Multimorbidity and integrated care

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2016

JONATHAN STOKES

Division of Population Health, Health Services Research, and Primary Care

Contents

Word count: 60,357

Figures

Main thesis:

Publication chapters:

Tables

Publication chapters:

Chapter	Table	Thesis page number	Publication page number
5	1: Outcome measures	63	7
5	2: Subgroup analyses	63	8
5	3: Demographics of included studies	63	10-11
5	4: Context of included studies	63	12-17
5	5: Details of interventions	63	18-23
5	6: Quality of included studies	63	24-26
5	7: Results of subgroup analyses	63	33
6	1: Overview of intervention	65	2
6	2: Individual baseline characteristics (before and after matching)	65	5
6	3: Average absolute outcomes per patient per month by PICT, pre-intervention/post-intervention	65	5
6	4: Individual-level adjusted model results	65	6
6	5: Practice characteristics (wave 1 compared to later joining practices)	65	6
6	6: Average absolute outcomes per 1000 patients per month by PICT (wave 1 compared to practices joining the intervention at a later date), pre-intervention/post-intervention	65	8
6	7: Practice-level adjusted model results	65	9
7	1: Number (and proportion) of patients positively identified by each dummy variable	67	11
7	2: Correlation between multimorbidity dummies used	67	12

Abbreviations

CBA: Controlled Before and After
CCG: Clinical Commissioning Group
CKD: Chronic Kidney Disease
COPD: Chronic Obstructive Pulmonary Disease
DALY: Disability Adjusted Life Year
DD: Difference-in-differences
DDD: Difference-in-Difference-in-Differences
DEPLESET: Demographics, Epidemiology, Political, Legal, Ecological, Socio-cultural, Economic, Technological
DES: Directed Enhanced Services
DoH: Department of Health
DRG: Diagnosis Related Groups
ED: Emergency Department
EPOC: Cochrane Effective Practice and Organisation of Care group
EU: European Union
FYFV: Five Year Forward View
GDP: Gross Domestic Product
GP: General Practitioner
HIV: Human Immunodeficiency Virus
ICP: Integrated Care Pilot
IT: Information Technology
ITS: Interrupted Time Series
MDT: Multidisciplinary team
NCD: Non-Communicable Disease
NHS: National Health Service
NICE: National Institute for Health and Care Excellence
nRCT: non-Randomised Controlled Trial
OECD: Organisation for Economic Co-operation and Development
P4P: Pay-for-Performance
PCP: Primary Care Practitioners
PHC: Primary Health Care
PICT: Practice Integrated Care Teams
PPI: Patient and Public Involvement
QALY: Quality Adjusted Life Year
QOF: Quality and Outcomes Framework
RCT: Randomised Controlled Trial
RD: Regression Discontinuity
SELFIE: Sustainable intEgrated care modeLs for multi-morbidity: delivery, FInancing and pErformance
UHC: Universal Health Coverage
UK: United Kingdom
US(A): United States (of America)
WHO: World Health Organization

Abstract

Background

Health systems internationally face a common set of challenges: ageing populations, increasing numbers of patients suffering from multiple long-term conditions (multimorbidity) and severe pressure on health and care budgets. 'Integrated care' is pitched as the solution to current health system challenges. But, in the literature, what integrated care actually involves is complex and contested.

Aims

1. What does 'integrated care' currently look like in practice in the NHS?
2. What is the effectiveness of current models of 'integrated care'?
3. To what extent are there differential effects of 'integrated care' for different types of multimorbidity?

Methods

The thesis utilises routinely collected data, systematic review and meta-analysis, combined with quasi-experimental methods (difference-in-differences, and subgroup analysis, difference-in-difference-in-differences).

Results

The current implementation of the concept of integrated care is predominantly carried out through multidisciplinary team (MDT) case management of 'at risk' (usually of secondary-care admissions) patients in primary care. This approach, however, has not proven capable of meeting health outcome and utilisation/cost aims. Patient satisfaction, though, is consistently improved by the approach. There might also be positive spill-over effects of increased team-working through MDTs for the wider practice population. There does not appear to be a multimorbidity subgroup which benefits significantly more than others in terms of secondary-care utilisation or cost. However, patients at the end of life and/or those with only primary-care sensitive conditions might benefit slightly more than others.

Conclusions

Integrated care, in its current manifestation, is not a silver bullet that will enable health systems to simultaneously accomplish better health outcomes for those with long-term conditions and multimorbidity while increasing their satisfaction with services and reducing costs. The current financial climate might mean that other means of achieving prioritised aims are required in the short-term, with *comprehensive* primary care and population health strategies employed to better prevent/compress the negative effects of lifestyle-associated conditions in the longer-term.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

Thanks to my supervisors Pete Bower, Kath Checkland, Søren Rud Kristensen and Sudeh Cheraghi-Sohi and to my friends and family for their support over the last three years, and to the National Institute for Health Research (NIHR) for funding me.

The Author

Qualifications

Master of Public Health (MPH) – Imperial College London (2013)
BSc Medical Sciences (Honours) – The University of Edinburgh (2011)

Publications

Whittaker, W, Anselmi, L, Kristensen, SR, Lau, Y-S, Bailey, S, Bower, P, Checkland, K, Elvey, R, Rothwell, K, **Stokes, J** & Hodgson, D. 2016, 'Associations between Extending Access to Primary Care and Emergency Department Visits: A Difference-In-Differences Analysis' *PLoS Medicine*, vol 13, no. 9, pp. e1002113. DOI: 10.1371/journal.pmed.1002113

Stokes, J, Checkland, K & Kristensen, S. 2016, 'Integrated care: theory to practice' *Journal of Health Services Research & Policy*, vol 21, no. 4, pp. 282-285. DOI: 10.1177/1355819616660581

Atun, R, Gurol-Urganci, I, Hone, T, Pell, L, **Stokes, J**, Habicht, T, Lukka, K, Raaper, E & Habicht, J. 2016, 'Shifting chronic disease management from hospitals to primary care in Estonian health system: analysis of national panel data' *Journal of Global Health*, vol 6, no. 2, 020401. DOI: 10.7189/jogh.06.020401

Stokes, J, Kristensen, S, Checkland, K & Bower, P. 2016, 'Effectiveness of multidisciplinary team case management: difference-in-differences analysis' *BMJ Open*, vol 6, no. 4. DOI: 10.1136/bmjopen-2015-010468

Stokes, J, Panagioti, M, Alam, R, Checkland, K, Cheraghi-Sohi, S & Bower, P. 2015, 'Effectiveness of Case Management for 'At Risk' Patients in Primary Care: A Systematic Review and Meta-Analysis' *PLoS One*, vol 10, no. 7, e0132340. DOI: 10.1371/journal.pone.0132340

Panagioti, M, **Stokes, J**, Esmail, A, Coventry, P, Cheraghi-Sohi, S, Alam, R & Bower, P. 2015, 'Multimorbidity and Patient Safety Incidents in Primary Care: A Systematic Review and Meta-Analysis' *PLoS One*, vol 10, no. 8, e0135947, pp. e0135947. DOI: 10.1371/journal.pone.0135947

Stokes, J, Gurol-Urganci, I, Hone, T & Atun, R. 2015, 'Effect of health system reforms in Turkey on user satisfaction' *Journal of Global Health*, vol 5, no. 2, pp. 020403. DOI: 10.7189/jogh.05.020403

Anselmi, L, Bailey, S, Bower, P, Checkland, K, Elvey, R, Hodgson, D, Rud Kristensen, S, McBride, A, Lau, Y-S, Parkin, S, Rothwell, K, **Stokes, J**, Sutton, M & Whittaker, W. 2015, NHS Greater Manchester Primary Care Demonstrator Evaluation. NIHR CLAHRC Greater Manchester, Manchester, UK. DOI: 10.13140/RG.2.1.4795.5043

Stokes, J. 2014, 'CCG implementation of integrated care in the NHS' Health Services Research: Evidence-based practice London, UK. 1-3 July 2014, King's College, London, UK, 1/07/14 - 3/07/14, pp. 1-65.

Thesis Overview

This thesis takes the 'alternative' format, with the bulk of the chapters presented in the style of journal publications. To date, two of three (Chapters 5 and 6) of the papers presented within this thesis have been published in peer-reviewed journals (together with a perspective article which summarises the PhD findings as a whole (Stokes et al., 2016)). The paper presented in Chapter 7 is currently under review with the journal *BMC Health Services Research*. Furthermore, the papers have a logical order as they were published/planned in relation to the key objectives of the aims of the thesis, and so together they constitute a coherent piece of work, which further justifies this 'alternative thesis' format.

The first three chapters of the thesis provide a general introduction. I begin by outlining 'the challenge' that health systems around the world face as they encounter demographic and epidemiological transitions, with the burden of disease shifting from acute to chronic conditions (frequently many accumulating in the same individual over the life-course, i.e. multimorbidity).

I then examine the health system response to attempt to cope with this transition, so-called 'integrated care', a concept with a number of different definitions and implementations, and I introduce the context in which the PhD research took place (the NHS in England). Following this, I focus on introducing the dominant model of 'integrated care' in practice, multidisciplinary team (MDT) case management.

The fourth chapter gives a short overview of the methods employed during the research. Each of the methods is then discussed in more detail as part of each of the journal articles presented in Chapters 5-7.

From the literature summarised in the introductory sections, three overarching research questions emerged:

1. What does 'integrated care' currently look like in practice in the NHS?
2. What is the effectiveness of current models of 'integrated care'?

3. To what extent are there differential effects of 'integrated care' for different types of multimorbidity?

These questions are addressed in the empirical research. These three papers and how the work links together are outlined below:

Effectiveness of case management systematic review (Chapter 5)

A small initial pilot study conducted as part of the PhD (outlined in Chapter 2) took a random sample of 10% of the 211 Clinical Commissioning Groups (CCGs), which had been mandated to integrate care in the NHS in England at the time (2013), and found that a clearly dominant model (81% of CCGs) of integrated care was apparent: multidisciplinary team case management (MDT). However, despite its wide adoption in practice, the effectiveness of the case management intervention was unclear in the literature, and so led to the first piece of research for the PhD project.

This first empirical paper, which emerged from this research, set out to systematically identify the evidence of effectiveness of the dominant model of integrated care identified in the English NHS. Using systematic review and meta-analysis methods, effectiveness was measured across multiple health system outcomes (**Health** – *self-assessed health status*, *mortality*; **Cost** – *total cost of care*, *healthcare utilisation* [*primary and non-specialist care* and *secondary care* separately], and; **Satisfaction** –*patient satisfaction*). Meta-analysis showed little effect across all outcomes, with an indication of a small statistically significant increase in patient satisfaction the only major finding.

A number of gaps in the current evidence base were identified. For example, there was a complementary role identified for rigorous quasi-experiments in routine settings to better balance internal and external validity to the current heavy-reliance (78% of included studies in the review) on randomised controlled trials (RCTs). In addition, there was a limited perspective on outcomes - the majority of studies measuring effectiveness solely at an individual level, i.e. looking at average effects for individuals directly involved in the interventions and not at wider practice-level effects – potentially relevant with increased 'professional integration' through MDT working. These gaps subsequently shaped the next piece of work, detailed below.

Integrated care in Central Manchester (MDT case management - Chapter 6)

Here I aimed to contribute to the evidence base for MDT case management through an evaluation of a local integrated care intervention (Central Manchester CCG – Practice Integrated Care Teams (PICT)). I explored a number of health system utilisation outcomes using a robust quasi-experimental (difference-in-differences) study design. Effects were modelled distinctly at both the individual- (to capture direct effects of the intervention) and practice-levels (to capture any potential spill-over effects).

Consistent with the previous review findings, only slight, clinically minimal differences between case managed patients and usual care were identified at the individual-level (i.e. those high-risk patients directly receiving the intervention). At the practice-level, however, it seemed that there might be small beneficial spill-over effects of increased professional integration. These spill-over effects were not conclusive though, and they failed a robustness check.

With case management as an intervention generally aimed at complex multimorbid patients, the next piece of work assessed secondary analysis of this data, stratifying results by 'type' of multimorbidity. The aim was to identify any potential for better targeting the direct effects of the intervention to those who may benefit most.

MDT case management and 'type' of multimorbidity (Chapter 7)

There are multiple ways of conceptualising multimorbidity, and there is little evidence for what is effective for treating these complex patients. I carried out secondary analysis of the Central Manchester PICT data using difference-in-difference-in-differences analysis, to observe any subgroup effects of the intervention, stratifying results by:

1) Mental-physical comorbidity versus others;
2) 3+ chronic conditions versus <3;
3) Discordant versus concordant conditions;
4) Cardiovascular/metabolic cluster conditions only versus others;
5) Mental health-associated cluster conditions only versus others;

6) Musculoskeletal disorder cluster conditions only versus others;
7) Charlson index >5 versus others.

The majority of conceptualisations suggested little to no difference in effect between subgroups. Where results were significant, the vast majority of effect sizes identified in either direction were very small. The trend across the majority of the results appeared to show very slight increases of secondary-care admissions with treatment for the most complex patients (i.e. highest risk). The exceptions to this, patients with a Charlson index >5 may benefit slightly more from case management with decreased ambulatory care sensitive condition (ACSC) admissions (effect size (ES): -0.06) and inpatient re-admissions (30 days, ES: -0.05), and patients with only cardiovascular/metabolic cluster conditions may benefit slightly more with decreased inpatient non-elective admissions (ES: -0.12).

Only the three significant estimates for the musculoskeletal disorder cluster met the minimum requirement for at least a 'small' effect. Two of these estimates in particular were very large. This cluster represented only 0.5% of the total patients analysed, however, so is immensely vulnerable to the effects of outliers, and makes me extremely cautious of interpreting these as 'real' effects.

Following the presentation of this empirical work, I conclude the thesis with a discussion section, which synthesises the various outcomes of the individual papers. Here I also discuss overall strengths and limitations of the project as a whole, its relevance in the context of the wider literature and the relevance of the findings to researchers and policy-makers, and I indicate directions for future research opportunities. Finally, I present the overall conclusions of the PhD research.

Author contributions

First, as an alternative style thesis composed of papers that I have written with the input of co-authors, I will outline the role of each author in the core journal-style chapters that follow:

Effectiveness of case management review (Chapter 5)

Data acquisition: As first author, I led in designing the study protocol and search strategy, and ran the database searches for the systematic review. My PhD supervisors gave me feedback on the protocol and search strategy after the first draft was completed.

Analysis: I led the abstract and title screening process, and data extraction. For scientific rigour, a proportion of the screening was also independently carried out by a second author, and all data extraction of the final included papers was independently replicated and checked for consistency by a second author. I performed the meta-analysis, and this, too, was double-checked by the second author.

Writing: I led the manuscript writing on all drafts, and I responded to reviewer comments, and received and incorporated feedback from all co-authors.

Integrated care in Central Manchester (MDT case management) (Chapter 6)

Data acquisition: As above, as first author, I led in acquiring the data (after an initial introductory email to the CCG) and designing the protocol, with feedback from my PhD supervisors after the first draft.

Analysis: I led all data analysis, with feedback on results and summary statistics received from the second author on a regular basis throughout the analysis period.

Writing: I led the manuscript writing on all drafts, and responded to reviewer comments, and received and incorporated feedback from all co-authors.

MDT case management and 'type' of multimorbidity (Chapter 7)

Data acquisition: This paper used the data acquired for Chapter 6, and I designed the protocol with feedback from my PhD supervisors after the first draft.

Analysis: I led all data analysis, with feedback on results and summary statistics received from the second author on a regular basis throughout the analysis period.

Writing: I led the manuscript writing on all drafts, and received and incorporated feedback from all co-authors. We are currently awaiting peer review from the journal *BMC Health Services Research* on the initial manuscript submission.

1. Health system challenges

1.1 Chapter Introduction

As complex, dynamic systems, health systems are constantly evolving, and must adapt to new challenges (Atun, 2012; Lipsitz, 2012). The most important pressures currently are widely reported in the health systems literature (Nolte and McKee, 2008), and also regularly grab headlines in the wider popular media. These pressures are demographic and epidemiological transitions, multimorbidity and economic constraints. Understanding these pressures is vital for understanding changes that the health system could and should make in order to meet and improve on its so-called *triple aim*: improving the experience of care (user satisfaction), improving the health of populations (health gain), and reducing per capita costs of healthcare (cost-effectiveness) (Berwick et al., 2008).

1.2 Demographic and epidemiological changes

1.2.1 Expanding and ageing population

Globally, the population is expanding. Currently, the world's population stands at an estimated 7.3 billion people (up from 2.5 billion in 1950), and is projected to rise to 9.7 billion in 2050, and to 11.2 billion in 2100 (United Nations, 2015). There are a number of reasons for this expansion in population, and the majority relate to our previous successes at tackling key challenges that our societies and health systems have faced. These include major public health advances and improved medical understanding and technology (Bunker, 2001). These successes are particularly obvious in more developed countries.

Global life expectancy at birth is currently 70 years, projected to rise to 77 years in 2050, and 83 years in 2100 (projections dependent on further reductions in the spread of HIV, and successfully combatting other infectious and chronic diseases) (United Nations, 2015). Under-5 mortality has also decreased rapidly. These absolute declines have been particularly large in Sub-Saharan Africa and the least developed countries, as they catch up with the more developed countries in the West (United

Nations, 2015). However, there remains a large gap in both life expectancy and under-5 mortality between countries based on their GDP (Gross Domestic Product) levels (Houweling et al., 2005), which is a proxy for the countries' development level.

This increasing longevity has led to a re-structuring of the population pyramid. The segment of the population aged 60 or over currently comprises 12% of the global population, and this segment is also the fastest growing one, at a rate of 3.26% per year. Consequently, by 2050, with the exception of Africa, all major areas of the world are expected to have nearly 25% or more of their populations aged over 60 (United Nations, 2015). In addition, the fertility rate tends to vary to a large degree across regions, further exacerbating changes to the population spread. As countries become more developed, they tend towards major reductions in average family size, with middle-income countries currently at an intermediate-fertility level, and developing countries remaining at a fairly high-fertility level (United Nations, 2015). As a result, we now see proportionately (through restructuring of the population pyramid) and absolutely (by means of a general trend towards increasing global population) more elderly people in the population.

1.2.2 Changing lifestyles

With longer life and increasing wealth, a so-called modernisation of lifestyles has also played a role in these demographic and epidemiological changes (Nolte and McKee, 2008). For example, the type of work performed by the majority of the population has shifted from jobs in heavy industry involving manual labour typical in the past to more sedentary and safe office work in developed, post-industrial nations (Bell, 1976). In addition, as more of the population works outside the home, and there is consequently less time for meals to be cooked from scratch, processed foods are consumed more frequently, with time-saving benefits but also potential health risks (Bouvard et al., 2015). Many of these lifestyle changes can be protective against some forms of disease/illness. For example, injuries are reduced as the workplace becomes safer (Shahnavaz, 1987). But, these changes also act as risk factors for other injuries and diseases, for instance musculoskeletal disorders such as repetitive strain injuries (Keller et al., 1998), obesity and its related metabolic as well as cardiovascular problems (Bray et al., 2004; Hubert et al., 1983), and cancers (Bouvard et al., 2015).

1.2.3 Epidemiological transition

In the context of an overall expanding and ageing population, with changing lifestyle factors as countries become more developed, we observe an epidemiological transition. This transition sees overall absolute disease burden decrease, and the proportional disease burden shift from a majority of infectious diseases to a majority of lifestyle-related, chronic, long-term conditions as countries develop (Omran, 1971). Therefore, chronic disease is now a global problem.

As people live longer, their likelihood of developing a chronic disease increases (Nolte and McKee, 2008). The influences of the social determinants of health have been well documented (Wilkinson and Marmot, 2003), and an accumulation of these and other risk factors over the life-course acts to increase the likelihood of individuals developing not just one but several of these conditions over a longer lifespan (so-called multimorbidity). This is particularly true because many of our healthcare advances allow us to control but not necessarily cure these conditions (Nolte and McKee, 2008).

1.3 Multimorbidity

1.3.1 Definition

Multimorbidity is most simply defined in terms of quantity (i.e. multi) of chronic diseases (i.e. morbidity). One of the most cited definitions of multimorbidity is therefore Boyd and Fortin's; "the coexistence of two or more chronic conditions, where one is not necessarily more central than others" (Boyd and Fortin, 2010). What is included as a chronic condition is debatable, and many suggest that they include not only a number of non-communicable diseases (NCDs) but also some caused by infectious agents, which can now be controlled over a lifetime, for example HIV (Nolte and McKee, 2008). Others have also included risk factors (e.g. obesity) when applying the definition in practice (Prazeres and Santiago, 2015; Willadsen et al., 2016). However multimorbidity is defined, though, a common issue is that increasingly complex, extended, and coordinated responses are required by multiple practitioners to manage such patients (Nolte and McKee, 2008).

Boyd and Fortin's definition builds on van den Akker's separation of multimorbidity from a similar concept, that of comorbidity. Comorbidity differs in its

emphasis on a single index condition, with other diseases said to co-exist with this index condition simultaneously (Feinstein, 1970; van den Akker et al., 1996). This subtle but significant difference between the definitions of comorbidity and multimorbidity (i.e. the lack of a single index condition in multimorbidity) means that the latter term is more suitable for application with regard to the primary care setting. After all, the identification of an index disease bears little obvious or useful relevance in primary care as treatment focus is on the patient as a whole (Valderas et al., 2009). In addition, seeing all diseases as equally important better frames the problem, allowing researchers to approach analysis of multimorbidity through a systems perspective, e.g. by paying attention to interactions.

Simply counting the number of conditions is the most obvious practical implementation of the simplest definition of multimorbidity and can be easily applied in practice. That is why to date, this simple measurement technique has been used extensively to describe the extent of the multimorbidity problem (estimating prevalence rates; and identifying where the greatest burden may lie), as well as for outlining common associations with demographic factors for instance. This simple approach is clearly limited, however.

