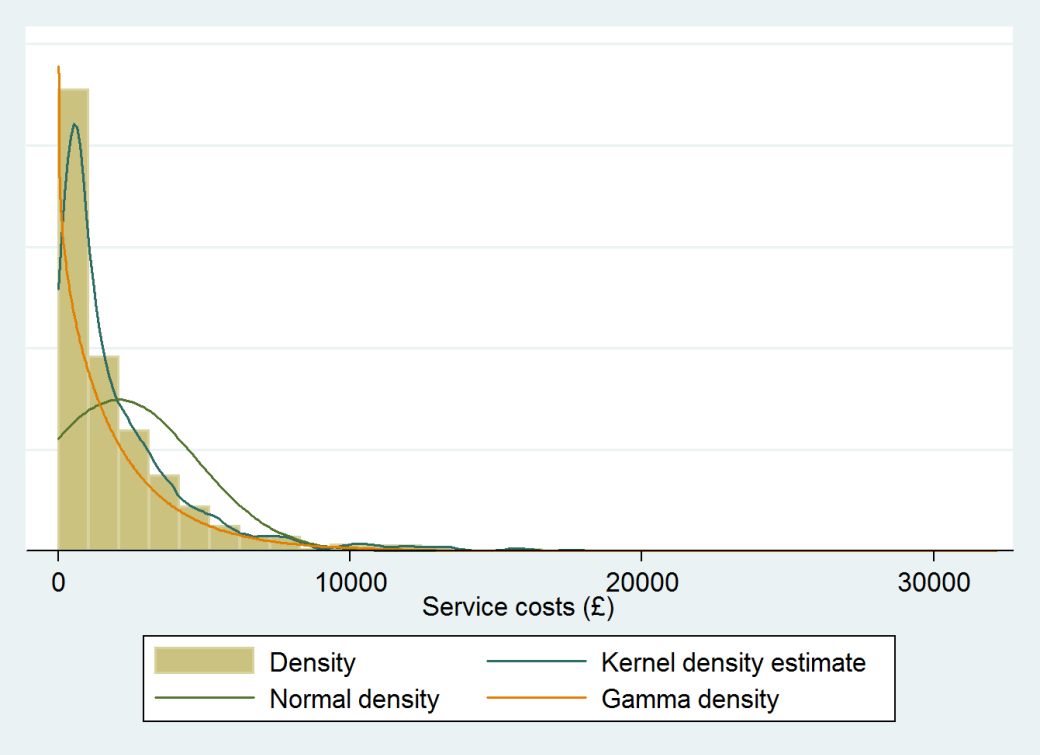
* $p<0.05$ on clustered t-test

** $p<0.01$ on clustered t-test

6.16.1 Distribution of Costs

The costs of health and social care (excluding the intervention) for study participants were right-skewed, as can be seen in Figure 6.4. The figure depicts their frequency distribution and also the corresponding density functions of the gamma and normal distributions over both time points. Examining the density probability plots for these distributions, it is evident that these data fit the gamma better than the normal distribution.

Figure 6.4 Frequency and theoretical distributions of service costs



6.16.2 Clustering Effects

Participants' mean health and social care costs (excluding the intervention) over the baseline and follow-up points, aggregated at participant and at general practice level, are presented in Table 6.21. The table also shows the overall, between- and within-cluster standard deviations of the means. Costs data are clustered within 191 general practices, almost equally divided between allocation groups (95 control and 96 intervention participants, with similar sizes of clusters (7.96 control vs. 7.81 intervention). Participants' costs (pooled across baseline and follow-up points) vary somewhat more within- than between-GP clusters, particularly in the control group. At participant level, there is more variation between- than within-participants in both groups (not surprising given that there are only two time points under observation).

The degree of between-cluster variation can be summarised by the intra-cluster correlation, or ICC (calculated here by one-way analysis of variance (Ukoumunne 2002)), also given in Table 6.21. Pooling observations over time, the ICC for general practice is slightly lower in the control than in the intervention group. The ICC at the participant level (with observations for two time points per cluster) is slightly higher in the control than in the intervention group, but the ICCs are broadly similar between groups. These statistics suggest that the extent of between-general practices variation in costs (excluding the intervention) and therefore homogeneity within practices is quite high (between 0.108 and 0.11) (see also Section 5.16.6). They do not, on the other hand, take into account the influence of participants' socio-demographic characteristics.

Table 6.21 Health and social care service costs: cluster means, counts and intra-cluster correlation coefficients (ICC)

Over time points				
Control	Mean	SD	Count	ICC[a]
	1,923	2,485	756	0.274 (0.181,0.368)
		between subject	n	
		1984	378	
		within subject	Mean n[b]	
		1498	2	
TC	Mean	SD	Count	ICC[a]
	2217	2854	750	0. 229 (0.133, 0.325)
		between subject	n	
		2237	375	
		within subject	Mean n[b]	
		1774	2	
Over participants				
	Mean	SD	Count	ICC[a]
Control	1,923[c]	1982	756	0.108 (0.044,0.172)
		between practice	n	
		1344	95	
		within practice	Mean n[b]	
		1608	7.96	
TC	Mean	SD	Count	ICC[a]
	2,217[c]	2236	750	0.110 (0.045, 0.174)
		between practice	n	
		1697	96	
		within practice	Mean n[b]	
		1788	7.81	

Note: imputed data; costs at baseline and follow-up, excludes costs of intervention
a Intra-cluster correlation, calculated by one-way analysis of variance; Searle's Confidence intervals report arithmetic mean cluster size for unbalanced data (Ukoumunne 2002)
b average number of units under observation
c participant costs pooled across time points

6.17 Model Results of Telecare Subgroup Analyses: Total Costs

I followed similar methods to those employed in examining the costs of telehealth participants (section 6.2.1). Results of models run can be found in Appendix 3, A3.2, Table 3.2). A three-level constant-only model (model (1)) of costs data demonstrates the proportion of variability in costs (including intervention costs) due to participant- and general practice-level factors. Between-participant variation ($\sigma_{\mu}^{2(2)}$ = 1.429; 95 per cent CI 1.294, 1.578) was greater than the variation between general practice clusters ($\sigma_{\mu}^{2(3)}$ = 1.222; 95 per cent CI 1.136, 1.315) (or 41 per cent vs. 35 per cent of the total variance). The within-participant standard deviation was 0.797 (95 per cent CI 0.734, 0.866) times the average cost. However on adding the DDD terms and other socio-demographic covariates into the random intercept model (Table 3.2, model (3)), the amount of unexplained third-level variation in costs decreased substantially so that the variance was no longer significantly different from zero ($\sigma_{\mu}^{2(3)}$ =1.017; 95 per cent CI 0.980, 1.056; p=0.371). The between-participant variance was higher and its standard error lower in the two- than the three-level model ($\sigma_{\mu}^{2(2)}$ = 1.256 (SE 0.047) vs. $\sigma_{\mu}^{2(3)}$ = 1.237 (SE 0.050)), while there was no difference in the within-participant variances estimated, as could be expected (Van den Noortgate, Opdenakker, and Onghena 2005). Most (not all) fixed effect coefficients became slightly larger. Nonetheless the impact of ignoring the third level appears to be small. Appendix 3 presents the results of the two-level model (with DDD terms and other socio-demographic covariates) of costs, both including and excluding intervention costs. Including direct intervention costs in the total costs, the coefficient for allocation to telecare (under ITT) was not significantly associated with increased spend. The coefficients on time, age bands, comorbidities, having ADL needs and site 2 dummy variable were significantly different from zero at either the 5 per cent or 1 per cent level. Age, and the passage of 9 months' time (follow-up) are associated with decreased costs; number of comorbidities, living in site 2 and having ADL needs are associated with increased costs. For instance, being in the "young old" category was associated with a 24 per cent decrease in costs relative to being under 65 years. Being unable to complete basic activities of daily living was associated with costs 170 per cent higher than those of being independent in such activities. As might be expected, given the substantial additional cost associated with the intervention, the interaction of time point and allocation was significantly different from zero at the 10 per cent level. The size and significance of the triple interaction term (DDD) was not much affected by the inclusion or exclusion of

intervention costs. The main effects of allocation and living arrangements and their interaction term were significantly greater than zero on joint conditional tests (p=0.001).

Excluding direct intervention costs produced similar results. The coefficient on the triple interaction term (DDD) was not significantly different from zero at the 5 per cent level (p=0.187). Joint conditional tests indicated that the coefficients on variables for allocation, time and living with others (the main effects) were significantly different from zero (p=0.01) and that coefficients on these variables and on the interaction terms (the main and interaction effects) were significantly different from zero (p=0.000), but the interaction effects were not (p=0.375).

The results of the random intercept model do not provide strong evidence of a telecare-related decrease in total health and social care expenditure at follow-up on those living with others. Results obtained from the population-averged models showed a pattern of similar results for total costs, including or excluding the intervention cost. As with the subject specific models, the main effects of allocation and living arrangements and their interaction term were significantly greater than zero at the 5 per cent level on joint conditional tests.

Table 6.22 Parameter estimates, subject specific (random intercept) and population-averaged (GEE) models of total and NHS costs in 3 months prior to baseline and 12-month follow-up

	Subject-specific			*Population average*		
	Excluding intervention costs	*Including intervention costs*	*NHS*	*Excluding intervention costs*	*Including intervention costs*	*NHS*
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
TC	1.106 (0.122)	1.111 (0.123)	0.956 (0.121)	1.155 (0.152)	1.157 (0.152)	0.991 (0.163)
Followup	0.804* (0.071)	0.813* (0.071)	0.711*** (0.073)	0.774** (0.076)	0.780* (0.077)	0.655*** (0.084)
Followup*TC	1.106 (0.149)	1.274^{+} (0.165)	1.201 (0.190)	1.126 (0.179)	1.244 (0.193)	1.301 (0.278)
Lives w/others	0.976 (0.102)	0.984 (0.103)	1.032 (0.123)	0.916 (0.098)	0.922 (0.098)	0.949 (0.128)
TC*Lives w/	0.985 (0.148)	0.986 (0.148)	1.188 (0.203)	1.032 (0.175)	1.031 (0.175)	1.241 (0.260)
Followup*Lives w/	1.033 (0.142)	1.030 (0.141)	1.002 (0.154)	1.190 (0.178)	1.186 (0.177)	1.161 (0.205)
TC*Followup*Lives w/	0.773 (0.151)	0.790 (0.149)	0.600* (0.133)	0.690^{+} (0.154)	0.702 (0.153)	0.519* (0.147)
Young old	0.760** (0.065)	0.763*** (0.063)	0.900 (0.086)	0.731*** (0.062)	0.734*** (0.061)	0.827^{+} (0.085)
Old old	0.799** (0.069)	0.791** (0.065)	0.895 (0.088)	0.781** (0.065)	0.775** (0.063)	0.846 (0.092)
Oldest old	0.784* (0.076)	0.780** (0.072)	0.737** (0.082)	0.789* (0.076)	0.788* (0.074)	0.724** (0.090)
Below-degree	0.994 (0.070)	0.996 (0.066)	0.942 (0.073)	1.004 (0.070)	1.002 (0.068)	0.928 (0.079)

	Subject-specific			Population average		
	Excluding intervention costs	*Including intervention costs*	*NHS*	*Excluding intervention costs*	*Including intervention costs*	*NHS*
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
Degree	1.070 (0.111)	1.062 (0.104)	1.134 (0.142)	1.037 (0.101)	1.030 (0.096)	1.136 (0.154)
Female	1.026 (0.066)	1.027 (0.063)	1.015 (0.072)	0.996 (0.064)	0.998 (0.063)	0.961 (0.077)
White-British	1.073 (0.125)	1.077 (0.119)	1.080 (0.132)	1.183 (0.141)	1.180 (0.136)	1.163 (0.154)
Comorb	1.146*** (0.021)	1.142*** (0.020)	1.197*** (0.026)	1.129*** (0.020)	1.128*** (0.020)	1.184*** (0.028)
Owns	0.958 (0.066)	0.964 (0.063)	1.010 (0.078)	0.996 (0.072)	0.996 (0.070)	1.019 (0.087)
IMD	0.997 (0.003)	0.997 (0.003)	0.994 (0.004)	0.998 (0.003)	0.998 (0.003)	0.994 (0.004)
Some ADL problems	1.478*** (0.099)	1.456*** (0.093)	1.406*** (0.100)	1.466*** (0.111)	1.453*** (0.106)	1.384*** (0.123)
Unable to wash/dress	2.829*** (0.258)	2.695*** (0.237)	1.761*** (0.184)	2.847*** (0.271)	2.755*** (0.255)	1.840*** (0.232)
Constant	1148.150*** (219.421)	1130.463*** (206.922)	798.314*** (173.842)	1158.410*** (228.635)	1158.095*** (222.116)	984.800*** (239.812)
σ	0.822*** (0.031)	0.791*** (0.031)	0.861*** (0.034)			
$\sigma[u]$	1.284*** (0.055)	1.256*** (0.047)	1.433*** (0.063)			
N_i	1506	1506	1506	1506	1506	1506
Interaction effects	F(4.000,3752233) =1.058 p=0.375	F(4.000,4230982) =1.584 p=0.175	F(4.000,3.087e+08)=2.976 p=0.018	F(4.000,342929)=1.191 p=0.313	F(4.000,377335)=1.268 p=0.280	F(4.000,2056192)=1.864 p=0.114

Note: lives w/= lives with others; comorb=number of comorbidities
+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Marginal Effects: Marginal effects of the interaction predicted on the basis of the subject-specific and population-averaged model were investigated to explore the average cost implications. I examined the partial effect of treatment allocation (Table 6.23) on total costs, decomposed by time and living arrangements, from the subject-specific model. At follow-up, expenditure on people in the intervention group who live alone was significantly higher than in controls; the expenditure on people in this group who live with others was also somewhat but not significantly higher than in controls at the 5 per cent level. The marginal effect of interest (the DDD) - the effect of telecare on three month-costs for participants living with others - was substantial; however the confidence intervals of the estimate were very wide and crossed zero (-£465; 95 per cent CI -£1275, £346). We also would need to consider the implications for expenditure on all participants allocated to the intervention. Adding together the partial effects of the intervention on both subgroups during the intervention phase, expenditure in the intervention period is non-significantly increased from baseline by £437 (95 per cent CI -£358, £1232).

I also examined total health and social care costs, after excluding direct costs of the intervention. The difference between baseline and follow-up costs for intervention participants living alone were somewhat higher than controls (£156; 95 per cent CI -£359, £671), while the (non-significant) difference between time points between intervention and control participants living with others is in the opposite direction (-£367; 95 per cent CI -£1002, £267). The total estimated expenditure on all intervention participants, across living arrangements, is a modest and non-significant decrease from baseline of £211 (95 per cent CI -£1022, £600).

The population average derived marginal effects results (Table 6.24) diverge somewhat from those derived from the subject-specific model. Differences in total costs between allocation groups at baseline for participants living with others are double that of the estimates from the SS model (and at follow-up are half that of the SS model), but are not significantly different from zero. The difference between allocation groups in costs, excluding intervention costs, of those living alone at follow-up is larger than in the SS model and significant (£362 SS p=0.057 vs. £457 PA p=0.015). However in other respects while the PA model estimates are larger and standard errors smaller, the conclusions remain the same as would be drawn from the SS model.

Table 6.23 Two-level subject-specific model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by living arrangement

	Excluding intervention costs				*Including intervention costs*			
	Live alone		Live with others		Live alone		Live with others	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	206	0.365	223	0.399	209	0.345	230	0.369
	(227)		(263)		(222)		(256)	
Follow-up	362	0.057	-146	0.547	660***	0.000	217	0.361
	(190)		(242)		(188)		(237)	
Follow-up-Baseline	156 (-359, 671)	0.553	-367 (-1002, 267)	0.257	450.852 (-55, 957)	0.081	-14 (-637, 609)	0.966
	Excluding intervention costs				*Including intervention costs*			
	Exp (β) (95% CI)		*p*		Exp (β) (95% CI)		*p*	
DDD	-523 (-1348, 301)		0.213		-465 (-1275,346)		0.261	
Total costs	-211 (-1022, 600)		0.609		437 (-358, 1232)		0.281	
N_i	1506				1506			

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Table 6.24 Population-averaged model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by living arrangement

	Excluding intervention costs				*Including intervention costs*			
	Live alone		Live with others		Live alone		Live with others	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	294.002 (273)	0.281	435.286 (267.728)	0.104	295.697 (271.029)	0.275	432.921 (266.046)	0.104
Follow-up	457.015* (187.099)	0.015	-158.416 (261.171)	0.544	667.164*** (187.142)	0.000	89.871 (260.203)	0.730
Follow-up-Baseline	163.013 (314.862)	0.605	-593.702 (356.335)	0.096	371.467 (313.58)	0.236	-343.05 (354.495)	0.333
	Excluding intervention costs				*Including intervention costs*			
	Exp (β) (95% CI)		*p*		Exp (β) (95% CI)		*p*	
DDD	-756.715 (-1695.63,182.2)		0.114		-714.517 (-1648.7, 219.666)		0.134	
Total costs	-430.689 (-1355.71, 494.333)		0.361		28.418 (-892.61,949.445)		0.952	
N_i	1506				1506			

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

These results suggest that while telecare increased total spending (including intervention costs) on all intervention participants, the magnitude of expenditure at follow-up was much greater in the case of those living alone than in those living with others, controlling for characteristics such as social care need. Also, setting aside the direct cost of the intervention, the results suggest that telecare substantially increased health and social care spending on those living alone, while the intervention had a weaker and opposite effect on the spend on those living with others. However, taking into account the differences between groups, and subgroups, over time, from model estimates, it would be difficult to conclude from the results that the intervention had any impact on total health and social care costs.

6.17.1 Costs by Sector

Health and social care costs were investigated separately to understand better the source of the variation in the costs of trial participants. At both time points, virtually all participants had received some form of NHS service in the prior three months; slightly over half (55 per cent at baseline, 54 per cent at follow-up) used some form of hospital service. Despite the trial taking place in a population with social care need, not all participants reported receipt of social care services (as some trial eligibility criteria were not related to existing receipt of care). At follow-up (baseline), 66 per cent (61 per cent) received some form of social care (as previously defined, see Table 4.2.

6.17.2 NHS Costs

In the random-intercept model, the interaction effects were significant at the 1 per cent level. Oldest-age and needs-related characteristics were significantly associated with decreased and increased costs respectively, as with the total costs. Age categories and ADL needs categories were also jointly significant on F-tests (age: F(3.000,25598553.804)=2.670, p=0. 046; ADLs: F(2.000,69366.118)=17.911, p=0.000). The coefficient on the DDD interaction term for living with others and allocation was significantly different from zero at the 5 per cent level. NHS costs at follow-up across the sample also decreased by 29 per cent. The results of the population-averaged (PA) model were broadly similar to those of the subject-specific (SS) model.

Marginal effects: While control participants living alone had higher costs than intervention participants in the pre-baseline period, the reverse occurred in those living with others (although in neither case were differences significant at the 5 per cent level) (see Table

6.25). Conversely, at follow-up, intervention group participants living alone had higher NHS costs, while the reverse occurred in those living with others (again in neither case were differences significant at the 5 per cent level). As a result, the change in the cost difference between experimental groups from the pre-baseline to the follow-up period was significantly lower in the living with others subgroup (at the 5 per cent level). In other words, for those living with others, the difference in NHS costs did vary depending on the allocation. This pattern was reversed in the living alone subgroup. The effect of telecare in the intervention period for participants living with others (the DDD), a savings of £560 (95 per cent CI -£1068, -£52), was significant at the 5 per cent level (Table 6.26). However, adding the partial effects of the intervention on both subgroups, while NHS spending in the last three months of the intervention period appears to be decreased compared to baseline by £201 (95 per cent CI -£696, £293), the total saving was not significantly different from zero at the 5 per cent level. Results of the population-averaged model (Table 6.27 and Table 6.28) were in line with those of the SS model, although the estimated DD in the group living with others was larger (-£481 PA vs. -£381); the DDD was also larger (-£725 PA vs. -£560).

Table 6.25 Two-level subject-specific model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by living arrangement

	NHS costs			
	Live alone		Live with others	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	-51.297	0.722	193.155	0.272
	(144.442)		(175.794)	
Follow-up	128.145	0.248	-187.437	0.108
	(111.029)		(116.619)	
Follow-up-Baseline	179.441 (-150.321,509.204)	0.286	-380.592 (-758.071,-3.113)	0.048
N_i	1506			
	NHS costs			
	Live alone		Live with others	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	-51.297 (144.442)	0.722	193.155 (175.794)	0.272
Follow-up	128 (111.029)	0.248	-187.437 (116.619)	0.108
Follow-up-Baseline	179.441 (-150.321,509.204)	0.286	-380.592 (-758.071,-3.113)	0.048
N_i	1506			

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Table 6.26 Two-level subject-specific model: difference-in-difference-in-difference (DDD) costs (£) estimates and total costs

	NHS costs	
	Exp (β) (95% CI)	*p*
DDD	-560 (-1068,-52)	0.031
Total costs	-201 (-696,293)	0.425
N_i	1506	

Table 6.27 Population-averaged model: Partial effect/discrete changes in costs (£), baseline to follow-up of ITT allocation, by living arrangement

	NHS costs			
	Live alone		Live with others	
Intervention vs. control	Exp (β) (SE)	*p*	Exp (β) (SE)	*p*
Baseline	-11 (195)	0.955	305 (190)	0.108
Follow-up	233 (132)	0.078	-176.071 (138.247)	0.203
Follow-up-Baseline	244 (235)	0.299	-481.238 (220.231)	0.029
N_i	1506		1506	

Table 6.28 Population-averaged model: difference-in-difference-in-difference (DDD) costs (£) estimates and total costs

	NHS costs	
	Exp (β) (95% CI)	*p*
DDD	-725.3 (-1366,-85)	0.026
Total costs	-237.176 (-860,385)	0.455
N_i	1506	

6.17.3 Hospital Costs: Two-part Models

A two-level model of participants' observations at baseline and follow-up was fitted. In the 'first' part of the model (a logistic model) (Table 6.29) the conditional odds of receipt of any kind of hospital care were nearly halved for the oldest-old compared to those under 65 years of age; the overall effect of age was significant at the 10 per cent level (F(3, 39706591)=2.291, p=0.076). Each additional chronic condition increased the odds of receipt of hospital care (by 29 per cent); having high ADL needs increased the odds of using hospital services by 65 per cent. Having a degree-level qualification increased the odds of receipt by

65 per cent (significant at the 10 per cent level). The estimate of the effect of the DDD interaction term was not significantly different from zero (p=0.168). Joint tests of interaction effects were not significantly different from zero.

Hospital care costs from the 'second' part of the model (those in receipt of any hospital care) are displayed in Table 6.30. The young-old had lower hospital costs compared to the under-65 group (p=0.014); the age band variables were jointly significantly different from zero at the 1 per cent level (F(3,6.577e+08)=4.586, p=0.003). High (but not moderate) ADL need was significantly associated with increased cost at the 10 per cent level. Each additional comorbid condition was associated with a 17 per cent increase in costs. The DDD interaction term was significantly different from zero at the 10 per cent level (p=0.085). However, as with the first part of the model, F-tests of the interaction effects indicated that these variables were not jointly significantly different from zero.

The exponentiated estimate of the covariance between random effects for subject in the two parts of the model was 0.933 (95 per cent CI 0.663,1.312); as the 95 per cent confidence limits cross one, it appears that the odds of receipt do not co-vary with the size of hospital costs.

Marginal effects: The marginal probability of receipt of hospital services was applied to the marginal mean hospital costs predicted by the subject-specific two-part model. For people living with others, the difference between time points in the difference between experimental groups' hospital costs was £389 (p=0.023). The difference in between-group differences (the DDD) was substantial (£494; 95 per cent CI -£953, -£36, p=0.035). The total saving made across those living alone and those living with others (£283; 95 per cent CI -£731, £165) was not significant at the 5 per cent level (p=0.215). The marginal effects derived from the two-parts of the population-averaged model were broadly similar but while the estimate of the DDD was larger, the 95 per cent confidence intervals of this estimate crossed zero.

6.17.4 Social Service Costs: Two-part Models

In the random-intercept two-part model (Table 6.29), the conditional odds of receipt of social care in site 2 were 4.7 times higher than those in site 1, while the odds of receipt in site 3 were about a third of those in site 1. Those with high ADL needs were 6.4 times more likely to receive care. The conditional odds of receipt increased by 79 per cent between baseline and follow-up periods (p=0.033), controlling for other factors. The main effects of time point,

allocation and living arrangement were not significantly greater than zero at the 5 per cent level, nor were their interaction effects. The DDD interaction effects were not jointly significantly greater than zero at the 5 per cent level.

Social care costs of those in receipt of social care were examined (Table 6.30). Costs were lower at the 5 per cent level in the young-old and old-old age-bands compared to the under-65s. The age band variables were jointly significantly different from zero ($F(3,858852)=2.619$, $p=0.049$). The effect of moderate ADL-need was to increase costs by a factor of 1.4 and the effect of severe need was to more than triple costs. Costs in site 2 were also significantly higher at the 5 per cent level than in site 1. The main effects, the DDD interaction terms and the interaction and main effects were not significant at the 5 per cent level.

The exponentiated covariance of the random effects of the first and second parts of the model was 1.966 (95 per cent CI 1.895, 2.040); the results suggest that receipt of social care and the costliness of social care are highly and positively related.

The population-averaged logistic and gamma regression models yielded comparable results to those of the subject-specific model.

Marginal effects: The results of the subject-specific and population-averaged models were similar (Table 6.31 and Table 6.32). In the case of people living alone, the intervention participants' marginal mean costs were somewhat higher than those of controls at both time points. There was little difference in costs between experimental groups in the case of people living with others. The estimated difference between the baseline and follow-up differences in the differences between intervention and control costs (the DDD) had wide confidence intervals that crossed zero. There was little evidence that the intervention had an impact on social care costs at follow-up either within or across living arrangements.

Table 6.29 Two-level model estimates of receipt of social care and hospital care from two-part subject-specific and population averaged models in prior 3 months, 12-month follow-up

	Subject-specific				*Population average*			
Parameter	*Social care* β (SE)	*Social care* Exp (β) (SE)	*Hospital care* β (SE)	*Hospital care* Exp (β) (SE)	*Social care* β (SE)	*Social care* Exp (β) (SE)	*Hospital care* β (SE)	*Hospital care* Exp (β) (SE)
TC	0.545^{+}	1.724^{+}	-0.352	0.703	0.432^{+}	1.540^{+}	-0.267	0.765
	(0.314)	(0.542)	(0.272)	(0.191)	(0.230)	(0.354)	(0.210)	(0.161)
Follow-up	0.584*	1.793*	-0.291	0.747	0.426*	1.531*	-0.225	0.799
	(0.274)	(0.491)	(0.239)	(0.178)	(0.200)	(0.306)	(0.184)	(0.147)
Follow-up*TC	-0.084	0.919	0.487	1.627	-0.058	0.944	0.378	1.459
	(0.405)	(0.373)	(0.341)	(0.555)	(0.296)	(0.279)	(0.264)	(0.386)
Lives w/ others	-0.016	0.984	-0.071	0.931	0.004	1.004	-0.052	0.950
	(0.331)	(0.326)	(0.278)	(0.259)	(0.242)	(0.243)	(0.215)	(0.205)
TC*Lives w/ others	-0.291	0.748	0.214	1.239	-0.242	0.785	0.164	1.178
	(0.467)	(0.350)	(0.389)	(0.481)	(0.345)	(0.271)	(0.301)	(0.355)
Follow-up*Lives w/ others	-0.398	0.672	0.294	1.342	-0.307	0.736	0.227	1.255
	(0.403)	(0.271)	(0.350)	(0.470)	(0.296)	(0.218)	(0.272)	(0.341)
TC* Follow-up*Lives w/ others	-0.473	0.623	-0.687	0.503	-0.349	0.705	-0.534	0.586
	(0.583)	(0.363)	(0.492)	(0.248)	(0.430)	(0.303)	(0.381)	(0.223)
Young old	-0.467^{+}	0.627^{+}	0.187	1.205	-0.342^{+}	0.710^{+}	0.143	1.154
	(0.269)	(0.169)	(0.228)	(0.275)	(0.200)	(0.142)	(0.175)	(0.202)
Old old	0.203	1.225	-0.058	0.944	0.133	1.142	-0.037	0.963
	(0.269)	(0.330)	(0.222)	(0.209)	(0.201)	(0.230)	(0.171)	(0.165)
Oldest old	0.412	1.510	-0.667**	0.513**	0.313	1.367	-0.514**	0.598**
	(0.310)	(0.468)	(0.240)	(0.123)	(0.230)	(0.315)	(0.183)	(0.110)
Below-degree	-0.021	0.980	-0.187	0.829	-0.016	0.984	-0.143	0.867
	(0.215)	(0.210)	(0.182)	(0.151)	(0.161)	(0.159)	(0.140)	(0.121)
Degree	0.430	1.538	0.500+	1.648+	0.272	1.313	0.382	1.465
	(0.354)	(0.544)	(0.304)	(0.501)	(0.257)	(0.338)	(0.235)	(0.344)

	Subject-specific				*Population average*			
Parameter	*Social care* β (SE)	*Social care* Exp (β) (SE)	*Hospital care* β (SE)	*Hospital care* Exp (β) (SE)	*Social care* β (SE)	*Social care* Exp (β) (SE)	*Hospital care* β (SE)	*Hospital care* Exp (β) (SE)
Female	0.305	1.357	0.235	1.265	0.225	1.253	0.182	1.200
	(0.194)	(0.264)	(0.162)	(0.205)	(0.144)	(0.181)	(0.125)	(0.150)
White-British	0.050	1.051	-0.094	0.910	0.019	1.019	-0.069	0.933
	(0.313)	(0.330)	(0.288)	(0.262)	(0.232)	(0.236)	(0.222)	(0.208)
Number of comorbidities	0.096	1.101	0.253***	1.288***	0.070	1.073	0.198***	1.219***
	(0.060)	(0.066)	(0.058)	(0.074)	(0.045)	(0.048)	(0.045)	(0.054)
Owner	-0.082	0.921	0.160	1.174	-0.081	0.923	0.123	1.131
	(0.221)	(0.203)	(0.175)	(0.205)	(0.166)	(0.153)	(0.135)	(0.153)
Site 2	1.546***	4.693***	-0.174	0.840	1.149***	3.155***	-0.135	0.873
	(0.244)	(1.144)	(0.208)	(0.175)	(0.179)	(0.564)	(0.161)	(0.141)
Site 3	-1.132***	0.322***	-0.005	0.995	-0.827***	0.437***	-0.001	0.999
	(0.326)	(0.105)	(0.286)	(0.285)	(0.238)	(0.104)	(0.221)	(0.220)
Mean IMD score	-0.003	0.997	-0.012	0.988	-0.002	0.998	-0.009	0.991
	(0.010)	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
Some ADL problems	0.673***	1.961***	0.280+	1.323+	0.502***	1.652***	0.215+	1.240+
	(0.186)	(0.365)	(0.167)	(0.221)	(0.137)	(0.227)	(0.129)	(0.160)
Unable to wash/dress	1.856***	6.401***	0.501*	1.650*	1.475***	4.371***	0.383*	1.466*
	(0.309)	(1.981)	(0.226)	(0.374)	(0.230)	(1.007)	(0.173)	(0.254)
Level 1 constant	-0.595	0.552	0.261	1.298	-0.442	0.643	0.194	1.214
	(0.520)	(0.287)	(0.470)	(0.610)	(0.383)	(0.246)	(0.363)	(0.440)
$\sigma^2[u]$	2.180***	8.850***	1.523***	4.587***				
	(0.059)	(0.520)	(0.375)	(1.721)				
N_i	1506	1506	1506	1506	1506	1506	1506	1506
Interaction effects	F(4.000,79039527)=2.259 p=0.060		F(4.000,3.729e+09)=0.643 p=0.632		F(4.000,516419)=2.341 p=0.053		F(4.000,3.984e+09)=0.652 p=0.626	

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

Table 6.30 Two-level model estimates of costs from two-part subject-specific and population averaged models in prior 3 months, 12-month follow-up

	Subject-specific		*Population average*	
Parameter	*Social Care* Exp (β) (SE)	*Hospital* Exp (β) (SE)	*Social Care* Exp (β) (SE)	*Hospital* Exp (β) (SE)
TC	1.271	1.121	1.273	1.197
	(0.193)	(0.249)	(0.202)	(0.362)
Follow-up	0.916	1.032	0.861	0.826
	(0.108)	(0.194)	(0.091)	(0.180)
Follow-up*TC	0.957	1.039	0.959	1.133
	(0.157)	(0.286)	(0.168)	(0.408)
Lives w/ others	0.904	0.819	0.927	0.848
	(0.152)	(0.185)	(0.144)	(0.218)
TC*Lives w/ others	0.741	1.315	0.757	1.227
	(0.166)	(0.413)	(0.170)	(0.489)
Follow-up*Lives w/ others	1.231	1.254	1.321	1.403
	(0.255)	(0.341)	(0.281)	(0.433)
TC* Follow-up*Lives w/ others	1.222	0.511^{+}	1.192	0.446^{+}
	(0.336)	(0.200)	(0.349)	(0.219)
Young old	0.725*	0.703*	0.730**	0.614**
	(0.098)	(0.101)	(0.089)	(0.098)
Old old	0.719*	0.824	0.702**	0.826
	(0.095)	(0.126)	(0.086)	(0.133)
Oldest old	0.817	0.707^{+}	0.794^{+}	0.824
	(0.110)	(0.136)	(0.104)	(0.168)
Below-degree	1.079	1.042	1.076	0.904
	(0.115)	(0.135)	(0.107)	(0.125)
Degree	0.915	0.946	0.859	0.929
	(0.117)	(0.178)	(0.104)	(0.192)
Female	1.060	0.850	1.041	0.857
	(0.100)	(0.101)	(0.093)	(0.110)
White-British	1.168	1.051	1.191	1.229
	(0.271)	(0.186)	(0.248)	(0.269)
Number of comorbidities	1.039	1.173***	1.023	1.165***
	(0.030)	(0.043)	(0.027)	(0.041)
Owner	0.988	0.960	1.006	0.991
	(0.101)	(0.123)	(0.105)	(0.137)
Site 2	2.290***	0.937	2.089***	0.865
	(0.249)	(0.142)	(0.223)	(0.139)
Site 3	1.243	0.945	1.500^{+}	0.941
	(0.282)	(0.199)	(0.326)	(0.251)
Mean IMD score	1.000	0.994	0.999	0.994
	(0.005)	(0.006)	(0.005)	(0.007)
Some ADL problems	1.365**	1.162	1.292*	1.249
	(0.143)	(0.139)	(0.146)	(0.176)
Unable to wash/dress	3.536***	1.371^{+}	2.892***	1.922***
	(0.424)	(0.243)	(0.360)	(0.378)
Level 1 constant	359.401***	703.520***	487.307***	921.535***
	(111.958)	(239.374)	(133.252)	(400.093)

	Subject-specific		*Population average*	
Parameter	*Social Care* Exp (β) (SE)	*Hospital* Exp (β) (SE)	*Social Care* Exp (β) (SE)	*Hospital* Exp (β) (SE)
σ	1.159*** (0.029)	0.917** (0.028)	963	822
$\sigma^2[u]$	1.233*** (0.011)	2.328*** (0.183)		
ρ_{12}	1.966*** (0.032)	0.933 (0.162)		
N_i	963	822		
Interaction effects	F(4.000,465216.906)=1.763 p=0.133	F(4.000,5.598e+09)=1.493 p=0.201	F(4.000,269831.871)=2.101 p=0.078	F(4.000,68209401.489) =1.464 p=0.210

Table 6.31 Two-level subject-specific model: Partial effect/discrete changes in costs, baseline to follow-up of ITT allocation, by living arrangement

	Hospital care				*Social care*			
	Live alone		Live with others		Live alone		Live with others	
Intervention vs. control	Exp (β) (SE)	*p*		*p*	Exp (β) (SE)	*p*		*p*
Baseline	-0.196 (119)	0.999	199 (138)	0.147	250 (131)	0.056	-23 (159)	0.885
Follow-up	105 (115)	0.359	-189 (119)	0.112	191 (114)	0.093	58 (205)	0.776
Follow-up-Baseline	105 (-201,411)	0.499	-389 (-723, -54)	0.023	-58 (-321, 204)	0.662	81 (-361, 524)	0.719

Table 6.32 Two-level subject-specific model: difference-in-difference-in-difference (DDD) estimates

	Hospital care		*Social care*	
	Exp (β) (95% CI)	*p*	Exp (β) (95% CI)	*p*
DDD	-494 (-953,-36)	0.035	140 (-375,655)	0.595
Across living arrangements	-283 (-732, 165)	0.215	22.882 (-491, 536)	0.87
N_i	1506		1506	

Note: partial effects calculated using estimates from two-level costs and receipt models

Table 6.33 Population-averaged model: Partial effect/discrete changes in costs, baseline to follow-up of ITT allocation, by living arrangement

Intervention vs. control	*Hospital care* Exp (β) (SE)				*Social care* Exp (β) (SE)			
	Live alone	*p*	Live with others	*p*	Live alone	*p*	Live with others	*p*
Baseline	44 (178)	0.807	198 (152)	0.190	331 (176)	0.061	3 (138)	0.981
Follow-up	184 (131)	0.159	-187 (113)	0.098	238 (138)	0.084	46 (182)	0.801
Follow-up-Baseline	141 (-279,560)	0.511	-386 (-752,-20)	0.039	-93 (-428,242)	0.588	42.565 (-348,433)	0.831

Table 6.34 Population-averaged model: difference-in-difference-in-difference (DDD) estimates

	Hospital care Exp (β) (SE)	*p*	*Social care* Exp (β) (SE)	*p*
DDD	-526.580 (-1097.735,44.576)	0.071	135.202 (-370.582,640.986)	0.600
Across living arrangements	-245.204 (-787.059,296.651)	0.375	-50.072 (-573.359,473.216)	0.851
N_i	1506		1506	

Note: partial effects derived from two-part population-averaged model estimates

6.18 Discussion of Telecare Subgroup Analyses Results

The treatment effect at follow-up in people living with others was to reduce combined health and social care expenditure, but this effect was not significantly different from zero at the 5 per cent level. The direction of the effect of treatment on expenditure did not vary depending on the inclusion or exclusion of direct intervention costs, nor on whether a subject-specific or population-averaged approach was taken. Controlling for other factors, health and social care costs were lower at follow-up regardless of allocation.

In a separate model of social care expenditure on those in receipt of care, the effect of trial allocation to the intervention was not associated with any trend in spending in terms of either living arrangements or time points. In separate models of NHS costs and costs of those using hospital services, the results indicate that the impact of the intervention on the costs of a given participant living with others differed from the impact on a given participant living alone. Any apparent reductions in expenditure related to telecare occurred in the NHS rather than in social care.