Other, more complex conceptualisations of multimorbidity exist in the literature. These can be broadly categorised into three additional groups (beyond the simple count concept): 1. Grouping chronic diseases by dyads or triads (e.g. common disease clusters); 2. Using an index of variable complexity, based, for example, on risk or on past healthcare utilisation identified as associated with substantial future care (e.g. Charlson index – detailed in Chapter 7), 3. Identifying homogeneous groups of people with common diseases and characteristics (e.g. those with both mental and physical comorbidities) (Lefèvre et al., 2014). With respect to the aims of adapting services and interventions to better manage patients with multimorbidity, however, it is unknown which of these approaches has the most practical applicability. Ultimately, measuring complexity with a simpler model may not be adequate, although it remains to be seen how much the existing alternatives can offer for the interpretation of subgroup analyses, or adaptation of interventions tailored to specific groups.

1.3.2 Prevalence

Estimates of the prevalence of multimorbidity vary widely in the literature, although its commonality is well established (Salisbury et al., 2011). This variation in estimates is most likely attributable to the lack of common definition discussed above, and the resultant variation in criteria used for measurement, e.g. differences in selection of the number and types of condition included (van den Akker et al., 1998). One of the largest, and a nationally representative sample (a cross-section of the Scottish population in 2007), however, estimated that 23.2% of the population are multimorbid (defined as two or more chronic conditions), with this proportion increasing with age (Barnett et al., 2012). A subsequent large study from the Basque country, furthermore, also found a commensurate proportion to Barnett et al (2012), with multimorbidity identified in 23.6% of the population (Orueta et al., 2014). But, these estimates, although large, are in fact fairly conservative compared to those of many other studies (Fortin et al., 2007b). The seminal Barnett et al. study also outlined the difference between the proportional prevalence (the proportion rises with age) and the absolute prevalence of multimorbidity (in absolute numbers, it is more common in those younger than 65) (Smith et al., 2012; Barnett et al., 2012). This highlights the importance of issues associated with multimorbidity for the population as a whole. In particular, though, the rate is higher in more deprived populations, where onset of multimorbidity was found to occur 10-15 years earlier than in the most affluent (and this was especially true for multimorbidity including at least one mental health disorder) (Barnett et al., 2012). These associations have also been replicated by other study findings (Salisbury et al., 2011; Orueta et al., 2014; Boyd and Fortin, 2010).

The overall prevalence can be broken down to two broadly described population groups (see Figure 1) according to the Department of Health (DoH) 'Comorbidities framework':

- Firstly, there is the group of older patients, whose multimorbidities arise mostly from increased life expectancy, and exposure to risk factors over their extended life course.

- Secondly, there is a younger group, whose multimorbidities can instead be perceived to arise from either congenital or early-life disease and/or shorter-term, but more intense exposure to risk factors (e.g. smoking, obesity, alcohol consumption, physical inactivity etc.) resulting from widening social inequalities and deprivation, as well as lifestyle changes.

The DoH suggests that these two sub-groups may require different approaches to action, e.g. more of a focus on prevention and tackling the wider determinants of health for the younger group, and a focus on maintaining functioning and quality of life, with coordination particularly important for the older group (Department of Health, 2014b).

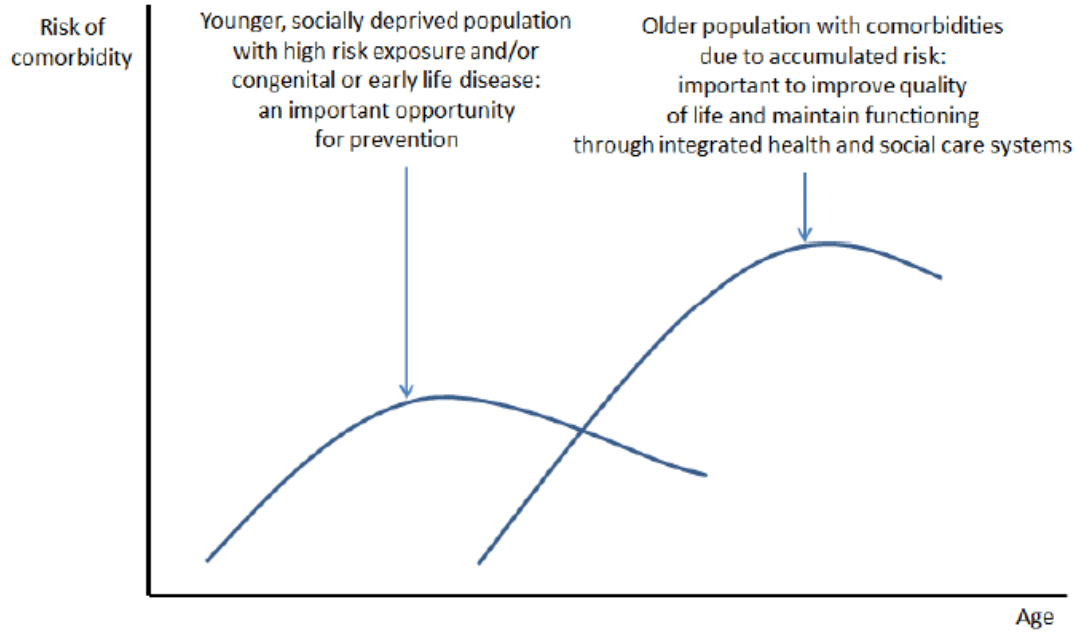


Figure 1: Two main population groups associated with multimorbidity (comorbidity) according to the Department of Health (DoH). Source: (Department of Health, 2014b)

1.3.3 The problem of multimorbidity

Multimorbidity is associated with a number of poor outcome measures. Decreased quality of life (Fortin et al., 2007b), increased hospital admissions (Smith et al., 2012) and time spent in hospital, increased complications following an operation, increased mortality rate (Fortin et al., 2007b), disability, frailty (Marengoni et al., 2011), and increased risk of developing further chronic mental health conditions (e.g.

depression) (Smith et al., 2012) are all more likely for patients with multimorbidity. These factors create additional healthcare needs for the patient, potentially creating self-management and additional treatment issues (e.g. potentially complex drug regimens - polypharmacy), the need to attend multiple appointments, and the added psychological distress caused by attempting to manage these added burdens (Wallace et al., 2015).

For physicians, these factors increase the clinical complexity of managing multiple conditions. Patients with multimorbidity constitute the majority of consultations for General Practitioners (GPs) (Salisbury et al., 2011). However, frequently multiple practitioners (across various levels of the health and care system) are required to manage these patients, which poses the potential for coordination and communication problems.

Evidence-based treatment of single long-term conditions has to date been led by guideline-based care, and chronic care pathways have also been targeted at single conditions. Physicians are given advice on the optimum drug and treatment regimens for a patient with that single condition only, and the guidelines that are frequently used tend to focus too narrowly on a single disease, with disease or treatment interactions rarely mentioned. This makes them unsuitable for adaptation to the multimorbid patient groups, as following multiple guidelines simultaneously has proven in practice to lead to unsafe drug interactions as well as likely duplication of efforts on the part of practitioners, which in turn results in a large burden of care for the patient (Boyd et al., 2005; Weiss et al., 2014). Recent guidelines for multimorbidity from the National Institute for Health and Care Excellence (NICE) suggest ways to address these issues, in particular the care burden (NICE, 2016b). The specificity of these recommendations is low, however, on account of the sparse evidence available for combination (Moffat and Mercer, 2015; Fraccaro et al., 2015).

The complexity of management, furthermore, can be expected to increase the opportunity for suboptimal management (Zulman et al., 2013). However, some studies have found that multimorbid patients' self-reported experiences of care are not necessarily poorer than those with single conditions (Bower et al., 2013). But as the authors of one such study state, a low response rate common in this type of study may miss many of those patients with multimorbidity receiving poor care (Bower et al., 2013).

In contrast, some qualitative studies have reported shortfalls in the delivery of care to patients with multimorbidity (Bayliss et al., 2008), and other studies have reported decreased quality of care for specific groups of conditions, particularly for those which are discordant/unrelated (Zulman et al., 2013). It may therefore be the case that there are sub-groups of patients at increased risk of poor care (Bower et al., 2013). This theory would accord with the findings of a recent systematic review, which concluded that the association between multimorbidity and patient safety incidents varies by 'type' of multimorbidity and 'type' of patient safety incident (Panagioti et al., 2015).

On the basis of the evidence summarised above, it is commonly assumed that patients with multimorbidity are at increased risk of receiving fragmented, inefficient, ineffective care, which increases their likelihood of preventable hospital admissions and complications (Boyd and Fortin, 2010). The complexity and frequency of interactions with multiple health services, in turn, increases vulnerability to coordination breakdowns (Bower et al., 2013), and presents organisational challenges for the health system (Fortin et al., 2007b). This indicates that multimorbid patients may have greater need in relation to less complex patients for well-coordinated, holistic, patient-centred care. However, to date there has been a lack of studies comparing interventions for effectively treating multimorbidity generally (Smith et al., 2016), and the effects of interventions on multimorbid patients through secondary analysis have similarly been neglected by those analysing interventions primarily aimed at other populations (van den Akker et al., 2015). There is therefore a gap in the literature concerning what is actually effective (or not) to manage these patients.

This gap in knowledge, furthermore, has major practical implications for delivery of healthcare. After all, it is frequently argued by health system researchers that current health systems have evolved to treat acute conditions and single chronic conditions, with a highly specialised, siloed approach working well to this end (with this success in turn leading to the epidemiological transition, as outlined in section 1.2.3 above). Treating patients with multiple chronic conditions in this arguably outdated health system setup, however, has major effectiveness and efficiency issues, with obvious cost implications.

1.4 Economic constraints

1.4.1 Increasing cost of care

In the majority of countries, healthcare expenditure (as a proportion of GDP) is increasing steadily (see Figure 2). Analysts frequently identify ever-advancing medical technology (e.g. drugs and medical devices) and diffusion of this technology through the health system as the major cost drivers. However, a number of other important factors also contribute: expanding and ageing populations with a more complex patient mix (as discussed above); increased public demand and expectations; personal income growth and labour costs for physicians; and inefficiencies in healthcare organisation (Sorenson et al., 2013).

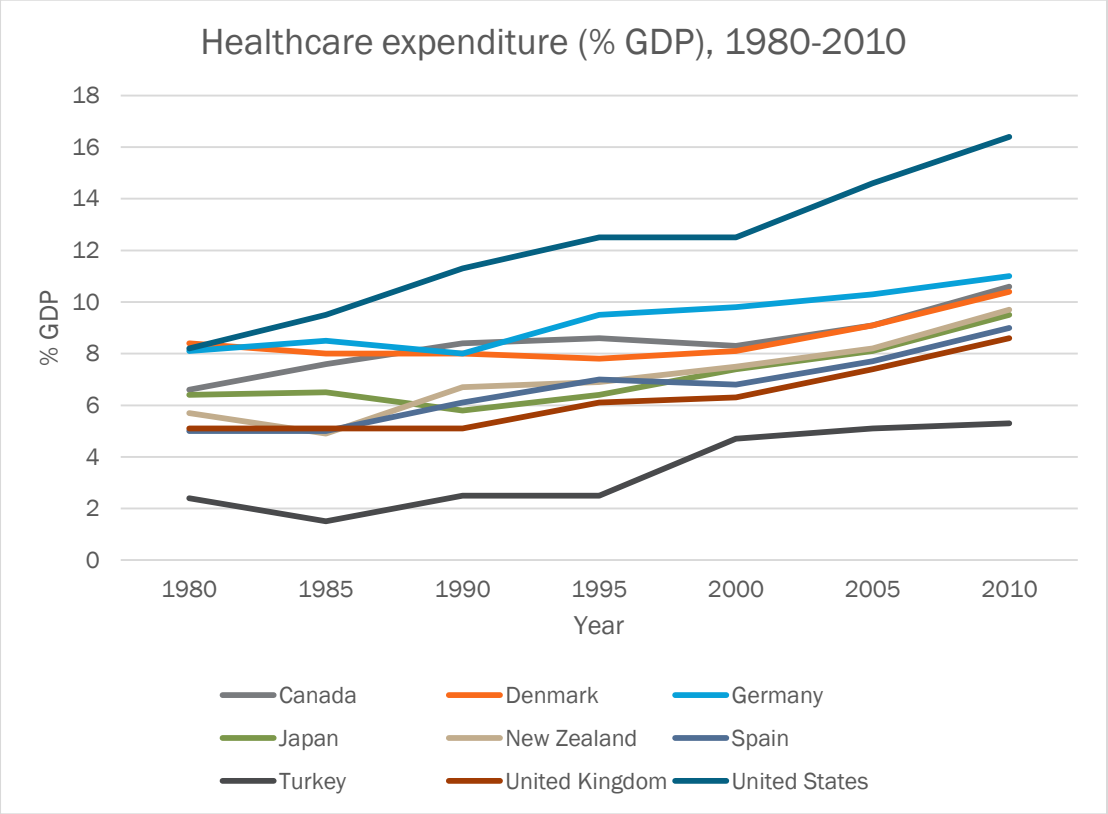


Figure 2: Share of GDP (%) for healthcare expenditure (1980-2010) in selected OECD countries. Source: http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT

As previously mentioned, managing multiple conditions requires increasingly complex interventions and multiple practitioners. Doing this over a long period of time for an increasing proportion of the population has obvious cost implications, unless there is a significant compression of morbidity, i.e. a reduced amount of time at the end of life in worse health (Fries, 1980). Evidence of whether this compression ever in fact occurs is mixed, and appears to be influenced by the type of health indicator chosen. Studies

using chronic disease indicators, for instance, tend to suggest an expansion of morbidity, while those using disability- or impairment-related measures tend to support the theory of compression (Chatterji et al., 2014).

There is, however, fairly clear evidence from multiple health system settings that treatment costs increase with increasing numbers of chronic conditions (thought to be expanding rather than compressing, as mentioned above). As the number of conditions increases, utilisation of health services (at all levels) also increases, and total healthcare costs subsequently increase too (Brilleman et al., 2014; Glynn et al., 2011; Bähler et al., 2015; Yoon et al., 2014). Multimorbidity, and not age, is thus found to be the key driver of health and social care costs (Kasteridis et al., 2015).

The majority of healthcare costs are consequently generally deemed to be driven by a small number of extremely high-cost patients (so-called *super utilisers* who tend to be multimorbid patients frequently with additional social complications) (Gawande, 2011). The majority of studies on the costs attributed to this group originate from the US and Canada (Chechulin et al., 2014), with a recent estimate by the United States Government Accountability Office stating that the most expensive 5% of Medicaid-only enrolees accounted for 48% of Medicaid-only health spending in each year from 2009 to 2011 (GAO, 2015). This group has therefore become the focus (in the US, Canada, and beyond) of many policy initiatives aimed at cutting health spending, i.e. 'high-risk' patients at the 'top of the population pyramid' (referring to the Kaiser pyramid shown in Figure 3) are being targeted (Boult et al., 2008).

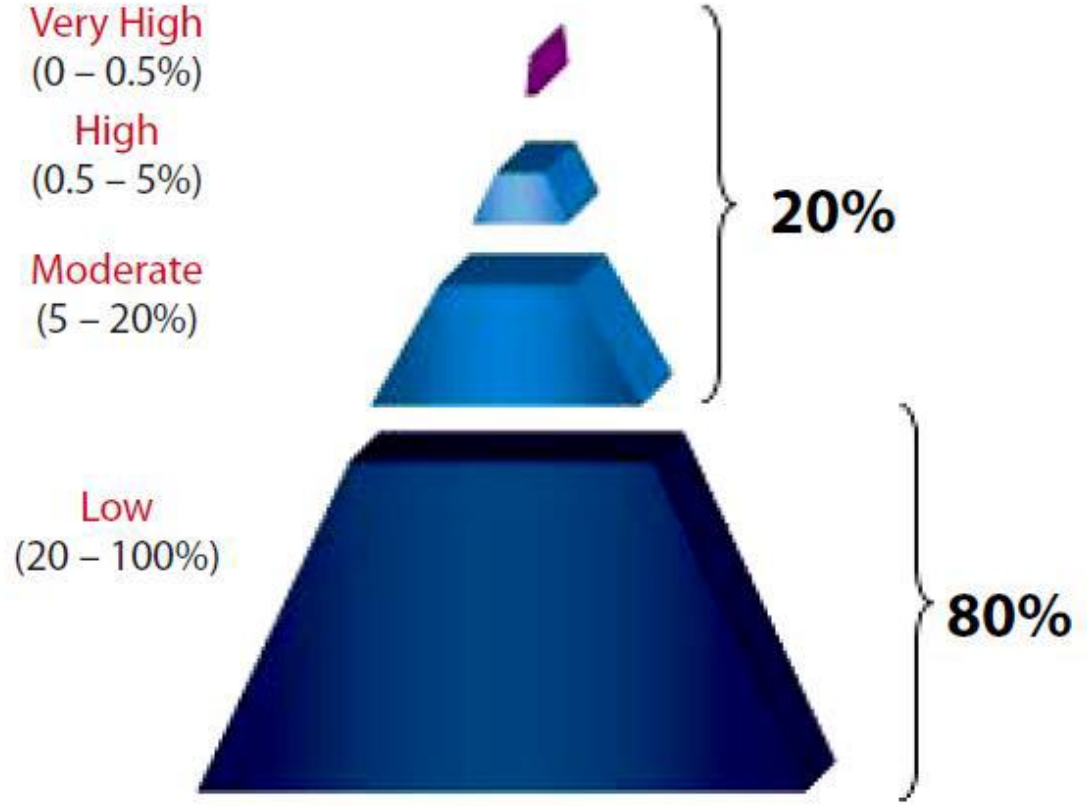


Figure 3: Kaiser pyramid showing the percentage of the population (right) and their relative risk (left). Source: (Lewis et al., 2011)

However, the assumption about the cost-saving benefits from targeting this high-risk group alone is conditional on two factors: 1) that the group is consistent over time, particularly in the case of a group with multiple ingrained long-term conditions, whose treatment is likely to be complex and take time; and 2) that it is actually possible to treat these patients in a way that prevents admissions/saves costs.

There is some evidence, though, that this group is not in fact consistent at the individual-level over time. A recent study in Denver, for instance, found, similarly to the evidence highlighted above, that 3% of their patients accounted for 30% of adult healthcare charges. However, fewer than half of these patients were identified in this high-risk 3% only seven months later, with just over a quarter still in this group after one year (Johnson et al., 2015). This may therefore be a prime example of the ecological fallacy, whereby population-level findings have been extrapolated to the individual-level inappropriately (Piantadosi et al., 1988).

Furthermore, there is little evidence that targeting this group has led to any significant reductions in emergency admissions to date (Purdy, 2010), nor that there is any theoretical basis for believing that doing so would have a major impact (Roland and Abel, 2012; Rose, 1981). The extremely small numbers identified as 'high-risk' by available risk tools also poses further questions. This is particularly true when the even smaller number of these patients whose expensive hospital admissions are potentially avoidable are considered alone (Wallace et al., 2016).

Nevertheless, the sustainability of the increasing healthcare expenditure at the country level is questionable (Appleby, 2013), as outlined at the start of this section. This is particularly true in the wake of the 2008 global economic crisis and the subsequent austerity measures that have been introduced in many parts of Europe in particular, leading to sometimes drastic cuts in government spending (McKee et al., 2012b). There is therefore a need for better use of resources to align increasing needs and expectations of the population with limited resources.

1.4.2 Three fronts: Demand, Efficiency and Funding

The NHS's Five Year Forward View outlines the changes that NHS England plans to implement to address the challenges outlined in this initial section (NHS England, 2014b). The document calls for action on 'three fronts' to maintain a comprehensive and high-quality service moving forward, tackling:

1) Demand - calling for preventative services as the focus for decreasing demand per capita;

2) Efficiency – calling for a 2% net efficiency gain per year through both 'catch up' (matching best performance already present in the system) and 'frontier shift' (using innovation to increase efficiency); and,

3) Funding – calling for the necessary funding to be invested to achieve the necessary actions.

New models of care (particularly addressing demand and efficiency) are called for to achieve appropriate results in health systems that have evolved to tackle acute

conditions (Nolte and McKee, 2008), with ‘integrated care’ the most favoured alternative.

1.5 Chapter Conclusions

In this chapter, I have shown the challenges that health systems are currently facing, and why they must find a solution to deliver effectively on their goals. Next, I will introduce a potential theoretical solution.

2. The preferred solution: Integrated care (in theory)

2.1 Chapter Introduction

There is a preferred solution to adapt to the health system challenges previously outlined, 'integrated care'. In this chapter, I introduce this concept, as well as the relevant context where this change examined in this thesis is being introduced, the NHS in England. The chapter concludes with the research questions that the thesis sets out to answer.

2.2 Integrated Care

The NHS England Five Year Forward View and international health system advisory groups like the World Health Organization call for the development of new models of care which 'integrate' services across different sectors (World Health Organization, 2015; NHS England, 2014b; Goodwin et al., 2012). As illustrated above, better 'integration' may be especially important for the increasing numbers of patients with multimorbidity. These patients have more complex and multiple needs (Fortin et al., 2007b), and are at increased risk of patient safety failures as a result of fragmented care (Panagioti et al., 2015).

This first section examines how integrated care is defined theoretically, and what it is expected to achieve.

2.2.1 Definition

There is no consensus on what integrated care actually means. Indeed, one hundred and seventy-five varying definitions of the concept were identified in a recent literature review (Armitage et al., 2009). A multitude of associated terms (e.g. coordinated care, seamless care, etc.) further complicate the matter, as do differences between definitions deployed in each of the myriad of health systems where integrated care is expected to be implemented (Nolte and McKee, 2008). The multiple stakeholders within each health system (the user, ground-level provider, management, policy makers, etc.), will each have a different perspective on how integrated care may be defined in line

with their own perceptions of and interaction with the system, and the goals they wish to achieve from it (Lloyd and Wait, 2006). However, the definitions of 'integrated care' can be broadly categorised into outcome-based definitions, or process-based definitions.

Outcome-based definitions vary by stakeholder perspective. Examining the literature, the patient perspective is the most common, in keeping with the move towards *person-centred* health systems (World Health Organization, 2015). The definition adopted by the NHS in England (BMA, 2014) exemplifies such an outcome-based definition. It consists of a 'narrative for person-centred coordinated care'. A series of 'I' statements that define integrated care in terms of the patient's experience of 'joined-up care', centred around the statement: "*I can plan my care with people who work together to understand me and my carer(s), allow me control, and bring together services to achieve the outcomes important to me*" (Department of Health, 2013). Many of the statements are also normative. For example, the statements call for integrated care to include:

- personal budgets: "*I know the amount of money available to me for care and support needs, and I can determine how this is used (whether it's my own money, direct payment, or a 'personal budget' from the council or NHS)*";

- patient accessible and writable care records: "*I can see my health and care records at any time. I can decide who to share them with. I can correct any mistakes in the information*";

- and case management: "*I work with my team to agree a care and support plan…I have regular reviews of my care and treatment, and of my care and support plan*".

Process-based definitions appear to be far less common in the literature, but when they are deployed focus on the process of "integration of health and social services" (Gröne and Garcia-Barbero, 2001). For example: "Integrated Care is a concept bringing together inputs, delivery, management and organization of services related to

diagnosis, treatment, care, rehabilitation and health promotion" (Gröne and Garcia-Barbero, 2001).

One of the most frequently cited definitions of integrated care, however, combines both process and outcome:

> *"Integration is a coherent set of methods and models on the funding, administrative, organizational, service delivery and clinical levels designed to create connectivity, alignment and collaboration within and between the cure and care sectors. The goal of these methods and models is to enhance quality of care and quality of life, consumer satisfaction and system efficiency for patients with complex, long term problems cutting across multiple services, providers and settings. The result of such multi-pronged efforts to promote integration for the benefit of these special patient groups is called 'integrated care'."* (Kodner and Spreeuwenberg, 2002)

This definition itself highlights the main issue with integrated care, though: it is not a single 'thing' at all. Instead, it is an umbrella term for changes to the health system (across all health system compartments e.g. funding/organisation/service delivery (Atun et al., 2013)) that better adapt it to deliver care to complex, multimorbid patients. The 'integration' focus suggests that this should be accomplished through closer working between different health and traditionally non-health sectors and services. It aims to achieve outcomes across multiple aims simultaneously: health outcomes, increased satisfaction and cost-effectiveness.

2.2.2 'Types' and assumptions of integrated care

While integrated care is inherently difficult to define because of its broad focus and aims, there have been various categorisations and frameworks that have been derived for classifying *types* of integration. The different taxonomies of integrated care have been identified in a literature review by Nolte & McKee (Nolte and McKee, 2008):

- Breadth of integration: **horizontal** integration (between organisations at the same level of healthcare, e.g. primary care)/ **vertical** integration (between organisations at different levels of healthcare, e.g. between primary and secondary care)

- Type of integration: **functional** integration (e.g. management, planning and quality improvement) / **organisational** integration (e.g. networks, mergers, contracting) / **professional** integration (e.g. joint working, group practice) / **clinical** integration (e.g. coordination of patient services)

- Degree of integration: Leutz (Leutz, 1999) describes integration on a scale from **Linkage → Coordination → Full integration**. The degree of integration required corresponds to the complexity of the user's need

- Process of integration: **structural** (integration of tasks and functions) / **cultural** (integration of values and norms) / **social** (development of relationships)

Others have used different classifications, for instance:

- **Virtual** (alliances, networks etc.) / **Real** (mergers) integration (Curry and Ham, 2010)

- **System level** / **Coordination of services or programs** / **progressive or sequential models** (Armitage et al., 2009)

- Levels of integration: **micro** / **meso** / **macro** (Curry and Ham, 2010)

- Strategies of integration: **Funding** / **Administrative** / **Organisational** / **Service delivery** / **Clinical** (Kodner and Spreeuwenberg, 2002)

- Main focus of integration: **Changed relationships between service providers** / **Coordination of clinical activities** / **Improving communication between service providers** / **Support for clinicians** / **Information systems to support coordination** / **Support for health/social care service users** (RAND Europe and Ernst & Young LLP, 2012; Powell Davies et al., 2008)

- Combining primary care and integrated care: Person-focused care (**Clinical integration**) / Population based care (**Professional integration**, **Organisational integration**, **System integration**) with **Normative/Functional integration** spanning both (Valentijn et al., 2015; Valentijn et al., 2013)

While many of these classification systems may function in theory, it can be difficult to apply many of them to the multi-faceted, complex approaches to integrated care in practice, i.e. to actually implement change. It is argued that integrated care will generate significant health gains alongside improvements in patient satisfaction and cost-effectiveness (Department of Health, 2013; World Health Organization, 2015). Better linkage and coordination are expected to decrease duplication of effort and so improve the efficiency and cost-effectiveness; better management by teams of multidisciplinary professionals is expected to improve management, particularly of multimorbid patients, and so health outcomes; and improved coordination combined with providing care as far as possible in the primary care setting is expected to improve patient experiences.

2.2.3 Primary health care at the centre of integrated care

Primary health care (PHC) is most notably defined as "first-contact, continuous, comprehensive, and coordinated care provided to populations undifferentiated by gender, disease, or organ system" (Starfield, 1994). The 1978 Alma-Ata meeting of health ministers from around the world, and resulting declaration, was a turning point for global health systems. The declaration combined a comprehensive philosophy for development, with PHC at the core – the aim being to provide 'health for all'. It also identified key principles to work towards, based on the WHO's broad definition of health ("a state of physical, mental and spiritual well-being, not merely an absence of disease or infirmity" (World Health Organization, 1948)): equity, and community participation (Lawn et al., 2008).

In spite of the declaration, however, the *comprehensive* primary care approach was rarely implemented. A move towards siloed, disease-specific interventions to manage chronic diseases became the norm instead (Lawn et al., 2008). Indeed,

Wagner's Chronic Care Model, one of the most widely used chronic disease treatment paradigms implemented since Alma Ata, has further emphasised the single-disease focus when the conceptual framework has been implemented in practice (Bodenheimer et al., 2002a; Bodenheimer et al., 2002b).

The most recent practical efforts towards realising these Alma Ata principles, however, follow the Millennium Development Goals, in the form of the move towards Universal Health Coverage (UHC). Again, comprehensive PHC (encompassing prevention of disease and health promotion) is identified as the fundamental priority focus to deliver functioning health systems, and achieve this vision in an economically sustainable way (World Health Organization, 2010). PHC is considered to be at the core of a well-functioning health system, which by improving health outcomes (Atun, 2012; World Health Organization, 2007) contributes to economic and social development (World Health Organization, 2010), and wealth creation (McKee et al., 2012a).

It has been argued that primary care is the most efficient level of the health system at which to invest resources (World Health Organization, 2010). The evidence for this at the population level is fairly strong, with health systems with a strong primary-care base generally delivering better results across improved health outcomes, improved patient satisfaction, and increased cost-effectiveness compared to weaker systems (Atun, 2004; Stokes et al., 2015a; Starfield et al., 2005). There are a number of possible mechanisms by which this cost-saving and efficiency might occur.

The first possibility is that PHC will be able to deliver effective preventative care, with upstream interventions preventing disease, or at least disease exacerbation (Starfield et al., 2005). This would in theory save costs associated with treating established disease with the combination of more expensive interventions and increased contacts with the health system. Where this cost saving can be optimised, however, may not be what we imagine as traditional primary care (e.g. GP-led care), since the cost-effectiveness of medical interventions (carried out in the traditional healthcare settings, e.g. primary, GP-led, or secondary, hospital-led care) varies significantly. Whereas the vast majority of public health interventions have been shown to be highly cost-effective – for example, a recent study synthesised the cost-effectiveness data from NICE public health guidance and found 85% of the public health interventions were cost-effective at

a threshold of £20,000 per quality-adjusted life year (QALY), and 15% were shown to be cost saving (Owen et al., 2011). Likewise, more 'basic' (and even less likely to be considered traditionally primary health care interventions), affecting water and sanitation, are found to be always cost-effective, with an estimated return on a $1 investment of between $5 and $28 (Hutton and Haller, 2004). While not traditional primary care examples, these interventions would conform to what Alma Ata called *comprehensive primary care*.

The second possibility is that PHC is a less expensive setting in which to deliver an intervention, with fewer overheads compared to the hospital setting (Monitor, 2013). As a consequence, a number of interventions attempt to substitute secondary care with care in this theoretically cheaper and more effective setting (particularly in the case of multimorbid patients, for example, perhaps generalists are better suited to manage the patient holistically). While there is some evidence to date that this substitution may be possible (Roberts and Mays, 1998), it is far from conclusive. This is particularly true for 'high-risk' (complex) groups of patients, where more evidence of what is effective for reducing hospital admissions is needed with little effectiveness to date (Purdy, 2010; Wallace et al., 2016). There is also little theoretical justification for and evidence of the same actions being cheaper in primary as opposed to secondary care, unless the action is conducted by someone from a lower paid profession. If the hospital remains open and running while the actions are conducted in primary care, then the expensive overheads will remain, while demand for primary care services is increased.

In any case, while there is a consensus that primary care is the best place to coordinate care from, there is no single version of primary care, just as there is no single version of a health system. These different system contexts will affect the outcome of interventions within them (Craig et al., 2008). Any additional integration of care activity will therefore be affected by the context in which it takes place.

2.3 Context

2.3.1 Importance of reporting context

It is clear, then, that context will play an important role when implementing health service interventions, and likewise when attempting to analyse or measure outcomes within the health system setting. This recognition of the significance of context in turn will affect both the methods used to conduct the analysis, and how the findings are reported.

When reporting findings of complex interventions/systems analyses, it is thus important to explicitly detail aspects of context, so that relevant statistical adjustments can be made by further researchers (e.g. in meta-analysis), or, so that generalisability can be hypothesised by policymakers who may wish to transfer an intervention to their own context.

Some health system analytical frameworks include a structured approach to reporting context specifically relevant to the health system. One notable example in the literature that focuses on context is Atun's health systems framework (Atun et al., 2013), which uses the acronym DEPLESET (Demographics, Epidemiology, Political, Legal, Ecological, Socio-cultural, Economic, Technological) to structure context reporting (Atun and Menabde, 2008). This DEPLESET model is utilised in the section below to describe the context (the NHS in England) directly relevant to the thesis, together with a brief description of key aspects of the health system itself.

2.3.2 NHS in England

Each of the four countries in the UK (England, Scotland, Northern Ireland, and Wales) has a tax-funded NHS with universal coverage (a Beveridge health system). In a tax-funded health system, as in any other such system, resources are limited, although the way in which these resource limitations are realised differs depending on funding source. So, for instance, while in a number of insurance-based health systems, services (resources) are rationed by ability to pay (either directly/through individual's insurance coverage), in the NHS, resources are free to everyone at the point of delivery, but are rationed through waiting times. This rationing system is naturally politically fraught (like any other), and can conflict with political target-making/-breaking.

Each of the devolved health systems in the UK differs in some respects. In the NHS in England, there has been greater marketisation of healthcare. For example, the division of purchasing from providers that was introduced by the Conservative government UK-wide in 1991 (Boyle, 2011) was abolished in Scotland (in 2004) and Wales (in 2009). In England, however, there has been a greater emphasis placed on developing patient choice and competition between providers, including expansion of private providers to deliver publicly funded healthcare (Bevan et al., 2014). The impact of this emphasis on competition on the integration of care is, however, debatable, as these concepts at face value appear to be conflicting.

UK-wide, the NHS has always had a strong primary-care focus, with the GP acting as 'gatekeeper' to the rest of the health system, coordinating the patient's interaction with other services (Starfield and Shi, 2002). The Health and Social Care Act 2012 in England attempted to further emphasise the strength of primary care for the health system as a whole, putting GPs in charge of CCGs tasked with purchasing services for their local populations (Department of Health, 2012). The success of this reorganisation, however, may be constrained by continued top-down management of CCGs (Coleman et al., 2015). The strength of primary care in the UK renders the health system highly regarded internationally (Davis et al., 2014). However, being fairly unique in its 'gatekeeping' position, this organisation of health services may distort the outcomes of services adopted from other systems. For example, interventions aimed at strengthening primary care in weaker systems may not translate effectively when adopted in the UK – i.e. the UK health system is likely to benefit only from very marginal gains ('ceiling effects') in this respect compared to others, as it already possesses an especially strong primary-care system.

The relationship of the health system with allied healthcare organisations (e.g. social care) again differs in England compared to the rest of the UK. In Scotland, for example, there is free personal social care for the over-65s (Bevan et al., 2014), whereas in England many adults pay for some or all of their social care with local authorities paying for individual care packages only on assessment of high needs and limited means (Department of Health, 2014a). These organisational and funding relationships between the two sectors, moreover, may have further implications for methods and effectiveness of integration between the two.

The Government in England, however, has attempted to use a number of policies to improve integration. The CCGs have a contracted duty to 'promote integration' (Department of Health, 2012). How this is accomplished is not prescribed nationally, but expected to arise locally in each of the over 200 CCGs. This mandate at the CCG-level reflects an apparent 'bottom-up' approach to integrated care in the NHS, with the central NHS body attempting to encourage local innovation and adaptation around the concept of integration (Darzi and Howitt, 2012). This approach is similarly illustrated by the numerous additional funding schemes (e.g. the Integrated Care Pilots, the Pioneers, the Vanguards) that the Government has introduced to support this innovation and to attempt to illustrate best practice. The Better Care Fund has also been established, a single pooled £5.3 billion health and care budget to support integration (NHS England, 2016a).

In recent years, health system planning in England has revolved around addressing a large healthcare funding gap, estimated at £30 billion a year by 2020/21. The 'Five year forward view' strategic document outlines how the NHS plans to tackle this large deficit (outlined above, in section 1.4.2) – tackling demand, efficiency and funding (NHS England, 2014b).

Box 1 summarises some of the important contextual factors surrounding the NHS in England.

Box 1: Context of the NHS in England (DEPLESET)

- Demographic – The population of England is currently ~54 million (ONS, 2014). This is growing steadily, with a projection of just over 61 million by 2032 (~55% of the increase from natural growth i.e. births minus deaths, and ~45% from net migration). The population is becoming more diverse, with the ethnic population estimated to make up 15% of the population by 2031. Life expectancy is increasing, as is healthy life expectancy, suggesting that the extra years of life will not necessarily be in ill health. The population is ageing, with those over 65 growing at a faster rate than those under it. Over the next 20 years, there will be an estimated 39% rise in those aged 65-84, and a 106% rise in those over 85 (The King's Fund, 2012).

- Epidemiology – As a high-income country, the major disease burden in England is now chronic diseases (Omran, 1971). In 2013, the number one cause of years of life lost was ischaemic heart disease, and the leading cause of disability-adjusted life years (DALYs) was back and neck pain.

Known risk factors (leading causes were diet and tobacco use) were estimated to account for 39.6% of DALYs (Newton et al., 2015).

- **Political** – Margaret Thatcher's neo-liberal Conservative government vastly changed the political agenda when elected in 1979. Privatisation of a number of state-owned assets occurred with the belief in the power of the market rather than a government to efficiently allocate resources. The NHS, however, was kept from marketisation during Thatcher's time in office. John Major succeeded Thatcher in 1990, and immediately introduced a purchaser-provider split and market forces in healthcare in 1991. In 1997, the New Labour Prime Minister, Tony Blair was elected, with similar neo-liberal aims. A number of top-down re-organisations of the NHS took place under the New Labour governments in the next decade, relying on the market and with power placed firmly in the hands of the providers, where strong foundation hospitals emerged as dominant, with weaker primary care trusts alongside (Toynbee, 2007). When the Coalition (Conservative and Liberal Democrats, led by Prime Minister David Cameron) took office in 2010, a top-down re-organisation introduced in the Health and Social Care Act 2012 attempted to re-emphasise the place of primary care with GP-led CCGs as major purchasers. The Act also re-emphasised market forces, stressing the importance of competition for provider contracts (Department of Health, 2012), which has subsequently led to an enormous increase in private providers, who have been awarded an estimated share of one-third of contracts tendered since the Act came into place (Iacobucci, 2014).

- **Legal** – The NHS was founded in 1948 after the Second World War, based on the central principle that health services would be available to all and financed entirely from taxation (NHS Choices, 2015). More recently, the NHS Constitution in England, formalised in 2009, has provided a common set of principles and values for the NHS, together with the rights and responsibilities for the stakeholders involved in healthcare (patients, staff, NHS organisations). For patients and the public, this concerns rights around access to health services; quality of care and environment; nationally approved treatments, drugs and programmes; respect, consent and confidentiality; informed choice; involvement in healthcare and the NHS; and complaint and redress. All bodies supplying NHS services are required by law to take account of this document in their decisions and actions (Department of Health, 2009). Organisation of NHS services most recently changed in law with the Health and Social Care Act 2012, with overall responsibility for running the NHS handed from the Department of Health to the arms-length body, NHS England. The formation of 'clinically-led' Clinical Commissioning Groups (CCGs) adopted the majority of the purchasing role, and provision of NHS services opened to wider competition regulation (overseen by Monitor, responsible for preventing anticompetitive behaviour and simultaneously the promotion of integration) and so increased private provision (Checkland et al., 2013).

- **Ecological** – The vast majority of the population (~80%) in England lives in predominantly urban settings, with the remaining (~20%) in local

authorities defined as predominantly rural (Government Statistical Service, 2015). With high population density in much of the country, air pollution is thought to contribute to ~40,000 deaths per year and may play a role in many chronic conditions (e.g. cancer, asthma, heart disease, and neurological changes linked to dementia) (Holgate et al., 2016).

- Socio-cultural – The population of England is very ethnically and culturally diverse. While the majority of the population identify with being White, there are large numbers from mixed-race, Asian, Black/Caribbean and other ethnic groups (ONS, 2015). In relation to healthcare, different ethnic groups may be genetically pre-disposed to various chronic diseases (e.g. relation of central obesity and insulin resistance to high diabetes prevalence and cardiovascular risk in South Asians), or cultural norms may alter risk (McKeigue et al., 1991). Additionally, there may be language issues during interactions with health services for some patients. These factors may contribute to differential treatment and prevention indications.

- Economic – England as part of the UK is ranked as a high-income OECD country (The World Bank, 2016). However, the distribution of wealth and economic activity is highly skewed and centralised in the south of the country (i.e. London). There are wide inequalities in income from North to South (Jenkins, 1995). There was a recent attempt to redress this balance through the creation of a 'northern powerhouse' (Osborne, 2014). Nationally, political choices since the 2010 general election have introduced austerity measures with wide-reaching cuts to public services (McKee et al., 2012b). Healthcare services have remained fairly protected (but, nevertheless, spending per capita has remained flat during this period and much lower than the OECD high-income average). Services complementary to healthcare (e.g. public health spending and social care services) have been cut in line with the general trend (McKee et al., 2012b).

- Technological – As a high-income country England is generally technologically advanced. In terms of healthcare, there is wide coverage of electronic patient records (helped by the need to implement electronic records to meet pay for performance targets e.g. the Quality and Outcomes Framework (QOF)). However, patient records remain largely unlinked between primary and secondary/social care settings. There are multiple providers of both primary and secondary care electronic recording systems (Parliamentary Office of Science and Technology, 2016). Previous national attempts to introduce a single joined up record for the population (e.g. 'Connecting for Health' in 2002) have failed spectacularly, wasting billions of pounds of taxpayers money (Cross, 2006). Current government plans include patient access to their own electronic health records, which are planned to be joined up between services by 2020 (Parliamentary Office of Science and Technology, 2016).

2.4 Aims and Objectives

The above discussion reveals some important research gaps: 'integrated care' is the mandate of CCGs in the NHS, but 'integrated care' is an extremely vague concept in the literature; and, despite the increase in multimorbid patients being one of the main drivers towards 'integrated care', multimorbidity itself can be measured in a number of ways; and, finally, there is little evidence of what is effective for treating these complex patients.

These gaps have informed the overarching research questions of this thesis:

1. What does 'integrated care' currently look like in practice in the NHS?
2. What is the effectiveness of current models of 'integrated care'?
3. To what extent are there differential effects of 'integrated care' for different types of multimorbidity?

2.5 Chapter Conclusions

In this chapter, I have highlighted the vague and rather all-encompassing nature of the concept of integrated care in theory. I have shown that primary care is at the centre of integrated care plans, and introduced the context in which the PhD research is set, the NHS in England. Gaps in the literature reviewed in this and the previous chapter have raised three key research questions that the remaining chapters will seek to answer.

3. Integrated care (in practice)

3.1 Chapter Introduction

The previous chapter introduced integrated care in theory, and now we turn to what this concept can look like when enacted in practice. I first introduce the wider examples of integrated care across the health system, and then focus in on what I identified in the early phase of my research as the dominant form of integrated care in the NHS to date, multidisciplinary team (MDT) case management, including the definition, historical context and logic model of this form of integration.

3.2 Integrated care in practice

In practice, unsurprisingly, the broadly defined concept of integrated care can take many forms. Box 2 shows some common examples of integrated care in practice across the health system (adapted from (Kodner and Spreeuwenberg, 2002)). Successfully implemented integrated care programmes frequently combine multiple changes simultaneously in different areas of the health system (Powell Davies et al., 2008).

Box 2: Examples of integrated care in practice, across the health system (adapted from (Kodner and Spreeuwenberg, 2002))

Financing:

- Pooling of funds
- Prepaid capitation

Governance & Organisation:

- Joint commissioning
- Discharge and transfer agreements

Resource Management:

- Co-location of services

- Integrated IT

Service Delivery:

- Case/care management
- Multidisciplinary teamwork

Evidence of effectiveness of integrated care interventions more generally is limited across the broadly defined outcomes of satisfaction, health outcomes and costs of care. The synthesis of evidence related to all three outcomes, furthermore, is hampered by the lack of conceptual clarity (Ouwens et al., 2005; Nolte and Pitchforth, 2014). The polymorphous nature of the integrated care concept and its frequent introduction as interventions with multiple components cause further issues with traditional research methods (Nolte and Pitchforth, 2014).

Satisfaction: While still limited, there is relatively good emerging evidence that patient experience can be enhanced by integrated care, particularly where coordination and communication is improved (Cameron, 2016). By the NHS's accepted definition of integrated care (i.e. the series of 'I' statements noted in the previous chapter illustrating the patients' feeling of better integration of care), then, current interventions appear to meet the primary aim of improving the patient's satisfaction with care.

Health outcomes: Addressing fragmented care of patients is expected to increase quality of care and lead to better health outcomes. Several less rigorous study designs (e.g. before-after/qualitative) have frequently reported improvements in quality of life, health, and ability to cope with everyday living. More robust study designs (i.e. comparing 'integrated care' intervention to a control group), however, report much more limited findings with only marginal differences identified between treated and non-treated groups. Significantly, studies with a rigorous design that looked at integrated care changes for older adults (i.e. those patients most likely to have multimorbidity and to be those most immediately in need) found no statistical differences (Cameron, 2016).