An examination of the marginal effects of telecare at baseline and follow-up by type of living arrangement yielded some interesting findings. It appears that there were savings to the NHS in the group of telecare participants living with others at the follow-up. However once the lack of savings made in providing services for telecare recipients living alone was taken into account, we cannot be certain of making savings, given the lack of statistical significance even at the 10 per cent level. A similar pattern was seen in the two-part model of hospital costs. In contrast, the results of the two-part model of social care expenditure suggest that costs were not greatly affected by the introduction of telecare, whether participants lived alone or with others.

Subject-specific and population-averaged modelling approaches were taken in order to examine the impact of telecare on individuals in the sample and also to be able to make inferences about the impact of telecare over the groups under comparison (experimental groups in different living arrangements) that would be useful in a policy context. The results of the two approaches yielded generally comparable findings.

Across sectors, time points and models, number of comorbidities and higher levels of ADL need were found to be associated consistently with higher costs. There were also associations between higher costs and site: higher total costs and higher odds of receipt and costs of social care were associated with site 2, relative to site 1. Site 3 was associated with lower odds of social care receipt than site 1. While older age was associated with lower total and NHS costs and lower odds of receipt of hospital care, this covariate was associated with higher hospital and social care costs.

The results presented here require careful interpretation. Telecare is advocated as a technology that promotes independence and thereby reduces costs associated with the care and support of its users. People living alone may be expected to particularly benefit from the technology because it is thought to reduce the risks of falling or other accidents. The results suggest that people in receipt of telecare who live alone are the subject of more health care expenditure than those in receipt of telecare who are living with others. This could be for a number of reasons. The need for care may not have been completely captured, for instance in that the level of cognitive impairment was not measured by any study instrument, so that the groups did actually differ in terms of related needs. Alternatively, responders and also responses to telecare activations could differ. People living alone might be more likely to receive a formal (paid) service response that might raise further concerns about unmet needs, which would prompt health investigations. Those living with others might get an unpaid response from their co-resident carer so that formal services are not activated, or

professionals involved might be less concerned given the availability of informal carers. In other words, health professionals' responses might differ because they perceive clients living alone to be at more *risk* than those living with others, for the same assessed level of *need*.

There is some evidence from the UK and elsewhere to support this suggestion. Living alone has been estimated to increase the hospital costs of elective surgery (Turner, Nikolova, and Sutton 2016). One US study found that older people living alone were 60 per cent more likely to visit an emergency department over a 12-month period than people living with a spouse (Hastings et al. 2008). Another found that a population of older veterans living alone used more outpatient visits than those living with others (Guzman, Sohn, and Harada 2004). Two Finnish studies provide evidence in the opposite direction. An analysis of older people participating in a trial of rehabilitation suggested that living alone was a strong predictor of *social care* use, as were indicators of availability of informal care, older age, mood, self-reported health, ADL-related need and cognition (Kehusmaa et al. 2012). Also, functional ability and living in the community were related to lower health care costs, controlling for other needs and socio-demographic characteristics. Older people living alone in Finland are more likely to receive health and social care services, and be the subject of higher health and social care expenditure, than those living with others (Noro, Häkkinen, and Laitinen 1999). In the UK, such allocation decisions may depend on implicit assumptions that unpaid co-resident carers can substitute for paid carers, particularly in times of cuts to councils' budgets (Milne et al. 2013). It is possible that while professionals are not comfortable with adjusting services to those living alone and relying on telecare to 'manage the risks of independent living' (Bower et al. 2011), they are more comfortable with adjusting services for those living with others. There is some evidence here that NHS and particularly hospital services were directed towards individuals living alone in receipt of telecare rather than those in receipt of telecare living with others; in contrast there is no evidence for this pattern of response by social services. Perhaps social care assessors are less risk-averse than health professionals (Cameron 2006). It is also possible that co-resident carers' behaviour was changed by telecare so that they did not seek health care for the participant where they otherwise would have done. It is also possible that telecare does more to address health-related rather than social care-related need than we might assume. Agboola et al. (2014) found that people admitted after activating their personal emergency response system were more likely to be admitted for chronic illness or infection reasons than for falls.

Another possibility is that the telecare sensors of people living alone were more likely to be activated for some reason. There is some evidence also that those living alone both are

more likely to use an alarm and to have more, and more serious, injuries if they have a fall (Elliott, Painter, and Hudson 2009). Although one rationale for advocating the use of personal/pendant alarms (not necessarily telecare) is to mitigate the consequences of falls, those falling do not necessarily use their alarms (Fleming and Brayne 2008) and false alarm-related call-outs to the emergency services can occur (Johnston et al. 2010). Other studies have found that personal alarms may not decrease anxiety, fear of falling or use of health services (Lee et al. 2008). Whether a similar phenomenon occurs with telecare (where sensor activations may create additional false alarms and additional visits and therefore more input and more expenditure) would require further research. Data on alarm activations was not available for analysis and so could not be controlled for in the analyses.

6.19 Limitations

The analysis does not adjust for the characteristics of others in the household, who may also have social care needs and be benefiting in some way from the telecare service, including less reliance on care and support services. Data was not collected on the relationship between the other members of the household and the participant so different caring relationships could not be controlled for. If the respondents differed in their likeliness to recall service use by living arrangement – for instance in that a co-resident carer could prompt a participant to recall an appointment that would have been otherwise forgotten by someone living alone – the difference between groups in terms of living arrangements could be exaggerated. However the relationship of the differences between allocation groups within these subgroups would not be affected.

The observations available for each participant were limited to three-month cost snapshots at baseline and follow-up. This limited the amount of information available on within-participant variation in costs over time. Making inferences about change over time must be considered in this light.

6.20 Implications for Policy and Research

The research presented here raises some questions that would benefit from further research. More work is needed to understand the decision-making that takes place when telecare systems bring people living alone into contact with health professionals. Perhaps a different approach to managing risk is needed, for instance to allow health professionals to ‘live with’ the perceived risks run by people living alone. If risk-aversion results in bad and expensive

outcomes for people living alone using telecare, then reducing risk-averse decision-making could shift resources into more effective activities. For example, the system could redistribute some funds towards those living with others who are perhaps benefiting from the reassurance provided to their carers to be able to go out and leave the person unaccompanied.

Chapter 7

Cost-effectiveness of Telehealth

There is enormous policy interest in technologies that help people with long-term conditions to manage disease and the symptoms of disease effectively and cost-effectively. Whether technologies such as telehealth can make a difference to the lives of the population with chronic disease has been the focus of much research since the 2000s. The relationship between health and social care expenditure and quality of life and psychosocial outcomes requires careful investigation. Further research is needed to allow policymakers to make an informed decision when considering whether telehealth represents a useful and effective route for delivering health care and equally whether it represents the best use of available public funds. This chapter focuses on the question: is telehealth in addition to standard support and treatment cost-effective compared to standard support and treatment alone?

In this chapter I present the findings of an analysis of data from the WSD cluster-randomised controlled trial, designed to address the question of cost-effectiveness of a telehealth intervention. I begin with a recapitulation of the context of the study and statistical methods employed in the economic evaluation. I present the results and discuss the implications of the findings for policymakers and address the limitations and strengths of the analysis.

7.1 Context

Telehealth was defined in the WSD trial as "the remote exchange of data between a patient and health care professional to assist in the diagnosis and management of a health care condition" (Bower et al. 2011). Participants used telemonitoring equipment to collect and transmit vital signs data. These data were classified into risk-related alerts (for example, using a traffic light system), according to parameters that would be set initially on the basis of clinical guidelines or by a clinician responsible for that patient's care. The parameters were reset by a clinician (general practitioner, telehealth nurse, or community matron) as required, after an initial settling-in period. The exact response to the alert depended on the risk level associated with the readings, clinical judgment and local protocols that were usually based on clinical guidelines (Bower et al. 2011).

Monitoring staff were also able to transmit health-related questions, messages, or videos to educate patients on their conditions, using the telehealth base unit or set-top box. Participants in the telehealth group were not charged for using telehealth services (for instance, freephone numbers were provided for calls to central monitoring teams or for transmitting vital signs data). On the other hand, participants were expected to have or to arrange for a telephone line, power points and electricity, and in one site they were expected to have a television available.

The telehealth systems employed within the trial included elements of both telemonitoring and telephone support. The trial was not designed to investigate the effect of individual service configurations or technologies (Bower et al. 2011). Rather, the evaluation examined whether "telehealth" as a class of technologies added to standard support and treatment, is cost-effective compared with standard care alone. As discussed in Chapter 4, each study site had different suppliers and service models, which evolved over the course of the trial.

7.2 Methods

The primary outcome of interest in the cost-effectiveness analysis was the total cost to health and social care per QALY gained by implementing the telehealth intervention. The evaluation also explored several secondary outcomes (state anxiety, depression, well-being and QALY derived from SF-6D utilities). As data were missing in variables to be used in the analysis (see Chapter 4), ten complete datasets were imputed as has been described in Chapter 4, Section 4.22. Seemingly unrelated regressions (SUR) were then applied, adjusting standard errors for cluster (general practice). Outcome equations adjusted for baseline outcome, site, demographic and individual characteristics (age, sex, ethnicity, IMD, index condition, number of chronic conditions). Cost equations adjusted for baseline costs, site, demographic and individual characteristics (age, sex, ethnicity, IMD, index condition, number of chronic conditions). This modelling approach allowed the estimation of the impact of the intervention while controlling for clustering and accounting for the correlation between the cost and outcome variables.

Sensitivity analyses explored the robustness of results to variations in the costs of the intervention in terms of lower input prices for equipment and lower costs of telehealth monitoring support.

The results from the SUR regressions were used to estimate incremental cost-effectiveness ratios (ICERs) – the additional cost per unit of outcome from the addition of telehealth to standard care – and net monetary benefit. Cost-effectiveness acceptability curves were plotted, depicting the likelihood that telehealth is cost-effective given different assumptions about willingness to pay for outcomes.

7.3 Results

At baseline, service use and costs data were available for 841 intervention and 728 control participants. At the 12-month follow-up, outcomes data were available for 974 participants, of whom 969 had costs data available (538 intervention, 431 control). Costs and outcomes data at baseline and 12-month follow-up were available for 965 participants (534 intervention, 431 control).

7.3.1 Costs

Total annual equivalent costs are given in Table 7.1 for the 965 cases available at baseline and follow-up (consisting of 431 control and 534 intervention participants). Annual equivalent costs prior to baseline were similar between experimental groups but rather higher in the usual care group in the study year, if intervention costs are not included in the total. In contrast, the usual care group had rather lower costs, if intervention costs are taken into account. Nonetheless, the cluster-adjusted 95 per cent confidence intervals of the total cost differences at baseline and at 12-month follow-up were wide and crossed zero.

7.3.2 Outcomes

Descriptive statistics on outcome data at baseline and 12-month follow-up are summarised in Table 7.2. Controls had slightly higher scores at baseline on a number of measures, although on formal testing (using a clustered t-test) the differences between groups were not significantly different from zero at the 5 per cent level. The difference in mean EQ-5D-3L-derived QALY (non-significant at the 5 per cent level) was quite small. Comparing results from the two generic preference-based measures of health, EQ-5D-3L and SF-12, baseline utilities derived from EQ-5D-3L were lower than those derived from the SF-6D; however either instrument yielded a very small between-group difference in mean utility scores (-0.001 and -0.007 respectively). The difference between experimental groups at 12-month follow-up in mean EQ-5D-3L-derived utilities (non-significant at the 5 per cent level) was

much larger than the difference in utilities derived from the SF-6D (0.031 vs. 0.007). The (non-significant) difference in EQ-5D-3L-derived QALY was also small but larger than the SF-6D derived QALY (0.031 vs 0.000). The Brief STAI was the only outcome measure where the difference between the groups was significant at the 5 per cent level; the difference was nonetheless small (less than one point, on a 24-point scale).

Table 7.1 Mean service costs (£) across Telehealth sample, annual equivalent

Resource item	Usual care (SE) (n=431)	Telehealth (SE) (n=534)	Raw difference (95% CI)
Pre-baseline period			
Total costs	4431 (325)	4731 (298)	300 (-572, 1173)
Intervention period			
Total costs excluding Telehealth delivery and equipment	5575 (480)	4603 (445)	-973 (-2266, 320)
Telehealth equipment costs	15 (34)	673 (33)	658 (563, 752)**
Telehealth intervention costs	22 (17)	1156 (17)	1134 (1086, 1182)**
Total costs including Telehealth delivery and equipment	5613 (480)	6431 (445)	819 (-476, 2113)
- project management costs	5609 (480)	6240 (446)	631 (-664, 1926)
Sensitivity analyses			
50% reduction in equipment prices	5605 (479)	6095 (445)	490 (-803, 1782)
80% reduction in equipment prices	5601 (479)	5893 (444)	292 (-1000, 1585)
Operating at increased capacity	5596 (478)	5960 (444)	364 (-926, 1654)
Operating at increased capacity and 50% reduction in equipment prices	5595 (479)	5621 (444)	25 (-1267, 1317)
Operating at increased capacity and 80% reduction in equipment prices	5591 (479)	5419 (444)	-172 (-1464, 1120)

Note: Table reports the annual equivalent costs for 965 cases with baseline cost data available, (10 complete datasets) (10 complete datasets). Standard errors are cluster-adjusted.
*p<0.01 on clustered t-test
**p<0.05 on clustered t-test

Table 7.2 Outcomes at baseline and 12-month follow-up, Telehealth sample

Outcome measure	Usual care (SE) (n=431)	Telehealth (SE) (n=534)	Difference (95% CI)
Baseline			
Utility (EQ-5D-3L)	0.556 (0.021)	0.554 (0.02)	-0.001 (-0.059, 0.056)
STAI	10.019 (0.249)	9.803 (0.235)	-0.216 (-0.891, 0.46)
CESD-10	9.377 (0.331)	9.405 (0.304)	0.028 (-0.862, 0.917)
ICECAP-O	0.786 (0.011)	0.797 (0.01)	0.011 (-0.019, 0.041)
Utility (SF-6D)	0.648 (0.009)	0.641 (0.009)	-0.007 (-0.032, 0.018)
12-month follow-up			
Utility (EQ-5D-3L)	0.537 (0.02)	0.568 (0.019)	0.031 (-0.024, 0.086)
STAI	11.528 (0.201)	10.692 (0.18)	-0.836 (-1.369, -0.303)**
CESD-10	10.491 (0.36)	9.735 (0.331)	-0.757 (-1.723, 0.21)
ICECAP-O	0.753 (0.009)	0.767 (0.009)	0.014 (-0.011, 0.039)
QALY - EQ-5D-3L	0.546 (0.019)	0.561 (0.018)	0.015 (-0.038, 0.068)
Utility (SF-6D)	0.629 (0.009)	0.636 (0.009)	0.007 (-0.018, 0.032)
QALY - SF-6D	0.638 (0.009)	0.638 (0.008)	0 (-0.023, 0.024)

Note: Table reports results for 965 cases with baseline cost and outcome data available, (10 complete datasets). Standard errors are cluster-adjusted.
*p<0.01 on clustered t-test
**p<0.05 on clustered t-test

7.4 Cost-effectiveness Analyses

In the SUR model, there was a small but significant difference between the groups in the primary outcome (Table 7.3), QALY over the period to the 12-month follow-up (a mean difference 0.014). In terms of secondary outcomes, most differences were small and not significant at the 5 per cent level, except for the CESD-10 (the difference being small, at less than one point) (Steffens et al. 2002); and the STAI, again with less than one-point difference. Base case costs including intervention costs were non-significantly higher among the telehealth group than the usual care group.

The SUR model estimates yielded an ICER of £67,000 per QALY (Table 7.3). Excluding project management costs, the ratio fell to £54,200. Looking at the secondary outcome measures, the ratio for an improvement from highest to lowest levels of anxiety on the Brief STAI scale was £22,600; for the CESD-10 scale, the ICER was £6,900 for achieving a five-point reduction; the ICER for an improvement from no capability to full capability on the ICECAP-O scale was £233,700.

Whether telehealth can be considered to be cost-effective depends on the willingness to pay for the outcomes generated. Figure 7.1 presents the probability that telehealth would be seen as cost-effective as an addition to usual care, using an acceptability curve for different values of willingness to pay. At the £30,000 threshold (associated with NICE recommendations (National Institute for Health and Clinical Excellence 2008)), the probability of cost-effectiveness was 20 per cent. This probability only exceeded 50 per cent at threshold values of willingness to pay above £67,000. Figure 7.1 also shows the probability of cost-effectiveness if costs related to project management were excluded: at the £30,000 threshold, the probability of cost-effectiveness was 29 per cent. Excluding project management costs, the probability exceeded 50 per cent only at values above about £57,000.

7.4.1 Secondary Outcomes

There were significant but very modest differences between intervention and control groups in state anxiety and depression symptoms (Table 7.3). The probability of cost-effectiveness for a 100 per cent improvement from highest to lowest levels of anxiety on the Brief STAI only exceeded 50 per cent at willingness to pay levels above about £22,600 (Figure 7.2). The probability that the treatment was cost-effective in achieving a five-point reduction on the CESD-10 scale exceeded 50 per cent at levels of willingness to pay above about £7000, and reached 90 per cent at about £22,000 (Figure 7.3). In relation to an improvement from no capability to full capability on the ICECAP-O index, the probability of cost-effectiveness of telehealth was 15 per cent at a willingness to pay of £50,000 (Figure 7.4). On the QALY derived from the SF-12, the difference between groups was 0.005 (in favour of the intervention); the ICER was 178,600. The probability of cost-effectiveness at £30,000 was 9 per cent.

Table 7.3 Differences in cost (£) and effect between Telehealth and UC groups (12 months), annual equivalent

Values in means (CI) unless otherwise stated	**Outcomes / total costs Control=431**	**Outcomes / total costs Telehealth=534**	**Difference in outcomes / total costs or ICER (Control=431; Telehealth=534)**
QALY (adjusted mean, SUR model†)	0.547 (0.537 , 0.557)	0.561 (0.552, 0.570)	0.014 (0.001, 0.028)*
Cost (adjusted mean, SUR model‡)	5 530 (4 601,6 460)	6 498 (5 924, 7 072)	968 (-145, 2080)
ICER (£ per QALY) (SUR model§)	-	-	67 000 (-8 600, 1 406 900)
excluding project management costs			
Cost (adjusted mean, SUR model‡)	5525 (4594, 6456)	6308 (5731, 6885)	783 (-332, 1 898)
ICER (£ per QALY) (SUR model§)	-	-	54 200 (-21 100, 1 202 400)
Secondary outcomes analyses			
Brief STAI (adjusted mean, SUR model†)	11.49 (11.814 , 11.184)	10.72 (11.022 , 10.428)	-0.774 (-0.342 , -1.206) ||
Brief STAI ICER (£) (SUR model‡§)			22 600 (-3 200, 74 600)
CESD-10 (adjusted mean, SUR model†)	10.46 (10.91 , 10.02)	9.76 (10.17 , 9.35)	-0.705 (-0.095 , -1.315) ¶
CESD-10 ICER (£) (SUR model‡§)			6 900 (-900, 72 000)
ICECAP-O (adjusted mean, SUR model†)	0.759 (0.746, 0.771)	0.763 (0.753, 0.772)	0.004 (-0.012, 0.020)
ICECAP-O ICER (£) (adjusted mean, SUR model†)			233 700 (unbounded,unbounded)
QALY SF-6D (adjusted mean, SUR model†)	0.635 (0.629,0.641)	0.641 (0.636, 0.646)	0.005 (-0.003, 0.014)
QALY SF-6D ICER (£) (SUR model§)			178 600 (-21 000, 378 800)

Note: Table reports results for 965 cases with baseline cost and outcome data available, (10 complete datasets). Standard errors are robust cluster-adjusted.

† From SUR analyses (outcome equation), adjusted for allocation, baseline outcome, site, age, sex, ethnicity, IMD, index condition, number of chronic conditions

‡ From SUR analyses (cost equation), adjusted for allocation, baseline costs, baseline outcome, site, age, sex, ethnicity, IMD, index condition, number of chronic conditions

§ rounded to nearest hundred

|| Retransformed to original scale to enable comparison with raw mean difference; transformed mean=0.043 (0.19, 0.067)

¶ Retransformed to original scale to enable comparison with raw mean difference; transformed mean=0.141 (0.019, 0.263)

Table 7.4 Sensitivity analyses: differences in cost (£) and effect between Telehealth and UC groups (12 months), annual equivalent

Values in means (CI) unless otherwise stated	Outcomes / total costs Control=431	Outcomes / total costs Telehealth=534	Difference in outcomes / total costs or ICER (Control=431; Telehealth=534)
Variations in intervention costs:			
Equipment prices reduced by 50%			
Cost (adjusted mean, SUR model‡)	5 524 (4 593, 6 455)	6160 (5 585, 6 735)	636 (-478, 1750)
ICER (£ per QALY) (SUR model§)	-	-	44 100 (-32 300, 1 045 500)
Equipment prices reduced by 80%			
Cost (adjusted mean, SUR model‡)	5 520 (4 588, 6 453)	5958 (5 381, 6 534)	437 (-678, 1 553)
ICER (£ per QALY) (SUR model§)	-	-	30 300 (-51 000, 833 000)
Operating at increased capacity			
Cost (adjusted mean, SUR model‡)	5 516 (4 588, 6 445)	6 024 (5 451, 6 598)	508 (-600, 1 600)
ICER (£ per QALY) (SUR model†‡)	-	-	35 167 (-43 500, 905 400)
Equipment prices reduced by 50% & operating at increased capacity			
Cost (adjusted mean, SUR model†)	5 464 (4 572 , 6 356)	5771 (5 183 , 6 359)	167 (-946, 1 280)
ICER (£ per QALY) (SUR model§)	-	-	11 600 (-87 500, 557 700)
Equipment prices reduced by 80% & operating at increased capacity			
Cost (adjusted mean, SUR model†)	5 517 (4 586, 6 448)	5684 (5 111, 6 257)	-32 (-1 100, 1 100)
ICER (£ per QALY) (SUR model§)	-	-	-2 200 (-134 100, 373 000)

Note: Table reports results for 965 cases with baseline cost and outcome data available, (10 complete datasets). Standard errors are robust cluster-adjusted.

† From SUR analyses (outcome equation), adjusted for baseline outcome, site, age, sex, ethnicity, IMD, index condition, number of chronic conditions

‡ From SUR analyses (cost equation), adjusted for baseline costs, baseline outcome, site, age, sex, ethnicity, IMD, index condition, number of chronic conditions

§ rounded to nearest hundred

Figure 7.1 Cost-effectiveness acceptability curve: QALY

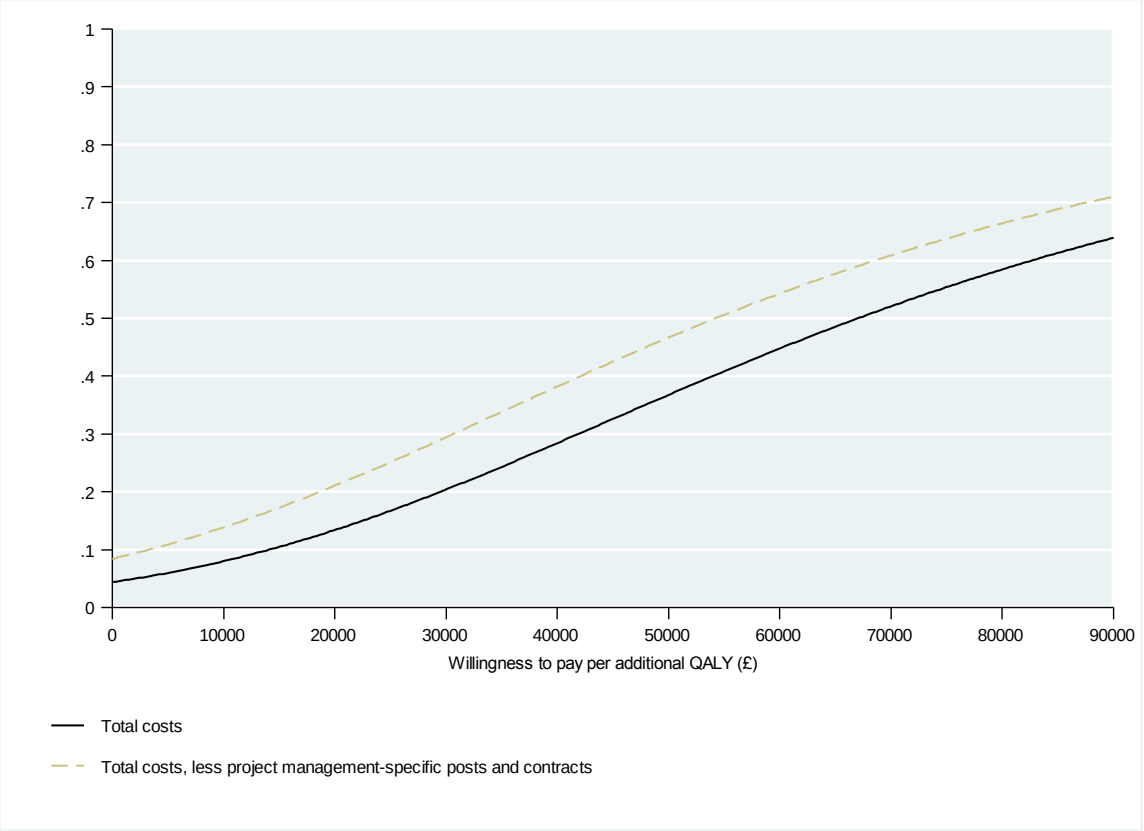


Figure 7.2 Cost-effectiveness acceptability curve: Brief STAI

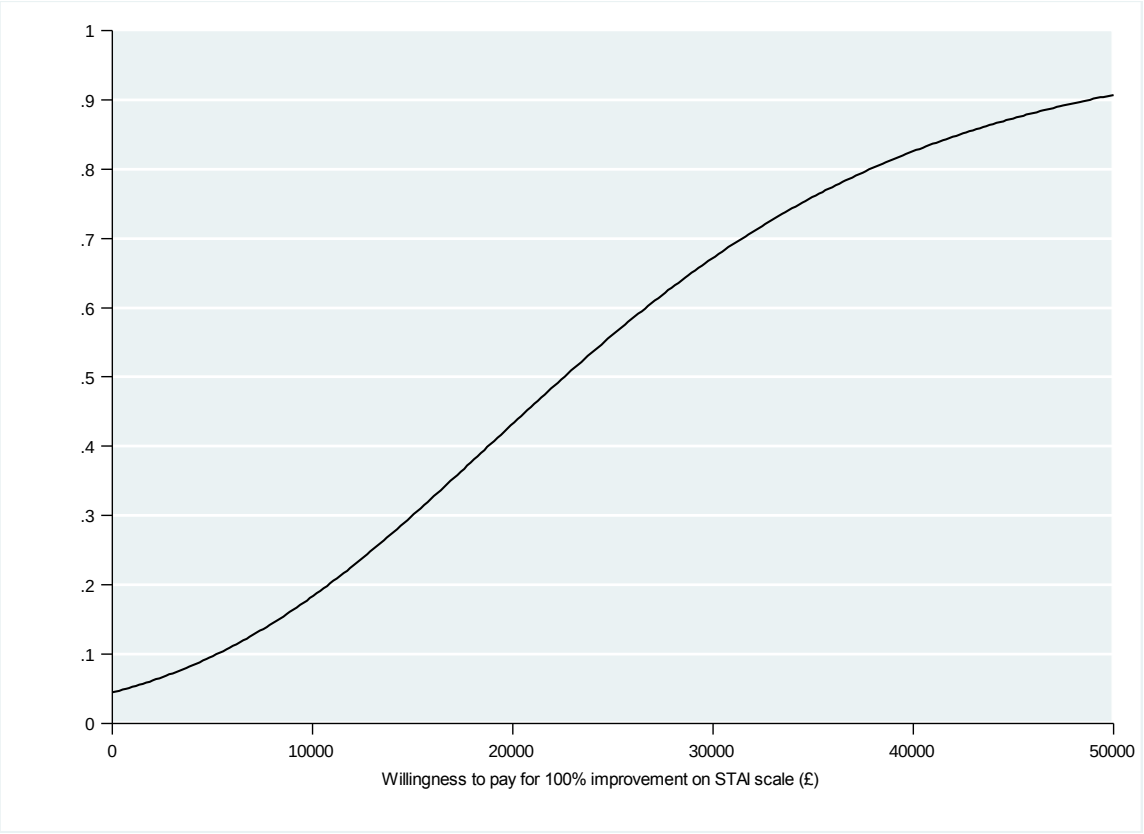


Figure 7.3 Cost-effectiveness acceptability curve: CESD

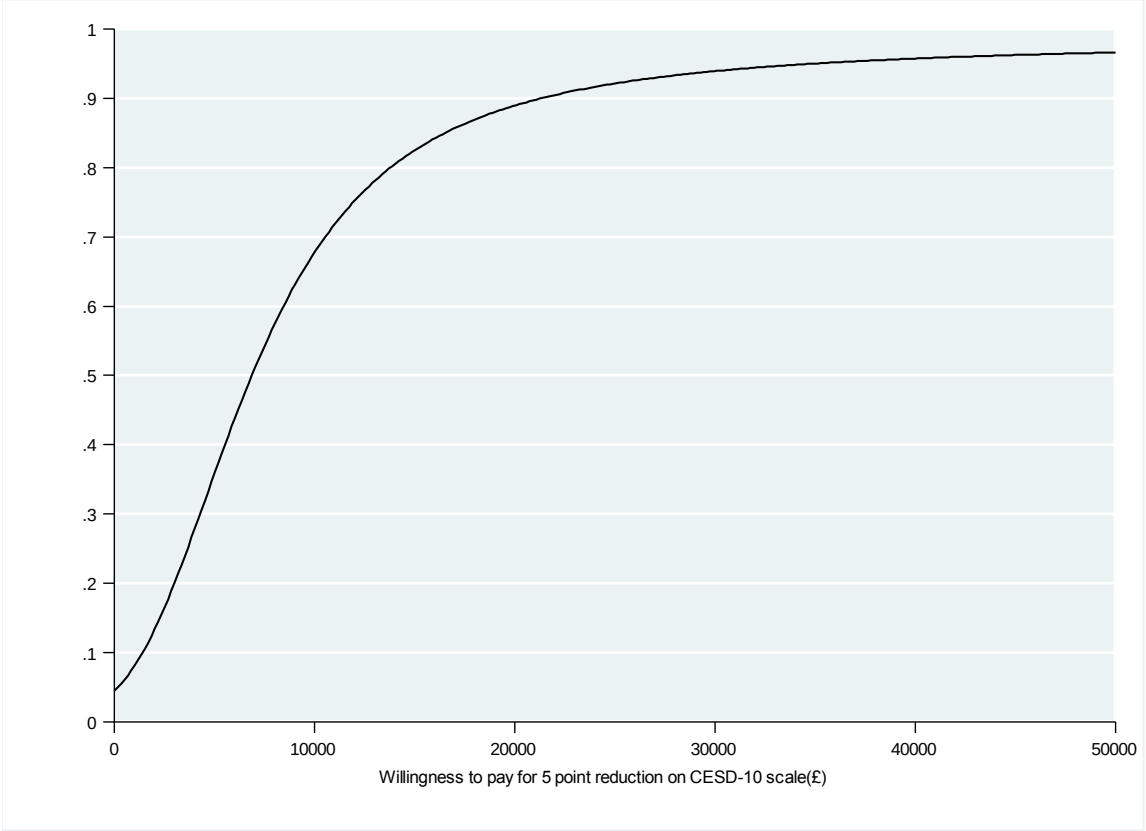


Figure 7.4 Cost-effectiveness acceptability curve: ICECAP-O

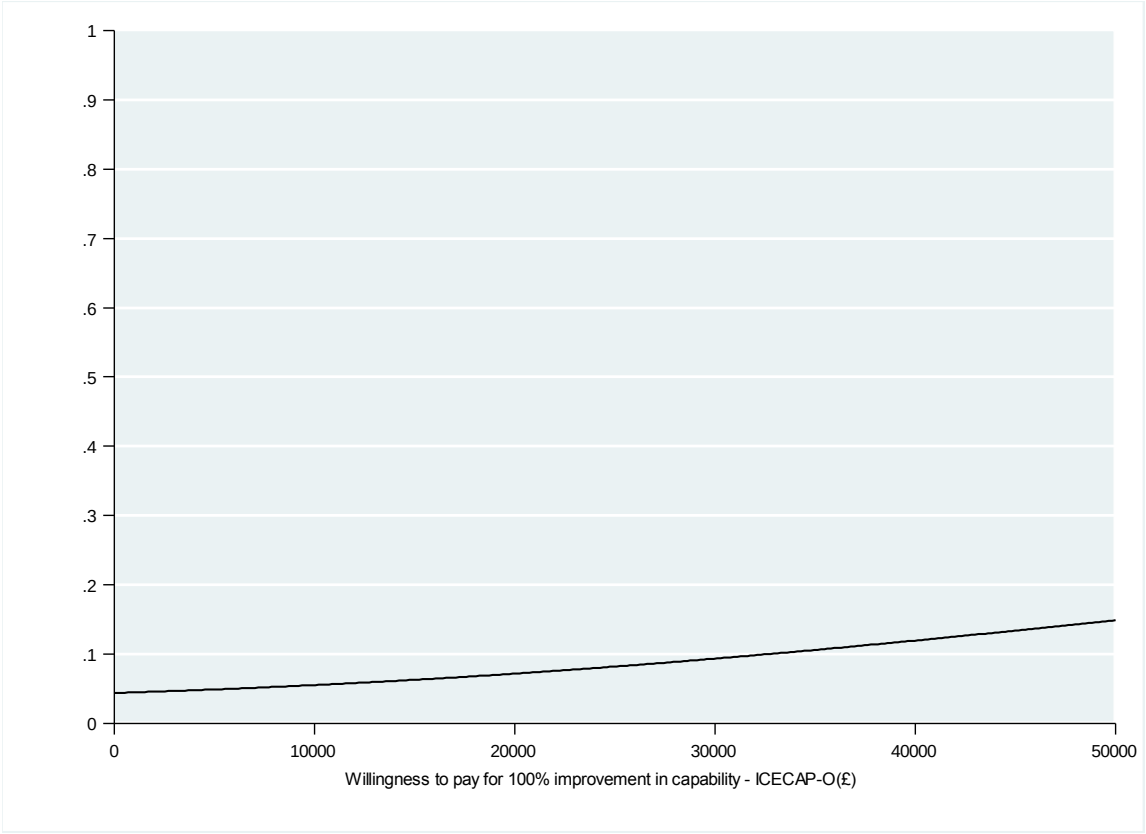
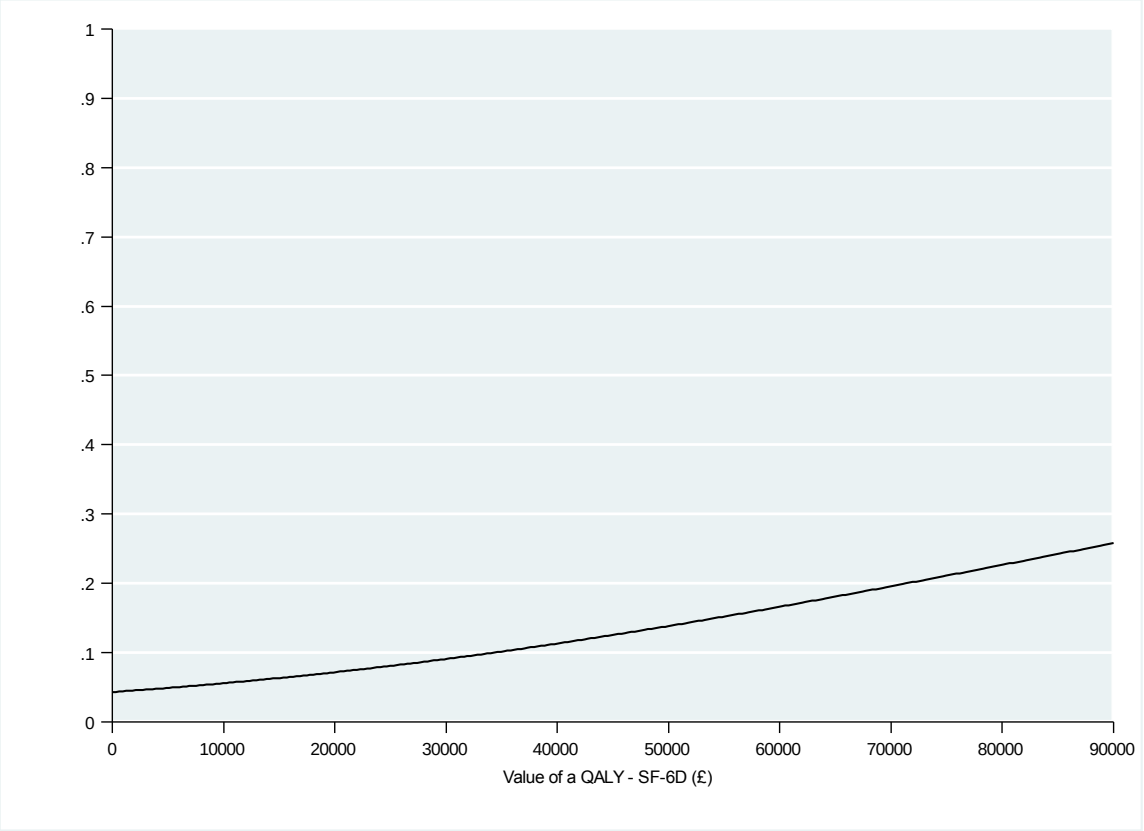


Figure 7.5 Cost-effectiveness acceptability curve: QALY (SF-6D)



7.4.2 Sensitivity Analyses

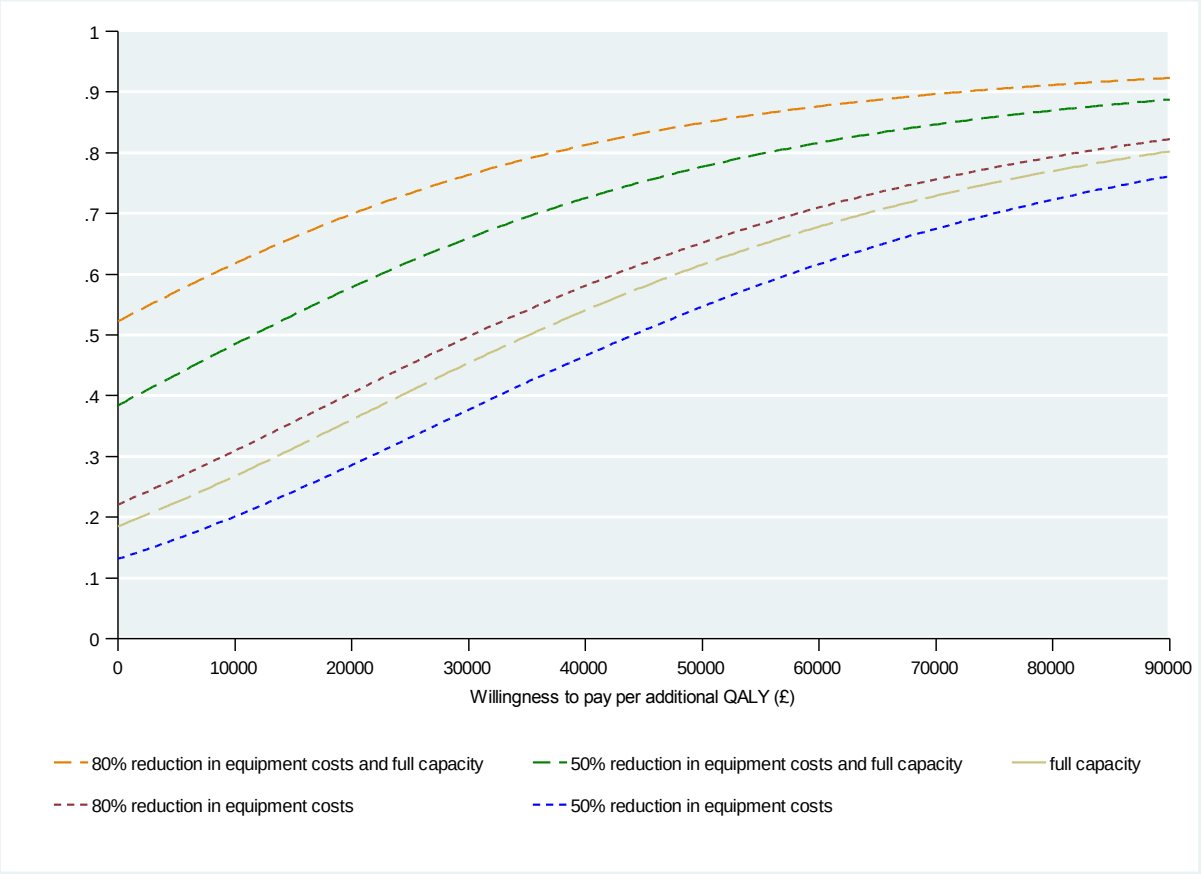
Equipment costs contributed a sizeable proportion of direct costs per person for the telehealth group (Table 7.1). Table 7.1presents the three-month costs estimated for the sensitivity analyses (presented here multiplied by 4 to give the yearly equivalent). If equipment prices fell by 80 per cent, estimated mean costs per year (unadjusted) for the telehealth group fell from £6431 (SE £445) to £5893 (SE £444). However, total costs of the telehealth group remained slightly higher than those of the usual care group (difference £292 (-£1000, £1585)). If equipment prices decreased by 50 per cent, total costs for the telehealth group were also higher than for the usual care group (£490, 95 per cent CI -£803, £1782)) Under the 80 per cent reduction in equipment costs scenario, the ICER fell to £30,300 per QALY (Table 7.4). In the scenario (Table 7.1) where the service was working ‘at increased capacity,’ the raw annual mean costs of the telehealth group fell to £5960 (standard error £444).