Costs of care: As noted above, integrated care initiatives are frequently driven by policy-makers who aim to contain costs within the healthcare system. However, there is little evidence of economic gain when implementing integrated-care interventions. A comprehensive literature review of systematic reviews found an overall shortage of robust evidence of cost-effectiveness, with the evidence available generally weak (frequently based on a very small number of studies and/or before-after study designs), making conclusions of causation impossible. The authors identified 19 reviews for inclusion, with eight reporting on cost-effectiveness. Most of the evidence, however, was for single-condition approaches to integration. There was no clear evidence of cost-effectiveness across the board. On the current evidence, therefore, the authors of the review question whether integrated care by implication would be cost-effective and support cost sustainability, or whether it should simply be considered a complex strategy to change the way health and care services are delivered, but not necessarily addressing the system challenges. They incline towards the latter conclusion (Nolte and Pitchforth, 2014).

At the outset of my research, I examined what was being implemented in practice as integrated care amongst a random selection of CCGs (contractually mandated to promote integration) in the NHS in England. Taking a random selection of 10% (n = 21) of the 211 CCGs in existence at the time (2013), publicly available documents were examined in an attempt to determine what each CCG was branding as 'integrated care' (in the public arena at least). Although the source of information limited the detail of what could be extracted, there was a clear dominance (n = 17/21, 81%) of a single particular model of integrated care present as the primary practice. This model can be described as multi-disciplinary team (MDT) case management of high-risk (of adverse outcomes and/or secondary-care utilisation) patients, and has tended to focus on achieving reduced use of acute, secondary-care services and total healthcare costs for these patients (Stokes, 2014). This approach is similarly at the forefront of the specially supported integrated care sites in England. Examination of the 'Integrated Care Pilots' (ICPs, 2009-2012); the 'Integrated Care and Support Pioneers' (wave 1 began in 2013, and wave 2 in 2015, both ongoing); the 'Vanguards' (2015 and ongoing); and most recently announced, 'Devolution' of health and social care (first to Greater Manchester, starting in April 2016) as part of my involvement in the EU-funded SELFIE (Sustainable

intEgrated care modeLs for multi-morbidity: delivery, FInancing and pErformance - http://www.selfie2020.eu/) project has subsequently revealed this (see Appendix 1 for more detailed findings from the CCG and government scheme analysis of integrated care approaches).

Internationally, moreover, this same case management focus appears to hold. A recently updated Cochrane review examined interventions for improving outcomes in patients with multimorbidity and similarly identified MDT case management as the predominant approach (Smith et al., 2016). Likewise, a large literature review in the recently drafted NICE guidelines on management of multimorbidity identified the MDT case management approach as key to integrated care (NICE, 2016a). Consequently, the evidence for the MDT and wider case management approach has international relevance.

3.3 MDT/ Case management

MDT case management is a central focus of integrated care when implemented in practice. It additionally forms a key part of popular and widely accepted recommendations for the treatment of chronic disease and multimorbidity (NICE, 2016a; Bodenheimer et al., 2002a).

This section outlines what (MDT) case management is, how it has evolved historically, and how it functions in theory to affect specific outcomes (the logic model). In Chapter 5, the evidence for effectiveness across a variety of outcome domains is then systematically reviewed.

3.3.1 Definition

Case management can take a variety of forms, but there are common elements by which it can be defined (Ross et al., 2011). Case management involves:

- Case finding – using one of a variety of means to identify specific patients to case manage. Frequently this is assisted by a population risk stratification tool, but it can also (or instead) rely on clinical judgement (Lewis et al., 2011).

- Individual assessment and care planning – the assessment can be more or less 'holistic' (e.g. it can involve a biopsychosocial assessment, as opposed to solely a medical one), depending on the type of intervention. The care plan is the centrefold of case management. Its main function is to outline the goals and processes of care, and to coordinate all of the service providers.

- Care coordination with regular review, monitoring and adaptation of the care plan – case management is an ongoing process, with regular patient contact to re-assess and adapt the care plan to the case as it evolves. Generally, the case manager acts as a single point of contact to coordinate care, and frequently to feed in to self-management and other support.

There are a number of variables differentiating each case management intervention. For example, the primary location of the case management (in home/in clinic/over telephone), the intensity of the intervention (frequency of patient contacts), the depth of the assessment and care planning process, the reimbursement method (incentives), etc. can all vary. Focusing on integration of care, one of the most important variables is likely to be who delivers the intervention. A multidisciplinary team may offer an additional 'professional integration' component to the coordination activity compared to delivery by a single case manager, for instance (Valentijn et al., 2013). The specific make-up of the team may also alter the degree of integration between sectors. The involvement of non-traditional primary-care workers, social workers, for instance, may alter working/referral patterns between health and allied sectors as the professionals begin to work more closely together.

Case management, then, clearly meets many (if not all) of the differential definitions of 'integrated care'. Case management takes place directly at the service-delivery level, and so shows clear potential for directly influencing patient outcomes (across health, satisfaction and cost). This satisfies outcome-oriented definitions in particular (such as that widely adopted in the NHS – as discussed in the previous chapter). The obvious coordination activity in all variants, and particularly the inter-professional working of MDT case management clearly alters processes of care, satisfying process-based definitions. These factors, together with case management's

historical context, discussed below, may explain why it has been so widely adopted as the primary integration activity to date.

3.3.2 Historical context

The concept of case management is not a new one. It has a long history, and this is important for understanding its current standing, and the expectations of its utilisation.

The first recognisable implementation was reported in 1863 in the United States. The first board of charities was formed in Massachusetts at this time, and its purpose was to coordinate public services for immigrants, the sick and the poor while conserving public funds. The charitable organisations that ensued from the establishment of these boards used an index card system to record each family's needs (including environmental and social circumstances) and plans to coordinate relevant services to address these. This individual attention was used primarily to reduce duplication of effort in the fragmented system, and, from some perspectives at the time, to stop the poor and immigrants 'cheating the system', and from others, to enhance the experience of those in need of the services (Kersbergen, 1996).

The term 'case management' was not coined until the 1960's, however. During this time, the Civil Rights Movement and President Johnson's 'War on Poverty' began. These movements resulted in the emergence of a number of new services, a highly complex and fragmented arena, which was almost impossible for the 'client' to navigate. (It was also during this time that the patient moved from a passive to active role, now seen as the consumer of services). The deinstitutionalisation of mental health care around the same time likewise created a fragmented array of services offering care in the community, and case management was subsequently introduced to this group of patients in the 1970's (Kersbergen, 1996).

In the 1980's, the rapidly increasing costs of care demanded new models of care, and case management was once again expanded to other health areas. Following the introduction of the Pay-for-Performance (P4P) Diagnosis Related Groups (DRG) payment schedule in hospitals, case management was primarily used during discharge planning to ensure that excess funds were not spent on lengthy hospital stays. Health insurance

companies also began utilising case management at this time to monitor the efficiency of care delivered in relation to that required by its members (Kersbergen, 1996).

In the UK, case management was adapted primarily to social care in the 1980's, where it was known as 'care management' (Drennan et al., 2011). It had a primary role in the Community Care Act 1991 under this name, as the focus of changes to community care at that time (Drennan and Goodman, 2004).

More recently, case management has frequently played a central role in implementation of Wagner's Chronic Care Model (Wagner, 1998; Bodenheimer et al., 2002a). This Model has subsequently emerged as one of the most popular treatment paradigms for chronic disease. It emphasises the use of clinical information systems, risk stratification, self-management support, and interdisciplinary teamwork. The patient group focus has shifted with the use of these risk stratification tools and IT systems with more complete patient records. A popular focus is generally 'at-risk' patients, those identified as being high-cost, or 'super-utilizers'. Unsurprisingly, then, focus within the context of current health system challenges (highlighted in Chapter 1) and within this chronic disease paradigm (the elements of the Chronic Care Model highlighted above) frequently draw on MDT case management as a central component to manage these complex (generally multimorbid) patients.

Over the course of its history, then, the target groups of case management have evolved with the pressures of the time. Delivery began with nurses and social workers, but more recently an interdisciplinary MDT approach has been adopted. Consistently, though, the concept has been implemented both to reduce costs and to improve patient outcomes. In the following section, the logic model of the currently popularised case management approach will be presented.

3.3.3 The logic model of (MDT) case management

A logic model is a pictorial representation of how a process or organisation works. It links the inputs and activities of the programme through to the expected outputs, outcomes, and impact. Logic models are theoretical in nature and can be used at different stages of a programme, from initial planning through to evaluation, where they can further aid programme evolution (WK Kellogg Foundation, 2004). The logic model

only provides key details of the model, and is therefore a simplification of the more complex realised form (McLaughlin and Jordan, 1999). The model 'reads' from left to right, following the chain of reasoning (WK Kellogg Foundation, 2004).

The logic model I have derived in Figure 4 below draws on current case management literature, described below, relating to each column of the model. Logic models build on a number of underlying assumptions. Highlighted in Box 3 are some of the key assumptions of the case management model.

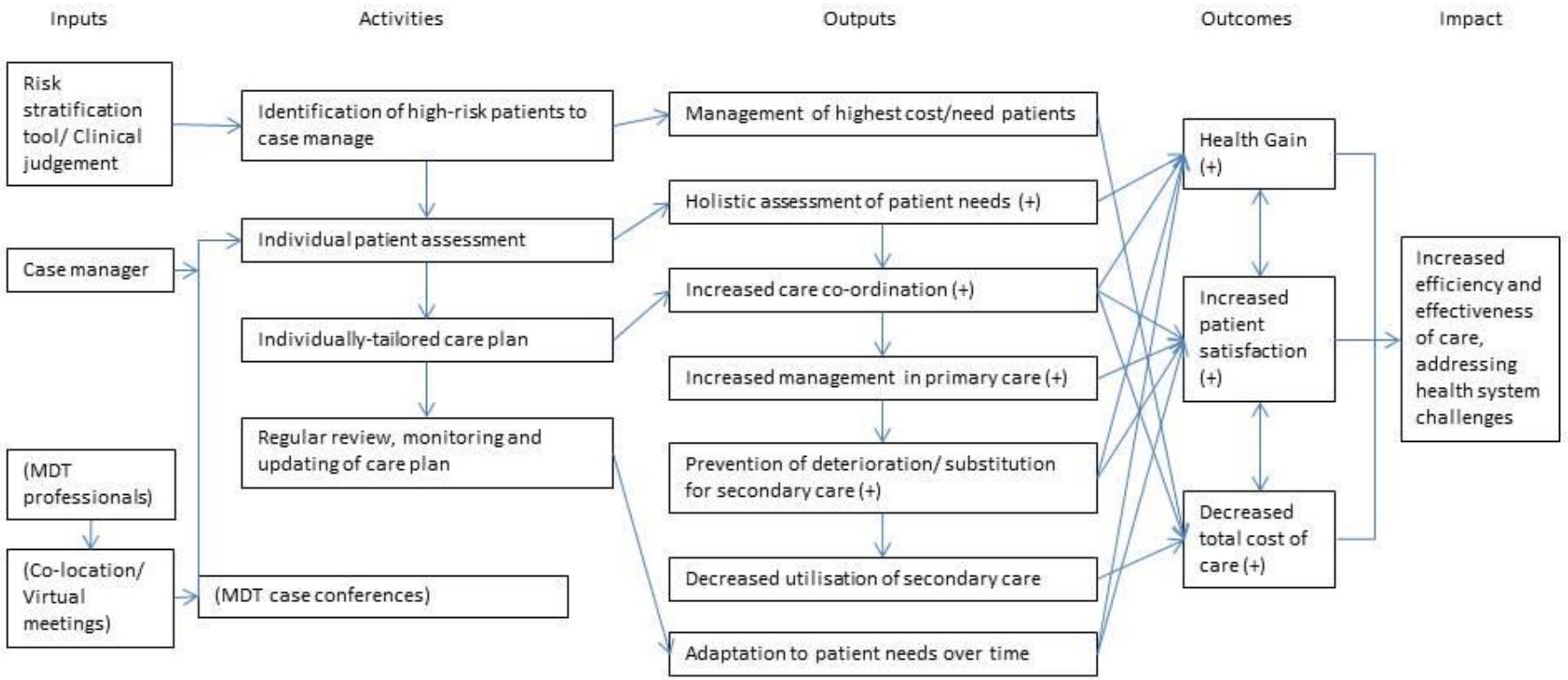


Figure 4: Logic model of (MDT) case management. In brackets are the specifics related to MDT case management, and '(+)' represents where this MDT is expected to provide a positive modifying effect

Box 3: Assumptions of logic model

The case management model shown in Figure 4 above rests on a number of assumptions. Key assumptions to highlight include:

- Treatment in primary care is less expensive than in secondary care
- Increased coordination and holistic care in primary care can substitute for/prevent secondary care
- Targeting high-cost, high-risk patients has scope for preventing secondary-care admissions

The key inputs to the case management model are the professional roles - the case management role (which a nurse, social worker, or a new professional can fill, for instance), and, for MDT case management, the additional health and care professionals involved (these roles can vary from intervention to intervention). The additional training requirements, resources to accommodate these professional(s), salaries and opportunity costs are excluded from the figure above for the sake of simplicity, but will have additional resource input implications. Moreover, if a risk stratification tool is used to select patients, additional time and training in electronic medical record development and software may be required and will have to be factored in. If clinical judgement is used instead, physician time to carry out this additional task is also required.

Key activities generalisable across case management interventions are selection of the patients to manage, individual patient assessment, care planning, and regular review, monitoring and adaptation of the care plan (Ross et al., 2011).

The outputs highlighted above illustrate the historical expectations of case management discussed in the previous section (increased coordination and a decrease in duplication of effort, which improves efficiency). Combined with these are some of the underlying assumptions detailed in Box 3. These are common assumptions in the case management model, as well as wider integrated care initiatives and other primary care access and policy initiatives implemented in recent years (e.g. extended primary care opening hours as a means to minimise secondary care utilisation) (Greater Manchester NIHR CLAHRC, 2015).

The literature suggests that the case management approach might affect three main outcome categories: health, patient satisfaction, and total cost of care (through decreased utilisation) (Ross et al., 2011).

The addition of an MDT to the case management approach is predicted to act as an effect modifier to these outcomes, potentially increasing the holistic nature of the care and the ability of coordination through greater professional integration of the health and care professionals involved (Valentijn et al., 2013; Wagner, 2000). In theory, this greater professional integration and closer interdisciplinary working may also 'spill-over' and positively affect non-treated patients in the intervention practices (not shown in the above model, but described in more detail in Chapter 6) (Bower et al., 2003).

3.4 Chapter Conclusion

In this chapter, I have shown how varied the broad integrated-care concept can appear when enacted in practice, with examples across the health system. I have, however, illustrated that the dominant form enacted by CCGs (mandated to integrate care in the NHS in England) has been MDT case management. Finally, I have defined this form of integration, its historical evolution and the logic model, showing that it is expected to have positive effects across health system goals.

4. Methods

4.1 Chapter Introduction

As an alternative format thesis, the detailed methods for each of the following chapters is contained within the journal articles (Chapters 5-7) and their additional appendices. This chapter, however, sets out the overarching 'evaluation problem', how it is managed in experimental and non-experimental studies generally, and outlines briefly the main methods used in this PhD research to deal with this problem and to address the research aims.

4.2 The evaluation problem

The fundamental 'evaluation problem' faced in this thesis is how to obtain the 'true causal effect' of an intervention such as 'case management'.

To obtain this 'true causal effect', we would need to simultaneously observe individuals affected by the intervention in two independent states: 1) having received the intervention and 2) as unaffected controls. This is, of course, impossible, but it is effectively what randomisation simulates at the population level by ensuring that assignment to treatment group is purely by chance, and so balancing observable and unobservable characteristics between the two treatment groups (assuming sufficient numbers of participants, and effective randomisation is achieved) (Roberts and Torgerson, 1999). This is why randomised controlled trials (RCTs) are considered the 'gold standard', according to the traditional 'hierarchy of evidence' (Guyatt et al., 1995) at least.

However, when interventions are rolled out in usual practice, they are rarely done so randomly (occasionally this is done in random clusters, but this is not common practice). Therefore, we tend to have no observable counterfactual (the opposite state to what we actually observe), so must use other study design or statistical methods to approximate the true causal effect as closely as possible.

Also, although RCTs are traditionally considered the gold standard of evidence, they nevertheless have their own flaws, particularly in the context of multimorbid patients and for assessing interventions such as integrated care, which frequently have multiple components. First, the eligibility criteria of RCTs can be problematic. In trying to isolate the effects of the treatment/intervention of interest, frequently the most complex patients (e.g. multimorbid patients) are excluded from trials. This usually means use of a non-representative population sample to test effectiveness in a trial (see Figure 5, which illustrates the tiny proportion of the total relevant population that might be selected in a trial, and their relation to the wider population in terms of criteria, including likely responsiveness). In addition, this means that intervention effects will tend to be mis-estimated, either tested on a population that is most likely to respond positively to treatment and so the effects may be over-estimated (Richardson and Doster, 2014), or tested on those who are likely to get better in any case, potentially under-estimating effects. While these study eligibility criteria can be improved in trial design, RCTs will nevertheless always take place under somewhat idiosyncratic conditions. For example, the practitioners and patients are likely to be more engaged than in routine practice, having been consented and having volunteered for the trial; they may be receiving additional funding/incentives for implementation and/or, have added input, assistance and motivation from a research team. Therefore, in routine practice any treatment effects might be quite different when compared to the RCT estimate (Davies et al., 2015).

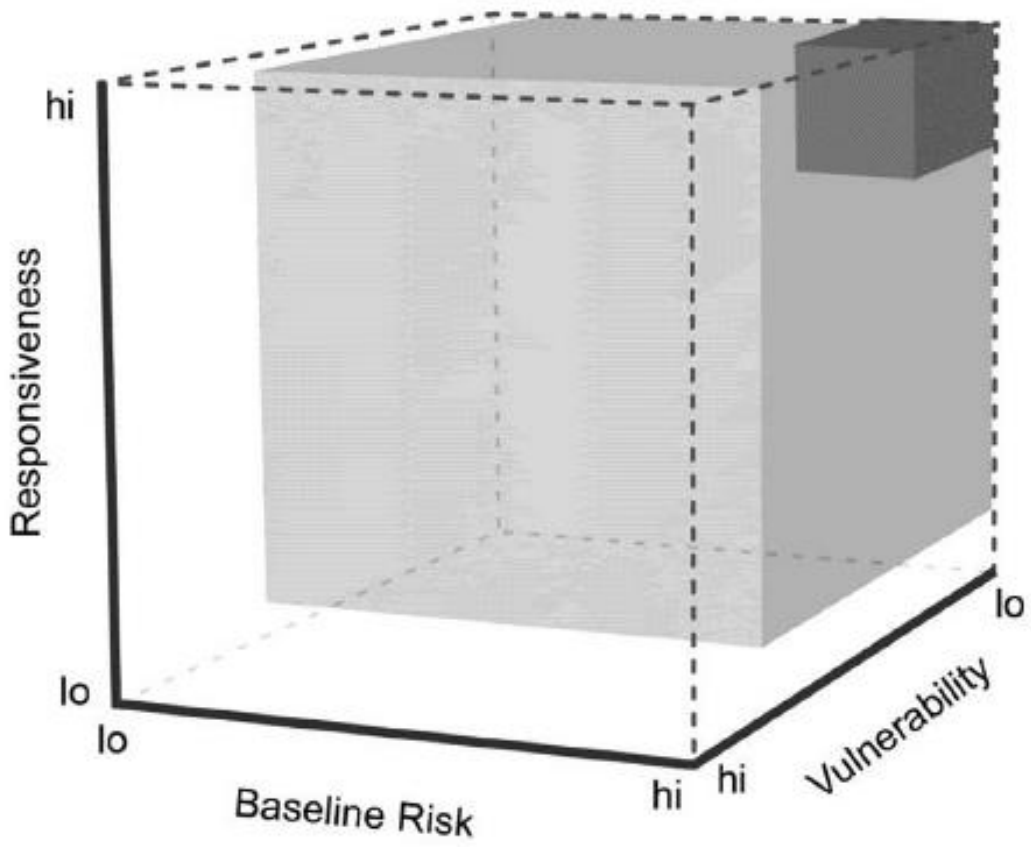


Figure 5: How those included in an RCT study design (dark grey – high responsiveness, high baseline risk, and low vulnerability) tend to differ from the majority of the population (light grey – with individuals lying across the spectrum of the three axes indicated). Source: (Richardson and Doster, 2014)

Furthermore, when using an RCT, a range of additional practical issues needs to be considered. For instance, cost (RCTs tend to require a lot of time and money to conduct robustly), strength of evidence needed for a decision to be made (i.e. what evidence might be considered 'good enough' for a policy decision), ability to recruit and randomise the intervention and potential harms of the treatment under trial (e.g. if none are likely, then a non-experimental method could be more than adequate instead) (Black, 1996).

An alternative to generating additional data via an RCT is to harness the abundant routinely collected data available, and use retrospective methods of analysis (Raine et al., 2016). Routinely collected data constitute a particularly strong resource in the NHS. Drawing on this data not only eases access to data sources, but since the data tend to be standardised nationally, this allows comparison between different localities. And, although traditionally there were fears of routinely collected data lacking quality (McKee, 1993), more recently the quality has improved. For example, primary-care data,

particularly those related to chronic diseases, have improved significantly, in large part thanks to the richness of data available due to payments associated with the Quality and Outcomes Framework (QOF) (Sigfrid et al., 2006). Similar data also exist for hospital episode statistics in the UK at the individual level. Use of these large, world-leading databases is now widely accepted in health services research (Fulop et al., 2001). This rich resource is particularly useful for the analysis of already 'in practice' interventions, such as integrated care in the NHS setting.

In the following chapters, therefore, routinely collected data and quasi-experimental methods are used to examine the research aims set out in Section 2.4. The detailed individual study methods are introduced in each of the empirical chapters (Chapters 5-7). First, in the sections below, I briefly introduce the three overarching research methods individually.

4.3 Systematic review and meta-analysis (Chapter 5)

The basic idea of the meta-analysis approach is that each single study contributes to the estimate of the true effect of the intervention. The individual study is then weighted, based on the sample size (and so, too, the margin of error with which it accurately estimates this true effect), and then pooled. It thus builds on the strengths of the individual studies by combining the sample sizes into a pooled estimate with a smaller margin of error in comparison to each individual primary study, and a more accurate representative effect with higher statistical power.