7.4.3 Reduction in Equipment Costs and Full Utilisation Combined Scenario

The two sensitivity analyses were also combined. At an 80 per cent reduction in equipment costs and a reduction of support costs associated with working at full capacity, the raw

difference between groups decreased (Table 7.1). Total mean costs of telehealth per year (unadjusted) per participant were non-significantly less for telehealth (-£172 (95% CI-£1464, £1120) than for usual care (Table 7.1). At a 50 per cent reduction in equipment costs with the same decreased labour costs, the corresponding cost was non-significantly more (£25, 95% CI -£1267, £1317) for the telehealth group than for the usual care group. In the adjusted model of costs derived from the SUR analyses (Table 7.4), the cost differences were not significant, being a little higher for the telehealth group than for the usual care group, assuming 50 per cent reductions in input price and higher working capacity (increase of £167, 95% CI -946, 1280); and a little lower assuming 80 per cent reductions in input price and higher working capacity (decrease of £32, 95% CI -1100, 1100). With an 80 per cent reduction in equipment costs and operating at the higher capacity, the cost-effectiveness ratio was negative (-£2,200 per QALY), with ICER confidence intervals that crossed zero. Figure 7.6 shows cost-effectiveness acceptability curves for all sensitivity analyses. No substantial changes to the results were seen: assuming an 80 per cent reduction in equipment costs, the probability that telehealth was cost-effective was 50 per cent at a willingness to pay level of £30,000 per QALY. Results from the sensitivity analyses based on operating at full capacity were similar. However, combining the two scenarios (at an 80 percent reduction in equipment costs) increased the likelihood that telehealth was cost-effective, to 76 per cent for a willingness to pay of £30,000 per QALY.

Figure 7.6 Cost-effectiveness acceptability curve: QALY, sensitivity to input prices



7.5 Discussion

The WSD telehealth questionnaire study was nested in a large-scale pragmatic, randomised controlled trial of telehealth in England. Costs and outcomes data at 12 months were available for 969 participants. The mean costs of self-reported service use, combined with telehealth intervention costs, were somewhat greater for the group randomised to telehealth in addition to standard care than for the group randomised to usual care alone. In a model adjusting for demographic characteristics and level of need, this difference in costs was also somewhat greater. For the primary outcome measure, the probability that telehealth was cost-effective was relatively low, only exceeding 50 per cent at willingness-to-pay values above £67,000 per QALY. The probability of cost-effectiveness measured in terms of anxiety and depression symptoms rose above 50 per cent at willingness to pay values in excess of £22,600 and £7,000, respectively.

7.5.1 Strengths and Limitations

One limitation of self-reported data on service use is that respondents may under-report services that they frequently use (Bhandari and Wagner 2006, Richards, Coast, and Peters 2003). Notwithstanding, relying on self-reported service use remains an important method of collecting data for a wide range of health and social care services, since administrative data are agency- or service-specific. Furthermore, some administrative data may under-report the patient's receipt of services not directly provided by the data owner (for instance GP records on patients' use of social care or other community health services) (Byford et al. 2007, Mistry et al. 2005). It has been recommended that a shorter period of recall is used for frequently used services in order to minimise this issue (Bhandari and Wagner 2006) and a three-month time-frame was used in this study.

An assumption was made that participants' costs between nine and 12 months could be multiplied up to a yearly cost. This method of estimation may have made the findings of this analysis more conservative; longitudinal hospital data have shown that initial differences between groups in bed days narrowed over the period of the intervention (Steventon et al. 2012). However, the pattern associated with acute hospital services cannot be assumed to hold with services that are more frequently used and less episodic, such as community nursing or home care.

The extent to which the costs and outcomes differed between those participants who completed the 12-month follow-up and those who did not is unknown. The analysis adjusted for demographic and cost covariates at baseline that might influence the decision to complete long-term follow-up, and so went some way towards addressing imbalances caused by dropout between intervention and control groups. The analyses took account of both clustering and the correlation between the cost and outcome variables.

The telehealth interventions under study were complex (Craig et al. 2008), involving both human services and advanced assistive technologies. A number of issues are likely to arise in the economic evaluation of complex interventions: users might be a heterogeneous group; users could be highly involved in the production of care; the more active the user involvement, the more complicated the association is between inputs and outputs; and multiple agencies could be involved in delivering the intervention (Byford et al. 2007). The intervention also involved coproduction by teams that varied in composition from site to site.

As a consequence of the pragmatic nature of the trial, heterogeneity arose from differences in the way that the interventions were delivered. Pragmatic trials, if they are to be

useful in directing policy and practice, must be representative of the real-world clinical population to which new interventions will be applied (Roland and Torgerson 1998).

There might have been variations in the mix and balance of mainstream services within and between the health and social care providers in the sites. While introducing the telehealth intervention into multiple sites improved generalisability, it was correspondingly more difficult to specify the exact nature of the intervention to be used and identify which features might have been more helpful in improving health-related quality of life. The study was neither intended nor powered to examine differences in outcomes between specific service delivery models. On the other hand, certain core features of the telehealth intervention were in evidence across the sites: store and forward systems, patient education protocols, computerised risk-based classification of vital signs data and central monitoring teams. Whether implementation of this "disruptive" technology (Coye, Haselkorn, and DeMello 2009) at these sites caused any system-wide change in the delivery of local health and social care services was beyond the scope of the economic evaluation.

Other research in the WSD research programme has examined the effects of telehealth on organisations and professionals (Bower et al. 2011). Unsurprisingly the impacts of the trial from these perspectives were ambiguous. The demands of the RCT itself may have hindered more organic processes of adopting remote care within organisations. After the trial, the sites began to review the population that had been recruited in order to satisfy the sample size requirements of the trial, seeking to target patients according to their own local service objectives (Hendy et al. 2012). Nursing professionals were largely supportive of the benefits of telehealth to empower patients to manage their conditions whereas GPs had more guarded attitudes towards the usefulness of the technology (MacNeill et al. 2014).

The economic evaluation focused on self-reported outcomes and service use, and did not include surrogate measures of outcome such as levels of glycated haemoglobin (HbA1c) (Park et al. 2008), blood pressure readings or mortality (although mortality was examined elsewhere in the WSD evaluation (Steventon et al. 2012)). Recent reviews and studies have identified promising results from trials of telehealth in a variety of long-term conditions including diabetes, heart failure, chronic obstructive pulmonary disease and asthma (Barlow et al. 2007, Clark et al. 2010, Inglis et al. 2010, Pare, Janna, and Sicotte 2007, Polisena, Coyle, et al. 2009, Polisena, Tran, et al. 2009). The bulk of this evidence concerns results measured by surrogate and mortality outcomes, rather than by self-reported data on health-related quality of life. Systematic reviews have reported rather mixed evidence in favour of telehealth in terms of outcomes of health-related quality of life for people with diabetes

(Polisena, Tran, et al. 2009) and respiratory conditions (Pare, Janna, and Sicotte 2007, Polisena, Coyle, et al. 2009). Evidence has also favoured telemonitoring for people with coronary heart failure (Inglis et al. 2010), not least because of the diversity of generic and condition-specific measures reported.

It is also important to consider the country context when comparing these results with previous studies, many of which were US-based. That health care is free at the point of use in the UK may mean that participants had better access to appropriate primary care services than a comparable population of users in the US; thus, there is less potential to reduce the use of the more expensive services in secondary care here. It should be noted that in this evaluation, there was a non-significant reduction in secondary care costs in the telehealth group. Another way in which the population might have had less room to show improvement was in terms of the level of need, or severity, of the index condition.

One question arising from these results would be that the timeframe of the evaluation may have been too short to show improvements in health-related quality of life, a potential weakness shared with many published economic evaluations of telemedicine (Mistry 2012). By the same token there is no evidence base to show that a longer time horizon leads to improved outcomes.

This study raises some questions for further research. The extent to which telehealth should be targeted towards specific patient populations and subpopulations should be further investigated in future studies. Also, specific models of TH delivery should be investigated to understand their relationship with variations in outcomes and costs (McLean et al. 2013). It would be helpful to understand variations in frequency and intensity of response to "breaches" of vital signs protocols (data that were not collected within the telehealth trial) and their relationship with quality of life outcomes.

7.5.2 Comparison with Other Studies

Few telehealth evaluations have examined the association between outcomes and costs (Bensink, Hailey, and Wootton 2006, Whitten et al. 2002). Recent reviews have found telehealth to be cost saving; however, the quality of the evaluations reviewed has generally been described as poor (Bergmo 2009, Polisena, Coyle, et al. 2009, Vergara Rojas and Gagnon 2008). Some reviews have found telehealth to decrease use of acute hospital services (Inglis et al. 2010, Polisena, Tran, et al. 2009, Polisena et al. 2010a), but there is less evidence in terms of use of primary care (Polisena, Tran, et al. 2009). In this analysis there

was a pattern of reduced use of health and social care services by the telehealth group, if intervention specific costs were excluded, although the differences were small.

Information on the costs of providing telehealth in the form of telemonitoring has been scarce. Direct intervention costs of telehealth (whether by telephone support or telemonitoring) reported in the literature range widely, and come from a variety of health systems and countries. Inglis and colleagues (Inglis et al. 2010) identified a small number of studies of telemonitoring for heart failure that gave such details. One (Balk et al. 2007) noted that the costs of telemonitoring increased the total costs for the intervention group, but did not give the actual intervention cost; another (Giordano et al. 2009) provided a mean annual cost per patient for telemonitoring of €185 (Inglis et al. 2010). Barlow and colleagues (Barlow et al. 2007) provided UK-based estimates of telehealth equipment costs of about £700-900 and monitoring costs of £260-520 per year (2007 prices). The estimated annual costs of telehealth monitoring, support, and equipment in our study varied between sites (about £1500-2000), reflecting the heterogeneity in models of telehealth delivery.

Because there are no societal thresholds for ICERs involving ICECAP-O, Brief STAI or CESD-10, we can only interpret any positive findings related to these instruments with caution. ICECAP-O is a relatively new instrument and little empirical information currently exists on the average values expected in a population with long-term conditions, or on its use in economic evaluations (Davis et al. 2012, Petrou and Gray 2011).

7.5.3 Implications for Clinicians and Policymakers

These results suggest that the QALY gain by people using telehealth in addition to standard support and treatment was similar to those receiving usual care, and that total costs for the telehealth group were higher than for the usual care group. The probability of cost-effectiveness judged by reference to this QALY measure was relatively low over a range of values of willingness to pay. Total costs were sensitive to the costs of the intervention, reducing the point estimate of the cost per QALY substantially such that it became negative (assuming that returns to scale could be achieved without altering outcomes). However, because the difference in total costs between treatment groups was not significant even with these assumed reductions, the probability of cost-effectiveness was only about 76 per cent at the £30,000 threshold of willingness to pay, used as a reference by NICE. These results take into account costs to both health and social care systems, to give a picture of the consequences to costs and quality of life from investment in telehealth across the health and

social care agencies. If investment in telehealth falls mainly to primary and social care purchasers, while most savings accrue to the acute sector - for which there is some weak evidence here - then reinvestment into community health and social care services would be vital.

Chapter 8

Cost-effectiveness of Telecare

Literature on the outcomes of telecare (as distinct from telehealth and telemedicine) is scarce (Barlow, Bayer, and Curry 2006) and cost-effectiveness studies nearly non-existent (Graybill, McMeekin, and Wildman 2014). In this chapter, I present the results of a cost-effectiveness analysis of the WSD telecare intervention, drawing on the WSD Telecare questionnaire study dataset. I first review the analytic methods employed, then review the results, and end with a discussion of the implications, strengths and limitations of the analysis, and future directions for research.

8.1 Methods

In this section I briefly summarise the analytic methods described in Chapter 4, Section 4.20. The primary outcome of interest in the cost-effectiveness analysis was total cost to health and social care to gain a QALY by implementing the telecare intervention. I also explored a number of other outcomes (state anxiety, SF-12 components, ICECAP and QALY derived from SF-6D utilities). Multiple imputation of missing variables was carried out (Chapter 4, Section 4.22). Seemingly unrelated regressions (SUR) were then applied, adjusting standard errors for cluster (general practice). Outcome equations adjusted for baseline utility (or baseline outcome in the case of other measures), site, demographic and individual characteristics (age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household); cost equations adjusted for baseline costs, baseline outcome, site, demographic and individual characteristics (age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household). This modelling approach allowed the estimation of the impact of the intervention while controlling for clustering and accounting for the correlation between the cost and outcome variables. Sensitivity analyses explored the robustness of results to (i) variations in the costs of the intervention (lower input prices for equipment and telecare monitoring support), and (ii) to the non-normality of the data. In the latter sensitivity analysis, the data were two-stage bootstrapped in R and the SUR model fitted; results from each imputation were combined in NORM (Schafer 1999) . Lastly, an interaction term for

allocation and living-arrangement subgroup (living alone or living with others) was included in the SUR models, in order to explore the cost-effectiveness of telecare within these subgroups.

8.2 Results

At 12-month follow-up, outcomes data were available for 379 telecare and 384 control participants (69 per cent vs. 60 per cent of the baseline sample respectively). Costs data were available at 12-month follow-up for 381 telecare and 376 control participants (69 per cent vs. 59 per cent of the baseline sample, respectively). Cost and outcome data across baseline and 12-month follow-up were available for 375 intervention and 378 control participants.

8.2.1 Costs

All costs considered in the main cost-effectiveness analyses and in the sensitivity analyses are summarised in Table 8.1, in terms of mean costs over the three months prior to baseline and 12-month follow-up, with cluster-adjusted standard errors. The annual-equivalent costs for the pre-baseline period were rather higher in the telecare group (£1499, 95% CI -£563, £3561). Total annual-equivalent costs over the intervention period were rather higher in the intervention than in the control group but not significantly different between groups on cluster-adjusted t-tests, whether including or excluding intervention costs, and under alternative assumptions about the intervention costs (excluding project-management posts and contracts or dedicated responder service costs from the intervention costs).

8.2.2 Outcomes

Raw mean outcome scores at both time points are summarised in similar fashion in Table 8.2. Outcomes did not differ greatly between groups at baseline with the exception of MCS-12 scores, where the telecare group had lower scores at baseline than controls. Differences between groups at the 12-month follow-up were small and not significantly different. The direction of effect was in favour of the intervention group in the STAI and the PCS-12 measures, and in favour of the control group in the EQ-5D-3L-generated utility and -derived QALY, ICECAP-O, MCS-12 and SF-6D-derived QALY measures (there was no difference between groups in SF-6D-generated utility scores). Utility scores derived from the SF-6D were substantially higher than those derived from the EQ-5D-3L at both baseline and follow-

up; the standard errors of the EQ-5D-3L mean utilities were three times larger than those of the SF-6D mean utilities.

Table 8.1 Mean service costs (£) across Telecare sample, annual equivalent

Resource item	Usual care (SE) (n=378)	Telecare (SE) (n=375)	Raw difference (95% CI)
Pre-baseline period			
Total costs	8152 (742)	9651 (736)	1499 (-563, 3561)
Intervention period			
Total costs excluding telecare delivery and equipment	7232 (667)	8088 (664)	856 (-1001, 2713)
Telecare equipment costs	4 (2)	82 (2)	78 (72, 84)**
Telecare intervention costs	35 (26)	710 (26)	676 (603, 748)**
Total costs including telecare delivery and equipment	7271 (676)	8880 (672)	1610 (-270, 3490)
-less project management posts & contracts	7266 (675)	8778 (671)	1512 (-366, 3389)
-less dedicated responder costs	7266 (675)	8809 (672)	1543 (-336, 3422)
Sensitivity analyses			
-at 50% reduction in equipment prices	7269 (676)	8839 (672)	1571 (-309, 3450)
-£5 cost per week	7248 (668)	8421 (664)	1173 (-685, 3032)
-£5 cost per week + 50% reduction in equipment prices	7246 (668)	8380 (664)	1134 (-723, 2992)

Note: Table reports the annual equivalent costs for 753 cases with baseline cost data available, (10 complete datasets). Standard errors are cluster-adjusted.
*p<0.01 on t-test
**p<0.05 on t-test

Table 8.2 Outcomes at baseline and 12-month follow-up, Telecare sample

Resource item	Usual care (SE) (n=378)	Telecare (SE) (n=375)	Difference (95% CI)
Baseline			
Utility (EQ-5D-3L)	0.338 (0.022)	0.317 (0.022)	-0.021 (-0.083, 0.041)
MCS 12	43.754 (0.75)	41.613 (0.746)	-2.141 (-4.227, -0.055)*
PCS 12	30.257 (0.537)	30.864 (0.535)	0.607 (-0.888, 2.101)
STAI	10.758 (0.274)	11.091 (0.272)	0.332 (-0.429, 1.094)
ICECAP-O	0.686 (0.012)	0.674 (0.012)	-0.012 (-0.045, 0.021)
Utility (SF-6D)	0.568 (0.007)	0.558 (0.007)	-0.011 (-0.03, 0.008)
12-month follow-up			
Utility (EQ-5D-3L)	0.333 (0.02)	0.321 (0.02)	-0.013 (-0.068, 0.042)
MCS 12	42.279 (0.658)	42.3 (0.658)	0.02 (-1.816, 1.856)
PCS 12	28.691 (0.546)	29.08 (0.543)	0.388 (-1.131, 1.907)
STAI	12.344 (0.256)	12.138 (0.256)	-0.206 (-0.919, 0.507)
ICECAP-O	0.649 (0.011)	0.638 (0.011)	-0.011 (-0.042, 0.02)
QALY - EQ-5D-3L	0.336 (0.019)	0.319 (0.019)	-0.017 (-0.07, 0.036)
Utility (SF-6D)	0.551 (0.007)	0.551 (0.007)	0 (-0.018, 0.019)
QALY - SF-6D	0.559 (0.006)	0.554 (0.006)	-0.005 (-0.023, 0.012)

Note: Table reports results for 753 cases with baseline cost and outcome data available, (10 complete datasets). Standard errors are cluster-adjusted.
*p<0.01 on t-test **p<0.05 on t-test

8.3 Cost-effectiveness Results

Examining the results of the SUR model, QALYs were very slightly higher in the telecare group (by 0.003 (95% CI -0.018, 0.024)) (Table 8.3). Adjusted annual costs of telecare participants were £1,000 (CIs -535, 2536) higher than those of control participants. The ICER was £368,000, with undefined confidence intervals. Considering costs excluding project management and excluding dedicated telecare responder costs, the ICERs were slightly lower (£332,000 and £343,000 respectively). The probability of cost-effectiveness at the higher end of the NICE willingness-to-pay threshold range, £30,000, was 16 per cent (Figure 8.1). Examining higher willingness-to-pay thresholds made little difference to the probability that telecare could be found to be cost-effective, reaching just under 30 per cent at a threshold of £90,000 per QALY. Varying the costs of the intervention by excluding the costs of (a) project management-specific posts and contracts or (b) dedicated response services in sites 2 and 3 produced relatively similar results, with probabilities of cost-effectiveness respectively 3 per cent and 2 per cent higher at a willingness-to-pay of £30,000 per QALY than that produced by the main estimate.

Table 8.3 Differences in cost and effect between Telecare and usual care groups (12 months), annual equivalent

Values in means (CI) unless otherwise stated	Outcomes / total costs Control=378	Outcomes / total costs Telecare=375	Difference in outcomes / total costs (Usual care=378; Telecare=375)
QALY (adjusted mean, SUR model†)	0.326 (0.312 , 0.340)	0.329 (0.313, 0.344)	0.003 (-0.018 , 0.024)
Cost (adjusted mean, SUR model‡)	7574 (6 535, 8612)	8574 (7 490, 9 658)	1000 (-535, 2536)
ICER (£ per QALY) (SUR model§)	-	-	368 000 (undefined, undefined)
Excluding project management costs			
Cost (adjusted mean, SUR model‡)	7568 (6 530, 8 607)	8472 (7 388, 9 557)	904 (-632, 2 440)
ICER (£ per QALY) (SUR model§)	-	-	332 000 (undefined, undefined)
Excluding dedicated response costs			
Cost (adjusted mean, SUR model‡)	7570 (6 531,8609)	8503 (7 419, 9 588)	933 (-603, 2470)
ICER (£ per QALY) (SUR model§)	-	-	343 00 (undefined, undefined)
Secondary outcomes analyses			
ICECAP-O (adjusted mean, SUR model†)	0.644 (0.629, 0.660)	0.642 (0.625, 0.659)	-0.002 (-0.026, 0.021)
STAI (adjusted mean, SUR model†)	12.426 (12.822 , 12.012)	12.066 (12.444 , 11.67)	-0.36 (0.198 , -0.918) ||
STAI ICER (£) (SUR model‡§)			50000 (undefined , undefined)
MCS-12 (adjusted mean, SUR model†)	41.796 (40.78 , 42.81)	42.788 (41.81 , 43.76)	0.992 (-0.42, 2.4)
MCS-12 ICER (£) (SUR model§)			4 000 (undefined , undefined)
PCS-12 (adjusted mean, SUR model†)	28.731 (28.02 , 29.44)	29.04 (28.28 , 29.81)	0.312 (-0.741 , 1.362)
PCS-12 ICER (£) (SUR model§)			10 000 (undefined, undefined)
QALY SF-6D (adjusted mean, SUR model†)	0.555 (0.551,0.559)	0.559 (0.554, 0.564)	0.004 (-0.002, 0.011)
QALY SF-6D ICER (£) (SUR model§)			240 000

Note: Table reports results for 753 cases with baseline cost and outcome data available (10 complete datasets). Standard errors are robust cluster-adjusted.
†from SUR analyses (outcome equation), adjusted for baseline utility, site, age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household
‡ from SUR analyses (cost equation), adjusted for baseline costs, baseline outcome, site, age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household
§ rounded to nearest thousand
|| re-transformed to original scale to enable comparison with raw mean difference; transformed mean= 0.020 (-0.011, 0.051)

Table 8.4 Sensitivity analyses: differences in cost and effect between Telecare and usual care groups (12 months), annual equivalent

Values in means (CI) unless otherwise stated	Outcomes / total costs Control=378	Outcomes / total costs Telecare=375	Difference in total costs (Usual care=378; Telecare=375)
Variations in intervention costs:			
Equipment prices reduced by 50%			
Cost (adjusted mean, SUR model‡)	7572 (6533,8611)	8534 (7450,9618)	962 (-574, 2497)
ICER (£ per QALY) (SUR model§)	-	-	353 000 (undefined, undefined)
Mainstream support package of £5 per week			
Cost (adjusted mean, SUR model‡)	7545 (6504 , 8586)	8121 (7032 , 9211)	576 (-964, 2117)
ICER (£ per QALY) (SUR model†‡)	-	-	212 000 (undefined, undefined)
Equipment prices reduced by 50% & mainstream support package of £5 per week			
Cost (adjusted mean, SUR model†)	7543 (6 502 , 8584)	8081 (6 993 , 9170)	538 (-1003, 2079)
ICER (£ per QALY) (SUR model§)	-	-	197 000 (undefined, undefined)
Two-stage bootstrapped estimates+ SUR:			
QALY (unadjusted group means‖, adjusted mean difference, SUR model‡¶)	0.336 (0.295, 0.377)	0.319 (0.278, 0.360)	-0.001 (-0.038, 0.036)
Cost (unadjusted group means, adjusted mean difference‖, SUR model‡¶)	7 285 (5965, 8605)	8 859 (7444, 10274)	911 (-942 , 2764)

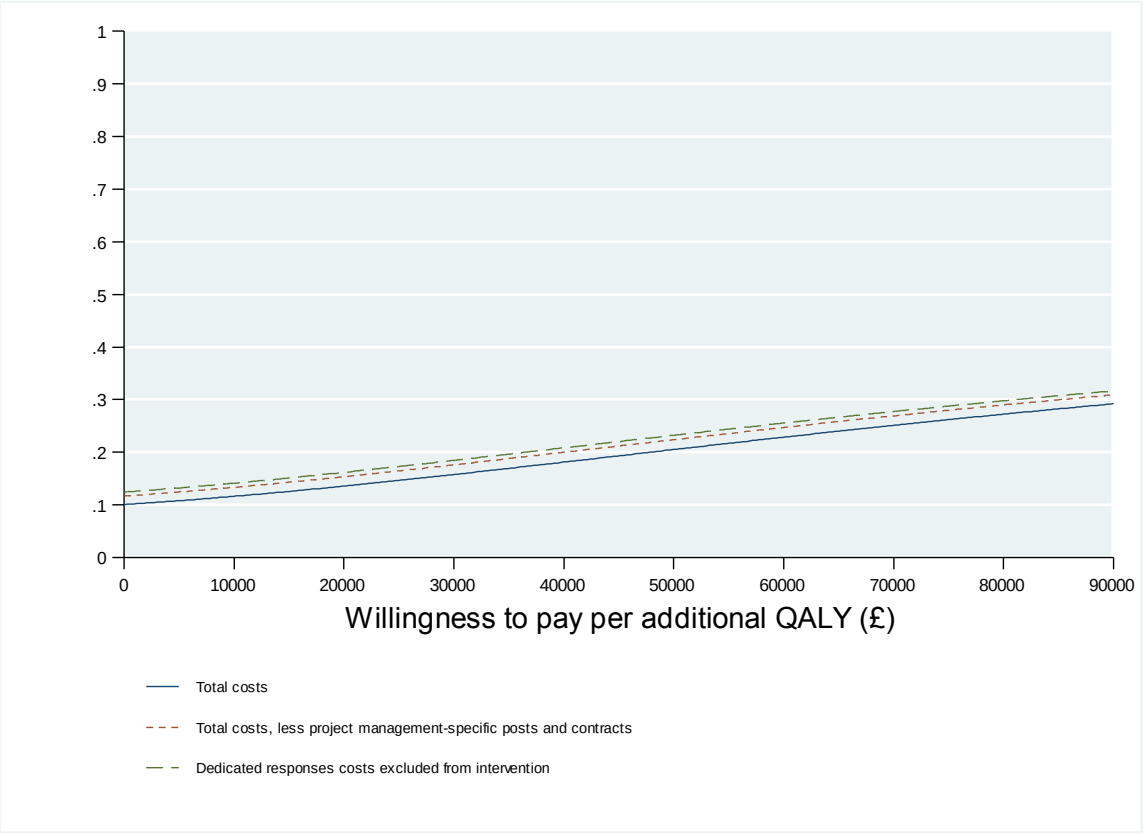
Note: Table reports results for 753 cases with baseline cost and outcome data available, (10 complete datasets). Standard errors are robust cluster-adjusted.
†from SUR analyses (outcome equation), adjusted for baseline utility, site, age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household
‡ from SUR analyses (cost equation), adjusted for baseline costs, baseline outcome, site, age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household
§ rounded to nearest thousand
‖ from two-stage bootstrapped estimates of group means, 3000 replications
¶ two-stage bootstrapped estimates of SUR coefficient on allocation, 3000 replications

Table 8.5 Subgroup analyses: Differences in cost and effect between Telecare and usual care groups (12 months), participants living together or alone, annual equivalent

Values in means (CI) unless otherwise stated	Outcomes / total costs Control=378	Outcomes / total costs Telecare=375	Difference in outcomes / total costs (Usual care=378; Telecare=375)
Living with others			
QALY (adjusted mean, SUR model†)	0.312 (0.292 , 0.332)	0.315 (0.295 , 0.335)	0.003 (-0.025 , 0.031)
Cost (adjusted mean, SUR model‡)	8582 (6 879 , 10 284)	8309 (6 801 , 9818)	-272 (-2 716 , 2 171)
ICER (£ per QALY) (SUR model§)	-	-	-91 000 (undefined, undefined)
excluding project management costs			
Cost (adjusted mean, SUR model‡)	8579 (6 876 , 10 282)	8 206 (6697 , 9715)	-373 (-2818 , 2071)
ICER (£ per QALY) (SUR model§)	-	-	-125 000 (undefined, undefined)
excluding dedicated response costs			
Cost (adjusted mean, SUR model‡)	8579 (6 876 , 10281)	8237 (6728 , 9745)	-342 (-2786 , 2102)
ICER (£ per QALY) (SUR model§)	-	-	-114 000 (undefined, undefined)
Living alone			
QALY (adjusted mean, SUR model†)	0.339 (0.318 , 0.36)	0.341 (0.319 , 0.363)	0.002 (-0.027 , 0.032)
Cost (adjusted mean, SUR model‡)	6 674 (5 531 ,7 816)	8 862 (7 580 , 10 144)	2 188 (568, 3 809)
ICER (£ per QALY) (SUR model§)	-	-	884 000 (33 000 , -77 000)
Excluding project management costs			
Cost (adjusted mean, SUR model‡)	6666 (5 524 , 7807)	8761 (7 479 , 10 044)	2 096 (476 , 3716)
ICER (£ per QALY) (SUR model§)	-	-	847 000 (29 000, -72 000)
Excluding dedicated response costs			
Cost (adjusted mean, SUR model‡)	6668 (5 526 , 7811)	8792 (7 509 , 10 075)	2124 (503 , 3 744)
ICER (£ per QALY) (SUR model§)	-	-	858 000 (30 000, -73 000)

Note: Table reports results for 753 cases with baseline cost and outcome data available (10 complete datasets). Standard errors are robust cluster-adjusted.
†from SUR analyses (outcome equation), adjusted for baseline utility, site, age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household
‡ from SUR analyses (cost equation), adjusted for baseline costs, baseline outcome, site, age, sex, ethnicity, IMD, number of chronic conditions, EQ5-D self-care score, previous community alarm, one-person household
§ rounded to nearest thousand

Figure 8.1 Cost-effectiveness acceptability curve: QALY



8.3.1 Secondary Outcomes

There was a small difference in mean adjusted ICECAP-O scores in favour of the control group (-0.002) and small differences in Brief STAI, MCS-12, PCS-12 and SF-6D in favour of the telecare group. The ICER for a movement from worst to best on the STAI scale was £50,000, for a 3-point increase in the PCS-12 was £10,000 and for a 4-point increase in the MCS-12 was £4,000[20]. The difference between groups on ICECAP-O favoured the control group. In terms of this outcome, telecare was dominated by the usual care alternative, being both (marginally) less effective and more expensive (Table 8.3). The cost per QALY gained (where QALYs were derived from SF-6D utilities) was £240,000. The probability of achieving a reduction from maximum to lowest level of state anxiety, as measured by the STAI, at levels of WTP of £10,000 to £20,000, ranged between 16 per cent and 24 per cent (Figure 8.2). The probability of achieving a 3-point increase in PCS-12 ranged between 51 per cent and 61 per cent (Figure 8.3), and the probability of achieving a 4-point increase in MCS-12, between 77 per cent and 86 per cent, at the same WTP levels (Figure 8.4). While

[20] The choice of effect size is explained in Chapter 4, Section 4.8.1.

the ICER produced by the SF-6D QALY was lower, the probability of cost-effectiveness was very similar to that produced by the EQ-5D-3L at the £30,000 range (13 per cent); the shape of the CEAC produced using the two instruments were similar at lower values but using QALY derived from the SF-6D (Figure 8.5) produced a somewhat flatter curve at higher WTP values.

Figure 8.2 Cost-effectiveness acceptability curve: Brief STAI

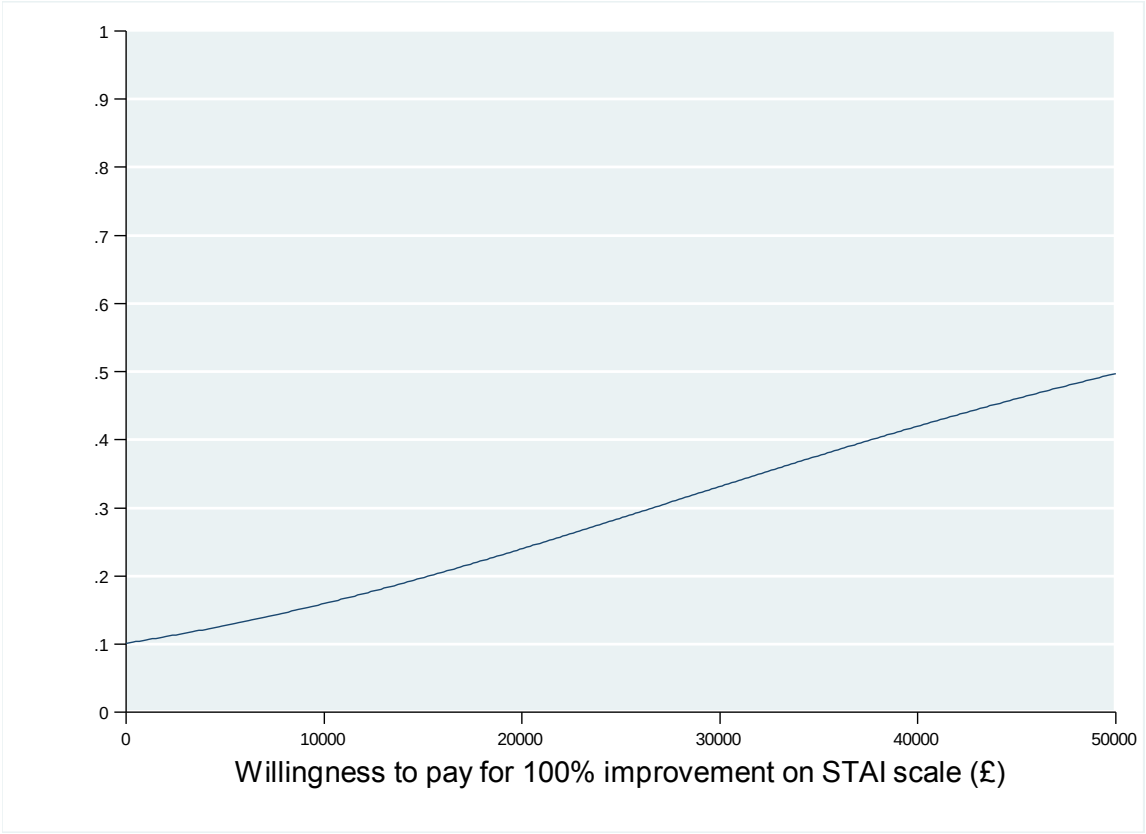


Figure 8.3 Cost-effectiveness acceptability curve: PCS-12

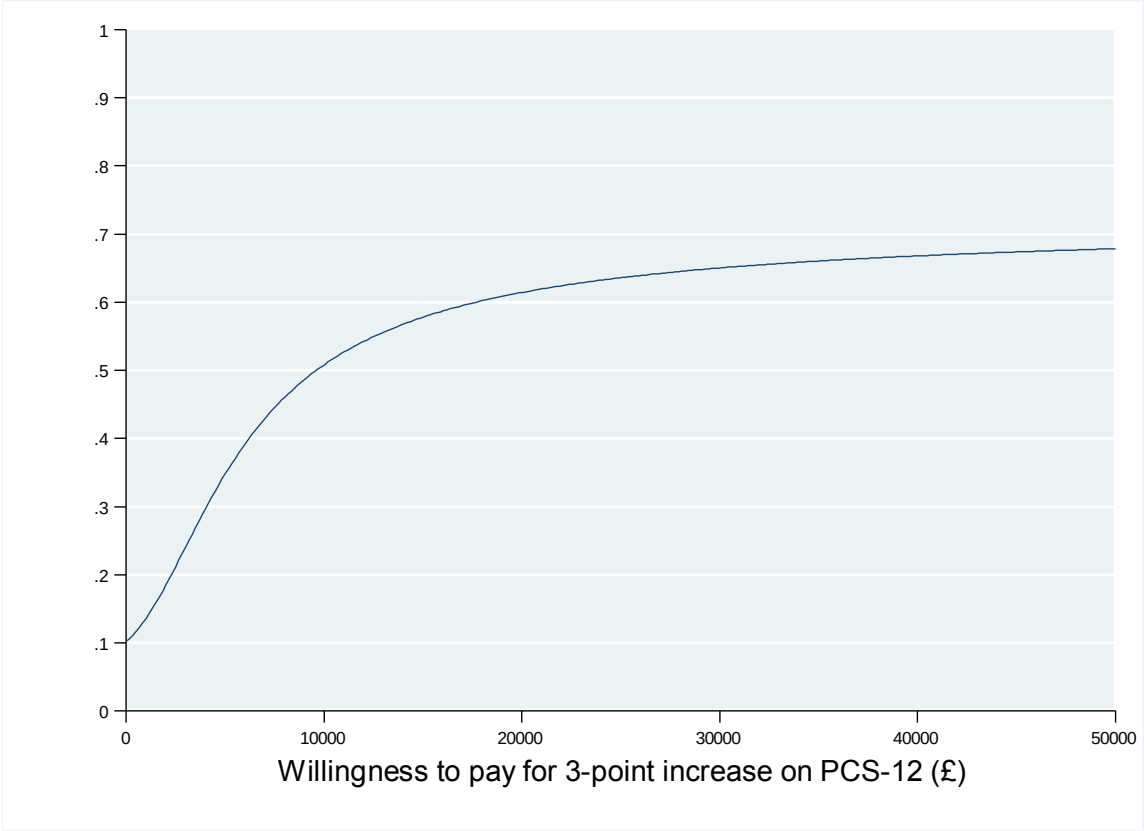


Figure 8.4 Cost-effectiveness acceptability curve: MCS-12

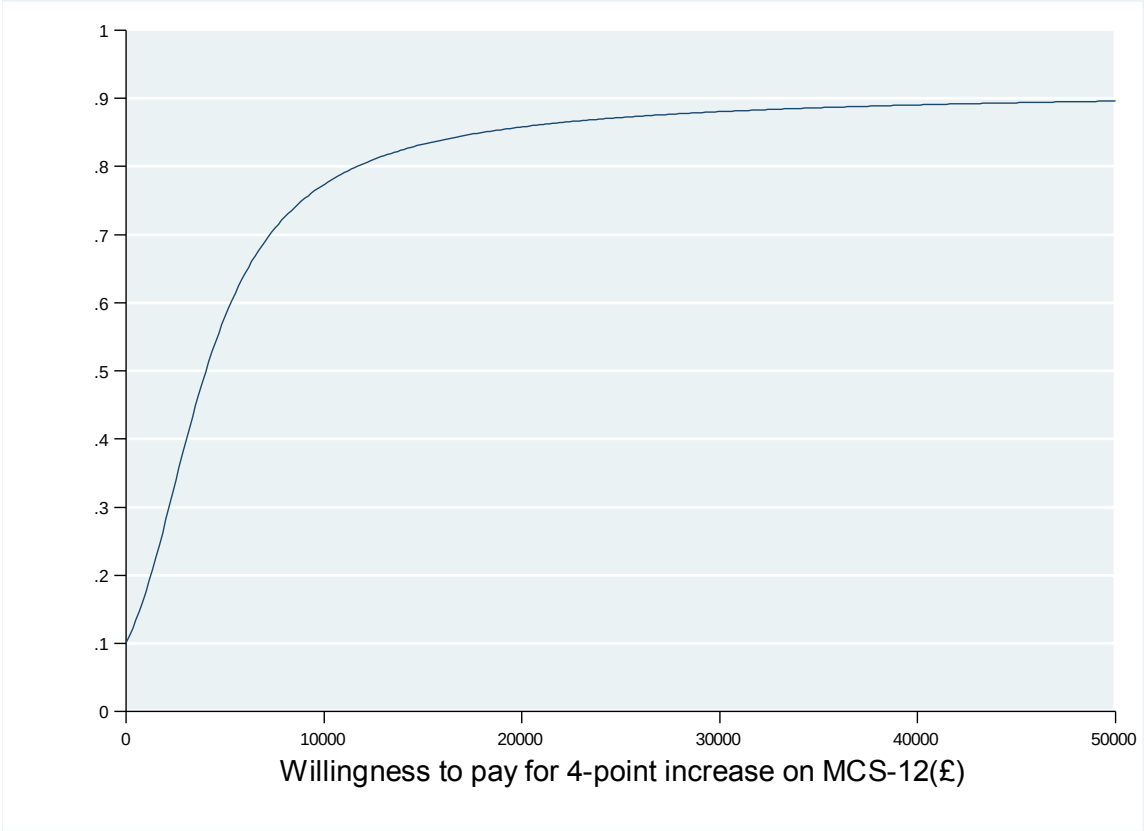
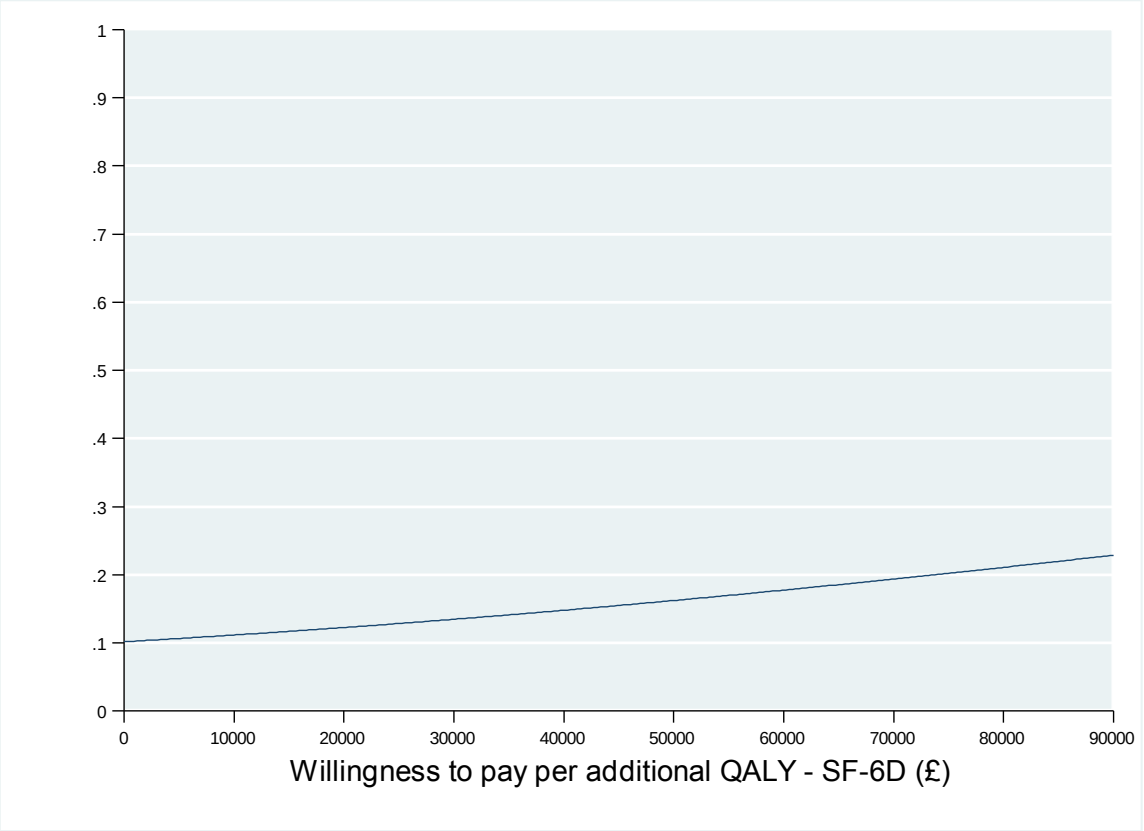


Figure 8.5 Cost-effectiveness acceptability curve: QALY (SF-6D)



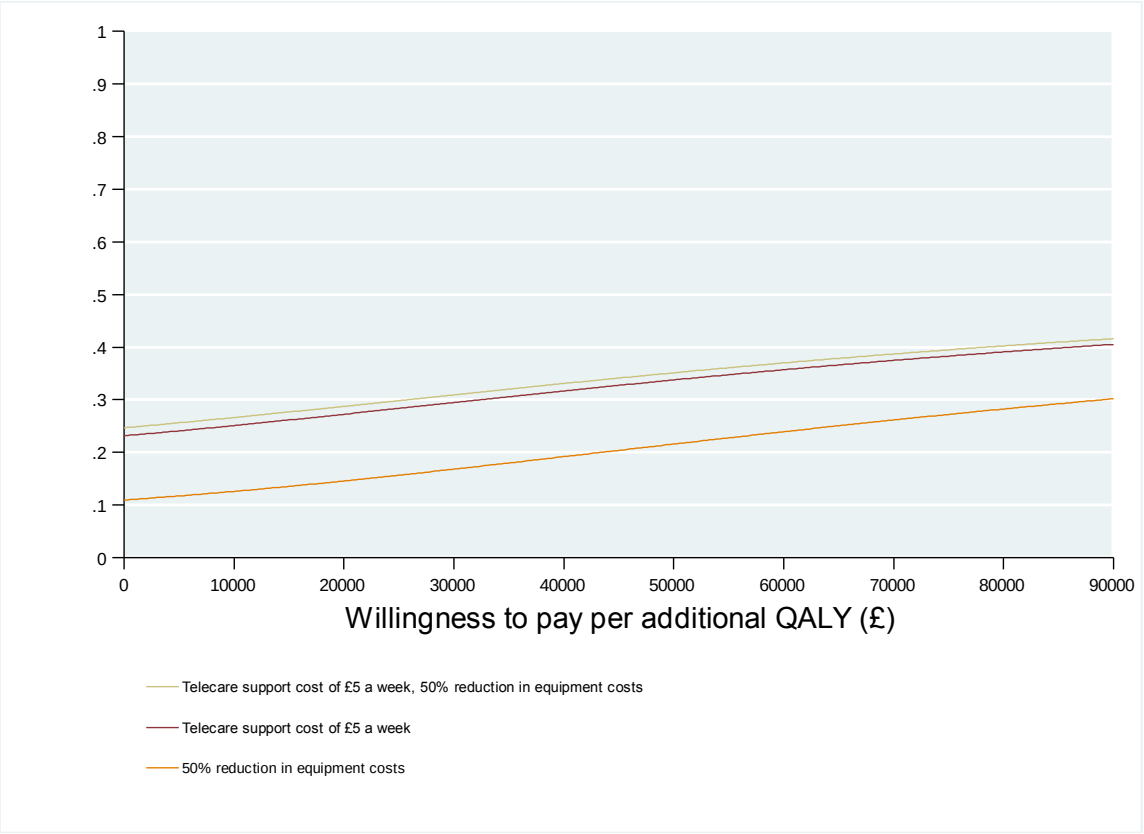
8.3.2 Sensitivity Analyses

The first set of sensitivity analyses explored the robustness of results to different assumptions about the costs of the intervention. One possibility explored was that the cost of telecare support could be lower than estimated. A 'mainstream' cost of £5 per week was taken from a published report on telecare costs in Welsh councils by Bayer and Barlow (2010). Using this estimate for the cost of telecare support, the probability of telecare being cost-effective was 30 per cent at the £30,000 WTP threshold (Figure 8.6). The cost per QALY gained was £212,000 (Table 8.4). Assumptions about the input prices for telecare equipment were also tested by substituting the cost of equipment if purchased at half the price paid in the trial. The probability of cost-effectiveness was 17 per cent assuming the same threshold WTP. The two assumptions were combined in a further scenario, which increased the probability that telecare was cost-effective to 31 per cent, again assuming a WTP of £30,000, and the ICER was £197,000.

In addition, a two-stage bootstrap procedure was performed, running the SUR on 3000 bootstrapped replications. The QALY difference was very small and negative (-0.001) (suggesting that telecare was slightly less effective than usual care) and the cost difference a

little lower than the main estimate (£911); neither the cost nor the outcome difference was significant at the 5 per cent level (p=0.954 and p=0.335 respectively).

Figure 8.6 Cost-effectiveness acceptability curve: QALY, sensitivity to input prices



8.3.3 Subgroup Analysis

In Chapter 6, I explored differences in the costs of telecare between living-arrangement subgroups (participants living alone and with others). Here I examine the cost-effectiveness of telecare within these subgroups. People living alone had very different costs and outcomes from those living with others. Mean QALY was 0.027 (95% CI 0.005, 0.049) higher for those living alone than for those living with others. As discussed in Chapter 6, the costs of those living with others in receipt of telecare had lower costs than in controls, whereas participants living alone in receipt of telecare had slightly higher costs. The results of the SUR model interacting the allocation and living arrangement variables (Table 8.5) indicate that there was a small difference in costs at follow-up between the experimental groups in participants living with others (with wide 95 per cent confidence intervals, crossing zero). There was a large and statistically significant difference in costs at follow-up between experimental groups in participants living alone (about £2180 higher in the intervention than in the control group). The intervention participants, whether living alone or with others, had very slightly

higher mean QALY than controls (0.003 and 0.002, respectively). The results were similar if excluding costs of project management or of dedicated response services. The cost per QALY was very large in the group living alone (£884,000) and smaller and negative (-91,000) in the group living with others. The probability of cost-effectiveness at the £30,000 NICE threshold was much lower in the living alone subgroup (2 per cent) than in the living with others group (60 per cent) (Figure 8.7). A 95 per cent cost-effectiveness ellipse plot drawing on the SUR results estimates (Figure 8.8) is useful in demonstrating that the costs and QALY for each subgroup overlap substantially. The difference in costs of those living alone are higher than those living with others so that all the points on the ellipse lie above the x-axis; the difference in costs of those living with others straddle the x-axis. The intervention is more costly than usual care; it is possible but not at all certain that the intervention is more effective. From the confidence intervals of the ICER for those living alone it can be seen that the intervention is less cost effective than usual care below £33 000 and that above that WTP, there is no certainty that telecare is more cost-effective than usual care (or the other way around). For people living with others, at a WTP of £30,000, while the probability of cost-effectiveness is 60 per cent, it is evident that there is sufficient uncertainty around the point estimate that no 95 per cent confidence intervals can be constructed.

Figure 8.7 Cost-effectiveness acceptability curve: QALY, subgroup analyses

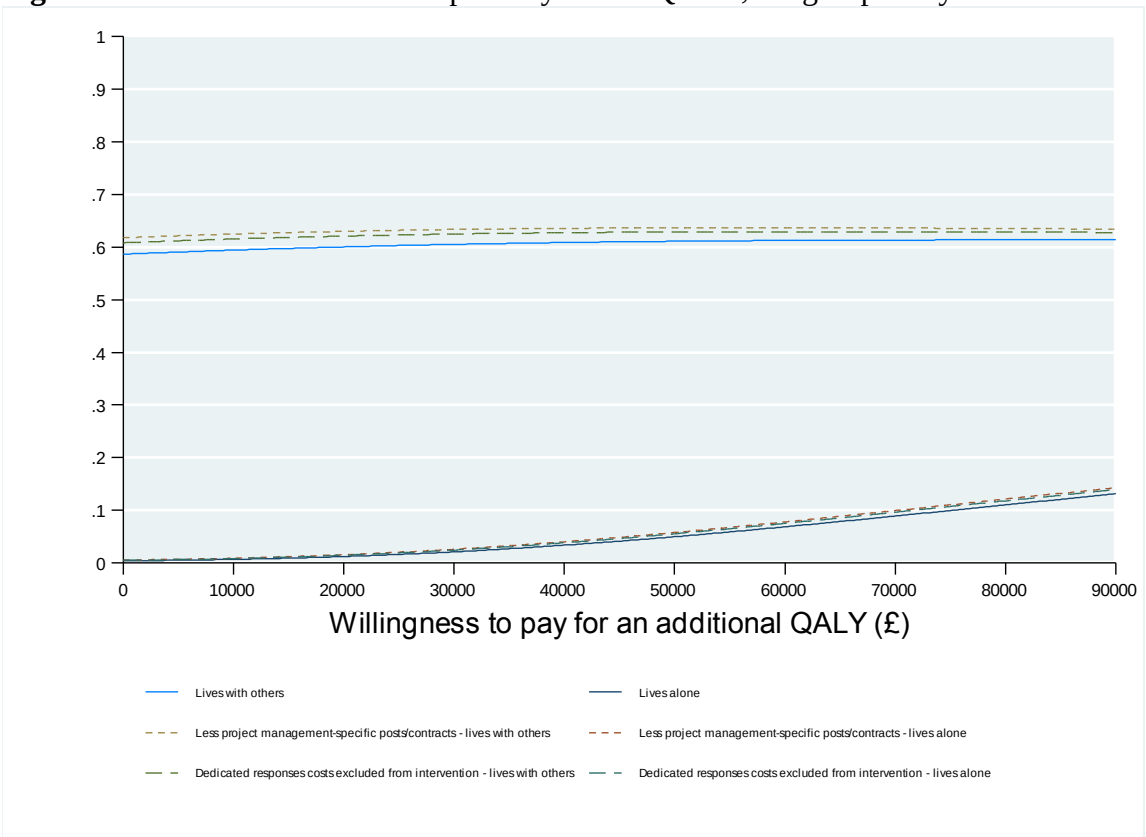
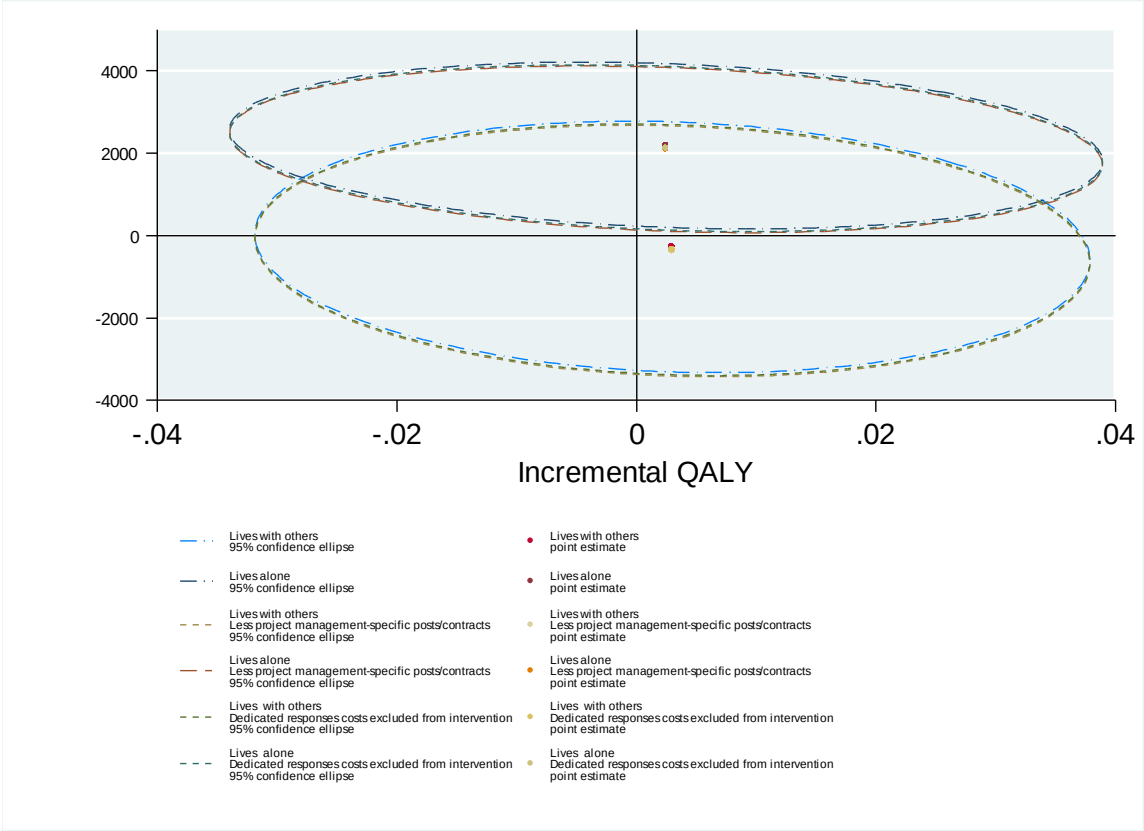


Figure 8.8 Cost-effectiveness ellipse: QALY, subgroup analyses



8.4 Discussion

The WSD telecare study involved a much larger number of participants than any previous study of its kind. The randomised-controlled trial design enabled a formal evaluation of the impact of telecare on costs and outcomes in a population with social care need. In these ways, the research reported here has been important in contributing to the very limited international evidence base on the cost-effectiveness of telecare.

The size and RCT design of the study constitute a major strength of this analysis. The use of cost-per-QALY as an outcome enables the technical efficiency of the WSD telecare intervention to be compared to future telecare interventions, and to that of other health and social care interventions. However a number of limitations must be acknowledged.

The numbers of participants who decided to take part in the questionnaire study were not entirely balanced, so that 48 per cent of participants in the telecare arm opted in, while a smaller proportion, 43 per cent of participants in the usual care arm opted in (Hirani et al. 2013). Also, there were more controls at baseline than intervention participants (634 vs. 548 respectively, or 16 per cent more) and the possibility that there was some self-selection after cluster-randomisation cannot be discounted. Such issues are well-known in cluster-

randomised trials (Puffer, Torgerson, and Watson 2003). There was substantial loss to follow-up at 12 months, of 40 per cent in the control group, and 32 per cent in the intervention group. That the attrition rates differed between the allocation groups opens up the possibility of bias, for instance in retaining a larger portion of the intervention group with more favourable outcomes relative to controls. On the other hand, baseline characteristics did not differ substantially within allocation groups for the cases with data available at baseline and 12-month follow-up. The analyses did adjust for confounders that could have influenced attrition, compensating to some extent for imbalances between the groups at follow-up. This does not rule out the possibility that other, unmeasured characteristics could have differed between groups at baseline or follow-up, or between those who did and did not complete the study.

The analysis did not take into account costs that might have arisen after participants dropped out from the trial because they were admitted to residential or nursing care, and did not treat dropout due to death differently from other sources of loss to 12-month follow-up. Also, I estimated the costs of health and social care over the study period by multiplying the 3-month costs prior to the 12-month follow-up by four. This relied on the assumption that costs were relatively constant over the year across the categories of service use. It is useful here to refer to another stream of work within the WSD research programme, which examined longitudinal administrative data on mortality, along with long-term residential or nursing care admissions and a (restricted) range of health and social care costs of the trial population over the 12-month study period (Steventon et al. 2013). In that research the mortality rates were similar (being 8.9 per cent in the control group and 8.7 per cent in the intervention group). Proportions admitted permanently to care homes were also similar in both groups (3.2 per cent usual care vs. 3.1 per cent telecare).

In the trial, given that the design was pragmatic, there was no prescription or standardisation of the processes for assessing potential participants' need for telecare. This flexibility had implications for external validity and reproducibility. The way that telecare services in other parts of the country assess potential clients' need for telecare could differ from the assessment practices in the trial sites. The scale of the trial precluded the collection and detailed examination of (largely qualitative) assessment documentation, to understand what variations might have existed in assessment practices in the trial sites. It should also be said that if a range of assessment models existed within the sites, then some of those models are likely to have existed in other places. A related point is that the variety of telecare equipment combinations and functions and of potential responses to sensor activations may

have increased the complexity of the relationships between trial inputs, outputs and outcomes (Byford and Sefton 2003). If different types and intensities of package have distinctly different impacts on outcomes and costs, the overall outcome-cost relationship may be difficult to interpret without in some way delineating the purpose of the package.

Lastly, data were not available on the number of sensor alerts or false alarms associated with individual participants or on participant-specific call centre responses to sensor alerts. In addition, there were no data on numbers of visits made by dedicated response teams to individual participants. The need to cost the telecare support element of the intervention at a site level narrowed the amount of potential variability in intervention costs between participants; this may have diminished the sensitivity of the analysis to detecting the impact of the intervention.

8.4.1 How Does Telecare Create Benefits for Service Users? Were These Adequately Measured? And Who Benefits?

The intervention appears to have had a minimal impact on a range of quality of life and psychosocial outcomes. A question arises: by what means would telecare systems act to improve outcomes? The assumptions underlying the expected impact need to be unpacked. This is a difficult task given that delivery systems of telecare support are variable; also telecare technology is ever-evolving. The scope of the technology evaluated here encompassed remote, automatic and passive monitoring systems that went beyond older and more basic forms of telecare largely focused on summoning assistance (which in this trial could also form an element of standard support and care). One reason given for telecare systems improving quality of life for their users is that it creates a sense of reassurance (Beale et al. 2010, Hirani et al. 2013, Roush and Teasdale 2011). How much *more* reassurance is provided by additional sensors, over and above that provided by the ability to summon help, is not easy to quantify.

A further issue is that the primary outcome measure was derived from the EQ-5D-3L, yet EQ-5D-3L may not be able to capture entirely the improvements brought about by telecare. The EQ-5D-3L was chosen because it is a generic measure of health-related quality of life that can be used as a basis for comparing alternative technologies (National Institute for Health and Clinical Excellence 2008); also, it is suitable for use with older people (cf. Haywood, Garratt, and Fitzpatrick 2005, Hawton et al. 2011). The EQ-5D-3L dimensions of health (self-care, anxiety/depression, usual activities, pain/discomfort and mobility) are

certainly relevant to the sorts of benefits that are expected of telecare. On the other hand, the EQ-5D-3L focuses on health and restoration of function rather than achievement of benefit through the more compensatory mechanisms provided by much of social care (Forder and Caiels 2011), including telecare. Thus it is possible that while EQ-5D-3L is sensitive to change in situations where changes in health are expected to be substantial (Haywood, Garratt, and Fitzpatrick 2005), this may not be the case with telecare.

Lastly there is a question as to where the benefits of telecare primarily lie. They may accrue mainly to telecare users' families and carers rather than the users themselves. It must be said that evidence on the impact on quality of life for carers of people using telecare is scarce; however the use of telecare may reduce carer strain (Davies, Rixon, and Newman 2013). Nevertheless, if there were beneficial impacts on carers as a result of the introduction of the intervention, these would not have been captured in the analysis presented here. It is worth considering that there can be dis-benefits associated with telecare; potential and actual telecare users' concerns about threats to privacy or to identity may impinge upon their willingness to use or continue to use telecare (Sanders et al. 2012). Benefits may accrue to telecare users with certain characteristics. As I have explored in the Chapter 6, higher costs may be associated with telecare users with particular characteristics such as living arrangements; the result of closer monitoring may have prompted *additional* service responses for some people.

The sub-group analysis of cost-effectiveness by participants living alone or with others proved equivocal but it seems that telecare would not be recommended for people with social care need who are living alone. More research is needed to examine this issue, given that outcomes were not powered to investigate sub-group effects.

8.5 Conclusion

There is great deal of interest amongst policy-makers in the potential of telecare to improve quality of life while containing or perhaps decreasing the use and costs of health and social care and support. However in this study a package of second-generation telecare equipment and associated monitoring service combined with (in two sites) a dedicated response service did not constitute a cost-effective alternative to usual care, assuming a commonly accepted willingness to pay for QALYs. The evidence in favour of telecare remains underwhelming from the cost-effectiveness perspective and so policy-makers should for now avoid characterising the technology as a 'magic bullet' (Poole 2006).

Chapter 9

Conclusion

In this thesis I have examined the costs and outcomes of implementing telehealth and telecare in England, in the context of the Department of Health-funded Whole Systems Demonstrator evaluation. Two large-scale cluster-randomised controlled trials of telehealth and telecare formed the core of the WSD programme. In each trial, the research team recruited General Practices within three English local-authority areas, to be randomised to either intervention or control. The analyses presented here drew on person-level data from the WSD Telehealth and Telecare Questionnaire Studies. The WSD Questionnaire studies collected self-report data from individuals participating in the respective WSD telehealth and telecare trials. In addition, during the period of the study, I collected information from each site about the production of the telehealth and telecare interventions, in order to calculate intervention specific-costs. My original contribution was to plan and conduct the costs and cost-effectiveness analyses of the WSD Questionnaire Studies data.

The overall objective was to examine the costs and benefits of introducing telehealth and telecare in England. The question was broken into four sub-questions.

9.1 What are the Patterns of Service Use for People with and without Telecare or Telehealth Support?

Data from the WSD Telehealth Questionnaire study was analysed to examine patterns of service use and costs by telehealth and usual care participants at baseline and follow-up assessments. Because of issues of quality in the 4-month data, comparisons presented were between baseline and 12-month follow-up. There were small differences between groups across the individual items of health and social care service use at both baseline and 12-months. The telehealth group had somewhat lower mean use of hospital services relative to usual care, particularly at the follow-up. Use of primary care services and community and specialist nurses was higher within both groups (but not different between groups) at follow-up. Both groups used some community care services such as home care and home help. In terms of the intervention, sites offered different configurations of telemonitoring services, but all involved a central monitoring call-centre, and two also offered access to monitoring data

by community-based nurses. Data on telehealth monitor alerts and responses to alerts by monitoring personnel were not available. Participants used 2.8 (SD 0.6) items of telehealth peripheral equipment. The most common combination of items was a BP monitor, pulse oximeter and weighing scales. The proportions of participants in receipt of different types of monitors varied more distinctly by index condition in site 1 compared to the other two sites (e.g. low proportions of patients with COPD and heart failure received glucometers).

In the WSD Telecare sample, differences between individual service use items (including community alarms) were small at baseline. Both groups' resource use at 12 months was also similar, apart from certain community-based health and care categories: home care, social work and community nursing visits, all greater in the telecare group. While use of community alarms was very high in the telecare group at the 12-month follow-up, as would be expected given the nature of the intervention, there were substantial rises in the use of pendant alarms between baseline and 12-month (and also 4-month) follow-up in the usual care group. In terms of the intervention, telecare services across the sites offered a call-centre monitoring service for any sensor and user-initiated alerts. There were differences between delivery models between sites, for instance in installation and response arrangements. Two of three sites offered a dedicated service to respond to alerts, which could include home visits. The average number of telecare equipment items used in the telecare group was 4.7 (SD 1.77). Most participants had 'functional monitoring' items (e.g. bed occupancy alarms), whereas few had sensors with security functions (e.g. bogus caller button). There were many possible combinations of equipment: the most frequent combination of devices observed involved those with functional (e.g. bed and chair occupancy sensors), stand-alone (e.g. key safes) and environmental (e.g. carbon monoxide detectors) functions.

9.2 What are the Total and Component (Service-specific) Costs per Person of the Support/Treatment Received?

In the WSD Telehealth Questionnaire sample, the costs of health and social care in the three month period prior to baseline were similar between allocation groups (£1289 (SE £71) intervention vs. £1273 (SE £66) control). At 12-month follow-up, costs of care, excluding intervention costs, between allocation groups also did not differ significantly (£1150 (SE £110) intervention vs. £1394 (SE £119) control). Hospital costs comprised nearly half (47 per cent) of the total costs (excluding intervention costs), with primary care and medication costs making up 18 per cent of costs each; social care costs comprised 16 per cent of the total. For

participants completing the 12-month assessment point, the mean three-month cost (in 2009/10 prices) of a package of telehealth support was £289 (SE £4) and for equipment, £168 (SE £8). Equipment costs made up a tenth of the overall costs of intervention participant costs. Counting in the costs of the intervention, three-month total costs in the intervention group were slightly higher than in the control group ((£1608 (SE £110) vs. £1403 (SE £120); a difference of £205 (95 per cent CI -£114, £524)). Practice-level clustering of total costs was similar within the allocation groups at each time point.

In the WSD Telecare sample, costs at baseline were similar (£2411 (SE £166) vs. £2484 (SE £174)). At the 12-month follow-up, costs (excluding intervention costs) of both groups were lower than at baseline; costs were somewhat but not significantly higher in the telecare group compared in the control group (£1801 (SE £167) vs. £2021 (SE £166)). Hospital costs comprised about a quarter of the total costs (excluding intervention costs), and community-based social care costs comprised slightly more than a third of the total. The cost per quarter of telecare support for intervention participants was £177 (SE £6) and the cost of telecare equipment only £20 (SE £1). The extent of clustering of total costs in the total baseline sample was much higher in the intervention than the control group, but not in the 12-month follow-up sample.

9.3 What Patient/User Characteristics are Associated with Cost Variations?

In Chapter 6, I presented two difference-in-difference-in-difference analyses: first, of data from the WSD Telehealth Questionnaire study and second of data from the WSD Telecare Questionnaire study. Two approaches were taken to working with clustered data: an explicitly multilevel approach, exploring how an individual participant's costs changed in response to the intervention; and a population-averaged approach more useful in the policy context to examine how costs across the intervention group differed from those of controls. In both cases, substantive modelling involved two levels (person at level 2 and time at level 1). Because costs were highly skewed, the models were fitted to a gamma distribution with a log-link. A two-part approach was taken to substantiate inferences about social and hospital care costs, given substantial zero costs in these categories.

9.3.1 Telehealth findings

I examined whether the three-month total, NHS, hospital and social care costs of the TH Questionnaire study participants allocated to telehealth or usual care differed between

baseline and long-term follow-up time points, depending on their index long-term condition (COPD, diabetes, or heart failure). An examination of raw (unadjusted) costs showed that in participants with COPD and heart failure, these were in general similar in both allocation groups at baseline; they were noticeably higher in the people with diabetes allocated to telehealth. At 12-month follow up, total costs (including intervention costs) were higher in the telehealth participants, across index conditions; the differences between allocation groups by index condition, including or excluding intervention costs, were not significant.

I used multilevel modelling to examine between-group cost differences in subgroups of participants with COPD, heart failure and diabetes. Controlling for socio-demographic and needs-related variables and taking both skewed distribution of the costs and within-person clustering effects into account, I found little evidence of variations in total, NHS-wide, hospital or social care costs based on long-term index condition. Both subject-specific and population-averaged models and marginal effects results suggested that Telehealth participants with diabetes and with heart failure had somewhat but not significantly lower costs (not including intervention costs) in the last three months of participating in the trial than those of usual care participants with those conditions (taking into account differences in costs in the three months prior to baseline). Nonetheless, the between-group differences in these costs were not significantly different between conditions. On the other hand, across conditions, those in the intervention group had on average lower 3-month NHS and overall costs than controls- if excluding the costs of the intervention - at 12-month follow-up. It is therefore important to consider that the estimate of overall cost savings could change considerably if intervention costs were to decrease markedly in response to falls in the price of the technology.

The costs of providing a Telehealth intervention were high, so that across the three conditions, total costs (including intervention costs) in the three months at the end of the study were greater in the intervention than in the control group. Also marginal effects results suggested that the COPD participants had significantly higher costs, including intervention costs, in the last three months of participating in the trial than those of usual care participants with those conditions after adjusting for the baseline difference.

Across time points, sectors and models, characteristics associated with chronic disability and ill-health - ADL need, older age and number of comorbid conditions - were consistently associated with higher costs. Higher levels of education were associated with increases in total costs and NHS and hospital costs, and odds of receipt of hospital care. Being female was associated with increased total costs (subject-specific model only) and with

higher social care costs. Only one socio-economic characteristic, owner-occupation, was associated with lower costs (total, NHS-wide, social care). Costs across sectors, time points and models did not differ by site.

9.3.2 Telecare findings

I examined whether the costs of participants allocated to telecare or usual care differed over time depending on their living arrangements (living alone or with others). Total unadjusted health and social care costs (excluding intervention costs) over the three months prior to follow-up of those living alone in receipt of telecare were significantly greater than of those receiving usual care (by £460), while costs of those living with others in receipt of telecare were somewhat lower (by £92) than of those living with others receiving usual care. Pre-baseline costs, in contrast, were somewhat greater in the telecare group, regardless of living arrangement.

I then used multilevel modelling to examine this question, controlling for other socio-demographic characteristics. Looking at the overall costs of health and social care (including or excluding the costs of the intervention), there was little evidence that these were affected by the telecare intervention, regardless of modelling approach (subject-specific or population-averaged). Controlling for other factors, health and social care costs were lower at follow-up regardless of allocation. Once costs were disaggregated, there were indications of differences between groups. Costs of social care, as with total costs, were apparently unaffected by the introduction of the telecare intervention. NHS costs and hospital services costs, on the other hand, differed depending on allocation and living arrangement. Examining the marginal effect of telecare at baseline and follow-up by type of living arrangement, costs to the NHS in the group of telecare participants living with others were lower at the follow-up. However, taking costs of telecare recipients living alone into account, it was no longer certain that the sum of health care costs across living arrangement groups would be lower.

Number of comorbidities and higher levels of ADL need were consistently associated with higher costs. There were some site effects in costs: higher total costs and higher odds of receipt and costs of social care being associated with site 2, relative to site 1, while site 3 was associated with lower odds of social care receipt than site 1. Older age was associated with lower total and NHS costs and lower odds of receipt of hospital care and lower hospital costs.

The influence of sites was evident in the general social care costs of participants, where site 2 had higher and site 3 had lower costs than site 1. This was in contrast to the

absence of site effects seen in the telehealth analyses. In the Telecare study population, with social care need and relatively high use of social care services, much greater local variation between social care than health care provision is to be expected, reflecting the fundamental differences in access (means tested vs. universal) between these services.

9.4 Are Telehealth and Telecare Cost-effective Compared to Standard Support/treatment?

9.4.1 Cost-effectiveness of Telehealth

In an evaluation using WSD telehealth questionnaire study data, costs and outcomes data at baseline and 12 month follow-ups were available for 965 participants (534 intervention, 431 control). Mean health and social care costs (including intervention costs) were somewhat greater in the telehealth group than for the usual care group at follow-up, but the difference on a clustered t-test did not reach significance at the 5 per cent level. Findings were similar from a SUR model that adjusted for demographic characteristics and level of need.

Differences in outcomes were small and no differences between groups were significant on clustered t-tests apart from the brief STAI where the difference was less than one point in favour of telehealth. In SUR models adjusting for baseline characteristics, between-group differences were also small. However there was a significant difference of 0.014 in EQ-5D-3L-derived QALY between groups at 12 month follow-up (in the direction of the intervention group). The between-group differences on other outcomes were not significant at the 5 per cent level, except for the CESD-10 (small at -0.705, less than one point on a 30-point scale) and the STAI (small at -0.774, less than one-point on a 24-point scale), in favour of the telehealth group.

In terms of EQ-5D-3L-derived QALY gain, the probability that telehealth was cost-effective was relatively low, only exceeding 50 per cent at willingness to pay values above £67,000.

On secondary outcomes, ICERs were: £22,600 for an improvement from highest to lowest levels of anxiety on the Brief STAI; £6,900 for achieving a five-point reduction on the CESD-10; £233,700 for an improvement from no capability to full capability on the ICECAP-O. The probability of cost-effectiveness in terms of anxiety and depression symptoms was 50 per cent at willingness to pay values of over £22,600 and £7,000, respectively. The probability of cost-effectiveness of telehealth on the ICECAP-O index was 15 per cent at a willingness to pay of £50,000. Mean QALY derived from the SF-6D were

0.005 higher in the intervention than in the control group; the ICER was £178,600. The probability that telehealth was effective was lower than that yielded by EQ-5D-3L and did not exceed 35 per cent over a range of willingness to pay from £0 to £90,000.