Although most often associated with RCTs, meta-analyses are not restricted to these in principle, which is important taking into account the potential problems identified with RCT estimates alone (in relation to multimorbidity and integrated care, as illustrated above). The meta-analysis in Chapter 5 also includes estimates from robust quasi-experiments recognised by the Cochrane Effective Practice and Organisation of Care (EPOC) group. These study types include: randomised controlled trials (RCTs), non-randomised controlled trials (nRCTs), controlled before and after studies (CBA), and interrupted time series (ITS).

In order to pool the study estimates, a common outcome measure must be present across studies. This commonality must be two-fold:

a) conceptual (e.g. to pool a *costs of secondary care* outcome, we must be satisfied that each of the studies are contributing a measure of costs of secondary care, and not, for example, costs of primary care, which is a different concept)

b) statistical (e.g. all studies must be measuring the outcome in comparable statistical units), although problems here can be overcome by converting the measures to a common unit (i.e. the effect size, for example, standardised mean difference)

There are also issues relating to whether the trials captured in the review are representative of all of those done. This publication bias can occur because there is a tendency for positive results to be published in higher-impact, more easily accessible journals than negative results (if these are published at all). If this bias is present, then the research is in danger of drawing the wrong conclusions about the intervention effects (Rothstein et al., 2006).

To summarise the method's issues discussed above:

- We assume that there is some single true effect of the studies/outcomes being pooled that we can measure;
- We require theoretical and statistical commonality to pool results in a meta-analysis;
- We assume that a representative sample of studies measuring this effect is included in the analysis, and that the pooled measure is not subject to publication bias.

4.4 Quasi-experiments (Difference-in-differences – DD, Chapter 6)

Difference-in-differences (DD) analysis of observational data is now a common way of evaluating healthcare interventions (Dimick and Ryan, 2014), and will be the main approach used in the evaluation of integrated care in this thesis (Chapters 6 and 7).

DD compares the difference in outcomes from before to those after an intervention between an intervention and a control group. Under a set of assumptions (discussed below), if the first difference is the difference in outcomes from before to after the intervention for the intervention group, and the second difference is the difference in outcomes from before to after the intervention for the control group, the difference in differences represents the intervention effect:

$$\widehat{\beta} = (\bar{y}_{I,2} - \bar{y}_{I,1}) - (\bar{y}_{C,2} - \bar{y}_{C,1})$$

where

$\bar{y}$ = average outcome
1 = pre-intervention period
2 = post-intervention period
I = Intervention group
C = Control group

The method controls for time-invariant unobserved differences between the two groups being compared, as well as any common shocks (i.e. other events occurring during the period of the analysis that might affect the outcome of interest that has an equal effect on both groups). In practice, an intervention effect can be estimated in a regression equation (allowing further confounders to be controlled for) using an interaction term between a dummy variable indicating whether an observation belongs to the treatment group and a post intervention-period dummy (or, alternatively, time fixed-effects) (Angrist and Pischke, 2008). The design is illustrated in Figure 6 which demonstrates the importance of the 'parallel trends' assumption: that the outcomes for the intervention and control groups would have been parallel if not for the intervention. The other key assumption is the 'common shocks' assumption: that the intervention and control group are equally affected by any common shocks (other than the effects of the intervention).

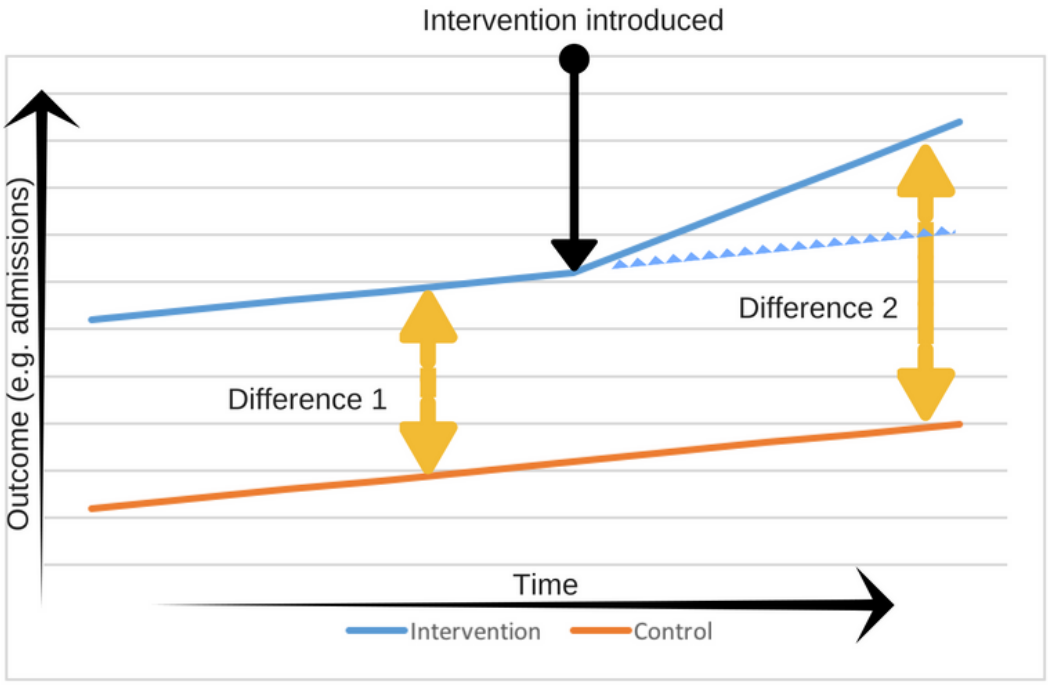


Figure 6: Difference-in-differences basics. The parallel outcome trend between the two groups (Intervention and Control) is assumed to continue (shown by the dotted blue line – the assumed conterfactual), except for the effects of the intervention on the Intervention group only. The difference-in-differences (i.e, Difference 2 – Difference 1) is thus attributed to the intervention effect

Ryan et al. have developed a checklist for robust DD analysis, shown in Box 4, below (Ryan et al., 2014).

Box 4: DD Checklist. Source: (Ryan et al., 2014)

1. Data on study outcomes is available both before and after the intervention was implemented for an exposed and control group
2. Trends in outcomes prior to intervention are parallel between exposed and control group
3. Outcome levels prior to intervention are not related to expectations for ability to change (e.g. no ceiling effects for one group, but not the other)
4. Test for violations of standard statistical assumptions (e.g. heteroscedasticity of standard errors)
5. Factors other than the intervention (occurring at time of treatment) should not differentially affect outcomes for exposed and control group
6. The composition of exposed and control group should not change over the time period of the study
7. Treatment effect does not 'spill-over' from exposed to control group

To summarise the method's issues discussed above:

- We assume that the intervention and control groups are comparable (but not necessarily the same), and that their outcome trends would have continued to be parallel over time if not for the intervention;
- We assume that no other effects on the outcome are affecting one group disproportionately (including their joining an intervention or ability to measure change, or other non-common shocks or spill-over effects).

4.5 Secondary analysis (difference-in-difference-in-differences – DDD, Chapter 7)

Secondary analysis of data can be used to explore effects of an intervention on specific subgroups in both randomised and non-randomised studies. Difference-in-difference-in-differences (DDD) analysis can be used to explore subgroup effects from DD analysis. As in any other regression equation, an additional interaction term can be added to represent the subgroup differentiation.

For example, subgroup analysis can be used to examine the effects of an integrated care intervention on those patients with specific 'types' of multimorbidity.

The simplest measure of multimorbidity, simply counting the number of diseases in a list (as discussed in section 1.3.1), has been criticised for not taking into account the severity of each condition. Diseases found in different organ systems are found to contribute differentially to health-related quality of life, for instance (Fortin et al., 2007a). Thinking at a single disease level, a patient with stage 1 chronic kidney disease (CKD) is likely to experience very different outcomes and have different care needs than a patient with stage 5 CKD (Noble and Lewis, 2008). Furthermore, differences in the list of diseases that is counted (e.g. one systematic review found the number of conditions included in multimorbidity counts ranged from 4-102, with an average of 18.5 (Diederichs et al., 2011)) will affect the results obtained (although guidelines have recently been drawn up to try and alleviate this problem – suggesting the use of a list of at least 12 of the most common chronic conditions for a given population) (Fortin et al., 2012). Secondary analysis can be used to explore whether the effect of an intervention varies through multiple measures of the multimorbidity concept.

As an extension of DD analysis, DDD shares the assumptions outlined in the section above. The DDD estimate starts with the time change in average outcome for the multimorbidity subgroup in the intervention group and removes the change in average outcome for that multimorbidity subgroup in the control group, plus removes the time change in average outcome for the counter multimorbidity subgroup in the intervention group (shown in equation below) (Wooldridge, 2007):

$$\widehat{\beta} = \left(\bar{y}_{I,M,2} - \bar{y}_{I,M,1}\right) - \left(\bar{y}_{C,M,2} - \bar{y}_{C,M,1}\right) - \left(\bar{y}_{I,N,2} - \bar{y}_{I,N,1}\right)$$

where

$\bar{y}$ = average outcome
1 = pre-intervention period
2 = post-intervention period
I = Intervention group
C = Control group
M = Complex multimorbidity subgroup
N = less complex multimorbidity subgroup

However, in taking this stratified approach, statistical power will inevitably be lost through comparing a smaller group of more similar patients than in the full analysis (Section 8.3 in Chapter 8 expands on this discussion further).

To summarise the method's issues discussed above:

- The method is valid only when the underlying DD assumptions are fulfilled (checked in the empirical Chapters, 6 and 7);
- We lose statistical power, as with any subgroup analysis.

4.6 Chapter Conclusions

In this chapter, I have shown the problems encountered when evaluating the type of healthcare interventions explored in this thesis, and outlined the main methods of analysis utilised in the following empirical chapters to answer the research aims. The next chapter is the first empirical piece using meta-analysis to determine the effectiveness of the case management intervention generally.

5. Effectiveness of case management review

5.1 Chapter Introduction

This empirical piece is a systematic review and meta-analysis examining effectiveness of the dominant form of integrated care in the NHS in England, case management, and across health-system goals. The article was published in *PLoS ONE* in July 2015.

Any supplementary material referred to in the text is available in Appendix 2.

PLOS | ONE

RESEARCH ARTICLE

Effectiveness of Case Management for 'At Risk' Patients in Primary Care: A Systematic Review and Meta-Analysis

Jonathan Stokes[1]*, Maria Panagioti[2], Rahul Alam[1], Kath Checkland[2], Sudeh Cheraghi-Sohi[1], Peter Bower[1]

1 NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom, **2** NIHR School for Primary Care Research, Centre for Primary Care, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom

* jonathan.stokes-3@postgrad.manchester.ac.uk

Citation: Stokes J, Panagioti M, Alam R, Checkland K, Cheraghi-Sohi S, Bower P (2015) Effectiveness of Case Management for 'At Risk' Patients in Primary Care: A Systematic Review and Meta-Analysis. PLoS ONE 10(7): e0132340. doi:10.1371/journal.pone.0132340

Editor: Terence J Quinn, University of Glasgow, UNITED KINGDOM

Received: April 11, 2015

Accepted: June 14, 2015

Published: July 17, 2015

Data Availability Statement: All relevant data are within the referenced included papers from the systematic review.

Funding: This work was funded by the National Institute for Health Research Greater Manchester Primary Care Patient Safety Translational Research Centre (NIHR GM PSTRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Background

An ageing population with multimorbidity is putting pressure on health systems. A popular method of managing this pressure is identification of patients in primary care 'at-risk' of hospitalisation, and delivering case management to improve outcomes and avoid admissions. However, the effectiveness of this model has not been subjected to rigorous quantitative synthesis.

Methods and Findings

We carried out a systematic review and meta-analysis of the effectiveness of case management for 'at-risk' patients in primary care. Six bibliographic databases were searched using terms for 'case management', 'primary care', and a methodology filter (Cochrane EPOC group). Effectiveness compared to usual care was measured across a number of relevant outcomes: **Health** – *self-assessed health status*, *mortality*; **Cost** – *total cost of care*, healthcare *utilisation* (*primary and non-specialist care* and *secondary care* separately), and; **Satisfaction** – *patient satisfaction*. We conducted secondary subgroup analyses to assess whether effectiveness was moderated by the particular model of case management, context, and study design. A total of 15,327 titles and abstracts were screened, 36 unique studies were included. Meta-analyses showed no significant differences in *total cost*, *mortality*, *utilisation of primary* or *secondary care*. A very small significant effect favouring case management was found for *self-reported health status* in the short-term (0.07, 95% CI 0.00 to 0.14). A small significant effect favouring case management was found for *patient satisfaction* in the short- (0.26, 0.16 to 0.36) and long-term (0.35, 0.04 to 0.66). Secondary subgroup analyses suggested the effectiveness of case management may be increased when delivered by a multidisciplinary team, when a social worker was involved, and when delivered in a setting rated as low in initial 'strength' of primary care.

Competing Interests: The authors have declared that no competing interests exist.

Conclusions

This was the first meta-analytic review which examined the effects of case management on a wide range of outcomes and considered also the effects of key moderators. Current results do not support case management as an effective model, especially concerning reduction of secondary care use or total costs. We consider reasons for lack of effect and highlight key research questions for the future.

Review Protocol

The review protocol is available as part of the PROSPERO database (registration number: CRD42014010824).

Introduction

Many health care systems currently face significant pressures resulting from both increasing numbers of older patients with multiple long-term conditions (multimorbidity), and pressure to reduce health care budgets or provide more efficient use of current resources [1].

To relieve these pressures, many policy makers and health system planners advocate 'integrated care' [1, 2].

Integrated care is a complex concept. Broadly, it is designed to "create connectivity, alignment and collaboration" [3]. A number of different methods can be used to achieve these inter-connections, and they can occur at multiple 'levels' of the health system (e.g. financing, resource management, service delivery—see Fig 1). Outcomes of effective integration of care are presumed to be better patient experience and outcomes, as well as greater efficiency [4] (i.e. patient satisfaction; health; and cost-effectiveness), therefore potentially addressing two of the major system pressures simultaneously.

A popular model of 'integrated care' at the service delivery level is 'case management' in primary care [6, 7]. Case management has been defined as:

> "a collaborative process of assessment, planning, facilitation, care coordination, evaluation, and advocacy for options and services to meet an individual's and family's comprehensive health needs through communication and available resources to promote quality, cost-effective outcomes" [8].

Variations exist in the delivery of case management. However, there are common components [6]:

- *case-finding* (identifying those 'at risk' who require case management, usually through prediction of high costs in the future [9, 10])
- *assessment* of the needs of the individual patient, and *care planning* (individualised care plan bringing together details of patient's personal circumstances with health and social care needs, and aiming to match these needs with service provision)
- *care co-ordination* (navigational role of case manager involving continual communication with patients, carers, professionals and services e.g. medication management, self-care support, care advocacy and negotiation; with regular review, monitoring and adaptation of the care plan)

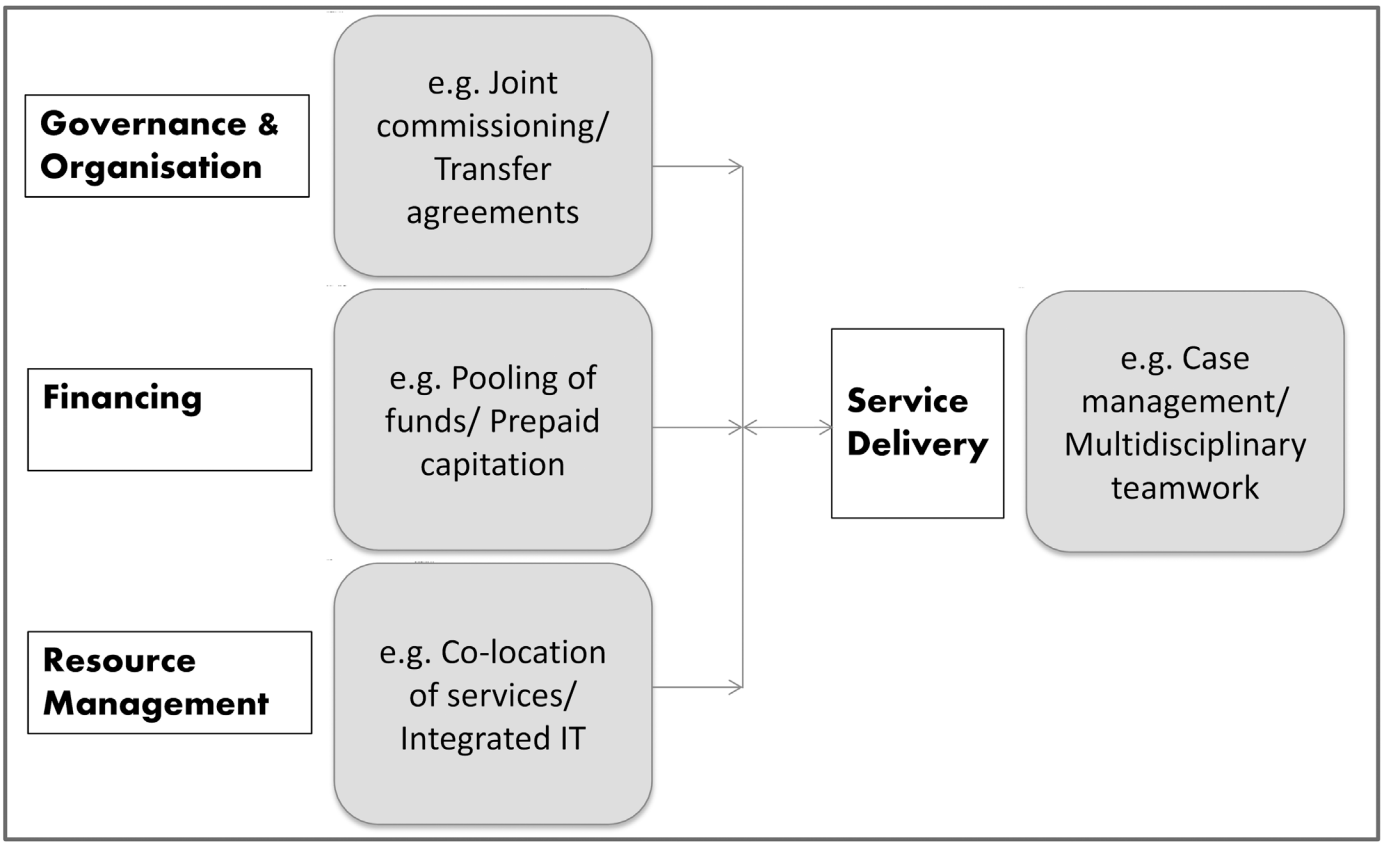


Fig 1. Examples of popular methods to 'integrate' care [3] within the health system [5].

doi:10.1371/journal.pone.0132340.g001

Case finding of 'at-risk' individuals can be done in three ways [11]:

1. clinical judgement (expert opinion),
2. threshold modelling (defining a set of rules e.g. number of previous hospital admissions, which alert the practitioner that the patient is at risk), or
3. using a predictive risk tool (where an algorithm is used to attempt to *predict* those patients who are at risk of a defined event)

In theory, the case management process may increase efficiency by reducing unnecessary contacts with the health system, including fragmented routine contacts, as well as emergency contacts caused by potentially preventable exacerbations. The goal is to better co-ordinate care, offering individually-tailored contacts and care planning.

Primary care is a suitable context for integrated care due to its place at the heart of the health system [12]. It is argued that increasing care in the community setting will facilitate cost savings compared to expensive hospital overheads [13]. In many health systems, primary care acts as a 'gatekeeper' to the rest of the system [14] and primary care practitioners should be particularly suited to managing and co-ordinating care for multiple health problems, compared to specialist physicians [15].

The potential benefits of case management have led to adoption in practice in many countries [6]. For example, in the United Kingdom, recent changes to the NHS GP contract (under the Unplanned Admissions Enhanced Service section), require a minimum of 2% of the

risk-identified population to be proactively case managed [16]. In the USA, a number of health insurers and health maintenance organisations offer case management to patients with long-term conditions, for example the 'Guided care' and similar programmes [6].

It is important that the provision of case management in primary care should be based on rigorous evidence. While many descriptive reviews exist examining specific types of case management in primary care (such as nurse-led case management [17]), there is no published systematic review of a range of current case management models for high risk individuals in primary care that provides a formal meta-analytic review of its effectiveness across a range of relevant outcomes.

Objectives

1. To synthesise the evidence for the effectiveness of case management in primary care for 'at risk' patients
2. To explore whether the effectiveness of case management in primary care is moderated by the particular model of case management implemented, context, and study design.

Methods

The methods and results for this review are reported in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The review protocol is available as part of the PROSPERO database (registration number: CRD42014010824).

Eligibility criteria

Studies were included in this review if they met the following criteria:

- Population: Adults (18+) with long-term condition(s)
  (While prevalence of multimorbidity (i.e. 'complex' cases) is highest in the elderly, the absolute numbers affected are greater in those below 65 [18])
- Intervention:
  - Adopting methods to identify 'at-risk' patients to receive the case management, with the aim of preventing acute exacerbations of symptoms, and/or secondary care utilisation among those at higher risk
  - Case management, including all of the following activities: case-finding; assessment; care planning; care co-ordination; regular review, monitoring and adaptation of the care plan
  - Primary care/community-based management (regardless of where the case was first identified)
- Comparison: usual care or no-case management
- Outcome categories: **Health**–*self-assessed health status*, *mortality*; **Cost**–*total cost of care*, healthcare *utilisation* (*primary and non-specialist care* and *secondary care* separately), and; **Satisfaction**–*patient satisfaction*
- Study design: Quantitative empirical research, meeting Cochrane Effective Practice and Organisation of Care (EPOC) Group study design criteria: randomised controlled trials

(RCTs), non-randomised controlled trials (nRCTs), controlled before and after studies (CBA), and interrupted time series (ITS)

Exclusion criteria:

- Case management targeted solely at care for patients with mental health problems, although mental health conditions could be included where they were co-morbidities alongside physical long-term conditions
- Hospital discharge planning (short-term management to facilitate the transition from hospital to home [19])
- Non-English language papers and grey literature

Search Strategy

Six main electronic bibliographic databases were searched for potential studies from inception until end of April 2014: MEDLINE (Ovid), EMBASE (Ovid), CINAHL, Cochrane Register of Controlled Trials (CENTRAL), Health Management Information Consortium (Ovid), and CAB Global Health (Ovid),. The search strategy used three key blocks of terms (including subject headings as well as text-words): 1) Case management2) EPOC methodology filter[20]3) Primary care filter[21]. S1 Appendix shows an example of the full search strategy for the MEDLINE database.

Hand searches of the reference lists of included papers, plus previous relevant systematic reviews [17, 22–30] supplemented the database searches.

Results from the above searches were combined in an Endnote library, and duplicates were removed (n = 2186) prior to study selection.

Study Selection

Study selection was carried out in two stages. First, titles and abstracts of the identified studies were screened in full by the first author. A proportion of these titles and abstracts (10%) were then independently screened by a second author (kappa coefficient = 0.78). Following this initial screening, the full texts of the identified articles were retrieved, and reviewed against the inclusion/exclusion criteria. Forty percent (n = 106/266) of the full text screening was carried out by two reviewers independently. Inter-rater reliability was high (kappa coefficient = 0.81), and any disagreements (n = 7) were resolved by group discussion (resulting in 4 included, and 3 excluded). The remaining full text screening was completed by the first author alone.

Data extraction

A data extraction form was formulated using Microsoft Excel. The form was initially piloted on two randomly selected studies. The following descriptive data were extracted for included studies:

- Patient: target population; total sample size (intervention/control); proportion of males; average age; average baseline number of long-term conditions; average baseline number of emergency department visits/ hospital admissions in previous year
- Intervention: name of the case management model; brief description of model; intensity of intervention; multidisciplinary team(and specific members) or single case manager; primary case manager; primary location of case management; risk stratification model used; whether

there was 24-hour availability of a case manager; caseload; whether the case manager received training in the intervention protocol; reimbursement method

- Context: country was used to define the 'strength' of primary health care, classified according to Starfield & Shi's work [31]
- Outcome categories: **Health**–*self-assessed health status*, *mortality*; **Cost**–*total cost of care*, healthcare *utilisation* (*primary and non-specialist care* and *secondary care* separately), and; **Satisfaction**–*patient satisfaction*
- Study design: design; study duration; unit of analysis; eligibility criteria; type of control group

On a separate sheet, relevant quantitative data for the meta-analysis were extracted (see quantitative analysis section below). Where adjusted and unadjusted results were both presented, the result adjusting for the most potentially confounding variables was extracted.

25 percent (n = 9 studies) of the data were extracted by two researchers working independently. The agreement was high (kappa coefficient = 0.85, across 326 data points), and the remainder of the data were extracted by the first author, and the accuracy of extraction verified by a second reviewer.

Quality Assessment

In the original protocol, we predicted having to use multiple measures of risk of bias to suit the various study types included in the eligibility criteria. However, having identified the full text articles and study designs represented, it became clear that it would be possible and preferable to use a single quality assessment tool, the EPOC risk of bias tool [32], better allowing comparison of quality across the included studies. The EPOC risk of bias tool encompasses nine standardised criteria to judge the quality of all RCTs, nRCTs, CBA and ITS studies. Each of the nine criteria is judged on a 3-point scale, corresponding to: low risk, unclear risk, and high risk. To ease comparison between studies, the total number of criteria met by each included study was also reported. Those studies at high risk of bias (fulfilling three or less criteria) were removed from the synthesis for sensitivity analysis.

Quantitative Analysis

Meta-analysis was carried out on six outcome categories related to the three main health system goals [5]. These were: **Health**–*self-assessed health status*, *mortality*; **Cost**–*total cost of care*, healthcare *utilisation* (*primary and non-specialist care* and *secondary care* separately), and; **Satisfaction**–*patient satisfaction*. Table 1 clarifies which measures were included within each of these outcome categories.

In addition to the outcomes specified in the original protocol, we also attempted to extract data related to the outcome category of 'patient safety': *admissions for ambulatory care sensitive conditions* [33], and *polypharmacy* (simple count of medications). However, none of the included studies reported these outcome measures, so results could not be synthesised.

Meta-analysis was carried out on each outcome, distinguishing between effects over the short-term (0–12 months), and longer-term (13+ months). Meta-analysis used the standardised mean difference measure, based on the mean of the case management group minus mean of the control group, divided by the pooled standard deviation [34]. When multiple measures were available for a single study within a certain outcome category, the median effect was used, as recommended in the literature [35] (e.g. for the outcome of *self-assessed health status*, if a measure of activities of daily living, of restricted activity days, and a measure of QALYs were all available for a given study, the effect size for each of these would be calculated, and the *median*

Table 1. Outcome measures.

Self-assessed health status	**Mortality**
- (Instrumental/) Activities of Daily Living	- Mortality within study period
- Physical/ mental health questionnaires	
- Bed days/ restricted activity days	
- Quality Adjusted Life Years (QALYs)	
Total cost of services	**Utilisation of primary and non-specialist care**
- Total cost	- Primary care physician visits
- Total insurance expenditure/ reimbursement	- Home care visits
	- Social worker visits
	- Nursing visits
Utilisation of secondary care	**Patient satisfaction**
- Emergency Department visits	- Patient satisfaction questionnaires
- Hospital admissions/ re-admissions/ days	- Patient quality of care ratings
- Inpatient/outpatient utilisation	
- Skilled nursing facility visits/ days	
- Ambulance calls	

doi:10.1371/journal.pone.0132340.t001

standardised mean difference would represent this studies' overall effect for this outcome). We adopted Cohen's rule of thumb for interpreting effect sizes, i.e. that 0.2 indicates a small effect, 0.5 a medium, and 0.8 a large effect [36].

Heterogeneity in the outcomes was assessed using the I^2 statistic, interpreted as the percentage of total variation in the study estimates due to heterogeneity [37]. A random effects model was chosen to present the pooled effect results based on the relatively high level of heterogeneity assumed between studies evaluating a complex intervention in a variety of service contexts.

Funnel plots were performed to assess small sample bias (which may bean indicator of publication bias), but only for those outcomes drawing on 10 or more studies, as recommended [38]. Egger's test of small-study effects was additionally performed to quantify observations in the funnel plots [39].

As a complex intervention, context may be of some importance when assessing case management [40]. Subgroup analyses were performed where 10 or more studies contributed effect size data. The pre-specified variables were:

- Context: strength of primary health care orientation of the health system (low versus intermediate/ high)
- Type of case management: multidisciplinary team (MDT) versus single case manager; type of risk tool used (judgement versus threshold/ predictive risk modelling); inclusion of a social worker in the case management (versus absence)
- Study design: RCT versus non-RCT

Table 2 discusses the justifications for these choices of subgroup.

Statistical significance between subgroups was judged by overlap of each subgroup's pooled effect (i.e. overlap of confidence intervals between subgroup effects indicates no significant difference) [47].

The majority of effect sizes were calculated using the Metaeasy software add-in for Microsoft Excel (version 1.0.4) [35]. The Metaeasy software allows standardisation of effect size from a variety of input parameters (dichotomous, continuous or both data types), according to eight possible methods described by the Cochrane Collaboration [48]. When multiple methods are

Table 2. Subgroup analyses.

Strength of primary care orientation: 'Case management' may be replacing some of the functions of well-co-ordinated, person-centred primary care [12]. The effects of case management may therefore be greater when it is delivered in contexts where routine primary care services are less well developed. To test this hypothesis, we stratified results by the assessed orientation to primary care of the study country's health system. The primary care orientation scores were developed by Starfield & Shi, and take into account—for each country—both characteristics of health system policy that are conducive to primary care, as well as characteristics of clinical practice [31].
Multidisciplinary team versus single case manager: The hypothesis that teams are more effective than individuals at problem solving and delivering services is established across a number of diverse organisational settings [41], and teams have also been advocated in the treatment of patients with long-term conditions [42]. We tested whether case management by teams was more effective than by individuals.
Type of risk tool used: Targeting the 'correct' patients will be vital to any effective case management programme, particularly when assessed on cost and utilisation outcomes [43]. To test whether identification of the 'correct' patients was more effective when carried out by a rule-based model, we compared clinical judgement with rule-based and predictive models.
Inclusion of a social worker in case management: Collaboration between health and social services is thought to be important for effective case management [6], particularly of multimorbid patients who frequently have a complex mix of health and social care issues [44]. It also provides an additional, 'professional' level of care integration to the intervention [45], encouraging the different disciplines to work more closely together. To test the relative effectiveness of inclusion of a social worker, we therefore stratified results by this variable.
RCT versus non-RCT: RCTs are theoretically less vulnerable to bias, and therefore may give slightly different estimates of effect compared to observational studies (smaller/larger/reversed) [46]. We therefore compared RCTs to non-RCTs to observe any potential inconsistencies.

doi:10.1371/journal.pone.0132340.t002

available for a single outcome, methods are prioritised according to expected statistical precision [35]. To maximise included results, the Metaeasy effect size calculation methods were supplemented by methods developed by Lipsey & Wilson [49], with a calculator available at http://www.campbellcollaboration.org/resources/effect_size_input.php. Effect directions were transformed so that a positive effect represented favouring case management for all outcome measures. These final effect sizes and their standard errors were then input to STATA together with relevant study information for the subgroup analysis. The final meta-analyses were then run on STATA (version 13) [50] using the *metan* command [51]. Funnel plots were prepared using the *metafunnel* command [52], and the Egger test with the *metabias* command [53].

Sensitivity analyses and Multiple comparisons

Two separate post-hoc sensitivity analyses were conducted in addition to the specified PROSPERO protocol. Studies were removed from analysis if they were:

1. at high risk of bias (meeting 3 or less of the criteria for assessment of study quality)
2. set in a Veteran's health setting, where over 90% of the patients were males.

With multiple comparisons, the chances of inflating type I errors is increased [54]. We therefore used the Holm-Bonferroni adjustment [55] for multiple comparisons to identify potential false positive results.

Results

Fig 2 shows the PRISMA flow diagram, with the studies included/excluded at each stage of the screening process. 36 unique studies were finally included in the meta-analyses.

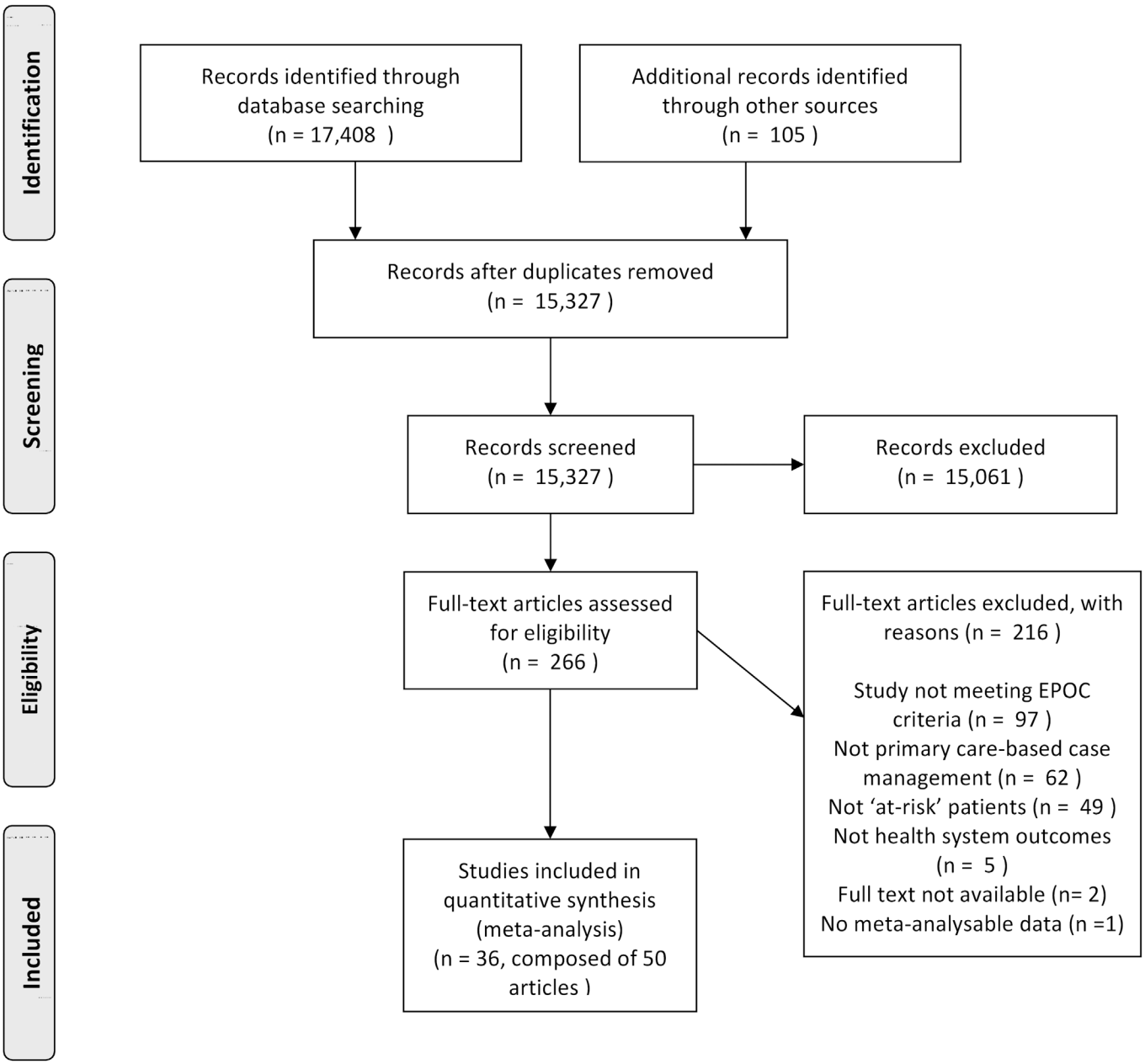


Fig 2. PRISMA flow diagram of study selection [56].

doi:10.1371/journal.pone.0132340.g002

Characteristics of included studies

Unsurprisingly, since the majority aimed at an elderly population, the average age in nearly all studies was high (mean age: 75.7, range of mean age: 49.0 to 87.3).

Table 3 shows the demographic characteristics of the studies. Of note, 5 (14%) provided case management to a population composed of over 90% males, carried out in veterans' settings. Inadequate information was provided across studies on baseline number of long-term conditions, and baseline utilisation of emergency and specialist services.

Table 3. Demographics of included studies. N/R = Not Reported; N/A = Not Applicable.

Study	Total (n)	Intervention (n)	Control (n)	% Male (controls)	Average age (controls) +-SD	Average no. of chronic conditions (controls) +-SD	Baseline average ED visits in previous year (controls)	Baseline average Hospital admissions in previous year (controls)
Beland 2006a[57]; Beland 2006b[58]	1309	656	653	28	82.3+-7.2	5.0+-2.3	N/R	N/R
Bernabei 1998[59]	199	99	100	29	81.3+-7.4	4.8+-1.7	N/R	N/R
Bird 2010[60]	COPD: 124; CHF: 89	COPD: 78; CHF: 67	COPD: 46; CHF: 22	COPD: 67; CHF: 63	COPD: 70 +-N/R; CHF: 76+-N/R	N/R	COPD: 4.8 +-3.0; CHF: 5.1+-1.8	COPD: 3.3+-2.1; CHF: 2.8+-1.4
Boult 2008[61]; Leff 2009[62]; Boyd 2010 [63]; Boult 2011[64]; Boult 2013 [65]	904	485	419	45	78.1+-N/R	4.3+-N/R	N/R	N/R
Boyd 1996[66]	54	27	27	30	81+-N/R	N/R	1.1+-N/R	1.6+-N/R
Burns 1995[67]; Burns 2000[68]	128	60	68	99	70.8+-3.7	2.0+-1.8	N/R	N/R
Coburn 2012[69]	1736	873	863	40	74.9+-6.5	3.8+-2.0	N/R	N/R
Counsell 2007[70]; Counsell 2009[71]	951	474	477	23	71.6+-5.8	2.6+-1.5	1.2+-2.4	0.4+-1.2
Dalby 2000[72]	142	73	69	23	78.1+-5.3	N/R	N/R	N/R
De Stampa 2014 [73]	428	105	323	28	87.3+-7.3	N/R	N/R	N/R
Dorr 2008[74]	3432	1144	2288	35	76.2+-7.1	N/R	N/R	N/R
Enguidanos 2006 [75]	452	TCM: 113; GCM: 117; POS: 124	98	36	N/R (65+)	N/R	N/R	N/R
Fan 2012[76]	426	209	217	96	65.8+-8.2	N/R	2.7+-2.2	N/R
Fitzgerald 1994[77]	668	333	335	100	64.6+-7.7	N/R	N/R	N/R
Fordyce 1997[78]	1090	326	764	45	N/R (65+)	N/R	N/R	0.24+-0.4
Gagnon 1999[79]	427	212	215	41	81.8+-6.7	N/R	0.9+-1.2	0.4+-0.7
Gravelle 2007[80]	7695 (practices)	62 (practices)	6960 (practices)	N/A	N/A	N/A	N/A	N/A
Hogg 2009[81]; Gray 2010[82]	241	120	121	37	72.8+-N/R	2.4+-N/R	N/R	N/R
Kruse 2010[83]	379	130	249	35	75.1+-6.8	N/R	N/R	N/R
Leung 2004[84]	260	130	130	52	75.3+-7.2	2.9+-1.5	0.3+-0.6	0.9+-1.2
Levine 2012[85]	298	156	142	36	80.6+-8.7	2.4+-1.5	N/R	N/R
Martin 2004[86]	93	44	49	65	69.1+-20	N/R	N/R	N/R
Metzelthin 2013[87]	346	153	193	31	76.8+-4.92	N/R	N/R	N/R
Morishita 1998[88]; Boult 2001[89]	568	294	274	58	78.7+-5.8	N/R	N/R	0.8+-1.0
Newcomer 2004 [90]	3079	1537	1542	40	N/R (65+)	N/R	N/R	N/R
Ploeg 2010[91]	719	361	358	46	81.3+-4.4	N/R	N/R	N/R
Rodenas 2008[92]	152	101	51	N/R	N/R (65+)	N/R	N/R	N/R
Rubenstein 2007 [93]	793	380	412	97	74.3+-6.1	N/R	N/R	N/R
Schraeder 2001[94]	941	530	411	25	75.4+-6.4	N/R	N/R	1.6+-0.94
Schraeder 2008[95]	677	400	277	40	76.4+-7.9	N/R	N/R	N/R
Shannon 2006[96]; Alkema 2007[97]	781	377	404	34	83.7+-7.36	N/R	0.51+-1.06	N/R

(*Continued*)

Table 3. (*Continued*)

Study	Total (n)	Intervention (n)	Control (n)	% Male (controls)	Average age (controls) +-SD	Average no. of chronic conditions (controls) +-SD	Baseline average ED visits in previous year (controls)	Baseline average Hospital admissions in previous year (controls)
Sledge2006[98]	96	47	49	41	49+-N/R	N/R	N/R	N/R
Stuck 2000[99]	791	264	527	29	81.5+-4.5	N/R	N/R	N/R
Sylvia 2008[100]; Boyd 2008[101]	127	62	65	54	75.8+-N/R	2.9+-N/R	N/R	N/R
Toseland 1996 [102]; Toseland 1997[103]; Engelhardt 1996 [104]; Engelhardt 2006 [105]	160	80	80	100	72.6+-5.75	2.6+-1.3	N/R	N/R
van Hout 2010[106]	651	331	320	31	81.5+-4.3	2.0+-1.4	N/R	1.6+-3.8

doi:10.1371/journal.pone.0132340.t003

Table 4 summarises the potentially relevant contextual factors. Of the 36 studies included, the majority were from the USA (n = 21, 58%). When classified according to relative strength of primary care orientation, 23 studies (64%) were set in a system with low strength of primary care, and 13 (36%) in an intermediate, or high strength system. Three studies (8%) were targeted at patients with specific conditions (COPD/chronic heart failure) while the majority targeted populations more broadly on frailty, chronic illness or high utilisation (92%).

A brief qualitative description of each intervention is also provided in Table 4. Table 5 compares some of the key attributes of each intervention more directly. Many criteria highlighted as key to understanding integrated care interventions [9] were inadequately reported which limited their utility for analysis. However, the type of risk tool, whether the case management was carried out by a MDT or single case manager, and the inclusion of a social worker in the case management could be recorded for all studies. The majority of studies used a 'threshold'/ 'predictive risk modelling' risk assessment tool (n = 32, 89%), with only 4 (11%) using clinical judgement. Twenty-one studies (58%) employed MDT case management. A social worker was involved in the case management in 12 studies (33%).

Methodological Quality

The majority of studies (n = 28, 78%) used an RCT. The length of follow-up in the studies varied, with a range of 6 to 60 months. Table 6 shows the methodological quality according to the nine criteria of the EPOC risk of bias tool. The studies were of variable quality, with 64% (n = 23) fulfilling seven or more criteria, 30% (n = 11) fulfilling between four and six criteria, and 6% (n = 2) fulfilling three or less.

Primary analyses

Figs 3–8 show the results of the primary meta-analyses for the six outcome categories assessed (both short- and long-term).

Health. A statistically significant effect favouring case management was found for *self-assessed health status* (Fig 3) in the short-term (0.07, 95% CI 0.00 to 0.14, I^2 = 35.1%, p = 0.094), but this effect was not present in the long-term (-0.01, 95% CI -0.08 to 0.05, I^2 = 12.8%, p = 0.327). No significant effect was found for *mortality* (short-term: 0.08, 95% CI -0.03 to 0.19, I^2 = 63.6%, p = 0.001; long-term: 0.03, 95% CI -0.04 to 0.09, I^2 = 40.0%, p = 0.067 –Fig 4).

Table 4. Context of included studies.