9.4.2 Cost-effectiveness of Telecare

In an economic evaluation of WSD telecare questionnaire study data, cost and outcome data of 753 participants were available across baseline and 12-month follow-up (375 intervention and 378 control). The health and social care cluster-adjusted costs of the intervention group were somewhat but not significantly higher than those of the control group (whether including or excluding intervention-specific costs). Between-group cluster-adjusted differences in outcomes at the 12-month follow-up were small and not significantly different. The results of a SUR model adjusting for baseline characteristics, ADL needs and previous use of community alarms in the intervention group showed that the telecare group had slightly higher EQ-5D-3L-3L-derived QALY than the control group. Differences between groups on other outcomes from the SUR models were small. Differences in brief STAI, MCS-12, PCS-12 and SF-6D were in favour of the telecare group; differences in ICECAP-O scores were in favour of the usual care group. Cost per EQ-5D-3L-derived QALY was £368,000. Cost-effectiveness at the willingness to pay threshold of £30,000 per QALY was 16 per cent. On other measures, ICER were: £50,000 for a movement from worst to best on the STAI; £10,000 for a 3-point increase in the PCS-12 and £4,000 for a 4-point increase in the MCS-12. On the ICECAP-O, telecare was dominated by the usual care alternative. The cost per SF-6D-derived QALY gained was £240,000. The probability of cost-effectiveness at the higher end of the NICE willingness-to-pay threshold range, £30,000, was 16 per cent. At levels of WTP of £10,000 to £20,000, the probability of achieving a reduction from maximum to lowest level of state anxiety ranged between 16 per cent and 24 per cent; of achieving a 3-point increase in PCS-12 ranged between 51 per cent and 61 per cent; of achieving a 4-point increase in MCS-12 ranged between 77 per cent and 86 per cent. The probability of cost-effectiveness on QALY derived from the SF-6D at the WTP of £30,000 was 13 per cent.

In a sub-group cost-effectiveness analysis, at 12 months, people who used telecare and were living with others (at baseline) had similar outcomes to people who used telecare and lived alone (at baseline). Total costs were not different between allocation groups in people living with others; whereas the costs were higher in the telecare group in people living

alone. There was little certainty of cost-effectiveness of telecare for people who were living with others at the baseline over a range of WTP from zero to £90,000; telecare for people who were living alone at baseline was less cost-effective than usual care at WTP values below £33,000 and there was no certainty of telecare being more or less cost-effective than usual care above that value.

Sensitivity analyses examining robustness of the analyses to assumptions about intervention-specific costs did not challenge any inferences made from the main analyses. An analysis examining robustness of assumptions of normality of distribution of costs and outcomes also did not challenge these inferences. The WSD telecare intervention was not a cost-effective alternative to usual care, at a commonly accepted willingness to pay for QALYs.

9.5 Relating the Findings to the Literature

I turn to the question of putting these findings in the context of the theoretical and empirical literature as reported in chapters 2 and 3. I also discuss relevant recent literature, including publications from the WSD research programme.

9.5.1 Telehealth

As discussed in chapter 7, at the time of the trial, relatively little information on the costs of telemonitoring was available in the literature (cf. Inglis et al. 2010). As discussed in chapter 5, service configurations and costs could vary substantially by site. It was not possible to collect data on responses to telehealth alerts, limiting the granularity of intervention support costs to site-level, although individual-level costs could vary depending on the device provided. These kinds of difficulties may explain why the telehealth intervention costs of telephone support or telemonitoring reported in the literature are so variable. With the passage of time, falls in prices of telemonitoring devices, and the penetration of smartphone-based medical applications (Cottrell, Chambers, and O'Connell 2012, NHS Stoke-on-Trent 2011), will have limited the generalisability of estimates from older studies. Cost estimates from the UK-based literature (section 7.5.2) ranged from about £960 to £1,420 per annum (2007 prices) (Barlow et al. 2007) and in the telehealth study here, between £1,500 to £2,000 per annum (2009/10 prices). A study on uncontrolled hypertension (Stoddart et al. 2013), contemporaneous with the WSD, used mobile phone technology to transmit readings from

blood-pressure monitors to attending physicians, estimating the 6-month cost at only £71 (2010 prices).

A cost-effectiveness study by Dixon, Hollinghurst, Edwards, Thomas, Foster, et al. (2016) examined a low-cost model of telehealth for raised cardiovascular risk in primary care (the NHS Direct Healthlines Service). The cost of the intervention was £129 (SD £56.33) (2012/13 prices). This was a web-based model whereby participants with telehealth entered blood-pressure readings manually into the Healthlines portal, with telephone-based behavioural and educational support from health care advisors (non-qualified clinical workers) in the now-defunct NHS Direct service. These intervention costs were much lower than those of the WSD, owing to lower labour costs (unqualified clinical support staffing based within NHS Direct) and manual inputting of vital signs by patients, using basic blood pressure monitors (Thomas et al. 2014)).

Witt Udsen et al. (2017) report a cost-effectiveness analysis from a Danish pragmatic trial of an asynchronous telehealth intervention delivered by municipality health care professionals[21] to people with COPD (578 in TH and 647 in UC). The intervention cost €704 (2014 prices). The adjusted total cost of health and social care was €728 (95% CI −€754, €2211) higher in the TH group; adjusted QALY were 0.0132 (95% CI −0.0083, 0.0346) higher in the TH than UC group. The ICER was €55,327 per QALY. The probability that the intervention was cost-effective reached 50 per cent at a WTP value of €55,000. The authors concluded that telehealth was not likely to be cost-effective for people with COPD over the range of NICE threshold values.

I examined variations in the costs of participants with COPD, diabetes and heart failure related to the introduction of the intervention, adjusting for other personal characteristics. Comorbidities, older age and ADL need were drivers of increased costs in all models, in line with other research on costs in chronic conditions (Tsiachristas and Rutten-van Mölken 2014), COPD (Hutchinson et al. 2010), (Hetlevik, Melbye, and Gjesdal 2016) and heart failure (in terms of ADL need)(Kang et al. 2016), while being female was a driver of overall costs, as has also been reported in the literature on costs of COPD (Menn et al. 2012, Hetlevik, Melbye, and Gjesdal 2016). It was not possible to capture alerts from the

[21] Vital signs data collected by oximeter, BP monitor and scales were communicated to monitoring nurses and assistants via computer tablet-based applications.

telemonitoring systems in the WSD trial, so there is no way to compare to the literature on service use (key events) related to alerts.

Compared to evidence from the systematic reviews of effectiveness of telehealth for long-term conditions presented in Chapter 3, the findings presented in Chapter 6 and Chapter 7 were less positive than some assessments (Barlow et al. 2007, Clark et al. 2010, Inglis et al. 2010, Pare, Janna, and Sicotte 2007, Polisena, Coyle, et al. 2009, Polisena, Tran, et al. 2009). Results presented in chapter 7 focused on self-reported outcomes. Clinical outcomes and reductions in service use rather than person-reported outcome measures are often presented as final outcomes in the literature. It is also striking that many reviews lump together several disparate outcomes (condition-specific and generic HrQoL, satisfaction with the technology, psychological outcome measures) as if they all measured the same concept. There is a need for systematic reviews to present outcomes informatively, for instance by considering how many studies found improvements in generic HrQoL measures. Where HrQoL outcomes were considered separately, systematic reviews reported rather mixed evidence in favour of telehealth for people with diabetes (Polisena, Tran, et al. 2009) and respiratory conditions (Pare, Janna, and Sicotte 2007, Polisena, Coyle, et al. 2009). A more recent review assessed the evidence of the impacts of telephone support and telemonitoring for heart failure on health-related quality life as strong (Inglis et al. 2015). The review also found evidence for the cost-effectiveness of telephone support but not for telemonitoring. In Bergmo (2009), there were surprisingly few economic evaluations giving ICER in terms of cost per QALY (only one of which conducted as part of an RCT, for asthma), making comparisons with that author's findings difficult.

In the planning stages of the WSD Questionnaire studies, EQ-5D-3L was chosen as the HrQoL measure to be used in the economic evaluation. However as the SF-12 was collected, it was possible to construct the SF-6D for comparative purposes. The evidence on the empirical validity of the EQ-5D-3L vs. SF-6D is fairly equivocal (Brazier 2007) and depends on the population examined. However, the EQ-5D-3L was chosen as a generic measure of health-related quality of life that can be used as a basis for comparing alternative technologies (National Institute for Health and Clinical Excellence 2008). Also, EQ-5D-3L can be expected to be sensitive to change where changes in health are expected to be substantial (Haywood, Garratt, and Fitzpatrick 2005). It is interesting that, compared to the gain in QALY derived from the EQ-5D-3L, gain in QALY derived from SF-6D was smaller (between-group differences in raw scores: 0.015 vs 0.005 respectively) (but neither difference was significant).

Grieve, Grishchenko, and Cairns (2009) observe that compared to SF-6D, EQ-5D-3L has lower utilities for health states (e.g. pain) that are more severe. The SF-6D descriptive system has two dimensions (vitality and social functioning) not found in the EQ-5D-3L. They suggest that people in relatively good health may have poorer health states in the SF-6D compared to EQ-5D-3L because it has two more levels. In this study, baseline EQ-5D-3L and 12-month follow up utilities were higher than SF-6D utilities regardless of allocation. If (as could reasonably be expected given the WSD telehealth study population) participants were not in good health, the results here are in line with this prediction.

While systematic reviews of telehealth for heart failure have detected decreases in the use of secondary care (Inglis et al. 2010, Polisena, Tran, et al. 2009, Polisena et al. 2010a), there has been less evidence for COPD and diabetes generally, and less evidence on use of primary care services in heart failure (Polisena, Tran, et al. 2009) Some more recent evidence from a trial of telemonitoring for Type 2 diabetes (Wild et al. 2016) suggests that telehealth does not reduce use of primary care services such as GP and practice nurse contacts (and see also Section 9.5.3 on other findings from the WSD programme). In the analyses presented in Chapter 5, use of some hospital and social care services by the telehealth group were slightly lower at the follow-up. In the costs analyses reported in Chapter 6, overall NHS costs (intervention specific costs excluded) were somewhat lower but there was no evidence of lower costs in the telehealth group, across the conditions, in terms of hospital care.

According to systematic reviews of economic evaluations of telehealth prior to 2013, telehealth studies have generally found that the technology reduces health care costs. Nonetheless, reviewers recommended interpreting these results with caution, given that the generally poor quality of the evaluations (Bergmo 2009, Polisena, Coyle, et al. 2009, Vergara Rojas and Gagnon 2008).

The Dixon, Hollinghurst, Edwards, Thomas, Foster, et al. (2016) study (N=641) found a web-based model of telehealth for raised cardiovascular risk (NHS Direct Healthlines Service) to be cost-effective over the NICE WTP threshold range (ICER of £10,859; probability of cost-effectiveness of 77% at £20,000 per QALY). Costs in the telehealth group were significantly higher at £138 (95 per cent CI £66, £211); QALY gain was 0.012 (95 per cent CI −0.001, 0.026). The authors, comparing their results to those of the published WSD Telehealth Questionnaire cost-effectiveness, noted the sensitivity of inferences on cost-effectiveness to costs of technology and the kinds of technology employed.

9.5.2 Telecare

I discuss here how some of the concepts presented in chapter 2 can shed light on the research findings in chapters 5, 6 and 8 and the implications of the findings for those concepts. The promotion and maintenance of independence in older people is often invoked as a reason for telecare provision. Policymakers tend to engage in 'modernist discourse', envisaging a compliant older population that can get on with technologies such as telecare and telehealth technologies, and hoping to reduce wasteful practices such as consuming services provided by human carers (Greenhalgh et al. 2012). This 'better outcomes' narrative links independence, efficiency savings and user quality of life (Glasby, Lynch, and Robinson 2018) although the chain of causality is never fully specified. Independence from this perspective equals not relying on state services (Glasby, Lynch, and Robinson 2018), or 'being able to look after yourself' (Peek, Aarts, and Wouters 2015, Sixsmith and Sixsmith 2008). The costs and cost-effectiveness analyses presented here examined cost and health-related quality of life, both elements of the 'better outcomes' narrative. On the question of total and component service costs, I concluded that telecare had little impact on these (as was also concluded in the trial evaluation paper (Steventon et al. 2013)).

In Chapter 6, I explored differences in costs between the telecare and usual care groups in people living alone and people living with others. This exploratory analysis suggested that in people living alone, the group ideally placed to benefit from telecare by staying 'independent', neither health nor social care costs decreased in the presence of telecare, whereas some health care costs of those living with others did. I discussed a set of possible reasons for these findings. It was possible that need was not adequately controlled for in the models by inclusion of the self-care covariate (for instance, other measures of need such as cognitive impairment were not measured in the study). Or, telecare exposed unmet health needs and thereby caused additional health service use by individuals living alone, while telecare for people living with others perhaps gave paid or unpaid carers confidence to substitute health care visits with remote monitoring. The results raise the possibility that cost savings from social care organisations' investments in telecare largely accumulate to health organisations (cf. Forder 2009). There should be sufficient funding to social care to allow any telecare-related reductions in health care spending to be sustained, which would require transfers from NHS to social care. A recent paper has identified reductions in hospital length of stay related to telecare use from Scottish routine health and social care date (which includes data on telecare equipment use) (Momanyi 2017a).

The analysis originally published (Henderson et al. 2014) provided a cost-per-QALY for telecare that was far in excess of the NICE threshold; the analysis reported here provides an even higher ICER (higher by £71,000 per QALY). This does not change any of the conclusions originally reported. The results of the subgroup analysis of cost-effectiveness by participants living alone or with others proved equivocal but it seems that telecare would not be recommended for people with social care need who are living alone. It is however important to be mindful of the context. There is a further consideration touched upon in chapter 2. Left to their own devices, older people living alone may postpone use of a pendant alarm until they see themselves at high risk for falls (Nyman 2014). In a general population of older people with social care need, the situation may be different to that observed in the trial population, who were allocated telecare.

The findings of chapters 6 and 8 taken together were in considerable contrast to the policy discourse on telecare, of fostering 'independence', decreasing social care service use and through some unexplained causal chain, improving quality of life.

On the question of health-related quality of life, I concluded that telecare had little impact on this outcome, although there was some trend to better mental health-related quality of life outcomes scores (on the MCS-12). The small improvement in mental health-related quality of life outcomes could be related to some sense of 'reassurance' (Beale et al. 2010, Hirani et al. 2013, Roush and Teasdale 2011). How much *more* reassurance was provided by additional sensors, over and above that provided by the ability to summon help, was not addressed in the trial design. I speculated that if WSD telecare did not vary substantially from first-generation telecare, the impact on outcomes for intervention participants might not have been much different than for the substantial proportion of controls (64%) reporting community alarm use by the end of the trial.

The study did not address independence as a psychological outcome per se (although the concept is captured as one of the five domains of the ICECAP-O). Concepts of independence such as not wanting to be dependent on technology or not wanting to be a burden on relatives (Peek, Aarts, and Wouters 2015, Sixsmith and Sixsmith 2008) were not directly addressed in the trial or study design; however qualitative work within the WSD programme examined reasons for declining telecare (Sanders et al. 2012). That research found that threats to privacy or to identity did impinge upon individuals' willingness to use or continue to use telecare.

Compared to the differences in QALY derived from the EQ-5D-3L, between-group differences in QALY derived from SF-6D were smaller (-0.017 vs -0.005 respectively). As

discussed in chapter 8, the EQ-5D-3L has been found to be suitable for use with older people (Haywood, Garratt, and Fitzpatrick 2005, Hawton et al. 2011). Nonetheless, it was not known whether EQ-5D-3L would be able fully detect any improvements brought about by telecare, given the EQ-5D-3L's concentration on health and restoration of function.

Returning to Grieve, Grishchenko, and Cairns (2009), it is interesting to see that people in the telecare sample were in apparently poorer health relative to the telehealth sample on the EQ-5D-3L. Yet on the SF-6D they had utilities that were, although lower, more similar to those of the telehealth sample. This does not appear consistent with the pattern suggested in the Grieve paper and observed in the Telehealth study. On the other hand, it may be that SF-6D was able to detect more positive aspects of quality of life in the vitality and social functioning dimensions, leading to overall higher scores than on EQ-5D-3L. Perhaps the sensitivity owing to these additional dimensions also led to a smaller negative raw difference in QALY between groups at the 12-month follow up. Ultimately, the inferences in terms of cost-effectiveness do not change, whichever instrument is used to measure HrQoL.

There are few studies of remotely comparable sample sizes with which to compare the findings of this work (except the WSD Telecare trial, see below). Morgenstern et al. (2015) reports results of a US-based trial of medical alert devices (first-generation telecare) in women with stroke. Analyses of data from 122 intervention and 112 control participants indicated no differences between groups at 90 days in health-related quality of life, depression, anxiety or pain, and perceived isolation.

9.5.3 WSD studies

The WSD research programme included examinations of the effects of telehealth and telecare on service use in the larger trial population (Steventon et al. 2012, Steventon et al. 2013), on psychosocial outcomes in the questionnaire population (Hirani et al. 2013, Henderson et al. 2013, Cartwright et al. 2013, Henderson et al. 2014, Rixon et al. 2015, Hirani et al. 2017) and from the perspectives of organisations (Hendy et al. 2012), health professionals (Sanders et al. 2012) and patients (MacNeill et al. 2014).

As there were many publications from the WSD programme, I have focused on discussing findings from the quantitative research programme that support or are at odds with the findings presented in the thesis. Publications of quantitative research on the WSD telehealth and telecare trial are summarised in Box 9.1 and Box 9.2 respectively.

Telehealth: The results presented in chapters 5, 6 and 7 are broadly in line with the findings of the service utilisation analyses conducted as part of the WSD study (Steventon et al. 2012, Bardsley, Steventon, and Doll 2013). Despite different approaches, source of data and sample sizes, hospital costs were found not to differ between groups over the trial period or three months prior to baseline in both the trial analysis and those presented here. In terms of use of primary care, despite the differences in the time period (12 months vs 3 months) between the trial and questionnaire on GP and practice nurse contacts, in neither that nor the analysis presented here were there substantial differences between groups at baseline or over the trial period. However practice nurse contacts in the questionnaire data at baseline were far lower than at 12 months and much lower than the figures (if divided by four) given in the Bardsley, Steventon, and Doll (2013) paper (see Box 9.1), suggesting that there might have been issues related to the means of administration of the questionnaires. An exploration of the impact of glycaemic control in the WSD trial sample (Steventon et al. 2014) may shed some light on an apparent trend in the subject-specific model to lower NHS costs at follow-up in the diabetic participants with telehealth. That study found a significant effect of telehealth on glycaemic control, albeit a small one, which could, if the trend to better control holds true for the smaller sample available in this analysis, have had some impact on reducing use of diabetes-related NHS services by this group.

The results in terms of outcomes presented in chapter 7 are broadly similar to those of Cartwright et al. (2013), whose overall conclusion was that while telelehealth did not result in better outcomes, there was no evidence of poorer outcomes either. The samples analysed in the (Cartwright et al. 2013) paper and the cost-effectiveness paper overlapped but were not the same, and analyses in the former included outcome measures at the 4-month follow-up. In other publications examining psychosocial outcomes by long-term condition (Hirani et al. 2017, Rixon et al. 2015, Newman and Whole System Demonstrator Programme Evaluation Team 2014), very little difference was found between allocation groups in each case. While I did not carry out a cost-effectiveness analysis to investigate the differences between groups with COPD, diabetes and heart failure, the results of those studies and the cost analyses in chapter 6 suggest that findings would not differ from the overall conclusions of chapter 7.

Box 9.1 Publications from the Whole Systems Demonstrator Telehealth studies

Health-Related Quality Of Life and Psychological Outcomes (WSD Telehealth Questionnaire Study)

Cartwright et al. (2013) examined the effect of second-generation telehealth on HrQOL and psychological distress in people with long-term conditions (COPD, diabetes and heart failure) over 12 months as part of the WSD Telehealth questionnaire study (1573 participants), nested in the WSD telehealth pragmatic, cluster-randomised trial (92 general practices were randomised to telehealth and 87 to usual care; 1605 telehealth and 1625 usual care participants). Outcomes were measured at baseline, four and 12 months. 759 participants completed measures at all time points (complete cases) and 1201 completed measures at baseline and one other point (available cases). Primary analyses were intention to treat (ITT); secondary per-protocol analyses examined efficacy (633 complete cases, 1108 available cases). In the ITT and per-protocol analyses, there were no significant differences between allocation groups in either complete or available cases samples. The authors concluded that the WSD telehealth intervention did not produce improvements in HrQOL or psychological distress for participants with long-term conditions; the intervention also did not worsen outcomes. Telehealth for the purposes of improving HrQoL or psychological outcomes was not recommended. Comparisons of the trial and questionnaire samples suggested there was some evidence of potential selection bias into the Questionnaire study, as a higher proportion of participants in the trial intervention group than of controls agreed to participate in the questionnaire study.

Hirani et al. (2017) examined generic and condition-specific HrQoL, and psychological distress (anxiety and depression) in the WSD telehealth questionnaire study participants with diabetes ((246 intervention, 209 control). Data from 167 intervention and 150 control (available cases, those completing baseline and at least one of the follow up assessments) were analysed. The groups did not differ significantly on any measure except the disinhibited eating scale of the Diabetes Health Profile, where the intervention group had higher (worse) scores than the controls; however the effect size on this measure was small and confidence intervals were wide and crossed zero. The required sample size to power the detection of a small effect was not met due to attrition. The intervention did not improve outcomes nor did it worsen outcomes for diabetic participants over 12 months.

Box 9.1 (continued)

Rixon et al. (2015) examine psychological distress (anxiety and depression) and generic and condition-specific measures of HrQOL in 447 participant in the WSD telehealth questionnaire study with COPD. Small improvements were found in 12-month condition-specific QOL (emotional functioning and mastery on The Chronic Respiratory Questionnaire (CRQ) (Guyatt 1987) in the intervention group compared to controls. The sample size did not reach adequate power because of attrition. No differences between in the telehealth and control groups were found in generic QOL or in terms of psychological distress.

WSD final report: In an analysis of outcomes in participants of the Telehealth Questionnaire trial with heart failure (Newman and Whole System Demonstrator Programme Evaluation Team 2014), there were 265 telehealth and 275 usual care participants with heart failure at baseline, 146 telehealth and 138 usual care in the complete cases sample and 228 telehealth and 209 usual care in the available cases sample. The analyses examined generic and condition-specific HrQOL and psychological distress outcomes. No robust effects over the 12 month period were reported in any outcome measures.

Service Utilisation and Mortality (WSD Telehealth Trial)

Steventon et al. (2012) reported the results of examining service utilisation and mortality in the 1605 telehealth and 1625 control participants recruited via 179 general practices into the WSD Telehealth trial. The analyses drew on linked longitudinal administrative data (Hospital Episodes Statistics, local commissioning data on emergency department visits, and general practice data). Results (using data for 1570 intervention and 1584 control participants) showed that the (raw) absolute difference in percentages admitted was 5.2 (42.9 intervention vs. 48.2 control). The (raw) absolute difference in percentage of deaths between groups was 3.7 (4.6 intervention vs. 8.3 control). Results of multilevel analyses, adjusting for case mix with predictive risk scores, indicated that the intervention group had a significantly lower proportion of hospital admissions over the 12 months from baseline, being 18% less likely to have an admission than controls. Unexpectedly, however, admissions in the control group rose at beginning of the trial, generating the significant overall between-group difference observed at the end. Also, the between-group difference in percentage admitted was lower than the 17.5% that the study had been designed to detect.

Box 9.1 (continued)

The analyses showed that the intervention group were 47% less likely to die than controls over the same period. There were significant differences between groups on secondary outcomes including numbers of bed-days (a reduction in the intervention group of 0.64 days) and emergency admissions but not on some other measures (outpatient visits and costs of hospital service use). Costs were calculated using the Department of Health PBR Tariff costs (costs to commissioners of NHS care). The raw costs of hospital care in the three months before baseline were £427 in intervention and £506 in the control group; the adjusted (geometric mean) difference in tariff costs between groups (not significant) was not large, at £188 less in the TH than the control group over the study period. The authors advanced some possible reasons for the rise of emergency admissions in the controls: that in the process of recruitment professionals detected unmet need in controls that they chose to treat; or that control patients were made more aware and concerned about their condition and so were more likely to present at hospital for emergency admission; or that there was a selection bias so that controls with higher risk, and intervention patients with lower risk, were selected into the trial. Comparisons of baseline characteristics between groups made the last of these possibilities unlikely.

Steventon et al. (2014) examined glycaemic control in participants in the WSD trial sample of patients with type 2 diabetes (N=513). Routine data for Hb A1c was available for 300 intervention and 213 control participants. Results suggested a modest but not clinically significant improvement in glycaemic control in the intervention group.

Steventon, Grieve, and Bardsley (2015) examined the generalisability of the WSD telehealth findings to routine clinical practice. Amongst other analyses, they examined the unexpected rise in admissions in the control group in the first months of the trial period, through a placebo test comparing the trial controls to matched local controls who had not participated in the trial. Placebo tests for this outcome and also for mortality failed. Furthermore, sensitivity analyses suggested (non-significant) increases in emergency admissions and higher mortality in the intervention group, which had important policy implications because "reductions in emergency admissions continue to be a major motivation to invest in telehealth." p.1033. The authors concluded that it was not possible to generalise the WSD telehealth trial results relating to emergency admissions and mortality to routine

Box 9.1 (continued)

NHS practice. The paper's authors (Steventon, Grieve, and Newman 2015) also observed that it was unlikely that clinical treatment of controls was altered during the trial period; but it was possible that the control participants' health seeking behaviours had been influenced by being recruited into the trial (for instance because of disappointment with allocation).

Bardsley, Steventon, and Doll (2013) examined the use of primary care by trial participants (1219 intervention, 1098 control). They also examined recorded tests/readings in parallel with the kinds of vital signs monitored within telehealth systems (e.g. HbA1C, oxygen levels, and weight). In the 12 months prior to the trial, contacts with GPs did not differ between experimental groups (TH: 8.8 (SD 6.8) vs. control: 9.0 (SD7.6) and Practice nurse: (TH 5.3 (SD 7.8) vs control 6.1 (SD 8.1). Over the 12 months of the trial, there were no significant differences in contacts with either GPs (TH 8.99 (7.00): vs. control 8.85 (8.16)) or practice nurses (TH 5.92 (9.83) vs control 6.28 (8.98). The authors noted that in the pre-baseline, across allocation groups, participants had a higher number of GP contacts than found in other studies of comparable populations. There was no difference between the groups in the numbers of recorded tests/readings over the period of the trial, suggesting that there had been no shift of testing out of general practice due to telemonitoring. They noted that system-level incentives, for instance payments for taking certain types of clinical readings, could mitigate against practices choosing to change how often they took readings. The conclusion arrived at was that there was no reason to suspect increased or decreased primary care workloads would be associated with telemonitoring.

Telecare: Results of the analyses of service use and costs in the telecare sample (chapters 5 and 6) are in line with the findings presented in the study of service use and costs in the Telecare trial by Steventon et al. (2013) This examined a somewhat more limited range of services than presented here. Despite a larger sample size and the use of routine data, resource use in the three months prior to baseline reported there was broadly comparable with similar items reported here, except in terms of proportions visiting GP surgeries (approximately twice the proportion reported here). Here similarly, the between-group differences in hospital and social care costs at follow-up were not significant, as was also found in the Trial results for the 12-month period, despite differences in overall sample sizes, data sources, and different costing approaches.

Results of outcomes for telecare and usual care participants presented in chapter 7 were similar to those presented in Hirani et al. (2013) (see Box 9.2), despite differences in the sample sizes examined. The most positive result noted in that paper was on the MCS-12, and this was reflected in the cost-effectiveness results for MCS-12 presented here.

Box 9.2 Publications from the Whole Systems Demonstrator Telecare studies

Health-related quality of life and psychological outcomes (WSD Telecare questionnaire study)

Hirani et al. (2013) assessed the effect of telecare on health-related quality of life and psychological outcomes of home-based telecare in people with social care needs over 12 months, as part of the WSD Telecare questionnaire study, nested in the WSD telecare pragmatic, cluster-randomised trial. The unit of randomisation was General Practice. There were 550 telecare participants (101 general practices) and 639 usual care participants (103 general practices) at baseline. Outcomes were measured at baseline, four and 12 months. 873 cases (430 intervention, 443 control) with data available at baseline and at least one other follow-up were available for analysis. Multilevel analyses indicated a significant difference between allocation groups on the adjusted SF-12 mental component scores (43.69 (SE 0.83) intervention vs. 40.52 (SE 0.88) control). The effect size estimate (Hedge's *g* of -0.177 (95 per cent CI –1.364, 1.009) was small. EQ-5D-3L scores declined significantly and depressive symptoms increased significantly in both groups from the 4-month to the 12-month follow-up. The intervention may have mitigated decline in mental health-related quality of life.

Service utilisation and mortality (WSD telecare trial)

Steventon et al. (2013) examined the impact of telecare on utilisation and costs of health and social care services in 1276 telecare and 1324 usual care participants recruited through 217 general practices into the WSD Telecare trial. The analyses drew on linked longitudinal administrative datasets (Hospital Episodes Statistics, local commissioning data on emergency department visits, general practice and local authority data). The primary outcome was proportion admitted to hospital over 12 months. Secondary outcomes were: mortality, emergency admissions, elective admissions, outpatient attendances, emergency department visits, falls admissions, bed days, GP contacts, practice nurses, proportion admitted to permanent residential or nursing care, home care weeks, hospital tariff costs, GP surgery costs, and social care costs.

Box 9.2 (continued)

Results for 1236 intervention and 1190 control participants with linked data indicated an unadjusted absolute difference in percentages admitted of –2.4 per cent (46.8 per cent intervention vs. 49.2 per cent control). Mortality rates were 8.7 per cent in the intervention group and 8.9 per cent in the control group. Proportions admitted permanently to care homes over the year were similar in both groups (3.1 per cent intervention vs. 3.2 per cent control). The unadjusted difference in 12-month hospital tariff costs was £242 (£2,804 intervention vs. £2,604 control) and in social care costs –£77 (£4,210 vs. £4,287 control). Multilevel models adjusted for baseline covariates and for predictive risk scores. Adjusted proportions admitted to hospital was significant in one model (baseline covariate adjustment: Odds ratio=0.83; 0.69 to 0.99) but not in the other (predictive risk score adjustment: Odds ratio=0.89; 0.74 to 1.07). There were no significant differences on other outcomes. The authors concluded that WSD telecare did not reduce service utilisation or costs over the 12-month study period.

9.6 Strengths and Limitations

According to systematic reviews of economic evaluations of telehealth prior to 2013, the quality of evaluations has left something to be desired in several ways .Recommendations have included using more diverse populations to improve external validity; using a standardised approach, such as an explicit economic evaluation framework, including all relevant costs and being clear about inclusions and exclusions (Bergmo 2009, Polisena, Coyle, et al. 2009).

The economic evaluations of the WSD Telehealth and Telecare Questionnaire studies adhered to good practice as defined by economic evaluation guidelines (cf. Husereau et al. 2013) including stating the research objectives, analytical viewpoint, choice of comparators, alternatives compared, outcomes, costs, details of currency and price, stating time horizon, giving details of statistical tests and confidence intervals, explaining the approach taken to and choice of variables for sensitivity analyses; comparing relevant alternatives, reporting an incremental analysis, presenting results for each outcome, answering the study question, and drawing conclusions with appropriate caveats. In addition, cost variations analyses took account of within-person clustering and the implications of two approaches to clustering. The cost-effectiveness analyses took account of both clustering and the correlation between the cost and outcome variables (cf. Gomes, Grieve, et al. 2012).

The economic evaluation of WSD telecare was important in contributing to the scarce evidence base on the cost-effectiveness of this technology. It differed from other RCTs of telecare in the published literature in several important ways. The data was collected from a study using a cluster-randomised-controlled trial design. The sample size available for analysis was on a much larger scale than other RCTs that are at all comparable (Tomita et al. 2007, Brownsell, Blackburn, and Hawley 2008, Morgenstern et al. 2015). As with the Telehealth trial, the approach conformed to economic evaluation guidelines. The economic evaluation formally assessed the impact of telecare on health and social care costs and on HrQoL outcomes.

These analyses have some limitations. I begin by discussing limitations relating to the economic data collection and specific analyses carried out for the dissertation. Some limitations are observable in the studies of both technologies: I begin with a discussion of these. Limitations specific to one or other evaluation are discussed subsequently, particularly in relation to threats to generalisability and uncertainties arising from sampling issues.

9.6.1 Biases and Issues Related to Self-Reported Service Use

As an evaluator of both the telecare and telehealth questionnaire studies, I faced several challenges. The data was to be collected by self-report. Service use can be subject to recall bias. It can suffer from 'telescoping' forward to take in use occurring prior to the period of retrospective recall, or backward, where imperfect recall excludes use within that period (Evans & Crawford 2000; Bhandari & Wagner 2006). The CSRIs that I devised for the study relied on a three-month retrospective period to try to minimise recall problems as has been recommended (Bhandari and Wagner 2006). However the three-month snapshot approach posed problems for the analyses of cost variations and cost-effectiveness evaluations in two ways. The participants' annual costs were estimated by multiplying the costs in the three months prior to 12-month follow-up by four. The samples completing four- and twelve-month datasets did not completely overlap. This, in combination with the relatively poor-quality data yielded by postal-questionnaire only administration led to the decision to concentrate on baseline and longer-term follow-up points. This may have been less important for less episodic service use such as community nursing than for hospital admissions (see Section 7.5.1). In examining the variations in costs, the analyses were limited to comparisons of pre-intervention and long-term post-intervention differences between allocation groups

and did not permit investigation of changes in costs between the short-term and long-term follow-ups.

In contrast to the self-reporting of service use in the questionnaire studies, the WSD trial evaluations (see Box 9.1 and Box 9.2) linked administrative data across health and social care organisations. Administrative data has tremendous potential for shedding light on service use over long periods of time, and for collecting service use as it would be defined for administrative purposes. For instance, a patient's ability to report that a hospital visit was technically an admission and not an A&E episode could be less accurate than data drawn from hospital records. However, it is difficult (and costly) to capture the full range of service use of relevance to people with relatively high use of health and social care through administrative means.

In the analyses, I limited multiple imputation of data to cases who had completed a follow-up assessment, rather than imputing data for non-completers. This decision was in line with the strategy across the original evaluations of the questionnaire data. On the other hand, improvements were made on the analyses originally reported (Henderson et al. 2014, Henderson et al. 2013) in that the missing telecare and telehealth cost data were imputed by predictive-mean matching to accommodate the skewness of the cost variables. Missing telehealth data was imputed separately by allocation group and used a linear mixed model to impute the data as recommended by Diaz-Ordaz, Kenward, and Grieve (2014) and Gomes et al. (2013).

While the telehealth and telecare interventions were in many ways dissimilar, both were complex (Craig et al. 2008). Conducting the economic evaluations was challenging for reasons enumerated by Byford et al. (2007): user heterogeneity, co-production of care by users and professionals, co-production of care by networks of service providers that varied by geographical area. Particularly in the case of telehealth, the associations between inputs and outcomes were potentially complicated by the degree of patient involvement in producing the intervention. The WSD telehealth and telecare trials were pragmatic, seeking to evaluate the impact of the interventions per se rather than to examine the impacts of different service configurations. While the design enhanced the generalisability of findings to real-world service settings, the pragmatic approach also allowed considerable heterogeneity in the delivery of telehealth and telecare between sites (see 5.7.1 and 5.16.2). In the telehealth trial, each site had different equipment and software procurement, installation, maintenance and monitoring arrangements. Each supplier of equipment and software offered somewhat

different options to the user and to the professionals monitoring vital signs data.[22] In the telecare trial, there were also differences in delivery models, in particular in the availability of dedicated telecare response teams. Nonetheless, in each study, the intervention featured certain core characteristics across all the sites. In the telehealth trial, all sites offered services such as central monitoring teams, patient education protocols and computerised risk-based classification of vital signs data. In the telecare trial, all sites featured call-centres that monitored sensor data and much of the telecare equipment was supplied by one company. The pragmatic design did impose some limitations on interpreting findings. For instance, the variability between sites in the ways in which the telehealth/telecare services were delivered meant that it was not possible to pinpoint which elements of the intervention were influencing outcomes such as HrQOL.

Other information that might be relevant to an economic evaluation was not available – for instance data on sensor alerts and false-alarms from providers' systems, as well as on dedicated response teams' visits in response to sensor alerts. In the case of telecare, information was not available on the telecare assessments that had been conducted and thus the goals of the prescribed telecare package were not known. Without having information on the intended purpose of the telecare package, there was no way to measure the impact of one combination of sensors over another on outcomes and costs. Costing the telehealth and telecare support elements of the intervention at a site level limited the granularity of the intervention support costs. This narrowed the amount of potential variability in intervention costs between participants and could have diminished the sensitivity of the analyses to detecting the impact of the intervention.

Neither evaluation was able to take mortality into account except as one aspect of sample attrition. It should be noted that this could be a consideration in the telehealth economic evaluation, where mortality in the trial sample was found to be significantly lower in the telehealth group (Steventon et al. 2012) (but also see Steventon, Grieve, and Bardsley (2015) and Box 9.1). On the other hand, there was little evidence of differential mortality in the telecare trial (see section 8.4). Also, whether outcomes, particularly in HrQoL would have shown improvements if the telehealth and telecare evaluation periods had been longer, is not known.

[22] Please note that Cartwright et al. (2013) gathered information on the clinical protocols in place in each site to respond to vital signs data from telehealth systems - see section 5.8.

There are several possible limitations to generalisability of these findings. There was substantial loss to follow-up between baseline and 12-month follow-up in both studies (in the Telehealth study 36 per cent in the intervention and 41 per cent in the control group; in the Telecare study, 32 per cent in the intervention and 40 per cent in the control groups). There were also imbalances in the size of the allocation groups, the intervention group being 16 per cent greater than the control group in both cases. It was possible for self-selection to have taken place after cluster-randomisation. There were a few differences within allocation groups in the characteristics of participants not completing 12-month follow-up. In both Telehealth and Telecare studies there were differences in proportions completing by site, also in terms of educational qualifications. Differences in characteristics between completing and non-completing samples appeared to be concentrated in the control group in the Telecare sample, whereas such differences were more concentrated in the intervention group in the Telehealth sample. The cost-effectiveness analyses adjusted for a number of baseline demographic and cost covariates that might influence the decision to complete the 12-month follow-up, to some extent mitigating imbalances between intervention and control groups caused by attrition.