Study	Country	Strength of primary care orientation (of country)*	Population	Study design (n participants)	Study length (months)	Brief description of model	Extracted outcomes for meta-analysis
Beland 2006a [57]; Beland 2006b [58]	Canada	intermediate	Elderly & functionally disabled	RCT; n = 1309	22	Community-based MDTs with full clinical responsibility for delivering and coordinating services. 24-hour availability via phone. Actively followed patients through care trajectory.	Utilisation (primary/secondary care)
Bernabei 1998 [59]	Italy	high#	Elderly & receiving home health services/ assistance	RCT; n = 199	12	MDT-designed care plan following assessment by GP/ case manager. Case manager followed-up every two months, and constantly available to deal with problems and monitor provision of services.	Mortality, Self-reported health status, Utilisation (primary/secondary care)
Bird 2010 [60]	Australia	intermediate	Frequent presenters for COPD/CHF	CBA; n = 124 (COPD)/n = 89 (CHF)	11	Patients allocated to disease-specific stream based on presentations. Results of initial case facilitator assessment discussed at case conference with MDT. Education, self-management, and coordination focus. Follow-up mostly at home	Mortality, Utilisation (secondary care)
Boult 2008 [61]; Leff 2009 [62]; Boyd 2010 [63]; Boult 2011 [64]; Boult 2013 [65]	USA	low	Elderly & high-risk multimorbid	cRCT; n = 904	32	Nurse responsible for assessing, planning care, monitoring, coaching self-management, coordination of services, and education for patient and family. Helped by team of physicians.	Total cost of services, Mortality, Patient satisfaction, Self-reported health status, Utilisation (primary/secondary care)
Boyd 1996 [66]	USA	low	Elderly & chronically ill	nRCT; n = 54	12	Community-based, integrating case management in patient's everyday life, with case manager available to monitor the patient's chronic illness (es). Developing care plan, coordinating services, and providing counselling support.	Mortality
Burns 1995 [67]; Burns 2000 [68]	USA	low	Frail elderly	RCT; n = 98	24	Consistent involvement of MDT (GEM team). Initially assess patient and provide ongoing management. Most appropriate team member for given patient served as main liaison.	Mortality, Self-reported health status, Utilisation (primary/secondary care)
Coburn 2012 [69]	USA	low	Elderly & chronically ill	RCT; n = 1736	60	Patients risk-stratified within intervention. Regardless of strata, nurse developed an individualised care plan. Group interventions were also provided by the care managers. Nurses collaborated with other healthcare professionals when required.	Mortality

(*Continued*)

Table 4. (*Continued*)

Study	Country	Strength of primary care orientation (of country)*	Population	Study design (n participants)	Study length (months)	Brief description of model	Extracted outcomes for meta-analysis
Counsell 2007 [70]; Counsell 2009 [71]	USA	low	Low income elderly	RCT; n = 951	24	Care plan developed in collaboration with MDT. Weekly team meetings to review team successes and problem-solve barriers to implementation. At least monthly home-based care management supported by an electronic medical record and web-based tracking system.	Total cost of services, Mortality, Patient satisfaction, Self-reported health status, Utilisation (secondary care)
Dalby 2000 [72]	Canada	intermediate	Frail elderly living in the community	RCT; n = 142	14	Nurse-led comprehensive assessment. Care plan developed in conjunction with primary physician. Follow-up visits and calls as needed. Nurse coordinates further community services	Mortality, Utilisation (primary/secondary care)
De Stampa 2014 [73]	France	low	Frail elderly	CBA; n = 428	12	Two-person team responsible for patient's care trajectory. The primary care manager developed care plan, ongoing role of physician to collaborate and share information. Support as needed from geriatricians.	Self-reported health status, Utilisation (secondary care)
Dorr 2008 [74]	USA	low	Elderly & chronically ill	nRCT; n = 3432	24	Case management aimed at addressing social, cognitive, and functional needs. Assisted by specialised IT software including structured protocols and guidelines. Co-creation of care plan with patients.	Mortality, Utilisation (secondary care)
Enguidanos 2006 [75]	USA	low	Frail elderly	RCT; n = 452	12	Study compares 4 strategies of care. Telephone case management (single case manager); Geriatric care management (GCM) (MDT involvement in care plan); GCM with purchase of service capability (addition of $2000 of designated paid services within first 6 months); Information and referral assistance (most basic, acts as control group).	Utilisation (primary/ secondary care)
Fan 2012 [76]	USA	low	Frequent presenters for COPD	RCT; n = 426	12	Initial individual educational programme, needs assessment, and an overview of COPD. Reinforced during group session, and with follow-up phone calls. Individualised plan for flare-ups, including prescriptions for prednisone and antibiotic.	Mortality, Patient satisfaction, Self-reported health status, Utilisation (secondary care)

(*Continued*)

Table 4. (*Continued*)

Study	Country	Strength of primary care orientation (of country)*	Population	Study design (n participants)	Study length (months)	Brief description of model	Extracted outcomes for meta-analysis
Fitzgerald 1994 [77]	USA	low	Inpatient medical service users	RCT; n = 668	12	Included instructing patients about their medical problems, facilitating access to usual care, and identifying and fulfilling unmet social and medical needs with standard or alternative sources of care. Periodic assessment of medical and social needs. Coordination of all appointments for patient. 24-hour telephone access	Mortality, Utilisation (primary/secondary care)
Fordyce 1997 [78]	USA	low	Frail elderly	RCT; n = 1090	36	Yearly health, functional, and social evaluation. Weekly team meetings where nurse presented cases for review. Medical-functioning profile worked up for each patient, acting as indication of intensity of follow-up, as needed. Follow-up mostly by telephone.	Utilisation (secondary care)
Gagnon 1999 [79]	Canada	intermediate	Frail elderly	RCT; n = 427	10	Coordination of all healthcare providers and implementation of a responsive plan of care. Monthly phone calls, and a home visit every 6 weeks were the minimum standard. Additional contacts when required. Specialist consultation available to nurses for complicated cases.	Patient satisfaction, Self-reported health status, Utilisation (secondary care)
Gravelle 2007 [80]	UK	high	Frail elderly	CBA; n = 7757 (practices)	48	Assessment, using structured assessment tools, a physical examination, which resulted in an individualised care plan. Patients were then monitored at a frequency determined by their classification of risk.	Mortality, Utilisation (secondary care)
Hogg 2009 [81]; Gray 2010 [82]	Canada	intermediate	Older & at-risk of adverse outcomes	RCT; n = 241	18	Nurses and pharmacist co-located at family practice, but delivered care almost exclusively at patient's home. Team-developed care plan. 22 patients also received a tele-health system for remote monitoring.	Total cost of services, Self-reported health status, Utilisation (primary/secondary care)
Kruse 2010 [83]	USA	low	Elderly & chronically ill, at-risk for catastrophic illness	nRCT; n = 379	60	Assessed patient's needs, provided education, coordinated referrals, provided first-access care and follow-up care following visits to doctor/ hospital on the telephone.	Mortality, Utilisation (primary/secondary care)

(*Continued*)

Table 4. (*Continued*)

Study	Country	Strength of primary care orientation (of country)*	Population	Study design (n participants)	Study length (months)	Brief description of model	Extracted outcomes for meta-analysis
Leung 2004 [84]	Hong Kong	intermediate^	Community-dwelling frail elderly	RCT; n = 260	6	Regular home-visits and telephone consultations. Care plan designed in discussion with patient and caregiver. Coordination of health and social services through referral plus case conference. Monitoring of health and hospitalisation patterns via computer programme. Counselling, health education, and supportive group services.	Self-reported health status, Utilisation (primary/secondary care)
Levine 2012 [85]	USA	low	Elderly & multimorbid, at-risk for hospitalisation	RCT; n = 298	12	Included early identification and treatment of illness exacerbation, patient-specific health education, self or caregiver management of disease, and advance care planning and other psychosocial issues. Team worked closely at all stages.	Total cost of services, Patient satisfaction, Utilisation (primary/secondary care)
Martin 2004 [86]	New Zealand	intermediate+	Acutely deteriorating COPD patients	RCT; n = 93	12	Generic care plan was individualised and signed off. Supplies of antibiotics and prednisone made available. Copies of plan held by each potential provider of care. Routine support and further education available.	Utilisation (primary/secondary care)
Metzelthin 2013 [87]	The Netherlands	high	Frail elderly	cRCT; n = 346	24	Core team (GP and nurse) cooperate closely with other health professionals as needed. Initial home-visit and assessment, meeting to design care plan, and treatment starts with protocol offering recommendations and guidelines.	Self-reported health status
Morishita 1998 [88]; Boult 2001 [89]	USA	low	Elderly & high-risk	RCT; n = 568	18	Consistent involvement of MDT (GEM team). Specialised GEM clinic introduced, where patients were followed-up. Individual team members saw patients approximately monthly, met to discuss. Regular telephone calls, and available 24-hours on telephone service	Total cost of services, Mortality, Patient satisfaction, Self-reported health status, Utilisation (primary care)
Newcomer 2004 [90]	USA	low	High-risk elderly	RCT; n = 3079	12	Patients triaged by risk category after initial assessment. Predominant method of contact was telephone, supplemented by monitoring utilisation. Nurse case manager distributed educational material and advice, coordinated services, but no direct role in treatment management.	Self-reported health status, Utilisation (primary/secondary care)

(*Continued*)

Table 4. (*Continued*)

Study	Country	Strength of primary care orientation (of country)*	Population	Study design (n participants)	Study length (months)	Brief description of model	Extracted outcomes for meta-analysis
Ploeg 2010 [91]	Canada	intermediate	Elderly & at-risk of functional decline	RCT; n = 719	12	Nurse-led comprehensive initial assessment, collaborative care planning, health promotion, and referral to community health and social support services. Assessments at baseline, 6 and 12 months. Additional health education and referrals to other health services.	Total cost of services, Mortality, Self-reported health status, Utilisation (primary/secondary care)
Rodenas 2008 [92]	Spain	high	Elderly & receiving home care	RCT; n = 152	12	Direct interaction with the patients was carried out by a MDT. The team took charge of: 1) assessing individual needs 2) designing and starting individual care itineraries 3) benefit quality assurance, and 4) monitoring and on-going review of the strategy. Extra health and social care resources were also available for the intervention group.	Patient satisfaction, Utilisation (primary/ secondary care)
Rubenstein 2007 [93]	USA	low	High-risk elderly	RCT; n = 793	36	Initial telephone assessment by physician assistant case manager. Some patients referred for further assessment and an interdisciplinary care plan at a geriatric assessment unit. Coordination of follow-up by phone, each patient mailed a copy of the care plan.	Self-reported health status, Utilisation (secondary care)
Schraeder 2001 [94]	USA	low	Community-dwelling elderly	RCT; n = 941	24	Team's goal was to provide enhanced primary care by providing assessments, flexible home office visits, detailed care planning, routine telephone monitoring, and coordination and procurement of supportive services. Nurse and care assistant co-located.	Total cost of services, Mortality, Utilisation (secondary care)
Schraeder 2008 [95]	USA	low	Community-dwelling, chronically ill elderly	nRCT; n = 677	36	Intervention emphasised collaboration between physicians, nurses and patients, risk identification, comprehensive assessment, collaborative planning, health monitoring, patient education, and transitional care. Nurse and care assistant co-located.	Utilisation (secondary care)
Shannon 2006 [96]; Alkema 2007 [97]	USA	low	Elderly & high utilisers	RCT; n = 781	12	Telephone-based management to coordinate services bridging medical and social care. Focus on referrals. Monthly follow-up calls.	Mortality, Utilisation (primary/secondary care)

(*Continued*)

Table 4. (*Continued*)

Study	Country	Strength of primary care orientation (of country)*	Population	Study design (n participants)	Study length (months)	Brief description of model	Extracted outcomes for meta-analysis
Sledge 2006 [98]	USA	low	Recent high use of inpatient services	RCT; n = 96	12	PIC intervention consisted of two components: 1) a comprehensive interdisciplinary medical and psychosocial assessment (2–3 hours on first visit), and 2) follow-up ambulatory case management for 1 year. Involvement differed by need, but minimum monthly call.	Total cost of services, Mortality, Patient satisfaction, Self-reported health status, Utilisation (primary/secondary care)
Stuck 2000 [99]	Switzerland	low#	In-home visits for disability prevention	RCT; n = 791	36	Annual nurse-led comprehensive assessments. Cases discussed with geriatrician and recommendations developed. In-home follow-up visits every 3 months. Nurses also provided health education, encouraged self-care, and attempted to improve communication with the physician. Interdisciplinary team available to discuss complex patients.	Mortality, Self-reported health status, Utilisation (secondary care)
Sylvia 2008 [100]; Boyd 2008 [101]	USA	low	Community-dwelling, chronically ill, elderly	nRCT; n = 127	6	At-home assessment, evidence-based care plan, promotion of self-management, monthly monitoring, coaching on healthy behaviours, coordination of transitions in care, and facilitating access to community resources.	Total cost of services, Patient satisfaction, Utilisation (primary/secondary care)
Toseland 1996 [102]; Toseland 1997 [103]; Engelhardt 1996 [104]; Engelhardt 2006 [105]	USA	low	Frail elderly	RCT; n = 160	48	Primary functions of the GEM team included: initial comprehensive assessment; development of a care plan; implementation of the care plan; periodic reassessment; monitoring and updating the care plan, and; referral to and coordination with other health and social service providers. Weekly team meetings to discuss.	Total cost of services, Mortality, Patient satisfaction, Self-reported health status, Utilisation (primary/secondary care)
van Hout 2010 [106]	The Netherlands	high	Community-dwelling frail elderly	RCT; n = 651	18	Assessment of health and care needs, recommended interventions based on guidelines, individually tailored care plans (copy left at patient's home for other care workers to see/add to). Home visits at least 4 times a year.	Mortality, Self-reported health status, Utilisation (secondary care)

* Source: Starfield et al 2002 [31], unless otherwise stated

Source: Macinko et al 2003 [107]

+ Source: Grant et al 1997 [108]

^ Source: Fry & Horder 1994 [109]

doi:10.1371/journal.pone.0132340.t004

Table 5. Details of interventions.

Study	Name of case management model	Intensity of intervention (patient contacts)	Risk Assessment Tool (judgement/ threshold/ predictive risk modelling)	MDT or single case manager (primary case manager in bold)	Primary location of case management	24-hour availability of case manager	Caseload per manager/ team	Training received by case manager	Case management reimbursement method
Beland 2006a [57]; Beland 2006b [58]	SIPA [French acronym for System of Integrated Care for Older Persons]	Not clear	**Threshold** *Functional Autonomy Measurement System (SMAF)*	MDT: **nurse/ social worker**, community nurses; occupational therapists, homemakers, staff family physicians, (consultant pharmacists), (community organisers)	Not clear	Yes	35–45	Yes	Family physician offered $400 per SIPA patient in addition to their usual FFS
Bernabei 1998 [59]	Integrated community care	Every 2 months	**Threshold** *previous use of home services*	MDT: **trained case manager**, general practitioner, geriatrician, social worker, nurses	Not clear	Not clear	Not clear	Yes	Not clear
Bird 2010 [60]	HARP [Hospital Admission Risk Programme]	4–7 times in 12 months	**Threshold** *previous hospital use*	MDT: **trained case facilitator**, N/S	Home	Not clear	Not clear	Unclear	Not clear
Boult 2008 [61]; Leff 2009 [62]; Boyd 2010 [63]; Boult 2011 [64]; Boult 2013 [65]	Guided care	Monthly	**Predictive risk modelling** *Hierarchical Condition Category (HCC)*	MDT: **Nurse**, physicians	Not clear	Not clear	50–60	Yes	FFS
Boyd 1996 [66]	Community-based case management	Averaged 4.45 hours per patient per month	**Threshold** *previous secondary care use*	Single: **nurse**	Home	Not clear	Not clear	Unclear	Not clear
Burns 1995 [67]; Burns 2000 [68]	GEM [Geriatric Evaluation and Management]	Not clear	**Threshold** *mixture of criteria judging frailty*	MDT: **GEM team** (physician, nurse, social worker, psychologist)	GEM clinic	Not clear	Not clear	Yes	Not clear

(*Continued*)

Table 5. (*Continued*)

Study	Name of case management model	Intensity of intervention (patient contacts)	Risk Assessment Tool (judgement/ threshold/ predictive risk modelling)	MDT or single case manager (primary case manager in bold)	Primary location of case management	24-hour availability of case manager	Caseload per manager/ team	Training received by case manager	Case management reimbursement method
Coburn 2012 [69]	Community-based nursing intervention	Minimum of monthly. Average 17.4 contacts per patient per year	**Predictive risk modelling** *Sutter Health Questionnaire/ numeric risk score*	Single: **nurse**	Various	Not clear	85–110	Yes	FFS + fixed fee per participant per month
Counsell 2007 [70]; Counsell 2009 [71]	GRACE [Geriatric Resources for Assessment and Care of Elders]	Minimum of monthly	**Threshold** *income level*	MDT: **nurse/ social worker**, geriatrician, pharmacist, physical therapist, mental health social worker, community-based services liaison	Home/ telephone	Not clear	Not clear	Unclear	Not clear
Dalby 2000 [72]	Visiting nurse	Not clear	**Threshold** *Questionnaire (functional impairment/past hospital use)*	Single: **nurse**	Home	Not clear	Not clear	Unclear	Capitation
De Stampa 2014 [73]	COPA [CO-ordinationPersonnesAgées]	Not clear	**Threshold** *Contact Assessment (CA) tool*	MDT: **Nurse**, primary care physician, (geriatrician)	Home	Not clear	40	Yes	Not clear
Dorr 2008 [74]	CMP [Care Management Plus]	Not clear	**Judgement** *clinical judgement*	Single: **nurse**	Not clear	Not clear	Not clear	Yes	Not clear
Enguidanos 2006 [75]	Kaiser Permanente Community Partners	TCM: 4–5 contacts per patient per 4-week period GCM: Approx 20 hours per case over 8–9 months	**Threshold** *functional/ utilisation criteria*	TCM- Single: **social worker** GCM- MDT: **nurse/social worker**, geriatrician, assistant department manager	TCM: Telephone GCM: Home/ telephone	Not clear	Not clear	Unclear	Not clear
Fan 2012 [76]	CCMP [Comprehensive Care Management Program]	Monthly for 3 months. Every 3 months thereafter.	**Threshold** *previous hospital use*	Single: **healthcare professional** (qualification varied by site)	Telephone	No	Not clear	Yes	Not clear

(*Continued*)

Table 5. (*Continued*)

Study	Name of case management model	Intensity of intervention (patient contacts)	Risk Assessment Tool (judgement/ threshold/ predictive risk modelling)	MDT or single case manager (primary case manager in bold)	Primary location of case management	24-hour availability of case manager	Caseload per manager/ team	Training received by case manager	Case management reimbursement method
Fitzgerald 1994 [77]	GMC [General Medicine Clinic] case management	Averaged 1.6 per patient per month	**Threshold** *previous hospital use*	Single: **nurse**	Clinic/ Telephone	Yes	Not clear	Unclear	Salaried nurse
Fordyce 1997 [78]	STAR [Senior Team Assessment and Referral programme]	Not clear	**Threshold** *STAR questionnaire (measuring frailty)*	MDT: **nurse**, geriatrician, health educator, geriatric psychiatrist	Telephone	Not clear	Not clear	Unclear	Not clear
Gagnon 1999 [79]	Community-based nurse case management	Minimum monthly call, and home visit every 6 weeks.	**Predictive risk modelling** *Boult assessment tool (40% or more probability of hospitalisation)*	Single: **nurse**	Home/ telephone	No	40–55	Yes	Not clear
Gravelle 2007 [80]	Evercare	Not clear	**Threshold** *previous emergency admissions*	Single: **nurse**	Not clear	Not clear	Not clear	Unclear	Not clear
Hogg 2009 [81]; Gray 2010 [82]	APTCare [Anticipatory and Preventive Team Care]	Not clear	**Judgement** *clinical judgement*	MDT: **nurse**, pharmacist, usual family physician	Home	Not clear	Not clear	Yes	FFS/ capitation
Kruse 2010 [83]	Nurse care coordination	Not clear	**Threshold** *previous outpatient use*	Single: **nurse**	Clinic/ telephone	Not clear	Not clear	Unclear	Not clear
Leung 2004 [84]	Case Management Project	Once every two weeks.	**Threshold** *previous hospital use*	MDT: **nurse/ social worker**, geriatricians, senior social workers, geriatric nursing specialist, clinical psychologist, rehabilitation therapists	Home/ telephone	Not clear	Not clear	Unclear	Not clear

(*Continued*)

Table 5. (*Continued*)

Study	Name of case management model	Intensity of intervention (patient contacts)	Risk Assessment Tool (judgement/ threshold/ predictive risk modelling)	MDT or single case manager (primary case manager in bold)	Primary location of case management	24-hour availability of case manager	Caseload per manager/ team	Training received by case manager	Case management reimbursement method
Levine 2012 [85]	CHA [Choices for Healthy Aging]	Minimum monthly	**Predictive risk modelling** *electronic risk assessment tool*	MDT: **team** (physician, nurse practitioner, nurse care manager, social worker)	Home/ telephone	Yes	Not clear	Unclear	Not clear
Martin 2004 [86]	Care plans for COPD	Visits at 0, 3, 6, and 12 months.	**Threshold** *previous COPD exacerbations requiring care*	MDT: **nurse**, respiratory specialist, GP	Not clear	Not clear	Not clear	Unclear	Not clear
Metzelthin 2013[87]	PoC [Prevention of Care]	Not clear	**Threshold** *Groningen Frailty Indicator*	MDT: **nurse**, GP, (occupational therapist), (physical therapist), (other health professionals as needed)	Home	Not clear	Not clear	Yes	Not clear
Morishita 1998 [88]; Boult 2001 [89]	GEM [Geriatric Evaluation and Management]	Monthly clinic visits + telephone availability	**Predictive risk modelling** *probability of repeated admission instrument*	MDT: **GEM team** (geriatrician, geriatric nurse practitioner, nurse, social worker)	GEM clinic/ telephone	Yes	Not clear	Unclear	FFS
Newcomer 2004 [90]	ECM [Enhanced Case Management]	Minimum monthly. Weekly until problem resolution. Average 7.7 hours per patient over 12 months.	**Threshold** *presence of chronic conditions (subsequently stratified by risk score obtained from assessment questionnaire)*	Single: **nurse**	Telephone	Not clear	250 (~60 actively case managed at any time)	Unclear	Not clear
Ploeg 2010 [91]	Preventive primary care outreach	Minimum 3 yearly visits + follow-up phone calls/ home visits.	**Threshold** *Sherbrooke postal questionnaire (assessing risk of functional decline)*	Single: **nurse**	Home/ telephone	Not clear	Not clear	Unclear	Capitation-based that includes some FFS
Rodenas 2008 [92]	Case management Valencia	Minimum once every 2 months.	**Judgement** *referral protocol of social and health cases*	MDT: **team** (physician, nurse, social worker)	Not clear	Not clear	Not clear	Yes	Not clear