The generalisabilty of these findings to other health and social care system settings and populations should be carefully considered. Much of the telehealth evidence base has emerged from the US. Free access to health care in the UK may lead to better access to primary care than would be the case for comparable telehealth users in the US, leaving less scope here for decreasing use of costlier secondary care.

Lastly, the costs of the telehealth and telecare interventions estimated in this thesis were based on systems in place in 2009/10. Telehealth and, to a lesser extent telecare, markets are expanding and prices of systems have been falling. Thus the costs presented here will not perfectly reflect current market conditions.

9.7 Implications for Policy and Practice

9.7.2 Telehealth

The cost-effectiveness results here suggested that the second-generation form of telehealth implemented in the WSD study – a telemonitoring approach with some elements of educational and behavioural support in addition to standard support and treatment – was slightly more effective and more costly than standard support and treatment. In spite of differences in service configurations, the sites all featured monitoring centres staffed by

qualified nurses and a range of peripherals that interfaced with some kind of base unit. This kind of telehealth was bound to be costlier than some other configurations of telehealth (Cottrell, Chambers, and O'Connell 2012, NHS Stoke-on-Trent 2011, Dixon, Hollinghurst, Edwards, Thomas, Gaunt, et al. 2016) because of the use of qualified staff and peripherals capable of transmitting data automatically. An evident lesson for policy and research would be to consider whether this form of telehealth service should be reserved for people with more severe conditions, while offering other, simpler forms of telemonitoring for routine use by patients' usual primary care providers. Evidence from other studies also suggests that targeting specific chronic health conditions in combination with specific clinical characteristics may be needed (e.g. people with very stable or very unstable readings may not benefit from monitoring (Vassilev et al. 2015)).

The evidence of the cost-effectiveness study should be used comparatively by commissioners to evaluate the benefits of telehealth for local populations, against other forms of disease management, such as self-management interventions. It will be important to keep the falling costs of equipment in mind in considering these findings – but that said, device costs are not the only costs involved in telemonitoring.

The wider implications for policy involve recognising that if there are improvements to quality of life that result from investments in telehealth across the health and social care systems, it is important to deploy mechanisms to re-invest funds from sectors that most benefit from telehealth (secondary care) to primary and community health services.

9.7.3 Telecare

The WSD questionnaire and trial studies provided evidence that second-generation telecare does not represent a panacea for the budgetary and demographic challenges facing policymakers or commissioners. This assessment is much in line with those of other commentators (Poole 2006, Glasby, Lynch, and Robinson 2018, Greenhalgh et al. 2013). Nonetheless, this research cannot be interpreted as evidence that *first-generation* telecare is ineffective. Also, patterns of telecare uptake in the general older population (Nyman and Victor 2014) suggest that outcomes and costs could be quite different than indicated here. For instance, people seeking to take up telecare, despite quite substantial user charges or private fees, may derive some benefit not seen in the WSD sample, who accepted the service free of charge.

9.7.4 Telehealth and Telecare Technologies

Given the passage of time since the WSD ended, the trajectory of government policy from that time is worth pointing out. WSD generated a great deal of evidence, but it took several years for that evidence to emerge into the public domain. Ettelt, Mays, and Allen (2015) included the WSD in their analysis of English 'policy experiments' (alongside the Individual Budgets and Partnerships for Older People Pilots). These experiments all in their separate ways threw up difficulties for policymakers who had commissioned the research as proof-of-concept, where the research subsequently did not provide the required proof. The history of the policy and press reactions to the findings of the WSD research suggests that "research rigour did not translate into policy impact" (Ettelt, Mays, and Allen 2015 p.303) as might have been hoped.

A recent report based on a survey of telecare managers in English local authorities gives some insight into the social care policy response to the evidence of the WSD telecare trial and questionnaire study (Woolham et al. 2018). Its authors observe that "[…] the WSD findings do not seem to have influenced local authorities and policy makers. The WSD remains an important study and its neglect is curious" (p.8). They report that 47 per cent of respondents asked were aware of the WSD findings. Respondents' opinions of the WSD findings were largely negative for reasons such as: these did not chime with their own local experience; that these were undermining good local work; that the study had been flawed or was outdated; or that telecare organisations did not trust them.

9.8 Implications for Research

The outcomes presented across economic evaluation studies of telehealth and (insofar as they exist) telecare have been multifarious (Vergara Rojas and Gagnon 2008, Bergmo 2009). The ability of future researchers and clinicians to evaluate the impact of telehealth and telecare in the light of the available evidence depends on the relevance, appropriateness and consistent measurement of outcomes (Gargon et al. 2014, Williamson et al. 2012). The wide variety of clinical applications of telehealth pose a challenge to consistency in the choice of trial outcomes, as there could be many possible clinical outcomes related to each condition. However there are methods for developing consistent sets of outcomes, or 'core outcome sets', for use in evaluating the effectiveness of interventions. Dodd et al. (2018) have proposed a taxonomy of outcomes to assist in consistent reporting of clinical trials and improve the efficiency of searching knowledge sources for systematic reviews and clinical

research. In this taxonomy, 38 outcome domains cover five core areas: death, physiological/clinical, life impacts, resource use and adverse events.

In the thesis, I focused on the relationship between health and social care costs and outcomes of the trial interventions. Core outcome sets related to telehealth or telecare of the kind deployed in the WSD trials do not currently appear in the COMET database (COMET Initiative 2018).[23] The outcomes examined within this thesis, alongside other studies and reviews discussed in Chapter 3 and this chapter, can be considered by future developers of core outcome sets for telehealth and telecare and thus can inform future economic evaluations and systematic reviews. Establishing core outcome sets for telehealth and telecare evaluations would extend the usefulness of trial data in this area. The outcome domains to be covered in these sets are likely to share some outcome domains in the core areas of death, life impacts, resource use and adverse events but to vary between telecare and telehealth in terms of outcome domains in the 'physiological/clinical' area. It would be important to first determine the scope of the outcome set in terms of the condition and population of interest, trial settings and intervention (Williamson et al. 2017), given the ambiguities in terminology used to describe telehealth and telecare as discussed in Chapter 2. Consideration should be given to delineating the type of 'telehealth' system to be examined, for instance whether the system operates on a store-and-forward and/or real-time basis, the communication technologies used (telephone, video, internet) and to the 'generation' of telecare systems. Settings should be defined: whether the telehealth system is employed primarily in secondary or primary care or across clinical settings; whether it is set up exclusively in hospitals/clinics or in the user's home; and whether the telecare system is used in the home, the wider community or in communal settings.

I end with some recommendations for future research specific to telehealth and telecare.

9.8.1 Telehealth

As discussed in chapter 2, there is a developing literature on the role of telehealth in disease management and particularly in self-management. Specific models of TH delivery should be investigated to understand their relationship with variations in outcomes and costs (McLean et al. 2013). I discussed literature (see Chapter 2, Section 2.6.4) suggesting that telehealth can

[23] A core outcome set exists for tele-emergency care, a telemedicine intervention operating in emergency department settings (Harris et al. 2017).

facilitate the complete feedback loop, that a (second-generation) telehealth system makes the condition visible to the patient, and that patients may benefit from seeing changes in vital signs graphically presented. However the components of telehealth that work to support self-management are not well described in existing studies (Hanlon et al. 2017). More granular information on the components of the telehealth intervention would be useful in order to understand which components are related to which kinds of outcomes (for instance self-monitoring might improve a person's sense of mastery over the condition, while receiving support by telephone might help with managing symptoms (Rixon et al. 2015)). In particular, it would be helpful to understand variations in frequency and intensity of response to telemonitoring triggers, data that were not collected within the telehealth trial, to investigate impacts on quality of life outcomes (Newman and Whole System Demonstrator Programme Evaluation Team 2014). An analysis of these data across several long-term conditions would shed light on whether there are variations in outcomes and key-event related costs by condition.

9.8.2 Telecare

Research should be conducted to investigate the costs of services used by recipients of telecare with different living arrangements (living alone and with others). This could usefully be combined with analysis of data from routine care settings rather than from experimental studies (Momanyi 2017a, b). The WSD telecare questionnaire and trial studies were not able to examine the question of targeting telecare to particular types of social care need or service population, nor the question of the intended purpose (at the assessment stage) of the telecare package: these should be addressed in future studies. Research into the effectiveness and cost-effectiveness of telecare in people with significant cognitive impairment is needed (cf. Leroi et al. 2013).

Lastly, the cost-effectiveness of telecare analysis raised some questions for further research. The WSD Telecare Study was not set up to examine by what means telecare systems act to improve outcomes. For instance, research could operationalise the concept of 'reassurance' and investigate whether this mediates or moderates the impact of telecare on quality of life. Another question that could address the link between telecare, utility and costs would be to examine whether targeting of telecare could be improved by assessing individuals' attitude to risk, the extent of risk-aversion and associations with health-related quality of life and use of health and social care services. Another question was left

unanswered by this study and deserves investigation. What is the *additional* benefit of remote sensors, above and beyond pendant alarms? For instance, future research could compare health-related quality of life and well-being provided by additional sensors (second-generation telecare), to that provided by simpler, first-generation forms of telecare.

Reference List

Aceros, Juan C., Jeannette Pols, and Miquel Domènech (2015) "Where is grandma? Home telecare, good aging and the domestication of later life" *Technological Forecasting and Social Change*, 93 (Supplement C):102-111. doi: https://doi.org/10.1016/j.techfore.2014.01.016.

Agboola, S., T. M. Hale, C. Masters, J. Kvedar, and K. Jethwani (2014) ""Real-world" practical evaluation strategies: a review of telehealth evaluation" *JMIR Research Protocols*, 3 (4):e75. doi: 10.2196/resprot.3459.

Agboola, Stephen, Sara Golas, Nils Fischer, Mariana Nikolova-Simons, Jorn op den Buijs, Linda Schertzer, Joseph Kvedar, and Kamal Jethwani (2017) "Healthcare utilization in older patients using personal emergency response systems: an analysis of electronic health records and medical alert data" *BMC Health Services Research*, 17 (1):282. doi: https://doi.org/10.1186/s12913-017-2196-1.

Age UK. (2016). "1.2m older people don't get the social care they need." Last Modified 17 November 2016, accessed 15-4-2018. https://www.ageuk.org.uk/latest-news/archive/12m-older-people-dont-get-the-social-care-they-need/.

Aldrich, F. (2003) "Smart homes past present and future" in R. Harper, eds. *Inside the smart home*. 17-39. London: Springer.

Allen, Caroline, and Jennifer Beecham (1993) "Costing Services: Ideals and Reality" in Ann Netten and Jennifer Beecham, eds. *Costing Community Care: Theory and Practice*. Avebury: Ashgate.

Andresen, E. M., J. A. Malmgren, W. B. Carter, and D. L. Patrick (1994) "Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale)" *American Journal of Preventive Medicine*, 10 (2):77-84.

Anker, S. D., F. Koehler, and W. T. Abraham (2011) "Telemedicine and remote management of patients with heart failure" *Lancet*, 378 (9792):731-9. doi: https://doi.org/10.1016/S0140-6736(11)61229-4.

Appleby, John. (2013) *Spending on health and social care over the next 50 years. Why think long term?*: The King's Fund.

Appleby, John, Rowena Crawford, and Carl Emmerson. (2009). *How cold will it be? Prospects for NHS funding: 2011-17.* London: King's Fund and Institute for Fiscal Studies.

Appleby, John, and Richard Humphries. (2010). *Estimate of future social care funding gap following Spending Review 2010. Supplementary written evidence from the King's Fund (PEX 12A). Health Select Committee. Public Expenditure.*

Astolfi, Roberto, Luca Lorenzoni, and Jillian Oderkirk (2012) "Informing policy makers about future health spending: A comparative analysis of forecasting methods in OECD countries" *Health Policy*, 107 (1):1-10. doi: http://dx.doi.org/10.1016/j.healthpol.2012.05.001.

Baldwin, S. A., G. W. Fellingham, and A. S. Baldwin (2016) "Statistical models for multilevel skewed physical activity data in health research and behavioral medicine" *Health Psychology*, 35 (6):552-62. doi: 10.1037/hea0000292.

Balk, A. H., W. Davidse, Pv Dommelen, E. Klaassen, K. Caliskan, P. van der Burgh, and C. M. Leenders (2008) "Tele-guidance of chronic heart failure patients enhances knowledge about the disease. A multi-centre, randomised controlled study" *European Journal of Heart Failure*, 10 (11):1136-42. doi: 10.1016/j.ejheart.2008.08.003.

Balk, Ahmm, C. M. Leenders, W. Davidse, C. Westerteicher, and G. Montfort Van (2007) "Personalized tele-guidance of heart failure patients. Effects of the MOTIVA interactive health care platform on hospital admissions, quality of life, knowledge of disease and self-care. A pilot study" *European Journal of Heart Failure*, 6:56-57.

Balta-Ozkan, Nazmiye, Rosemary Davidson, Martha Bicket, and Lorraine Whitmarsh (2013) "Social barriers to the adoption of smart homes" *Energy Policy*, 63:363-374. doi: 10.1016/J.ENPOL.2013.08.043.

Bardsley, Martin, Adam Steventon, and Helen Doll (2013) "Impact of telehealth on general practice contacts: findings from the whole systems demonstrator cluster randomised trial" *BMC Health Services Research*, 13 (1):395.

Barlow, J, S Bayer, and R Curry (2006) "Implementing complex innovations in fluid multi-stakeholder environments: experiences of 'telecare'" *Technovation*, 26:396 - 406.

Barlow, James, Richard Curry, Theti Chrysanthaki, Jane Hendy, and Nael Taher. (2012). *Developing the capacity of the remote care industry to supply Britain's future needs.* London: the Health and Care Research and Innovation Centre.

Barlow, James, Debbie Singh, Steffen Bayer, and Richard Curry (2007) "A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions" *J Telemed Telecare*, 13 (4):172-179. doi: 10.1258/135763307780908058.

Bartholomew, David J., Fiona Steele, Irini Moustaki, and Jane Galbraith. (2008) *Analysis of multivariate social science data*. 2nd ed: Routledge, U.K.

Bate, Alex. (2017). *Health and Social Care Integration.* London: House of Commons Library.

Bates, D., M. Machler, B. M. Bolker, and S. C. Walker (2015) "Fitting Linear Mixed-Effects Models Using lme4" *Journal of Statistical Software*, 67 (1):1-48.

Bayer, Steffen, and J Barlow. (2010). *Analysis of data on the impact of telecare provided under Telecare Capital Grant.* London: Imperial College Business School.

Beale, Sophie, Paul Truman, Diana Sanderson, and Jen Kruger (2010) "The Initial Evaluation of the Scottish Telecare Development Program" *Journal of Technology in Human Services*, 28 (1-2):60-73. doi: 10.1080/15228831003770767.

Beecham, J. K., and M. R. J. Knapp (2001) "Costing psychiatric interventions" in G. Thornicroft, C. Brewin and J. K. Wing, eds. *Measuring Mental Health Needs*. 220-224. London: Gaskell.

Beecham, Jennifer. (2000) *Unit Costs - Not Exactly Child's Play*. London: Department of Health, Personal Social Services Research Unit and Dartington Social Care Research Unit.

Bensink, Mark, David Hailey, and Richard Wootton (2006) "A systematic review of successes and failures in home telehealth: preliminary results" *Journal of Telemedicine and Telecare*, 12 (suppl_3):8-16. doi: 10.1258/135763306779380174.

Bentley, Claire L., Lauren A. Powell, Alison Orrell, and Gail A. Mountain (2016) "Making Telecare desirable rather than a last resort" *Ageing and Society*:1-28. doi: 10.1017/S0144686X16001355.

Bergmo, T (2009) "Can economic evaluation in telemedicine be trusted? A systematic review of the literature" *Cost Effectiveness and Resource Allocation,* 7:18.

Bhandari, A., and T. Wagner (2006) "Self-reported utilization of health care services: improving measurement and accuracy" *Medical Care Research and Review,* 63 (2):217-35.

Biddiss, E., S. Brownsell, and M. S. Hawley (2009) "Predicting need for intervention in individuals with congestive heart failure using a home-based telecare system" *Journal of Telemedicine and Telecare,* 15 (5):226-31. doi: 10.1258/jtt.2009.081203.

Blendon, Robert J., Cathy Schoen, Catherine DesRoches, Robin Osborn, and Kinga Zapert (2003) "Common Concerns Amid Diverse Systems: Health Care Experiences In Five Countries" *Health Affairs,* 22 (3):106-121. doi: 10.1377/hlthaff.22.3.106.

Bower, P, M Cartwright, SP Hirani, J Barlow, J Hendy, M Knapp, C Henderson, A Rogers, C Sanders, M Bardsley, A Steventon, R Fitzpatrick, H Doll, and S Newman (2011) "A comprehensive evaluation of the impact of telemonitoring in patients with long-term conditions and social care needs: protocol for the Whole System Demonstrator cluster randomised trial" *BMC Health Services Research*, 11:184.

Bowes, Alison, and Gillian McColgan. (2006). *Smart technology and community care for older people: innovation in West Lothian, Scotland.* Edinburgh: University of Stirling.

Brazier, J. E., and J. Roberts (2004) "The estimation of a preference-based measure of health from the SF-12" *Medical Care*, 42 (9):851-9. doi: 00005650-200409000-00004 [pii].

Brazier, John. (2007) *Measuring and valuing health benefits for economic evaluation*. Oxford; New York: Oxford University Press.

Brooks, Richard (1996) "EuroQol: the current state of play" *Health Policy,* 37 (1):53-72. doi: 10.1016/0168-8510(96)00822-6.

Brouwer, Werner B. F., Anthony J. Culyer, N. Job A. van Exel, and Frans F. H. Rutten (2008) "Welfarism vs. extra-welfarism" *Journal of Health Economics*, 27 (2):325-338. doi: https://doi.org/10.1016/j.jhealeco.2007.07.003.

Brouwer, Werner B. F., and Marc A. Koopmanschap (2000) "On the economic foundations of CEA. Ladies and gentlemen, take your positions!" *Journal of Health Economics*, 19 (4):439-459. doi: https://doi.org/10.1016/S0167-6296(99)00038-7.

Brown, Randall, Deborah Peikes, Arnold Chen, Judy Ng, Jennifer Schore, and Clara Soh. (2007). *The Evaluation of the Medicare Coordinated Care Demonstration: Findings for the First Two Years*. Princeton, New Jersey: Mathematica Policy Research, Inc.

Brownsell, S., S. Blackburn, and M. S. Hawley (2008) "An evaluation of second and third generation telecare services in older people's housing" *Journal of Telemedicine and Telecare*, 14 (1):8-12. doi: 10.1258/jtt.2007.070410.

Brownsell, S., D. Bradley, S. Blackburn, F. Cardinaux, and M. S. Hawley (2011) "A systematic review of lifestyle monitoring technologies" *Journal of Telemedicine and Telecare*, 17 (4):185-9. doi: jtt.2010.100803.

Buchanan, James, and Sarah Wordsworth (2015) "Welfarism Versus Extra-Welfarism: Can the Choice of Economic Evaluation Approach Impact on the Adoption Decisions Recommended by Economic Evaluation Studies?" *PharmacoEconomics*, 33 (6):571-579. doi: 10.1007/s40273-015-0261-3.

Byford, S., M. Leese, M. Knapp, H. Seivewright, S. Cameron, V. Jones, K. Davidson, and P. Tyrer (2007) "Comparison of alternative methods of collection of service use data for the economic evaluation of health care interventions" *Health Economics*, 16 (5):531-536.

Byford, Sarah, and Tom Sefton (2003) "Economic Evaluation of Complex Health and Social Care Interventions" *National Institute Economic Review*, 186 (1):98-108. doi: 10.1177/002795010300100114.

Cameron, Adrian Colin, and P. K. Trivedi. (2005) *Microeconometrics : methods and applications*. Cambridge ; New York: Cambridge University Press.

Cameron, Kirsteen Sarah. 2006. "Needs-led assessment in health and social care : a community-based comparative study." PhD, College of Medical Veterinary and Life Sciences, University of Glasgow.

Carpenter, James, and Mike Kenward. (2007). *Missing data in randomised controlled trials — a practical guide.* London: London School of Hygiene.

Cartwright, M., S. P. Hirani, L. Rixon, M. Beynon, H. Doll, P. Bower, M. Bardsley, A. Steventon, M. Knapp, C. Henderson, A. Rogers, C. Sanders, R. Fitzpatrick, J. Barlow, and S. P. Newman (2013) "Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial" *BMJ*, 346:f653. doi: https://doi.org/10.1136/bmj.f653.

Cerri, Karin H., Martin Knapp, and Jose-Luis Fernandez (2013) "Decision making by NICE: examining the influences of evidence, process and context" *Health Economics, Policy and Law*, FirstView:1-23. doi: doi:10.1017/S1744133113000030.

Chaudhry, S. I., J. A. Mattera, J. P. Curtis, J. A. Spertus, J. Herrin, Z. Lin, C. O. Phillips, B. V. Hodshon, L. S. Cooper, and H. M. Krumholz (2010) "Telemonitoring in patients with heart failure" *New England Journal of Medicine*, 363 (24):2301-9. doi: 10.1056/NEJMoa1010029.

Clark, Christopher E., Lindsay F. P. Smith, Rod S. Taylor, and John L. Campbell (2010) "Nurse led interventions to improve control of blood pressure in people with hypertension: systematic review and meta-analysis" *BMJ (Clinical Research Ed.)*, 341:c3995-c3995.

Coast, J. (2009) "Maximisation in extra-welfarism: A critique of the current position in health economics" *Social Science & Medicine*, 69 (5):786-92. doi: 10.1016/j.socscimed.2009.06.026.

Coast, J., T. N. Flynn, L. Natarajan, K. Sproston, J. Lewis, J. J. Louviere, and T. J. Peters (2008) "Valuing the ICECAP capability index for older people" *Social Science & Medicine*, 67 (5):874-82. doi: https://doi.org/10.1016/j.socscimed.2008.05.015.

Coast, J., T. Flynn, E. Sutton, H. Al-Janabi, J. Vosper, S. Lavender, J. Louviere, and T. Peters (2008) "Investigating Choice Experiments for Preferences of Older People (ICEPOP):

evaluative spaces in health economics" *Journal of Health Services Research & Policy*, 13 Suppl 3:31-7. doi: http://dx.doi.org/10.1258/jhsrp.2008.008024.

Coast, Joanna, Richard D. Smith, and Paula Lorgelly (2008) "Welfarism, extra-welfarism and capability: The spread of ideas in health economics" *Social Science & Medicine*, 67 (7):1190-1198. doi: https://doi.org/10.1016/j.socscimed.2008.06.027.

Comas-Herrera, Adelina, Raphael Wittenberg, and Linda Pickard (2010) "The Long Road to Universalism? Recent Developments in the Financing of Long-term Care in England" *Social Policy & Administration*, 44 (4):375-391.

COMET Initiative. (2018). "The COMET (Core Outcome Measures in Effectiveness Trials) Initiative." accessed 25th August 2018. www.comet-initiative.org.

Commission for Social Care Inspection. (2008). *Cutting the cake fairly. CSCI review of eligibility criteria for social care.* Commission for Social Care Inspection.

Commission for Social Care Inspection. (2009). *The State of Social Care in England 2007-08.* Commission for Social Care Inspection.

Cook, Erica J., Gurch Randhawa, Andy Guppy, Chloe Sharp, Garry Barton, Andrew Bateman, and Jane Crawford-White (2017) "Exploring factors that impact the decision to use assistive telecare: perspectives of family care-givers of older people in the United Kingdom" *Ageing and Society*:1-21. doi: 10.1017/S0144686X1700037X.

Corbett-Nolan, Andrew, and John Bullivant. (2012). *Care and support at home: An audit of telecare services in England.* Battle, East Sussex.

Cottrell, Elizabeth, Ruth Chambers, and Phil O'Connell (2012) "Using simple telehealth in primary care to reduce blood pressure: a service evaluation" *BMJ Open*, 2 (6). doi: 10.1136/bmjopen-2012-001391.

Coye, M. J., A. Haselkorn, and S. DeMello (2009) "Remote patient management: technology-enabled innovation and evolving business models for chronic disease care" *Health Affairs (Millwood)*, 28 (1):126-35. doi: http://dx.doi.org/10.1377/hlthaff.28.1.126.

Craig, Peter, Paul Dieppe, Sally Macintyre, Susan Michie, Irwin Nazareth, and Mark Petticrew. (2008). *Developing and evaluating complex interventions: new guidance.* London: Medical Research Council.

Cromarty, Hannah. (2017). *Adult Social Care Funding (England).* London: House of Commons Library.

Culyer, A. J. (1989) "The Normative Economics of Health Care Finance and Provision" *Oxford Review of Economic Policy*, 5 (1):34-58. doi: http://dx.doi.org/10.1093/oxrep/5.1.34.

Currell, R., C. Urquhart, P. Wainwright, and R. Lewis (2000) "Telemedicine versus face to face patient care: effects on professional practice and health care outcomes" *Cochrane Database of Systematic Reviews*, (2):CD002098. doi: http://dx.doi.org/10.1002/14651858.CD002098.

Curtis, Lesley, ed. 2010. *Unit Costs of Health and Social Care 2010*. Canterbury: Personal Social Services Research Unit.

Cutler, David M. (2002) "Equality, Efficiency, and Market Fundamentals: The Dynamics of International Medical-Care Reform" *Journal of Economic Literature*, 40 (3):881-906.

Cylus, Jonathan, Erica Richardson, Lisa Findley, Marcus Longley, Ciaran O'Neill, and David Steel (2015) "United Kingdom: Health system review" *Health Systems in Transition*, 17 (5):1-125.

Darnell, K., A. K. Dwivedi, Z. Weng, and R. J. Panos (2013) "Disproportionate utilization of healthcare resources among veterans with COPD: a retrospective analysis of factors associated with COPD healthcare cost" *Cost Effectiveness and Resource Allocation*, 11:13. doi: 10.1186/1478-7547-11-13.

Das, Saudamini, and Stephen C. Smith. (2012). *Awareness as an Adaptation Strategy for Reducing Mortality from Heat Waves: Evidence from a Disaster Risk Management Program in India.* Kathmandu, Nepal: South Asian Network for Development and Environmental Economics (SANDEE).

Datta, Santanu K., Eugene Z. Oddone, Maren K. Olsen, Melinda Orr, Felicia McCant, Pam Gentry, and Hayden B. Bosworth (2010) "Economic analysis of a tailored behavioral

intervention to improve blood pressure control for primary care patients" *American Heart Journal*, 160 (2):257-263. doi: 10.1016/j.ahj.2010.05.024.

Davies, Anna, Lorna Rixon, and Stanton Newman (2013) "Systematic review of the effects of telecare provided for a person with social care needs on outcomes for their informal carers" *Health & Social Care in the Community*, 21 (6):582-597. doi: 10.1111/hsc.12035.

Davies, Bleddyn, and Martin Knapp. (1981) *Old people's homes and the production of welfare*. London: Routledge & Kegan Paul.

Davis, Fred D. (1989) "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology" *MIS Quarterly*, 13 (3):319-340. doi: 10.2307/249008.

Davis, Fred D., Richard P. Bagozzi, and Paul R. Warshaw (1989) "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models" *Management Science*, 35 (8):982-1003.

Davis, J. C., T. Liu-Ambrose, C. G. Richardson, and S. Bryan (2012) "A comparison of the ICECAP-O with EQ-5D in a falls prevention clinical setting: are they complements or substitutes?" *Qual Life Res*. doi: 10.1007/s11136-012-0225-4.

de Miguel Diez, J., P. Carrasco Garrido, M. Garcia Carballo, A. Gil de Miguel, J. Rejas Gutierrez, J. M. Bellon Cano, V. Hernandez Barrera, and R. Jimenez Garcia (2008) "Determinants and predictors of the cost of COPD in primary care: a Spanish perspective" *International Journal of Chronic Obstructive Pulmonary Disease*, 3 (4):701-12.

De San Miguel, Kirsten, Gill Lewin, Elissa Burton, Christine Toye, Duncan Boldy, and Peter Howat (2015) "Exploring risk profiles and emergency frequency of purchasers and non-purchasers of personal emergency alarms: a prospective cohort study" *BMC Geriatrics*, 15 (1):140. doi: doi:10.1186/s12877-015-0139-4.

De San Miguel, Kristen, G. Lewin, E.L. Burton, P. Howat, D. Boldy, and C. Toye (2017a) "Personal emergency alarms: do health outcomes differ for purchasers and nonpurchasers?" *Home Health Care Services Quarterly*, 36 (3-4):164-177. doi: 10.1080/01621424.2017.1373718.

De San Miguel, Kristen, and Gill Lewin (2008) "Personal emergency alarms: What impact do they have on older people's lives?" *Australasian Journal on Ageing*, 27 (2):103-105. doi: 10.1111/j.1741-6612.2008.00286.x.

De San Miguel, Kristen, Gill Lewin, E lissa Burton, Peter Howat, Duncan Boldy, and Christine Toye (2017b) "Personal emergency alarms: do health outcomes differ for purchasers and nonpurchasers?" *https://doi.org/10.1080/01621424.2017.1373718*. doi: 10.1080/01621424.2017.1373718.

Deloitte Centre for Health Solutions. (2012). *Primary care: Working differently. Telecare and telehealth - a game changer for health and social care.* London: Deloitte Centre for Health Solutions.

Demiris, G., and B. K. Hensel (2008) "Technologies for an aging society: a systematic review of "smart home" applications" *Yearbook of medical informatics*:33-40.

Department of Health. (2001). "Equipment." [Website]. Department of Health, accessed 12th April 2012. http://www.info.doh.gov.uk/doh/finman.nsf/072561aa006322660725618c006b09a0/ac57fae9494e33c400256a42005c19bc?OpenDocument.

Department of Health. (2005a). *Building Telecare in England*

Department of Health. (2005b). *Independence, Well-being and Choice: Our vision for the future of social care for adults in England.*

Department of Health. (2005c). *Strategic Business Case: Models for Telecare.*

Department of Health. (2005d). *Supporting People with Long Term Conditions. An NHS and Social Care Model to support local innovation and integration.*

Department of Health. (2006). *Our health our care our say: a new direction for community services.*

Department of Health. (2007). "White Paper pilots: whole system long term conditions (LTC) demonstrator programme." accessed September 28, 2008.

Department of Health. (2008a). "Individual budgets." [Web page], Last Modified 12 February 2008 accessed 20th October 2008. http://www.dh.gov.uk/en/SocialCare/Socialcarereform/Personalisation/Individualbudgets/DH_4125774.

Department of Health. (2008b). *Raising the Profile of Long Term Conditions Care: a compendium of information.* . London: Her Majesty's Stationery Office (HMSO).

Department of Health. (2008c). *Transforming Social Care. Local Authority Circular.* London: Department of Health.

Department of Health. 2009. General opthalmic services: increases to NHS sight test fee and NHS optical voucher values from 1 April 2009.

Department of Health. (2010a). *Direct Payments for Health Care: Information for pilot sites.* London: Department of Health,.

Department of Health. (2010b). *A Vision for Adult Social Care: Capable Communities and Active Citizens.* London: Department of Health.

Department of Health. (2011). *Reference Costs 2009-10 Publication.* London: Department of Health.

Department of Health. (2012a). *A concordat between the Department of Health and the telehealth and telecare industry.* London.

Department of Health. (2012b). *Long Term Conditions Compendium of Information: Third Edition.*

Department of Health. (2015). *Policy paper: 2010 to 2015 government policy: long term health conditions.*

Department of Health. 2016. Factsheet 1: General responsibilities of local authorities: prevention, information and advice, and shaping the market of care and support services.

Department of Health Care Services Efficiency Delivery Programme (2010). "Transforming Community Equipment Services national catalogue and tariff for Simple Aids to Daily Living." accessed 19/2/2011.

Department of Health Change Agent Team. (2005). *Telecare - Getting Started.* Change Agent Team and the Integrating Community Equipment Services (ICES).

Diaz-Ordaz, K., M. G. Kenward, and R. Grieve (2014) "Handling missing values in cost effectiveness analyses that use data from cluster randomized trials" *Journal of the Royal Statistical Society Series a-Statistics in Society*, 177 (2):457-474.

Dixon, P., S. Hollinghurst, L. Edwards, C. Thomas, A. Foster, B. Davies, D. Gaunt, A. A. Montgomery, and C. Salisbury (2016) "Cost-effectiveness of telehealth for patients with depression: evidence from the Healthlines randomised controlled trial" *BJPsych Open*, 2 (4):262-269. doi: 10.1192/bjpo.bp.116.002907.

Dixon, P., S. Hollinghurst, L. Edwards, C. Thomas, D. Gaunt, A. Foster, S. Large, A. A. Montgomery, and C. Salisbury (2016) "Cost-effectiveness of telehealth for patients with raised cardiovascular disease risk: evidence from the Healthlines randomised controlled trial" *BMJ Open*, 6 (8):e012352. doi: 10.1136/bmjopen-2016-012352.

Dodd, S., M. Clarke, L. Becker, C. Mavergames, R. Fish, and P. R. Williamson (2018) "A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery" *Journal of Clinical Epidemiology*, 96:84-92. doi: 10.1016/j.jclinepi.2017.12.020.

Dolan, Paul (1997) "Modeling Valuations for EuroQol Health States" *Medical Care*, 35 (11):1095-1108.

Dolan, Paul, Claire Gudex, Paul Kind, and Alan Williams. (1995). *A social tariff for EuroQol: Results from a UK general population survey.* York: University of York.

Donner, Allan, and Neil Klar. (2000) *Design and analysis of cluster randomization trials in health research*. London: Arnold.

Doughty, K., A. Monk, C. Bayliss, S. Brown, L. Dewsbury, B. Dunk, V. Gallagher, K. Grafham, M. Jones, C. Lowe, L. McAlister, K. McSorley, P. Mills, C. Skidmore, A. Stewart, B. Taylor, and D. Ward (2008) "Telecare, telehealth and assistive technologies: do we know what we're talking about?" *Housing, Care & Support*, 11 (3):36-41.

Doughty, K., and C. Steele (2009) "The use of telecare and community equipment to improve joint working across health and social care" *Journal of Assistive Technologies*, 3 (4):36-43.

Drummond, Michael, Bernie O'Brien, Greg L. Stoddart, and George W. Torrance. (1999) *Methods for the Economic Evaluation of Health Care Programmes* second ed. Oxford Oxford University Press.

Drummond, Michael, Mark J. Sculpher, Karl Claxton, G. L. Stoddart, and George W. Torrance. (2015) *Methods for the economic evaluation of health care programmes*. Fourth edition / Michael F. Drummond, Mark J. Sculpher, Karl Claxton, Greg L. Stoddart, George W. Torrance. ed.

Drummond, Michael, Mark J. Sculpher, George W. Torrance, Bernie O'Brien, and Greg L. Stoddart. (2005) *Methods for the Economic Evaluation of Health Care Programmes* Third ed. Oxford: Oxford University Press.

Duan, Naihua, Willard G. Manning, Carl N. Morris, and Joseph P. Newhouse (1983) "A Comparison of Alternative Models for the Demand for Medical Care" *Journal of Business & Economic Statistics*, 1 (2):115-126. doi: 10.1080/07350015.1983.10509330.

Dunatchik, Allison, Rossella Icardi, Caireen Roberts, and Margaret Blake. (2016). *Predicting unmet social care needs and links with well-being: Findings from the secondary analysis.* London: NatCen Social Research & Ipsos MORI.

Elliott, S., J. Painter, and S. Hudson (2009) "Living alone and fall risk factors in community-dwelling middle age and older adults" *Journal of Community Health*, 34 (4):301-10. doi: 10.1007/s10900-009-9152-x.

Ettelt, Stefanie, Nicholas Mays, and Pauline Allen (2015) "Policy experiments: Investigating effectiveness or confirming direction?" *Evaluation*, 21 (3):292-307. doi: 10.1177/1356389015590737.

EuroQol Group (1990) "EuroQol-a new facility for the measurement of health-related quality of life" *Health Policy*, 16 (3):199-208.

Falkingham, Jane, Maria Evandrou, Teresa McGowan, David Bell, and Alison Bowes. (2010). *Demographic issues, projections and trends: Older people with high support needs in the UK.* Southampton: ESRC Centre for Population Change, University of Southampton and University of Stirling.

Fallis, Wendy M., Diane Silverthorne, Jonathon Franklin, and Susan McClement (2008) "Client and Responder Perceptions of a Personal Emergency Response System: Lifeline" *http://dx.doi.org/10.1300/J027v26n03_01*. doi: Home Health Care Services Quarterly, Vol. 26, No. 3, 2007: pp. 1–21.

Fieberg, John, Randall H. Rieger, Michael C. Zicus, and Jonathan S. Schildcrout (2009) "Regression modelling of correlated data in ecology: subject-specific and population averaged response patterns" *Journal of Applied Ecology*, 46 (5):1018-1025. doi: 10.1111/j.1365-2664.2009.01692.x.

Fieller, E. C. (1954) "Some Problems in Interval Estimation" *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 16 (2):175-185.

Fleming, J., and C. Brayne (2008) "Inability to get up after falling, subsequent time on floor, and summoning help: prospective cohort study in people over 90" *BMJ*, 337:a2227. doi: https://doi.org/10.1136/bmj.a2227.

Forder, J. (2009) "Long-term care and hospital utilisation by older people: an analysis of substitution rates" *Health Economics*, 18 (11):1322-38. doi: 10.1002/hec.1438.

Forder, Julien. (2007). *Self-funded social care for older people: an analysis of eligibility, variations and future projections.* Canterbury: PSSRU.

Forder, Julien E., and James Caiels (2011) "Measuring the outcomes of long-term care" *Social Science & Medicine*, 73 (12):1766-1774. doi: http://dx.doi.org/10.1016/j.socscimed.2011.09.023.

Garcia-Lizana, F., and I. Munoz-Mayorga (2010) "Telemedicine for depression: a systematic review" *Perspectives in Psychiatric Care*, 46 (2):119-26. doi: https://doi.org/10.1111/j.1744-6163.2010.00247.x.

Garcia-Polo, C., B. Alcazar-Navarrete, L. A. Ruiz-Iturriaga, A. Herrejon, J. A. Ros-Lucas, P. Garcia-Sidro, G. Tirado-Conde, J. L. Lopez-Campos, C. Martinez-Rivera, J. Costan-Galicia, S. Mayoralas-Alises, J. De Miguel-Diez, M. Miravitlles, and Epoc Group In (2012) "Factors associated with high healthcare resource utilisation among COPD patients" *Respiratory Medicine*, 106 (12):1734-42. doi: 10.1016/j.rmed.2012.09.009.

Gargon, E., B. Gurung, N. Medley, D. G. Altman, J. M. Blazeby, M. Clarke, and P. R. Williamson (2014) "Choosing important health outcomes for comparative effectiveness research: a systematic review" *PLoS One*, 9 (6):e99111. doi: 10.1371/journal.pone.0099111.

GBD 2016 DALYs and Hale Collaborators (2017) "Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016" *Lancet*, 390 (10100):1260-1344. doi: 10.1016/S0140-6736(17)32130-X.

Gee, M. Perry, A. Deborah Greenwood, A. Debora Paterniti, Deborah Ward, and Soederberg Lisa M. Miller (2015) "The eHealth Enhanced Chronic Care Model: A Theory Derivation Approach" *Journal of Medical Internet Research*, 17 (4):e86. doi: 10.2196/jmir.4067.

Giordano, A., S. Scalvini, E. Zanelli, U. CorrÃ , G. L. Longobardi, V. A. Ricci, P. Baiardi, and F. Glisenti (2009) "Multicenter randomised trial on home-based telemanagement to prevent hospital readmission of patients with chronic heart failure" *International Journal Of Cardiology*, 131 (2):192-199.

Glasby, Jon, Jennifer Lynch, and Suzanne Robinson (2018) "If telecare is the answer, what was the question? Storylines, tensions and the unintended consequences of technology-supported care" *Critical Social Policy*.

Glick, Henry. (2007) *Economic evaluation in clinical trials*, *Handbooks in health economic evaluation series*. Oxford ; New York: Oxford University Press.

Gokalp, H., and M. Clarke (2013) "Monitoring Activities of Daily Living of the Elderly and the Potential for Its Use in Telecare and Telehealth: A Review" *Telemedicine and E-Health*, 19 (12):910-923. doi: 10.1089/tmj.2013.0109.

Golant, S. M. (2017) "A theoretical model to explain the smart technology adoption behaviors of elder consumers (Elderadopt)" *Journal of Aging Studies*, 42:56-73. doi: 10.1016/j.jaging.2017.07.003.

Gomes, M., K. Diaz-Ordaz, R. Grieve, and M. G. Kenward (2013) "Multiple imputation methods for handling missing data in cost-effectiveness analyses that use data from

hierarchical studies: an application to cluster randomized trials" *Medical Decision Making*, 33 (8):1051-63. doi: 10.1177/0272989X13492203.

Gomes, M., R. Grieve, R. Nixon, and W. J. Edmunds (2012) "Statistical methods for cost-effectiveness analyses that use data from cluster randomized trials: a systematic review and checklist for critical appraisal" *Medical Decision Making*, 32 (1):209-20. doi: https://doi.org/10.1177/0272989X11407341.

Gomes, M., E. S. Ng, R. Grieve, R. Nixon, J. Carpenter, and S. G. Thompson (2012) "Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials" *Medical Decision Making*, 32 (2):350-61. doi: 10.1177/0272989X11418372.

Gray, Alastair M., and Thomas Wilkinson (2016) "Economic evaluation of healthcare interventions: Old and new directions" *Oxford Review of Economic Policy*, 32:102-121. doi: 10.1093/oxrep/grv020.

Graybill, Erin M., Peter McMeekin, and John Wildman (2014) "Can Aging in Place Be Cost Effective? A Systematic Review" *PLoS One*, 9 (7):e102705. doi: 10.1371/journal.pone.0102705.

Greenhalgh, T., J. Wherton, P. Sugarhood, S. Hinder, R. Procter, and R. Stones (2013) "What matters to older people with assisted living needs? A phenomenological analysis of the use and non-use of telehealth and telecare" *Social Science and Medicine*, 93:86-94. doi: 10.1016/j.socscimed.2013.05.036.

Greenhalgh, Trisha, Rob Procter, Joe Wherton, Paul Sugarhood, and Sara Shaw (2012) "The organising vision for telehealth and telecare: Discourse analysis" *BMJ Open*, 2 (4). doi: 10.1136/bmjopen-2012-001574.

Greenhalgh, Trisha, Joseph Wherton, Chrysanthi Papoutsi, Jennifer Lynch, Gemma Hughes, Christine A'Court, Susan Hinder, Nick Fahy, Rob Procter, and Sara Shaw (2017) "Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies" *Journal of Medical Internet Research*, 19 (11). doi: 10.2196/jmir.8775.

Grewal, I., J. Lewis, T. Flynn, J. Brown, J. Bond, and J. Coast (2006) "Developing attributes for a generic quality of life measure for older people: preferences or capabilities?" *Soc Sci Med*, 62 (8):1891-901.

Grieve, Richard, Marina Grishchenko, and John Cairns (2009) "SF-6D versus EQ-5D: reasons for differences in utility scores and impact on reported cost-utility" *The European Journal of Health Economics*, 10 (1):15-23. doi: 10.1007/s10198-008-0097-2.

Guyatt, G. H., Berman, L. B., Townsend, M., Pugsley, S. O., & Chambers, L. W. (1987) "A measure of quality of life for clinical trials in chronic lung disease" *Thorax, ,* 42 (10):773-778.

Guzman-Castillo, M., S. Ahmadi-Abhari, P. Bandosz, S. Capewell, A. Steptoe, A. Singh-Manoux, M. Kivimaki, M. J. Shipley, E. J. Brunner, and M. O'Flaherty (2017) "Forecasted trends in disability and life expectancy in England and Wales up to 2025: a modelling study" *Lancet Public Health*, 2 (7):e307-e313. doi: 10.1016/S2468-2667(17)30091-9.

Guzman, J. S., L. Sohn, and N. D. Harada (2004) "Living alone and outpatient care use by older veterans" *Journal of the American Geriatric Society*, 52 (4):617-22. doi: https://doi.org/10.1111/j.1532-5415.2004.52172.x.

Hanlon, Peter, Luke Daines, Christine Campbell, Brian McKinstry, David Weller, and Hilary Pinnock (2017) "Telehealth Interventions to Support Self-Management of Long-Term Conditions: A Systematic Metareview of Diabetes, Heart Failure, Asthma, Chronic Obstructive Pulmonary Disease, and Cancer" *Journal of Medical Internet Research*, 19 (5):e172. doi: 10.2196/jmir.6688.

Harris, Y., B. Gilman, M. M. Ward, J. Ladinsky, J. Crowley, C. Warren, and C. Caplan (2017) "Building the Evidence Base for Tele-Emergency Care: Efforts to Identify a Standardized Set of Outcome Measures" *Telemedicine and e-Health*, 23 (7):561-566. doi: 10.1089/tmj.2016.0190.

Hastings, S. N., L. K. George, G. G. Fillenbaum, R. S. Park, B. M. Burchett, and K. E. Schmader (2008) "Does lack of social support lead to more ED visits for older adults?" *Am J Emerg Med*, 26 (4):454-61. doi: 10.1016/j.ajem.2007.07.005.

Hawley-Hague, H., E. Boulton, A. Hall, K. Pfeiffer, and C. Todd (2014) "Older adults' perceptions of technologies aimed at falls prevention, detection or monitoring: a systematic review" *International Journal of Medical Informatics*, 83 (6):416-26. doi: https://doi.org/10.1016/j.ijmedinf.2014.03.002.

Hawton, A., C. Green, A. P. Dickens, S. H. Richards, R. S. Taylor, R. Edwards, C. J. Greaves, and J. L. Campbell (2011) "The impact of social isolation on the health status and health-related quality of life of older people" *Quality of Life Research*, 20 (1):57-67. doi: 10.1007/s11136-010-9717-2.

Haywood, K. L., A. M. Garratt, and R. Fitzpatrick (2005) "Quality of life in older people: A structured review of generic self-assessed health instruments" *Quality of Life Research*, 14 (7):1651-1668.

Heagerty, Patrick J., and Brenda F. Kurland (2001) "Misspecified maximum likelihood estimates and generalised linear mixed models" *Biometrika*, 88 (4):973-985. doi: 10.1093/biomet/88.4.973.

Health and Social Care Information Centre, Prescribing Support Unit. (2011). *Prescription Cost Analysis, England 2010.*

Henderson, C., M. Knapp, J. L. Fernandez, J. Beecham, S. P. Hirani, M. Beynon, M. Cartwright, L. Rixon, H. Doll, P. Bower, A. Steventon, A. Rogers, R. Fitzpatrick, J. Barlow, M. Bardsley, and S. P. Newman (2014) "Cost-effectiveness of telecare for people with social care needs: the Whole Systems Demonstrator cluster randomised trial" *Age and Ageing*, 43 (6):794-800. doi: https://doi.org/10.1093/ageing/afu067.

Henderson, C., M. Knapp, J. L. Fernandez, J. Beecham, S. P. Hirani, M. Cartwright, L. Rixon, M. Beynon, A. Rogers, P. Bower, H. Doll, R. Fitzpatrick, A. Steventon, M. Bardsley, J. Hendy, and S. P. Newman (2013) "Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial" *BMJ*, 346:f1035. doi: https://doi.org/10.1136/bmj.f1035

Henderson, Catherine, Jennifer Beecham, and Martin Knapp (2013) "The costs of telecare and telehealth" in Lesley Curtis, eds. *Unit Costs of Health and Social Care 2013*. Canterbury: Personal Social Services Research Unit.

Hendy, Jane, Theopisti Chrysanthaki, James Barlow, Martin Knapp, Anne Rogers, Caroline Sanders, Peter Bower, Robert Bowen, Ray Fitzpatrick, Martin Bardsley, and Stanton Newman (2012) "An organisational analysis of the implementation of telecare and telehealth: the whole systems demonstrator" *BMC Health Services Research*, 12 (1):403.

Henningsen, Arne, and Jeff D. Hamann (2007) "systemfit: A Package for Estimating Systems of Simultaneous Equations in R" *Journal of Statistical Software; Vol 1, Issue 4 (2008)*.

CLTEST: Stata modules for performing cluster-adjusted chi-square and t-tests.

Hetlevik, O., H. Melbye, and S. Gjesdal (2016) "GP utilisation by education level among adults with COPD or asthma: a cross-sectional register-based study" *NPJ Primary Care Respiratory Medicine*, 26:16027. doi: 10.1038/npjpcrm.2016.27.

Hirani, S. P., M. Beynon, M. Cartwright, L. Rixon, H. Doll, C. Henderson, M. Bardsley, A. Steventon, M. Knapp, A. Rogers, P. Bower, C. Sanders, R. Fitzpatrick, J. Hendy, and S. P. Newman (2013) "The effect of telecare on the quality of life and psychological well-being of elderly recipients of social care over a 12-month period: the Whole Systems Demonstrator cluster randomised trial" *Age and Ageing*. doi: https://doi.org/10.1093/ageing/aft185.

Hirani, Shashivadan P, Lorna Rixon, Martin Cartwright, Michelle Beynon, and Stanton P Newman (2017) "The Effect of Telehealth on Quality of Life and Psychological Outcomes Over a 12-Month Period in a Diabetes Cohort Within the Whole Systems Demonstrator Cluster Randomized Trial" *JMIR Diabetes*, 2 (2):e18. doi: 10.2196/diabetes.7128.

HM Government. 2014. The Care Act (c. 23).

Hoch, J. S., M. A. Rockx, and A. D. Krahn (2006) "Using the net benefit regression framework to construct cost-effectiveness acceptability curves: an example using data from a trial of external loop recorders versus Holter monitoring for ambulatory monitoring of "community acquired" syncope" *BMC Health Services Research*, 6:68. doi: 10.1186/1472-6963-6-68.

Hollis, S., and F. Campbell (1999) "What is meant by intention to treat analysis? Survey of published randomised controlled trials" *Bmj*, 319 (7211):670-4.

Howse, Kenneth. (2006). *Increasing life expectancy and the compression of morbidity: a critical review of the debate.* Oxford: Oxford Institute of Ageing.

Hox, J. J. (2010) *Multilevel analysis : techniques and applications*. 2nd ed. ed. London: Routledge.

Hu, F. B., J. Goldberg, D. Hedeker, B. R. Flay, and M. A. Pentz (1998) "Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes" *American Journal Epidemiology*, 147 (7):694-703.

Husereau, Don, Michael Drummond, Stavros Petrou, Chris Carswell, David Moher, Dan Greenberg, Federico Augustovski, Andrew H. Briggs, Josephine Mauskopf, and Elizabeth Loder (2013) "Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement" *BMJ*, 346.

Hutchinson, A., C. Brand, L. Irving, C. Roberts, P. Thompson, and D. Campbell (2010) "Acute care costs of patients admitted for management of chronic obstructive pulmonary disease exacerbations: contribution of disease severity, infection and chronic heart failure" *Internal Medical Journal*, 40 (5):364-71. doi: 10.1111/j.1445-5994.2010.02195.x.

Ikenwilo, Divine (2013) "A difference-in-differences analysis of the effect of free dental check-ups in Scotland" *Social Science & Medicine*, 83 (0):10-18. doi: http://dx.doi.org/10.1016/j.socscimed.2013.01.027.

Inglis, S. C., R. A. Clark, F. A. McAlister, J. Ball, C. Lewinter, D. Cullington, S. Stewart, and J. G. Cleland (2010) "Structured telephone support or telemonitoring programmes for patients with chronic heart failure" *Cochrane Database of Systematic Reviews*, (8):CD007228. doi: 10.1002/14651858.CD007228.pub2.

Inglis, Sally C., Robyn A. Clark, Riet Dierckx, David Prieto-Merino, and John G. F. Cleland (2015) "Structured telephone support or non-invasive telemonitoring for patients with heart failure" *Cochrane Database of Systematic Reviews*, (10). doi: 10.1002/14651858.CD007228.pub3.

Ipsos MORI. (2017). *Unmet Need for Care.* London: Ipsos MORI.

Jaana, M., G. Paré, and C. Sicotte (2009) "Home telemonitoring for respiratory conditions: a systematic review" *American Journal of Managed Care*, 15.

Jacobs, R., and E. Barrenho (2011) "Impact of crisis resolution and home treatment teams on psychiatric admissions in England" *British Journal Psychiatry*, 199 (1):71-6. doi: https://doi.org/10.1192/bjp.bp.110.079830.

Jagger, C., F. E. Matthews, P. Wohland, T. Fouweather, B. C. Stephan, L. Robinson, A. Arthur, C. Brayne, Function Medical Research Council Cognitive, and Collaboration Ageing (2016) "A comparison of health expectancies over two decades in England: results of the Cognitive Function and Ageing Study I and II" *Lancet*, 387 (10020):779-86. doi: http://dx.doi.org/10.1016/S0140-6736(15)00947-2.

Jenkinson, C., R. Layte, D. Jenkinson, K. Lawrence, S. Petersen, C. Paice, and J. Stradling (1997) "A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies?" *Journal of Public Health Medicine*, 19 (2):179-86.

Johnston, K, MJ Buxton, DR Jones, and R Fitzpatrick (1999) "Assessing the costs of healthcare technologies in clinical trials" *Health Technology Assessment*, 3 (6).

Johnston, Kylie, Anthea Worley, Karen Grimmer-Somers, Michele Sutherland, and Lorraine Amos (2010) "Personal alarm use to call the ambulance after a fall in older people: characteristics of clients and falls" *Journal of Emergency Primary Health Care*, 8 (4):1-9.

Kang, Y., M. D. McHugh, J. Chittams, and K. H. Bowles (2016) "Utilizing Home Healthcare Electronic Health Records for Telehomecare Patients With Heart Failure: A Decision Tree Approach to Detect Associations With Rehospitalizations" *Computers, Informatics, Nursing*, 34 (4):175-82. doi: 10.1097/CIN.0000000000000223.

Kehusmaa, S., I. Autti-Ramo, H. Helenius, K. Hinkka, M. Valaste, and P. Rissanen (2012) "Factors associated with the utilization and costs of health and social services in frail elderly patients" *BMC Health Services Research*, 12. doi: https://doi.org/10.1186/1472-6963-12-204.

Kim, Hee-Woong, Hock Chuan Chan, and Sumeet Gupta (2007) "Value-based Adoption of Mobile Internet: An empirical investigation" *Decision Support Systems*, 43 (1):111-126. doi: https://doi.org/10.1016/j.dss.2005.05.009.

Knapp, Martin. (1984) *The economics of social care*, *Studies in social policy*. London: Macmillan.

Kubitschke, Lutz, and Kevin Cullen. (2010). *ICT & Ageing: European Study on Users, Markets and Technologies. Final Report.* Brussels: Commission of the European Communities.

Lazarus, Richard S., and Susan Folkman. (1984) *Stress, appraisal, and coping*. New York: Springer Pub. Co.

Lee, A. H., Y. Zhao, K. K. Yau, and L. Xiang (2010) "How to analyze longitudinal multilevel physical activity data with many zeros?" *Preventive Medicine*, 51 (6):476-81. doi: https://doi.org/10.1016/j.ypmed.2010.09.012.

Lee, Jacques S., Hurley Mary Jane, Carew Debra, Fisher Rory, Kiss Alex, and Drummond Neil (2008) "A Randomized Clinical Trial to Assess the Impact on an Emergency Response System on Anxiety and Health Care Use among Older Emergency Patients after a Fall" *Academic Emergency Medicine*, 14 (4):301-308. doi: 10.1197/j.aem.2006.11.017.

Leroi, Iracema, John Woolham, Rebecca Gathercole, Robert Howard, Barbara Dunk, Chris Fox, John O'Brien, Andrew Bateman, Fiona Poland, Peter Bentham, Alistair Burns, Anna Davies, Kirsty Forsyth, Richard Gray, Martin Knapp, Stanton Newman, Rupert McShane, and Craig Ritchie (2013) "Does telecare prolong community living in dementia? A study protocol for a pragmatic, randomised controlled trial" *Trials*, 14 (1):349. doi: 10.1186/1745-6215-14-349.

Li, K. H., X. L. Meng, T. E. Raghunathan, and D. B. Rubin (1991) "Significance Levels from Repeated P-Values with Multiply-Imputed Data" *Statistica Sinica*, 1 (1):65-92.

Little, Roderick J. A., and Donald B. Rubin. (2002) *Statistical analysis with missing data*. 2nd ed. ed. Hoboken, N.J. ; [Chichester]: Wiley.

Liu, L., R. L. Strawderman, M. E. Cowen, and Y. C. T. Shih (2010) "A flexible two-part random effects model for correlated medical costs" *Journal of Health Economics*, 29 (1):110-123. doi: 10.1016/j.jhealeco.2009.11.010.

Lloyd, James. (2012a). *The Future of Who Uses Telecare.* London: The Strategic Society Centre.

Lloyd, James. (2012b). *Who uses telecare?* London: The Strategic Society Centre.

Local Government Association. 2015. Adult social care, health and wellbeing: A Shared Commitment 2015 Spending Review Submission. London: Local Government Association.

Local Government Association. (2016). *Adult social care funding: 2016 state of the nation report.* London: Local Government Association,.

Ludtke, O., A. Robitzsch, and S. Grund (2017) "Multiple Imputation of Missing Data in Multilevel Designs: A Comparison of Different Strategies" *Psychological Methods*, 22 (1):141-165. doi: 10.1037/met0000096.

MacNeill, Virginia, Caroline Sanders, Ray Fitzpatrick, Jane Hendy, James Barlow, Martin Knapp, Anne Rogers, Martin Bardsley, and Stanton P. Newman (2014) "Experiences of front-line health professionals in the delivery of telehealth: a qualitative study" *British Journal of General Practice*, 64 (624):e401-e407.

Manca, A., N. Hawkins, and M. J. Sculpher (2005) "Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility" *Health Economics*, 14 (5):487-496.

Manca, A., N. Rice, M. J. Sculpher, and A. H. Briggs (2005) "Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models" *Health Economics*, 14 (5):471-485.

Manning, W. G., A. Basu, and J. Mullahy (2005) "Generalized modeling approaches to risk adjustment of skewed outcomes data" *Journal of Health Economics*, 24 (3):465-488. doi: DOI 10.1016/j.jhealeco.2004.09.011.

Marmot, M., James Nazroo, James Banks, R. Blundell, B. Erens, Carli Lessof, and F. Huppert. 2008. English Longitudinal Study of Ageing: Wave 0 (1998, 1999 and 2001) and

Waves 1-3 (2002-2007). Tenth edition, SN5050 [computer file]. Colchester, UK: UK Data Archive [distributor].

Marteau, T. M., and H. Bekker (1992) "The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI)" *British Journal of Clinical Psychology*, 31 (Pt 3):301-6.

Martin, S., G. Kelly, W. G. Kernohan, B. McCreight, and C. Nugent (2008) "Smart home technologies for health and social care support" *Cochrane Database of Systematic Reviews*, (4):CD006412.

Mason, J. M., R. J. Young, J. P. New, J. M. Gibson, A. F. Long, T. Gambling, and T. Friede (2006) "Economic analysis of a telemedicine intervention to improve glycemic control in patients with diabetes mellitus: illustration of a novel analytic method" *Disease Management & Health Outcomes*, 14 (6):377-385.

McCabe, C., K. Claxton, and A. J. Culyer (2008) "The NICE cost-effectiveness threshold: what it is and what that means" *Pharmacoeconomics*, 26 (9):733-44.

McKenna, A. C., M. Kloseck, R. Crilly, and J. Polgar (2015) "Purchasing and Using Personal Emergency Response Systems (PERS): how decisions are made by community-dwelling seniors in Canada" *BMC Geriatrics*, 15. doi: 10.1186/s12877-015-0079-z.

McLean, Susannah, Aziz Sheikh, Kathrin Cresswell, Ulugbek Nurmatov, Mome Mukherjee, Akiko Hemmi, and Claudia Pagliari (2013) "The Impact of Telehealthcare on the Quality and Safety of Care: A Systematic Overview" *PLoS ONE*, 8 (8):e71238. doi: 10.1371/journal.pone.0071238.

Menn, P., J. Heinrich, R. M. Huber, R. A. Jorres, J. John, S. Karrasch, A. Peters, H. Schulz, R. Holle, and Kora Study Group (2012) "Direct medical costs of COPD--an excess cost approach based on two population-based studies" *Respiratory Medicine*, 106 (4):540-8. doi: 10.1016/j.rmed.2011.10.013.

Milligan, Christine, Celia Roberts, and Maggie Mort (2011) "Telecare and older people: Who cares where?" *Social Science and Medicine*, 72 (3):347-354. doi: 10.1016/j.socscimed.2010.08.014.

Milne, Alisoun, Charlotte Brigden, Ann Palmer, and Elina Konta (2013) "The intersection of employment and care: evidence from a UK case study" *European Journal of Social Work*, 16 (5):651-670. doi: 10.1080/13691457.2012.724388.

Mistry, H. (2012) "Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years" *Journal of Telemedicine and Telecare*, 18 (1):1-6. doi: 10.1258/jtt.2011.110505.

Mistry, H., M. Buxton, L. Longworth, J. Chatwin, and R. Peveler (2005) "Comparison of general practitioner records and patient self-report questionnaires for estimation of costs" *European Journal of Health Economics*, 6 (3):261-6.

Moher, D., S. Hopewell, K. F. Schulz, V. Montori, P. C. Gotzsche, P. J. Devereaux, D. Elbourne, M. Egger, and D. G. Altman (2010) "CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials" *BMJ*, 340:c869. doi: 10.1136/bmj.c869.

Momanyi, Kevin. (2017a) *An Econometric Analysis of The Impact of Telecare on The Length of Stay in Hospital*.

Momanyi, Kevin. 2017b. Telecare and Unplanned Hospitalisation in Scotland: Evidence From Linked Survey and Administrative Data. Presentation Abstract.

Monitor Deloitte. (2015). *Digital Health in the UK: An industry study for the Office of Life Sciences.*

Moreno, L., S. B. Dale, A. Y. Chen, and C. A. Magee (2009) "Costs to Medicare of the Informatics for Diabetes Education and Telemedicine (IDEATel) home telemedicine demonstration: findings from an independent evaluation" *Diabetes Care*, 32 (7):1202-4. doi: 10.2337/dc09-0094.

Morgenstern, Lewis B., Eric E. Adelman, Rebecca Hughes, Jeffrey J. Wing, and Lynda D. Lisabeth (2015) "The Women Independently Living Alone with a Medical Alert Device (WILMA) trial" *Translational Stroke Research*, 6 (5):355-360. doi: 10.1007/s12975-015-0411-0.

Morton, Katherine, Laura Dennison, Carl May, Elizabeth Murray, Paul Little, Richard J. McManus, and Lucy Yardley (2017) "Using digital interventions for self-management of chronic physical health conditions: A meta-ethnography review of published studies" *Patient Education and Counseling*, 100 (4):616-635. doi: https://doi.org/10.1016/j.pec.2016.10.019.

Mullahy, J. (1998) "Much ado about two: reconsidering retransformation and the two-part model in health econometrics" *Journal of Health Economics*, 17 (3):247-81. doi: S0167-6296(98)00030-7 [pii].

National Institute for Clinical Excellence. (2004). *Guide to the Methods of Technology Appraisal.* London: National Institute for Clinical Excellence.

National Institute for Health and Clinical Excellence. (2008). *Guide to the methods of technology appraisal.* London: National Institute for Health and Clinical Excellence.

Newman, Stanton, and Whole System Demonstrator Programme Evaluation Team. (2014). *The Whole Systems Demonstrator Evaluation of Telehealth and Telecare. Final report.* London: City University.

NHS England. (2017). *Personal health budgets and Integrated Personal Commissioning. National expansion plan.* London: NHS England.

NHS Stoke-on-Trent. 2011. "“Florence” the Simple Telehealth service." In: NHS Stoke-on-Trent. http://www.networks.nhs.uk/nhs-networks/simple-telehealth/documents/STH%20DS01%20-%20Patient%20v1%20draft%2009.pdf (accessed 18/6/2013).

NHS Supply Chain. (2010). "Aids for Daily living products and pricing."

Noble, M., D. McLennan, K. Wilkinson, A. Whitworth, and H. Barnes. (2008). *The English Indices of Multiple Deprivation 2007.* London: Communities and Local Government: Social Disadvantage Research Centre, University of Oxford.

NORM: Multiple imputation of incomplete multivariate data under a normal model (Version 2) [Software] 1999. University Park: The Methodology Center, Penn State.

Noro, Anja M., Unto T. Häkkinen, and Olli J. Laitinen (1999) "Determinants of health service use and expenditure among the elderly Finnish population" *European Journal of Public Health*, 9 (3):174-180. doi: 10.1093/eurpub/9.3.174.

Nyman, S. R., and C. R. Victor (2014) "Use of personal call alarms among community-dwelling older people" *Ageing and Society*, 34 (1):67-89. doi: 10.1017/S0144686X12000803.

O'Brien, B. J., and A. H. Briggs (2002) "Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods" *Statistical Methods in Medical Research*, 11 (6):455-68.

OECD. (2010) *Value for Money in Health Spending*. Paris: OECD Publishing.

OECD. (2013). *Public spending on health and long-term care: a new set of projections. OECD Economic Policy Papers No. 6.* In *OECD Economic Policy Papers*.

Office for Budget Responsibility. (2015). *Fiscal sustainability report. Presented to Parliament pursuant to Section 8 of the Budget Responsibility and National Audit Act 2011.*

Office for Budget Responsibility. (2016). *Fiscal sustainability and public spending on health.*

Office for Budget Responsibility. (2017). *Fiscal sustainability report. Presented to Parliament pursuant to Section 8 of the Budget Responsibility and National Audit Act 2011.*

Office for National Statistics. (2009). *National population projections, 2008-based.*

Office for National Statistics. (2011). *2011 Census Analysis: Do the Demographic and Socio-Economic Characteristics of those Living Alone in England and Wales Differ from the General Population?* Office for National Statistics.

Office for National Statistics. (2014). *National Population Projections, 2012-based Reference Volume: Series PP2.*

Office for National Statistics. (2016). *National Population Projections: 2016-based statistical bulletin.*

Older People's Inquiry, Norma Raynes, Heather Clark, Jennifer Beecham, and Joseph Rowntree Foundation. (2006) *The report of the Older People's Inquiry into 'That bit of help'*. York: Joseph Rowntree Foundation.

Pare, G, M Janna, and C Sicotte (2007) "Systematic review of home telemonitoring for chronic diseases: the evidence base" *Journal of the American Medical Informatics Assocation*, 14 (3):269 - 277.

Pare, G., K. Moqadem, G. Pineau, and C. St-Hilaire (2010) "Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review" *J Med Internet Res*, 12 (2):e21.

Park, P., R. K. Simmons, A. T. Prevost, and S. J. Griffin (2008) "Screening for type 2 diabetes is feasible, acceptable, but associated with increased short-term anxiety: a randomised controlled trial in British general practice" *BMC Public Health*, 8:350. doi: 10.1186/1471-2458-8-350.

Parker, S. L., S. K. Mendenhall, D. N. Shau, O. Adogwa, W. N. Anderson, C. J. Devin, and M. J. McGirt (2012) "Minimum clinically important difference in pain, disability, and quality of life after neural decompression and fusion for same-level recurrent lumbar stenosis: understanding clinical versus statistical significance" *Journal of Neurosurgy: Spine*, 16 (5):471-8. doi: 10.3171/2012.1.SPINE11842.

Patmore, Charles, and Alison McNulty. (2005). *Making home care for older people more flexible and person-centred - Factors which promote this* Social Policy Research Unit, the University of York.

Pauly, Mark V. (1995) "Valuing health care benefits in money terms" in Frank A. Sloan, eds. *Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies*. 99-124. Cambridge: Cambridge University Press.

Pearce, G., H. L. Parke, H. Pinnock, E. Epiphaniou, C. L. Bourne, A. Sheikh, and S. J. Taylor (2016) "The PRISMS taxonomy of self-management support: derivation of a novel taxonomy and initial testing of its utility" *Journal of Health Services Research and Policy*, 21 (2):73-82. doi: 10.1177/1355819615602725.

Peek, Sebastiaan T. M., Sil Aarts, and Eveline J. M. Wouters (2015) "Can Smart Home Technology Deliver on the Promise of Independent Living?" in Joost van Hoof, George Demiris and Eveline J. M. Wouters, eds. *Handbook of Smart Homes, Health Care and Well-Being*. 1-10. Cham: Springer International Publishing.

Peek, Sebastiaan T. M., Eveline J. M. Wouters, Joost van Hoof, Katrien G. Luijkx, Hennie R. Boeije, and Hubertus J. M. Vrijhoef (2014) "Factors influencing acceptance of technology for aging in place: A systematic review" *International Journal of Medical Informatics*, 83 (4):235-248. doi: http://dx.doi.org/10.1016/j.ijmedinf.2014.01.004.

Percival, J., and J. Hanson (2006) "Big brother or brave new world? Telecare and its implications for older people's independence and social inclusion" *Critical Social Policy*, 26 (4):888-909. doi: Doi 10.1177/0261018306068480.

Petrie, Keith J., and John Weinman. (1997) *Perceptions of health and illness: current research and applications*. Australia: Harwood Academic Publishers.

Petrou, S., and A. Gray (2011) "Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting" *BMJ*, 342:d1548. doi: 10.1136/bmj.d1548.

Piau, A., E. Campo, P. Rumeau, B. Vellas, and F. Nourhashemi (2014) "Aging society and gerontechnology: a solution for an independent living?" *Journal of Nutrition, Health and Aging*, 18 (1):97-112. doi: 10.1007/s12603-013-0356-5.

Pickard, Linda. (2008). *Informal Care for Older People Provided by Their Adult Children: Projections of Supply and Demand to 2041 in England.Report to the Strategy Unit (Cabinet Office) and the Department of Health.* London: PSSRU, London School of Economics.

Pickard, Linda, Raphael Wittenberg, Adelina Comas-Herrera, Derek King, and Juliette Malley (2007) "Care by Spouses, Care by Children: Projections of Informal Care for Older People in England to 2031" *Social Policy and Society*, 6 (03):353-366. doi: https://doi.org/10.1017/S1474746407003685.

Polisena, J., D. Coyle, K. Coyle, and S. McGill (2009) "Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations" *International Journal of Technology Assessment in Health Care*, 25 (3):339-49.

Polisena, J., K. Tran, K. Cimon, B. Hutton, S. McGill, and K. Palmer (2009) "Home telehealth for diabetes management: a systematic review and meta-analysis" *Diabetes, Obesity and Metabolism*, 11 (10):913-30.

Polisena, J., K. Tran, K. Cimon, B. Hutton, S. McGill, K. Palmer, and R. E. Scott (2010a) "Home telehealth for chronic obstructive pulmonary disease: a systematic review and meta-analysis" *Journal of Telemedicine and Telecare*, 16 (3):120-7.

Polisena, J., K. Tran, K. Cimon, B. Hutton, S. McGill, K. Palmer, and R. E. Scott (2010b) "Home telemonitoring for congestive heart failure: a systematic review and meta-analysis" *Journal of Telemedicine and Telecare*, 16 (2):68-76.

Poole, Teresa. (2006). *Telecare and Older People. Appendix to Wanless Social care review (Wanless D, Fernandez JL, Poole T, Beesley L, Henwood M, Moscone F. Securing good care for older people: Taking a long-term view. London: King's Fund, 2006.).* King's Fund.

Powell, John, Lee Gunn, Pam Lowe, Bart Sheehan, Frances Griffiths, and Aileen Clarke (2010) "New networked technologies and carers of people with dementia: an interview study" *Ageing & Society*. doi: doi:10.1017/S0144686X1000019X.

Puffer, S., D. Torgerson, and J. Watson (2003) "Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals" *BMJ*, 327 (7418):785-9. doi: 10.1136/bmj.327.7418.785.

Rabe-Hesketh, S., and Anders Skrondal. (2012) *Multilevel and longitudinal modeling using Stata*. 3rd ed. 2 vols. College Station, Tex.: Stata Press Publication.

Rabe-Hesketh, S., T. Toulopoulou, and R. M. Murray (2001) "Multilevel modeling of cognitive function in schizophrenic patients and their first degree relatives" *Multivariate Behavioral Research*, 36 (2):279-298. doi: Doi 10.1207/S15327906mbr3602_07.

Radhakrishnan, K., C. S. Jacelon, C. Bigelow, J. Roche, J. Marquard, and K. H. Bowles (2013a) "Use of a homecare electronic health record to find associations between patient characteristics and re-hospitalizations in patients with heart failure using telehealth" *Journal of Telemedicine and Telecare*, 19 (2):107-12. doi: 10.1258/jtt.2012.120509.

Radhakrishnan, K., C. S. Jacelon, C. Bigelow, J. P. Roche, J. L. Marquard, and K. H. Bowles (2013b) "Association of comorbidities with home care service utilization of patients with heart failure while receiving telehealth" *Journal of Cardiovascular Nursing*, 28 (3):216-27. doi: 10.1097/JCN.0b013e3182512331.

Rahimpour, Mohammadreza, Nigel H. Lovell, Branko G. Celler, and John McCormick (2008) "Patients' perceptions of a home telecare system" *International Journal of Medical Informatics*, 77 (7):486-498. doi: https://doi.org/10.1016/j.ijmedinf.2007.10.006.

Rawlins, M (1999) "In pursuit of quality: the National Institute for Clinical Excellence" *Lancet*, 353:1079-1082. doi: 10.1016/S0140-6736(99)02381-8.

Reeder, Blaine, Ellen Meyer, Amanda Lazar, Shomir Chaudhuri, Hilaire J. Thompson, and George Demiris (2013) "Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: A systematic review" *International Journal of Medical Informatics*, 82 (7):565-579. doi: https://doi.org/10.1016/j.ijmedinf.2013.03.007.

Richards, S. H., J. Coast, and T. J. Peters (2003) "Patient-reported use of health service resources compared with information from health providers" *Health Soc Care Community*, 11 (6):510-8.

Ridder, D. de, and K. Schreurs (1996) "Coping, social support and chronic disease: A research agenda" *Psychology, Health & Medicine*, 1 (1):71-82. doi: 10.1080/13548509608400007.

Rijken, M., M. Jones, M. Heijmans, and A. Dixon (2008) "Supporting self-management" in E. Nolte and M. McKee, eds. *Caring for people with chronic conditions: a health system perspective*. 116-142. Berkshire: Open University Press.

Rixon, L., S. P. Hirani, M. Cartwright, M. Beynon, H. Doll, A. Steventon, C. Henderson, and S. P. Newman (2015) "A RCT of telehealth for COPD patient's quality of life: the whole system demonstrator evaluation" *Clinical Respiratory Journal*. doi: 10.1111/crj.12359.

Roberts, Celia, Maggie Mort, and Christine Milligan (2012) "Calling for Care: 'Disembodied' Work, Teleoperators and Older People Living at Home" *Sociology*. doi: 10.1177/0038038511422551.

Rodrıguez, German (2008) "Multilevel Generalized Linear Models" in Jan de Leeuw and Erik Meijer, eds. *Handbook of multilevel analysis*. 335-376. New York: Springer.

Rogers, A, P Bower, Caroline Gardner, Claire Gately, Anne Kennedy, Victoria Lee, Elizabeth Middleton, David Reeves, and Gerry Richardson. (2006). *The National Evaluation of the Pilot Phase of the Expert Patients Programme. Final Report.* Manchester: National Primary Care Research & Development Centre.

Roland, Martin, and David J Torgerson (1998) "Understanding controlled trials: What are pragmatic trials?" *BMJ*, 316 (7127):285. doi: 10.1136/bmj.316.7127.285.

Roush, Robert E., and Thomas A. Teasdale (2011) "Personal Emergency Response Services: Do the Benefits Justify the Cost in Seniors Housing and Care Properties?" *Seniors Housing & Care Journal*, 19 (1):73-82.

Rubin, Donald B. (1987) *Multiple imputation for nonresponse in surveys*. New York ; Chichester: Wiley.

Salisbury, C., C. Thomas, A. O'Cathain, A. Rogers, C. Pope, L. Yardley, S. Hollinghurst, T. Fahey, G. Lewis, S. Large, L. Edwards, A. Rowsell, J. Segar, S. Brownsell, and A. A. Montgomery (2015) "TElehealth in CHronic disease: mixed-methods study to develop the TECH conceptual model for intervention design and evaluation" *BMJ Open*, 5 (2):e006448. doi: http://dx.doi.org/10.1136/bmjopen-2014-006448.

Samsa, G., D. Edelman, M. L. Rothman, G. R. Williams, J. Lipscomb, and D. Matchar (1999) "Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II" *PharmacoEconomics*, 15 (2):141-55.

Sanders, C., A. Rogers, R. Bowen, P. Bower, S. Hirani, M. Cartwright, R. Fitzpatrick, M. Knapp, J. Barlow, J. Hendy, T. Chrysanthaki, M. Bardsley, and S. P. Newman (2012) "Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study" *BMC Health Serv Res*, 12:220. doi: 10.1186/1472-6963-12-220.

Schafer, Joseph L. 1999. NORM Users Guide. Multiple imputation of incomplete multivariate data under a normal model. Version 2. The Pennsylvania State University.

Schermer, M (2009) "Telecare and self-management: opportunity to change the paradigm?". doi: 10.1136/jme.2009.030973.

Schulz, K. F., D. G. Altman, and D. Moher (2010) "CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials" *BMJ*, 340:c332. doi: 10.1136/bmj.c332.

Segar, J., A. Rogers, C. Salisbury, and C. Thomas (2013) "Roles and identities in transition: boundaries of work and inter-professional relationships at the interface between telehealth and primary care" *Health and Social Care in the Community*, 21.

Sixsmith, Andrew (1986) "Independence and home in later life" in, eds. *Dependency and interdependency in old age: Theoretical perspectives and policy alternatives*. 338-347. London: Croom Helm.

Sixsmith, Andrew, and Judith Sixsmith (2008) "Ageing in Place in the United Kingdom" *Ageing International*, 32 (3):219-235. doi: 10.1007/s12126-008-9019-y.

Standing, Craig, Susan Standing, Marie-Louise McDermott, Raj Gururajan, and Reza Kiani Mavi (2016) "The Paradoxes of Telehealth: a Review of the Literature 2000–2015" *Systems Research and Behavioral Science*, 35 (1):90-101. doi: 10.1002/sres.2442.

StataCorp. (2015a) *Multilevel mixed-effects reference manual. Release 14*. College Station, Texas.

StataCorp. (2015b) *Stata Statistical Software: Release 14.0*. College Station,Texas: StataCorp LP.

Steffens, D. C., K. R. Krishnan, C. Crump, and G. L. Burke (2002) "Cerebrovascular disease and evolution of depressive symptoms in the cardiovascular health study" *Stroke*, 33 (6):1636-44.

Steventon, A., M. Bardsley, J. Billings, J. Dixon, H. Doll, S. Hirani, M. Cartwright, L. Rixon, M. Knapp, C. Henderson, A. Rogers, R. Fitzpatrick, J. Hendy, and S. Newman (2012) "Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial" *BMJ*, 344:e3874. doi: 10.1136/bmj.e3874.

Steventon, Adam, Martin Bardsley, John Billings, Jennifer Dixon, Helen Doll, Michelle Beynon, Shashi Hirani, Martin Cartwright, Lorna Rixon, Martin Knapp, Catherine Henderson, Anne Rogers, Jane Hendy, Ray Fitzpatrick, and Stanton Newman (2013) "Effect

of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial" *Age and Ageing*. doi: 10.1093/ageing/aft008.

Steventon, Adam, Martin Bardsley, Helen Doll, Elizabeth Tuckey, and Stanton P. Newman (2014) "Effect of telehealth on glycaemic control: analysis of patients with type 2 diabetes in the Whole Systems Demonstrator cluster randomised trial" *BMC Health Services Research*, 14 (1):334. doi: 10.1186/1472-6963-14-334.

Steventon, Adam, R. Grieve, and S Newman. 2015. Why did the Whole Systems Demonstrator report reductions in emergency hospital admissions? Insights from new analyses. In *Electronic letter to the editor*. BMJ.

Steventon, Adam, Richard Grieve, and Martin Bardsley (2015) "An Approach to Assess Generalizability in Comparative Effectiveness Research: A Case Study of the Whole Systems Demonstrator Cluster Randomized Trial Comparing Telehealth with Usual Care for Patients with Chronic Health Conditions" *Medical Decision Making*, 35 (8):1023-1036. doi: 10.1177/0272989x15585131.

Stoddart, Andrew, Janet Hanley, Sarah Wild, Claudia Pagliari, Mary Paterson, Steff Lewis, Aziz Sheikh, Ashma Krishan, Paul Padfield, and Brian McKinstry (2013) "Telemonitoring-based service redesign for the management of uncontrolled hypertension (HITS): cost and cost-effectiveness analysis of a randomised controlled trial" *BMJ Open*, 3 (5). doi: 10.1136/bmjopen-2013-002681.

Suzman, Richard, David P. Willis, and Kenneth G. Manton. (1992) *The Oldest old*. New York: Oxford University Press.

Telecare Aware. "What is Telecare?" [Web page], accessed 7 March 2015. http://telecareaware.com/what-is-telecare/.

The King's Fund. (2005). "Briefings: International health care comparisons." [Web page], Last Modified 01.04.05, accessed 20th October 2008. http://www.kingsfund.org.uk/publications/briefings/international.html#section1.

The King's Fund. (2016). *Social care for older people: Home truths.* London: The King's Fund.

Thomas, Clare L., Mei-See Man, Alicia O'Cathain, Sandra Hollinghurst, Shirley Large, Louisa Edwards, Jon Nicholl, Alan A. Montgomery, and Chris Salisbury (2014) "Effectiveness and cost-effectiveness of a telehealth intervention to support the management of long-term conditions: study protocol for two linked randomized controlled trials" *Trials*, 15 (1):1-14. doi: 10.1186/1745-6215-15-36.

Thompson, S. G., R. M. Nixon, and R. Grieve (2006) "Addressing the issues that arise in analysing multicentre cost data, with application to a multinational study" *Journal of Health Economics*, 25 (6):1015-1028. doi: DOI 10.1016/j.jhealeco.2006.02.001.

Tomassini, C. (2006) "The oldest old in Great Britain: change over the last 20 years" *Popul Trends*, (123):32-9.

Tomita, M. R., W. C. Mann, K. Stanton, A. D. Tomita, and V. Sundar (2007) "Use of currently available smart home technology by frail elders process and outcomes" *Topics in Geriatric Rehabilitation*, 23 (1):24-34.

Tsiachristas, Apostolos, and Maureen P. M. H. Rutten-van Mölken (2014) "Exploring the variability of patient costs in disease management programs: a hierarchical modelling approach" *Applied Economics*, 46 (9):940-951. doi: 10.1080/00036846.2013.864044.

Turner, A. J., S. Nikolova, and M. Sutton (2016) "The effect of living alone on the costs and benefits of surgery amongst older people" *Social Science & Medicine*, 150:95-103. doi: 10.1016/j.socscimed.2015.11.053.

Ukoumunne, Obioha C. (2002) "A comparison of confidence interval methods for the intraclass correlation coefficient in cluster randomized trials" *Statistics in Medicine*, 21 (24):3757-3774. doi: 10.1002/sim.1330.

Van den Noortgate, W., M. C. Opdenakker, and P. Onghena (2005) "The effects of ignoring a level in multilevel analysis" *School Effectiveness and School Improvement*, 16 (3):281-303. doi: 10.1080/09243450500114850.

Vassilev, Ivaylo, Alison Rowsell, Catherine Pope, Anne Kennedy, Alicia O'Cathain, Chris Salisbury, and Anne Rogers (2015) "Assessing the implementability of telehealth interventions for self-management support: a realist review" *Implementation Science*, 10 (1):59. doi: doi:10.1186/s13012-015-0238-9.

Vazquez-Polo, F. J., M. Negrin, J. M. Cabases, E. Sanchez, J. M. Haro, and L. Salvador-Carulla (2005) "An analysis of the costs of treating schizophrenia in Spain: a hierarchical Bayesian approach" *Journal of Mental Health Policy and Economics*, 8 (3):153-65.

Venkatesh, V., M. G. Morris, G. B. Davis, and F. D. Davis (2003) "User acceptance of information technology: toward a unified view" *MIS Quarterly*, 27.

Vergara Rojas, S. , and M. P. Gagnon (2008) "A systematic review of the key indicators for assessing telehomecare cost-effectiveness" *Telemedicine J and e-Health*, 14 (9):896-904.

Vincent, C., D. Reinharz, I. Deaudelin, M. Garceau, and L. R. Talbot (2006) "Public telesurveillance service for frail elderly living at home, outcomes and cost evolution: a quasi experimental design with two follow-ups" *Health and Quality of Life Outcomes*, 4:41. doi: 10.1186/1477-7525-4-41.

Wagner, E. H., B. T. Austin, C. Davis, M. Hindmarsh, J. Schaefer, and A. Bonomi (2001) "Improving chronic illness care: translating evidence into action" *Health Affairs (Millwood)*, 20 (6):64-78. doi: 10.1377/hlthaff.20.6.64.

Wagner, E. H., B. T. Austin, and M. Von Korff (1996) "Improving outcomes in chronic illness" *Managed Care Quarterly*, 4 (2):12-25.

Wanless, Derek. (2002). *Securing our Future Health: Taking a Long-Term View. Final Report.* HM Treasury.

Ware, J, M Kosinski, JB Bjorner, DM Turner-Bowker, B Gandek, and ME Maruish. (2007). *User's Manual for the SF-36v2 Health Survey.* Lincoln, RI: Quality Metric Inc.

Weinman, J., and K. J. Petrie (1997) "Illness perceptions: a new paradigm for psychosomatics?" *Journal of Psychosomatic Research*, 42 (2):113-6.

West Lothian Council. (2004). *Smart Support at Home - The Impact of Telecare Technology on Primary and Community Care Systems in West Lothian.* Unpublished.

Whitten, P. S., F. S. Mair, A. Haycox, C. R. May, T. L. Williams, and S. Hellmich (2002) "Systematic review of cost effectiveness studies of telemedicine interventions" *BMJ*, 324 (7351):1434-7.

Wild, S. H., J. Hanley, S. C. Lewis, J. A. McKnight, L. B. McCloughan, P. L. Padfield, R. A. Parker, M. Paterson, H. Pinnock, A. Sheikh, and B. McKinstry (2016) "Supported Telemonitoring and Glycemic Control in People with Type 2 Diabetes: The Telescot Diabetes Pragmatic Multicenter Randomized Controlled Trial" *PLoS Medicine*, 13 (7):e1002098. doi: 10.1371/journal.pmed.1002098.

Willems, D. C. M., M. A. Joore, J. J. E. Hendriks, R. A. H. van Duurling, E. F. M. Wouters, and J. L. Severens (2007) "Process evaluation of a nurse-led telemonitoring programme for patients with asthma" *Journal of Telemedicine and Telecare*, 13 (6):310-317. doi: Doi 10.1258/135763307781644898.

Willems, Daniëlle C. M., Manuela A. Joore, Johannes J. E. Hendriks, Emiel F. M. Wouters, and Johan L. Severens (2007) "Cost-effectiveness of a nurse-led telemonitoring intervention based on peak expiratory flow measurements in asthmatics: results of a randomised controlled trial" *Cost effectiveness and resource allocation*, 5:10-10. doi: 10.1186/1478-7547-5-10.

Williams, Bridget, Phil Copestake, John Eversley, and Bruce Stafford. (2008). *Experiences and Expectations of Disabled People - A research report for the Office for Disability Issues.* Prepared by: GfK NOP Social Research, Office for Public Management, ppre, Icarus Collective and University of Nottingham.

Williams, V., C. R. Victor, and R. McCrindle (2013) "It is always on your mind: experiences and perceptions of falling of older people and their carers and the potential of a mobile falls detection device" *Curr Gerontol Geriatr Res*, 2013:295073. doi: 10.1155/2013/295073.

Williamson, P. R., D. G. Altman, H. Bagley, K. L. Barnes, J. M. Blazeby, S. T. Brookes, M. Clarke, E. Gargon, S. Gorst, N. Harman, J. J. Kirkham, A. McNair, C. A. C. Prinsen, J. Schmitt, C. B. Terwee, and B. Young (2017) "The COMET Handbook: version 1.0" *Trials*, 18 (Suppl 3):280. doi: 10.1186/s13063-017-1978-4.

Williamson, P. R., D. G. Altman, J. M. Blazeby, M. Clarke, D. Devane, E. Gargon, and P. Tugwell (2012) "Developing core outcome sets for clinical trials: issues to consider" *Trials*, 13:132. doi: 10.1186/1745-6215-13-132.

Witt Udsen, Flemming, Pernille Heyckendorff Lilholt, Ole Hejlesen, and Lars Ehlers (2017) "Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: results from the Danish 'TeleCare North' cluster-randomised trial" *BMJ Open*, 7 (5).

Wittenberg, Raphael, Linda Pickard, Juliette Malley, Derek King, Adelina Comas-Herrera, and Robin Darton. (2008). *Future Demand for Social Care, 2005 to 2041: Projections of Demand for Social Care for Older People in England.* PSSRU.

Wooldridge, J. 2008. A Course in Applied Econometrics Lecture 11: Difference-in-Differences Estimation. IRP Lectures. UW Madison.

Woolham, J. (2005) *The Effectiveness of Assistive Technology in Supporting the Independence of People with dementia: the Safe At Home project*. London: Hawker Publications.

Woolham, J. G., N. Steils, M. Fisk, J. Porteus, and K. Forsyth. (2018). *The UTOPIA project. Using Telecare for Older People In Adult social care: The findings of a 2016-17 national survey of local authority telecare provision for older people in England.* King's College London: Social Care Workforce Research Unit.

Zeger, S. L., K. Y. Liang, and P. S. Albert (1988) "Models for longitudinal data: a generalized estimating equation approach" *Biometrics*, 44 (4):1049-60.

Zorn, C. J. W. (2001) "Generalized estimating equation models for correlated data: A review with applications" *American Journal of Political Science*, 45 (2):470-490. doi: Doi 10.2307/2669353.

Appendices

Appendix 1

Unit Costs

Table A1. Unit costs

Resource item	Unit	Unit cost	Source/reference	Further description
Hospital use				
A&E	attendance	103 - 133	Department of Health (2011)	TPCTAandEMSNA and TPCTAandEMSAD tabs
Inpatient care	bed-day	116 - 1657	Department of Health (2011)	Weighted average cost of bed-day per HRG subchapter code, assigned based on the participant-reported specialty/reason given for using service
Inpatient care – unknown reason/specialty	bed-day	505	Department of Health (2011)	Weighted average cost of bed-day across all adult specialties, used when no specialty/reason was given for using service (TNEI_L and TPCTEI tabs)
Day hospital/day case	attendance	156 - 1496	Department of Health (2011)	Weighted average cost of day case per HRG subchapter code, based on the participant-reported reason given for using service (TPCTDC tab)
Day hospital/day case – unknown reason/specialty	attendance	660	Department of Health (2011)	Weighted average of cost of day case across all adult specialties when no specialty/reason was given for using service (TPCTDC tab)
Outpatient care and procedures	attendance	23 - 306	Department of Health (2011)	Weighted average cost of outpatient visit for the specialty (consultant and non-consultant visits), excluding first appointments (TPCTCLFUSFF, TPCTCLFUMFF,TPCTNCLFUSFF tabs), or procedure (TPCTOPROC tab) based on the participant-reported specialty/reason given for using service
Outpatient care – unknown reason/specialty	attendance	112	Department of Health (2011)	Weighted average cost of outpatient visit across specialties (consultant and non-consultant visits), excluding first appointments (TPCTCLFUSFF,TPCTCLFUMFF, TPCTNCLFUSFF tabs), used when no specialty/reason was given for using service

Resource item	Unit	Unit cost	Source/reference	Further description
Outpatient scans	attendance	117	Department of Health (2011)	Weighted average of cost of diagnostic imaging activity (Diagnostic Imaging Procedures, TPCTDIAGIM_OP tab) used where scan was reported as the reason for using service
Community				
Paramedic	per visit	192.00	Curtis (2010)	Average of all paramedic services (categories A, B & C)
Community matron	minute	1.31	Curtis (2010)	Nurse specialist (community). Excludes qualification costs.
Community matron (telephone)	minute	1.28	Curtis (2010)	Nurse specialist (community) average cost per minute of face to face client contact time. Excludes qualification costs.
Community or district nurse	minute	1.13	Curtis (2010)	Community nurse average cost per minute of home visit. Excludes qualification costs.
Community or district nurse (telephone)	minute	0.52	Curtis (2010)	Community nurse average cost per minute of nurse time. Excludes qualification costs.
Practice nurse	minute	0.52	Curtis (2010)	Nurse (General Practice) average cost per minute of direct contact time. Excludes qualification costs.
Night nurse	minute	0.50	Curtis (2010)	Rapid response nurse per delivered hour including travel
Specialist nurse	minute	0.95	Curtis (2010)	Nurse (advanced) cost of direct client contact time – used for telephone contact time. Excludes qualification costs.
		1.31		Nurse specialist (community) average cost of face-to-face client contact time including travel cost – used for home visit contact time. Excludes qualification costs.

Resource item	Unit	Unit cost	Source/reference	Further description
Physiotherapist or occupational therapist	minute	0.65	Curtis (2010)	NHS community occupational therapist/physiotherapist, cost per minute of home visit. Excludes qualification costs.
GP (home)	minute	4.00	Curtis (2010)	GP cost per minute of home visit. Excludes direct care staff and qualification costs
	visit	94.00		GP home visit lasting 23.4 minutes (including travel time). Excludes direct care staff and qualification costs
GP (surgery)	minute	2.40	Curtis (2010)	GP surgery/clinic minute. Excludes direct care staff and qualification costs.
	visit	28.00		GP surgery consultation lasting 11.7 minutes. Excludes direct care staff and qualification costs.
GP (telephone)	consultation	17.00	Curtis (2010)	GP telephone consultation lasting 7.1 minutes
Dentist	contact	86.85	Department of Health (2011)	Community Dental Services (CN20)
Chiropodist	contact	35.37	Department of Health (2011)	Community Podiatry Services (N910)
Optician	eye test	20.26	Department of Health (2009)	NHS sight test
Community mental health				
Psychiatrist	minute	4.72	Curtis (2010)	Consultant psychiatrist per minute of patient contact. Excludes qualification costs.
Mental health nurse	minute	0.83	Curtis (2010)	Nurse (mental health) per hour of face-to-face contact including travel. Excludes qualification costs.

Resource item	Unit	Unit cost	Source/reference	Further description
Community care services				
Social worker	minute	0.92	Curtis (2010)	Social worker (adult) cost of client-related work. Excludes qualification costs.
Council home help visit	minute	0.42	Curtis (2010)	Local authority home care worker weekday face-to-face
Home care/home help	minute	0.22	Curtis (2010)	Independently provided home care worker weekday, face-to-face
Paid night carer	minute	0.50	Curtis (2010)	Local authority home care worker weekday evenings, face-to-face
Meals on Wheels	meal	5.00	Curtis (2010)	Average of cost of a meal from Local Authority wheels (£6.00) and from the independent sector (£4.00)
Personal/ community alarm	item	13.50	Curtis (2010)	Median annual cost, annuitised over 10 years at 3.5%; cost over 3 months
Major and minor adaptations	adaptation	1.5 - 455	Curtis (2010)	Kitchen adaptations, stair lift, toilet relocation, low-level bath, electrical modifications, outdoor railings, joinery work, new bathroom. Median annual cost, annuitised over 10 years at 3.5%; cost over 3 months
Equipment including mobility aids and for daily living	item	0.1-97.5	Curtis (2010); Department of Health Care Services Efficiency Delivery Programme (2010); NHS Supply Chain (2010)	Manual wheelchair, electric wheelchair, shower chair, chair-raise, bedrail, commode, reacher, kitchen/perching stool, hoist, trolley, shoehorn, raised toilet seat, bath lift, toilet frame All items annuitised over 10 years at 3.5%, following PSSRU unit costing methods; cost over 3 months

Resource item	Unit	Unit cost	Source/reference	Further description
Care home				
Residential care home	day	63.72	Curtis (2010)	Mean costs, from PSS EX1 2008/09 returns uprated using the PSS Pay & Prices inflator
Nursing home	day	70.57	Curtis (2010)	Mean costs, from PSS EX1 2008/09 returns uprated using the PSS Pay & Prices inflator
Day services				
Day care and other day attendances	attendance	36.00-155.82	Curtis (2010) ; Older People's Inquiry, Raynes et al. (2006); Rogers, Bower et al. (2006); Department of Health (2011)	Local authority day care for older people, from PSS EX1, cost per service user per week, assuming attendance of three sessions per week, Voluntary day care for older people based on Age Concern 1999/2000 survey from 10 day centres, uprated using PSS Pay & Prices index; Costing the bakers' dozen: RISE lunch club; National Evaluation of the expert patient programme, course; National Schedule of Reference Costs 2009-10 weighted average over all services (stroke, elderly, other), PCTDCFRAD tab
Medications				
Medications	Standard Quantity Units	0.01-419.62	Health and Social Care Information Centre (2011)	Price per unit (Nic/Qty (£))

Appendix 2

Descriptive Statistics of the 4-month Follow-up Data

Questionnaires at the short-term follow-up were administered by post only. The quality of the CSRI data available was worse than at baseline, when all questionnaires had been administered by interviewers; data quality was also worse than that of the 12-month follow-up, when the questionnaires had been administered by a mixture of interview and postal methods. As an example of the issues arising with the postal-only administration, a number of respondents opted to provide details of services they had used in the 'other' boxes on the CSRI. As a result, intensive scrutiny of individual forms was needed; where the service reported was actually a category of service use given in the questionnaire, these responses were re-classified as such.

Table A2.1 Baseline characteristics of participants with economic data available at baseline and 4 month follow-up across Telehealth sample

	Total baseline sample			**Participants completing 4-month follow-up study instruments***			**Participants not completing 4-month follow-up study instruments†**		
	UC (n=728)	**TH (n=841)**	**Raw**	**UC (n=425)**	**TH (n=544)**	**Raw**	**UC (n=300)**	**TH (n=286)**	**Raw**
Mean years of age (SD)	70.6 (11.8)	70.1 (11.8)	–0.5	71.4 (15.6)	70.7 (16.8)	–0.6	69.5 (17.7)	68.9 (17.7)	–0.6
Under 65 (young)	215 (30%)	242 (29%)	-1%	113 (27%)	141 (26%)	-1%	102 (34%)	98 (34%)	0%
65-74 (young old) ‡	214 (29%)	288 (34%)	5%	133 (31%)	203 (37%)	6%	78 (26%)	83 (29%)	3%
75-84 (old old)	239 (33%)	243 (29%)	-4%	143 (34%)	167 (31%)	-3%	96 (32%)	72 (25%)	-7%
85+ (oldest old) §	60 (8%)	68 (8%)	0%	36 (8%)	33 (6%)	-2%	24 (8%)	33 (12%)	4%
Women	290 (40%)	347 (41%)	1%	169 (40%)	207 (38%)	-2%	121 (40%)	134 (47%)	7%
Mean IMD score (SD) ‡‖	28.6 (52.2)	27.7 (55.3)	-0.9 (-6.2, 4.5)	26.9 (38.7)	24.8 (43.3)	-2.1 (-7.3, 3.1)	31.1 (35.3)	33.2 (35.5)	2.1 (-3.7, 7.9)
1st quintile‡	130 (18%)	215 (26%)	8%	86 (20%)	163 (30%)	10%	43 (14%)	49 (17%)	3%
2nd quintile‡	164 (23%)	140 (17%)	-6%	108 (25%)	103 (19%)	-6%	55 (18%)	37 (13%)	-5%
3rd quintile‡	124 (17%)	155 (18%	1%	80 (19%)	112 (21%)	2%	43 (14%)	41 (14%)	0%
4th quintile‡	168 (23%)	165 (20%)	-3%	82 (19%)	101 (19%)	-1%	86 (29%)	62 (22%)	-7%
5th quintile‡¶	142 (20%)	166 (20%)	0%	69 (16%)	65 (12%)	-4%	73 (24%)	97 (34%)	10%

Index condition									
COPD	244 (34%)	334 (40%)	6%	144 (34%)	236 (43%)	10%	98 (33%)	94 (33%)	0%
Heart failure	275 (38%)	263 (31%)	–7%	171 (40%)	190 (3%)	–4%	103 (34%)	70 (24%)	-10%
Diabetes**	209 (29%)	244 (29%)	–0%	110 (26%)	118 (22%)	–4%	99 (33%)	122 (43%)	10%
No of comorbidities	2 (2.8)	1.8 (2.9)	–0.2	2.0 (2.5)	1.8 (2.7)	–0.2	2.1 (2.3)	1.9 (2.3)	–0.2
WSD site									
Site 1††	234 (32%)	256 (30%)	–2%	142 (33%)	192 (35%)	2%	91 (30%)	61 (21%)	–9%
Site 2‡‡	283 (39%)	342 (41%)	2%	183 (43%)	244 (45%)	2%	98 (33%)	94 (33%)	0%
Site 3***	211 (29%)	243 (29%)	–0%	100 (24%)	108 (20%)	–4%	111 (37%)	131 (46%)	9%
White British ethnicity‡	630 (87%)	735 (87%)	1%	388 (91%)	501 (92%)	1%	239 (80%)	225 (79%)	-1%
Living alone‡	195 (27%)	229 (27%)	0%	121 (28%)	139 (26%)	-3%	73 (24%)	85 (30%)	5%

Data are mean (cluster-adjusted standard deviation) or number (%) of patients.
UC=usual care; TH=telehealth; COPD=chronic obstructive pulmonary disease; SD=Standard deviation.
*cases where costs and outcomes data were available
† Outcomes instruments not completed and/or CSRI not completed
‡ Difference within TH: differences between completion/non-completion clustered χ^2= 4.591 and p<0.05
§ Difference within TH: differences between completion/non-completion clustered χ^2= 7.333 and p<0.01
‡ Imputed data
‖ Difference within TH: clustered *t*= 2.086 p<0.05
¶ Difference within TH: clustered χ^2=5.669 p<0.05

**Difference within UC: differences between completion/non-completion: z=2.084, P<0.05. Difference within TH: differences between completion/non-completion: clustered χ^2 6.470, P<0.05
††Difference within TH: differences between completion/non-completion *z*=-4.154, P<0.001
‡‡ Difference within TH: differences between completion/non-completion *z*=-3.340, P<0.001. Difference within UC: differences between completion/non-completion: *z*=-2.829, P<0.01
*** Difference within TH: differences between completion/non-completion z=7.847, P<0.001. Difference within UC: differences between completion/non-completion: *z*=3.933, P<0.001

Table A2.2 Number and size of clusters, participants with economic data available at baseline and 4 month follow-up across Telehealth sample

	Total baseline sample		Participants completing 4-month follow-up study instruments*	
	UC (N=73)	TH (N=81)	UC (N=64)	TH (N=74)
Cluster mean [min – max]	10 [1-44]	10.4 [1-48]	6.6 [1-28]	7.4 [1-35]

Data are mean [min – max]
UC=usual care; TH=telehealth; COPD=chronic obstructive pulmonary disease.
*where costs and outcomes data were available

Table A2.3 Mean service costs (standard errors) over previous 3 months across Telehealth sample, available cases at 4-month follow-up (imputed data)

Resource item	Control (n=425)	Telehealth (n=547)	Difference
	Mean (SE)	Mean (SE)	Mean (95% CI)
Hospital use*	552 (56)	467 (50)	-85 (-234, 64)
Community health services/primary care*	166 (21)	140 (18)	-27 (-82, 28)
Community mental health*	2 (1)	2 (1)	1 (-3, 4)
Community care services*†	118 (26)	94 (23)	-24 (-93, 45)
Care home respite*	7 (3)	8 (3)	2 (-6, 9)
Day services LA*	29 (8)	14 (7)	-14 (-35, 6)
Day services NHS*	4 (2)	0 (2)	-4 (-11, 2)
Medications*	301 (8)	322 (7)	21 (1, 41)*
Equipment/Adaptations LA*	3 (1)	2 (1)	-1 (-4, 2)
Equipment LA/Adaptations NHS*	0 (0)	0 (0)	0 (0, 0)
Total costs exc. telehealth delivery& equipment*	1182 (72)	1050 (63)	-132 (-321, 57)
Telehealth intervention	4 (9)	165 (9)	161 (136, 186)**
Telehealth equipment	6 (5)	289 (4)	283 (270, 295)**
Total costs inc. telehealth delivery& equipment*	1193 (72)	1504 (63)	311 (122, 501)**

Note: Includes cases where baseline cost data were missing. Imputed data (10 completed datasets).
† Includes community alarms

Table A2.4 Baseline characteristics of participants with economic data available at baseline and 4-month follow-up across Telecare sample

	Total baseline sample			Participants completing 4-month follow-up study instruments*			Participants not completing 4-month follow-up study instruments†		
	UC	TC	Raw	UC	TC	Raw	UC	TC	Raw
	(n=634)	(n=548)		(n=256)	(n=261)		(n=371)	(n=280)	
Mean years of age (SD)‡	74.3 (17.5)	74 (17.1)	-0.3	74.2 (14.9)	75.0 (15.1)	0.7	69.5 (17.7)	68.9 (17.7)	-0.6
Under 65 (young)	138 (22%)	129 (24%)	2%	59 (23%)	60 (23%)	0%	78 (23%)	69 (25%)	4%
65-74 (young old)	139 (22%)	116 (21%)	-1%	51 (20%)	47 (18%)	-2%	87 (23%)	68 (24%)	1%
75-84 (old old)	208 (33%)	168 (31%)	-2%	95 (37%)	188 (34%)	-3%	110 (30%)	76 (27%)	-3%
85+ (oldest old)	149 (24%)	135 (25%)	-1%	51 (20%)	66 (25%)	5%	96 (26%)	67 (24%)	-6%
Female	415 (65%)	344 (63%)	2%	173 (68%)	167 (64%)	-4%	237 (64%)	174 (62%)	-2%
Mean comorbidities (SD)	1.1 (1.6)	1.1 (1.6)	-0.0	1.1 (1.5)	1 (1.5)	-0.1	1 (1.7)	1.1 (1.7)	0.0
White-British‡§	561 (89%)	482 (88%)	-1%	233 (91%)	243 (93%)	0%	324 (87%)	232 (83%)	-4%
WSD site									
Site 1	137 (22%)	125 (23%)	1%	58 (23%)	58 (22%)	0%	75 (20%)	67 (24%)	4%
Site 2‖	309 (49%)	273 (50%)	1%	169 (52%)	145 (56%)	4%	176 (47%)	122 (44%)	-4%
Site 3¶	188 (30%)	150 (27%)	-2%	125 (25%)	158 (22%)	-3%	120 (32%)	91 (33%)	0%

	Total baseline sample			Participants completing 4-month follow-up study instruments*			Participants not completing 4-month follow-up study instruments†		
	UC	TC	Raw	UC	TC	Raw	UC	TC	Raw
	(n=634)	(n=548)		(n=256)	(n=261)		(n=371)	(n=280)	
IMD	28.8 (40.4)	27.8 (38.2)	-0.7	27.5 (28.0)	25.3 (28.9)	-2.2	29.7 (32.1)	30.3 (28.8)	0.6
1st quintile§	152 (24%)	127 (23%)	-1%	74 (29%)	76 (29%)	0%	81 (22%)	46 (16%)	-5%
2nd quintile§	82 (13%)	109 (20%)	7%	31 (12%)	50 (19%)	7%	49 (13%)	57 (20%)	7%
3rd quintile§	133 (21%)	100 (18%)	-3%	52 (20%)	51 (20%)	-1%	77 (21%)	47 (17%)	-4%
4th quintile§	120 (19%)	102 (19%)	0%	45 (18%)	43 (16%)	-1%	74 (20%)	62 (22%)	2%
5th quintile§	146 (23%)	110 (20%)	-3%	54 (21%)	41 (16%)	-5%	90 (24%)	68 (24%)	0%
Living alone§	340 (54%)	285 (52%)	-2%	137 (54%)	130 (50%)	-4%	196 (53%)	149 (53%)	0%

Data are mean (cluster-adjusted standard deviation) or number (%) of patients.
UC=usual care; TC=telecare
*Costs and outcomes data available
† Outcomes instruments not completed and/or CSRI not completed
‡ Within UC: differences between completion/completion sample $p<0.05$ on clustered t-test
§ Imputed data
|| Within TC: differences between completion/completion $p<0.05$ on z-test of proportions
¶ Within TC: differences between completion/completion $p<0.05$ on z-test of proportions

Table A2.5 Number and size of clusters, participants with economic data available at baseline and 4-month follow-up across Telecare sample

	Total baseline sample		Participants completing 4-month follow-up study instruments*	
	UC	TC	UC	TC
	(N=103)	(N=101)	(N=89)	(N=85)
Cluster mean [min – max]	6.2 [1-26]	5.4 [1-21]	2.9 [1-11]	3.1 [1-12]

Data are mean [min – max]
UC=usual care; TC=telecare
*where costs and outcomes data were available

Table A2.6 Mean service costs (standard errors) over previous 3 months across Telecare sample, available cases at 4-month follow-up (imputed data)

Resource item	Control (n=259)	Telecare (n=262)	Difference
	Mean (SE)	Mean (SE)	Mean (95% CI)
Hospital use*	682 (139)	621 (139)	-61 (-449, 327)
Community health services/primary care*	150 (26)	210 (26)	61 (-12, 134)
Community mental health*	25 (11)	25 (11)	1 (-29, 30)
Community care services*†	525 (89)	492 (90)	-33 (-283, 216)
Care home respite*	0 (5)	8 (5)	8 (-7, 23)
Day services LA*	135 (28)	163 (28)	28 (-50, 106)
Day services NHS*	24 (16)	18 (16)	-6 (-50, 38)
Medications*	190 (11)	188 (11)	-1 (-32, 30)
Equipment/Adaptations LA*	7 (1)	8 (1)	1 (-2, 5)
Equipment LA/Adaptations NHS*	1 (1)	1 (1)	-1 (-2, 1)
Total costs exc. telecare delivery and equipment*	1737 (208)	1734 (210)	-4 (-587, 579)
Telecare intervention	9 (7)	186 (7)	177 (158, 196)**
Telecare equipment	1 (1)	20 (1)	20 (18, 21)**
Total costs incl. TC delivery & equipment	1747 (209)	1940 (211)	193 (-393, 779)

Note: Includes cases where baseline cost data are missing. Imputed data (10 completed datasets).
† Includes community alarms

Appendix 3

Three-level Models of Total Costs: Results

The tables below present the results of the three-level models investigated and reported in Chapter 6, Sections 6.2.1 and 6.17: three-level null models, DDD models without covariates (presented here but not discussed in the chapter) and DDD models with covariates.

A3.1 Telehealth: Models

Table 3.1 Estimates, subject specific (random intercept) models of total costs (including intervention) in 3 months prior to baseline and 12-month follow-up

	(1)	(2)	(3)	(4)	(5)
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
TH		0.923	1.025		0.933
		(0.103)	(0.105)		(0.105)
Followup		0.965	0.964		0.977
		(0.115)	(0.117)		(0.109)
TH*Followup		1.595***	1.537**		1.575***
		(0.215)	(0.215)		(0.210)
HF		0.897	0.887		0.913
		(0.110)	(0.102)		(0.110)
Diab		0.824	0.861		0.857
		(0.107)	(0.096)		(0.109)
TH*HF		1.124	1.048		1.095
		(0.197)	(0.181)		(0.176)
TH*Diab		1.355	1.248		1.352+
		(0.255)	(0.211)		(0.238)
HF*Followup		1.223	1.254		1.208
		(0.189)	(0.194)		(0.181)
Diab*Followup		1.433**	1.376*		1.416*
		(0.191)	(0.193)		(0.214)
TH*Followup*HF		0.824	0.794		0.830
		(0.158)	(0.156)		(0.154)
TH*Follow*Diab		0.680*	0.718		0.684+
		(0.132)	(0.146)		(0.135)
Young old			1.107+		
			(0.067)		
Old-old			1.114		
			(0.075)		

	(1)	(2)	(3)	(4)	(5)
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
Oldest old			1.649***		
			(0.158)		
GCSE/O/A-level			1.191***		
			(0.058)		
Degree-level			1.122+		
			(0.074)		
Female			1.099*		
			(0.041)		
White-British			1.185		
			(0.123)		
Comorb			1.144***		
			(0.019)		
Owns			0.889*		
			(0.052)		
Site 2			1.075		
			(0.065)		
Site 3			1.075		
			(0.095)		
IMD			1.003		
			(0.002)		
Some problems			1.555***		
			(0.071)		
Unable wash/dress			2.460***		
			(0.359)		
Level 1 constant	1119.321***	1006.264***	580.169***	1094.552***	965.521***
	(38.810)	(88.142)	(85.309)	(31.568)	(85.581)

	(1)	(2)	(3)	(4)	(5)
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
σ	0.781***	0.749***	0.746***	0.782***	0.749***
	(0.035)	(0.037)	(0.035)	(0.035)	(0.037)
$\sigma^2[u\ ]$				1.414***	1.467***
				(0.050)	(0.053)
$\sigma^2[u^3]$	1.041*	1.033*	1.002		
	(0.019)	(0.016)	(0.005)		
$\sigma^2[u^2]$	1.366***	1.423***	1.234***		
	(0.053)	(0.055)	(0.040)		
N	1930	1930	1930	1930	1930

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001

A3.2 Telecare: Models

Table 3.2 Estimates, subject specific (random intercept) models of total costs (including intervention) in 3 months prior to baseline and 12-month follow-up

	(1)	(2)	(3)	(4)	(5)
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
TC		1.096	1.108		1.125
		(0.144)	(0.114)		(0.135)
Follow-up		0.836^{+}	0.812*		0.833*
		(0.076)	(0.076)		(0.070)
Follow-up*TC		1.304*	1.273^{+}		1.318*
		(0.172)	(0.174)		(0.165)
Lives w/		1.186^{+}	0.983		1.234^{+}
		(0.118)	(0.093)		(0.143)
TC*Lives w/		1.032	0.992		1.019
		(0.142)	(0.130)		(0.168)
Follow-up*Lives w/		1.049	1.032		1.047
		(0.128)	(0.135)		(0.137)
TC* Follow-up*Lives w/		0.766	0.793		0.763
		(0.128)	(0.142)		(0.140)
Young old			0.766**		
			(0.065)		
Old-old			0.797**		
			(0.062)		
Oldest old			0.785*		
			(0.077)		
Below-degree			0.996		
			(0.059)		
Degree			1.066		
			(0.104)		
Female			1.025		
			(0.066)		

	(1)	(2)	(3)	(4)	(5)
Parameter	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)	Exp (β) (SE)
White-British			1.082		
			(0.112)		
Number of comorbidities			1.141***		
			(0.020)		
Owns			0.967		
			(0.070)		
Site 2			1.513***		
			(0.111)		
Site 3			0.873		
			(0.106)		
Mean IMD score			0.998		
			(0.003)		
Some ADL problems			1.453***		
			(0.090)		
Unable to wash/dress			2.693***		
			(0.224)		
Constant	1688.180***	1530.968***	1124.082***	1638.815***	1451.222***
	(88.426)	(147.369)	(198.306)	(64.533)	(124.099)
	0.797***	0.794***	0.790***	0.800***	0.797***
σ	(0.034)	(0.034)	(0.031)	(0.034)	(0.034)
$\sigma^2[u\ \ ]$				1.728***	1.695***
				(0.098)	(0.093)
$\sigma^2[u^3]$	1.222***	1.202***	1.017		
	(0.046)	(0.045)	(0.019)		
$\sigma^2[u^2]$	1.429***	1.427***	1.237***		
	(0.072)	(0.072)	(0.051)		
N	1506	1506	1506	1506	1506

+ p<0.1 *, p<0.05 **, p<0.01, ***p<0.001