(*Continued*)

Table 5. (*Continued*)

Study	Name of case management model	Intensity of intervention (patient contacts)	Risk Assessment Tool (judgement/ threshold/ predictive risk modelling)	MDT or single case manager (primary case manager in bold)	Primary location of case management	24-hour availability of case manager	Caseload per manager/ team	Training received by case manager	Case management reimbursement method
Rubenstein 2007 [93]	Screening, case finding, referral	One month after first contact. Every 3 months thereafter.	Threshold *Geriatric Postal Screening Survey*	Single **Physician assistant**	Telephone	Not clear	Not clear	Unclear	Not clear
Schraeder 2001 [94]	Collaborative primary care nurse case management	Average 8 contacts per patient per year.	**Judgement/ Threshold** *clinical judgement/ presence of determined risk factors*	MDT: **nurse/ case assistant**, primary care physician	Various	Not clear	Not clear	Unclear	Not clear
Schraeder 2008 [95]	Collaborative primary care nurse case management	Minimum monthly	**Threshold** *health screening questionnaire*	MDT: **nurse**, case assistant, primary care physician	Various	Not clear	Not clear	Unclear	Not clear
Shannon 2006 [96]; Alkema 2007 [97]	Care Advocate Programme	Minimum monthly	**Predictive risk modelling** *health care utilisation algorithm*	Single: **social worker**	Telephone	Not clear	Not clear	Unclear	Not clear
Sledge 2006 [98]	PIC [Primary Intensive Care]	Minimum monthly	**Threshold** *previous hospital use*	MDT: **psychiatric nurse**, social worker, psychiatrist, general internist	Telephone	Not clear	21	Unclear	Not clear
Stuck 2000 [99]	In-home visits for disability prevention	Every 3 months.	**Threshold** *Scoring on 6 criteria generated from the literature*	Single: **nurse**	Home	Not clear	Not clear	Yes	Not clear
Sylvia 2008 [100]; Boyd 2008 [101]	Guided care	Minimum monthly	**Predictive risk modelling** *Adjusted Clinical Groups Predictive Model*	MDT: **nurse**, primary care physician	Not clear	Not clear	50–60	Yes	Capitated insurance system
Toseland 1996 [102]; Toseland 1997 [103]; Engelhardt 1996 [104]; Engelhardt 2006 [105]	GEM [Geriatric Evaluation and Management]	Not clear	**Threshold** *previous outpatient use + functional impairments*	MDT: **nurse**, geriatrician, social worker	GEM clinic	Not clear	Not clear	Unclear	Not clear

(*Continued*)

Table 5. (*Continued*)

Study	Name of case management model	Intensity of intervention (patient contacts)	Risk Assessment Tool (judgement/ threshold/ predictive risk modelling)	MDT or single case manager (primary case manager in bold)	Primary location of case management	24-hour availability of case manager	Caseload per manager/ team	Training received by case manager	Case management reimbursement method
van Hout 2010 [106]	Nurse home visits	Minimum 4 visits per patient per year	**Threshold** *frailty score (COOP-WONCA charts)*	Single: **nurse**	Home	Not clear	Not clear	Yes	Not clear

doi:10.1371/journal.pone.0132340.t005

Table 6. Quality of included studies.

Study	Was the allocation sequence adequately generated?	Was the allocation adequately concealed?	Were baseline outcome measurements similar?	Were baseline characteristics similar?	Were incomplete outcome data adequately addressed?	Was knowledge of the allocated interventions adequately prevented during the study?	Was the study adequately protected against contamination?	Was the study free from selective outcome reporting?	Was the study free from other risks of bias?	Criteria met
Beland 2006a [57]; Beland 2006b [58]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Bernabei 1998 [59]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Bird 2010 [60]	No	No	No	No	Yes	Yes	Yes	Yes	No	**4**
Boult 2008 [61]; Leff 2009 [62]; Boyd 2010 [63]; Boult 2011 [64]; Boult 2013 [65]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Boyd 1996 [66]	No	Unclear	Yes	Yes	Unclear	Yes	Unclear	Yes	Unclear	**4**
Burns 1995 [67]; Burns 2000 [68]	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	**8**
Coburn 2012 [69]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Counsell 2007 [70]; Counsell 2009 [71]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Dalby 2000 [72]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	**8**
De Stampa 2014 [73]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**7**
Dorr 2008 [74]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	**7**
Enguidanos 2006 [75]	Unclear	Unclear	Yes	Yes	Unclear	Unclear	Unclear	Unclear	Yes	**3**
Fan 2012 [76]	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	**8**
Fitzgerald 1994 [77]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**8**
Fordyce 1997 [78]	Unclear	Unclear	No	No	Yes	Yes	Unclear	Yes	Yes	**4**
Gagnon 1999 [79]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	**8**
Gravelle 2007 [80]	No	No	Yes	No	Unclear	Yes	Yes	Yes	Yes	**5**

(*Continued*)

Table 6. (*Continued*)

Study	Was the allocation sequence adequately generated?	Was the allocation adequately concealed?	Were baseline outcome measurements similar?	Were baseline characteristics similar?	Were incomplete outcome data adequately addressed?	Was knowledge of the allocated interventions adequately prevented during the study?	Was the study adequately protected against contamination?	Was the study free from selective outcome reporting?	Was the study free from other risks of bias?	Criteria met
Hogg 2009 [81]; Gray 2010 [82]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Kruse 2010 [83]	No	Yes	Unclear	No	Yes	Yes	Yes	Yes	Yes	**6**
Leung 2004 [84]	Unclear	Unclear	No	Yes	Unclear	Yes	Unclear	Yes	Yes	**4**
Levine 2012 [85]	Yes	Unclear	Unclear	Yes	Yes	Yes	Unclear	Yes	Yes	**6**
Martin 2004 [86]	Unclear	Unclear	Unclear	No	Unclear	Yes	Unclear	Yes	Yes	**3**
Metzelthin 2013[87]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Morishita 1998 [88]; Boult 2001 [89]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	**8**
Newcomer 2004 [90]	Unclear	Unclear	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	**6**
Ploeg 2010 [91]	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	**8**
Rodenas 2008 [92]	Yes	Yes	No	No	Yes	Yes	Unclear	Yes	Yes	**6**
Rubenstein 2007 [93]	Yes	Yes	Yes	Unclear	No	Yes	Yes	Yes	Yes	**7**
Schraeder 2001 [94]	Unclear	Yes	No	No	Yes	Yes	Yes	Yes	Yes	**6**
Schraeder 2008 [95]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**8**
Shannon 2006 [96]; Alkema 2007 [97]	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**8**
Sledge 2006 [98]	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	**8**
Stuck 2000 [99]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**9**
Sylvia 2008 [100]; Boyd 2008 [101]	No	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	**7**

(*Continued*)

Table 6. (*Continued*)

Study	Was the allocation sequence adequately generated?	Was the allocation adequately concealed?	Were baseline outcome measurements similar?	Were baseline characteristics similar?	Were incomplete outcome data adequately addressed?	Was knowledge of the allocated interventions adequately prevented during the study?	Was the study adequately protected against contamination?	Was the study free from selective outcome reporting?	Was the study free from other risks of bias?	Criteria met
Toseland 1996 [102]; Toseland 1997 [103]; Engelhardt 1996 [104]; Engelhardt 2006 [105]	Unclear	Unclear	Unclear	Yes	Yes	Yes	Unclear	Yes	Yes	**5**
van Hout 2010 [106]	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	**8**

doi:10.1371/journal.pone.0132340.t006

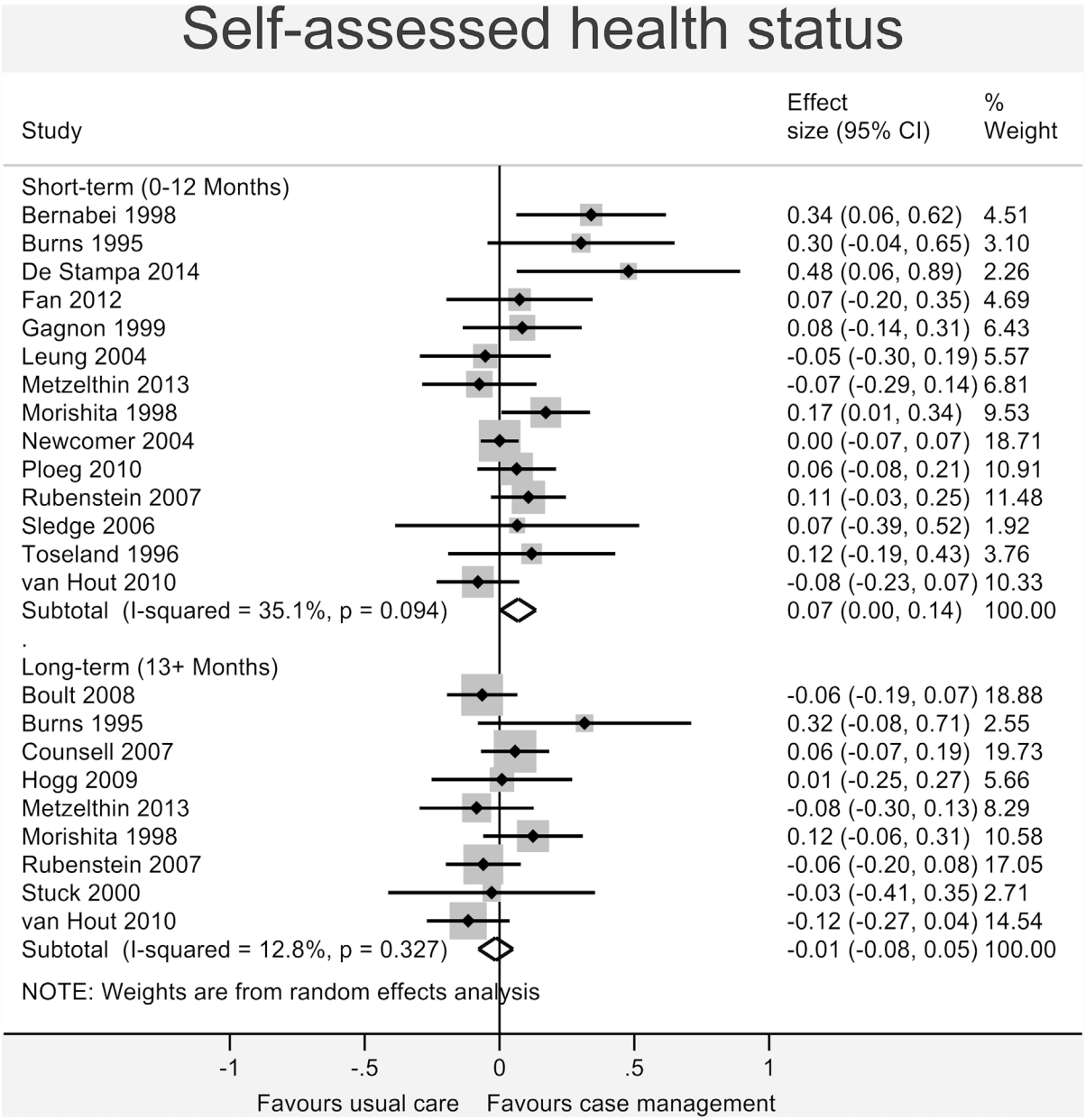


Fig 3. Forrest plot for self-assessed health status outcome. Effect estimates are the standardised mean difference, where the solid vertical line at 0 indicates no effect. Effect estimates are based on a random-effects model. Each subtotal shows the overall effect estimate for the time-period indicated.

doi:10.1371/journal.pone.0132340.g003

Cost. No significant effect was found for *total cost of services* (short-term: -0.00, 95% CI -0.07 to 0.06, I^2 = 0.0%, p = 0.784; long-term: -0.03, 95% CI -0.16 to 0.10, I^2 = 46.0%, p = 0.116 –Fig 5), *utilisation of primary and non-specialist care* (short-term: -0.08, 95% CI -0.22 to 0.05, I^2 = 79.2%, p<0.001; long-term: -0.10, 95% CI -0.29 to 0.09, I^2 = 78.6%, p<0.001 –Fig 6) or *secondary care* (short-term: 0.04, 95% CI -0.02 to 0.10, I^2 = 39.6%, p = 0.027; long-term: -0.02, 95% CI -0.08 to 0.04, I^2 = 22.8%, p = 0.194 –Fig 7).

Satisfaction. *Patient satisfaction* (Fig 8) showed a statistically significant beneficial effect in the case management group in the short-term (0.26, 95% CI 0.16 to 0.36, I^2 = 0.0%, p = 0.465), increasing in the long-term (0.35, 95% CI 0.04 to 0.66, I^2 = 88.3%, p<0.001).

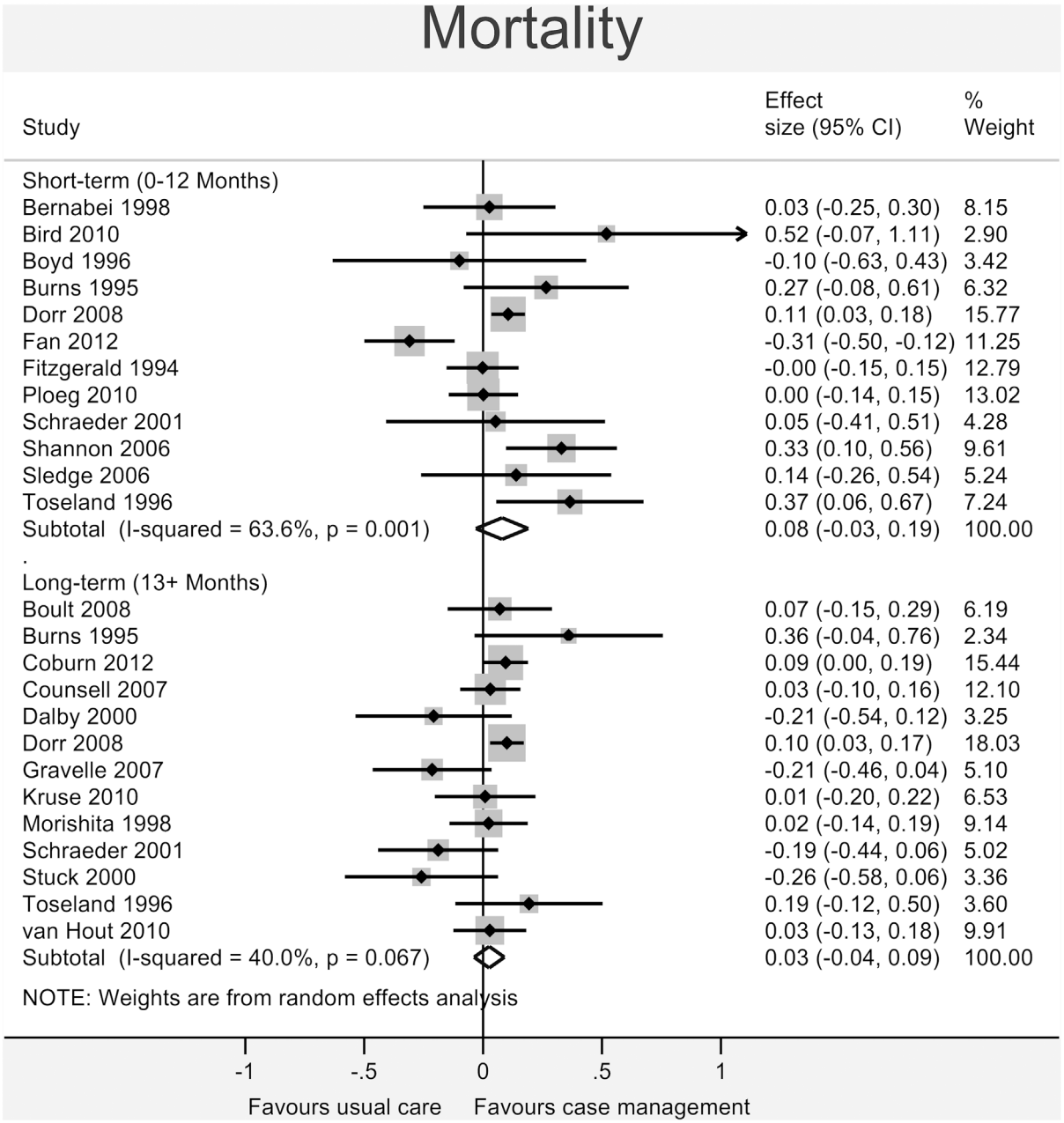


Fig 4. Forrest plot for mortality outcome. Effect estimates are the standardised mean difference, where the solid vertical line at 0 indicates no effect. Effect estimates are based on a random-effects model. Each subtotal shows the overall effect estimate for the time-period indicated.

doi:10.1371/journal.pone.0132340.g004

Heterogeneity, measured with the I^2 statistic, varied by outcome and time-period measured. Those with particularly high I^2 (over 75% [47]), included *utilisation of primary and non-specialist care* (short- and long-term), and *patient satisfaction* (long-term).

The funnel plots showed a fairly even distribution of small studies, suggesting no small study bias. The one exception was for *self-assessed health status*, which appeared slightly skewed towards favourable results for the intervention in smaller studies. However, results of the Egger test found no statistically significant small-study effects across any of the outcomes assessed.

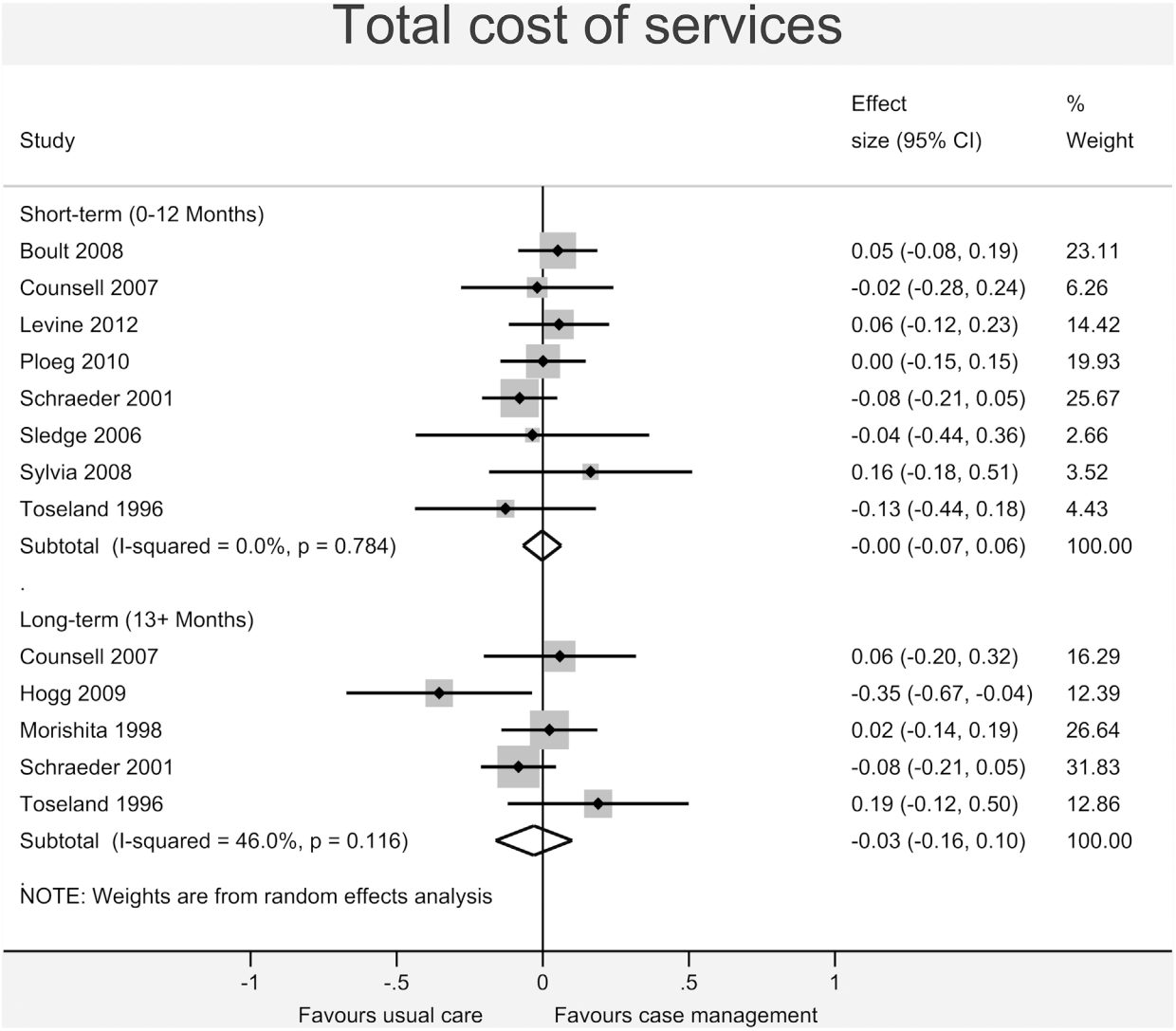


Fig 5. Forrest plot for total cost of services outcome. Effect estimates are the standardised mean difference, where the solid vertical line at 0 indicates no effect. Effect estimates are based on a random-effects model. Each subtotal shows the overall effect estimate for the time-period indicated.

doi:10.1371/journal.pone.0132340.g005

Subgroup analyses

The following outcome categories met the minimum criteria of 10 studies contributing to the primary analysis: *mortality* (short-, and long-term), *self-assessed health status* (short-term), *utilisation of primary and non-specialist care* (short-term), and *utilisation of secondary care* (short-, and long-term).

The results for each of the subgroup analyses are summarised in Table 7, below (the forest plots can be found in S2 Appendix).

Power to determine differences in subgroup analyses is limited, the large number of comparisons risks inflating rates of Type I error, and there may be other differences between studies that have not been taken into account in these univariate comparisons. Therefore, these results should be treated with appropriate caution. When interpreting subgroup effects, significant difference between subgroups is the important comparative factor. Importantly, no statistically significant differences were found when comparing between subgroups.

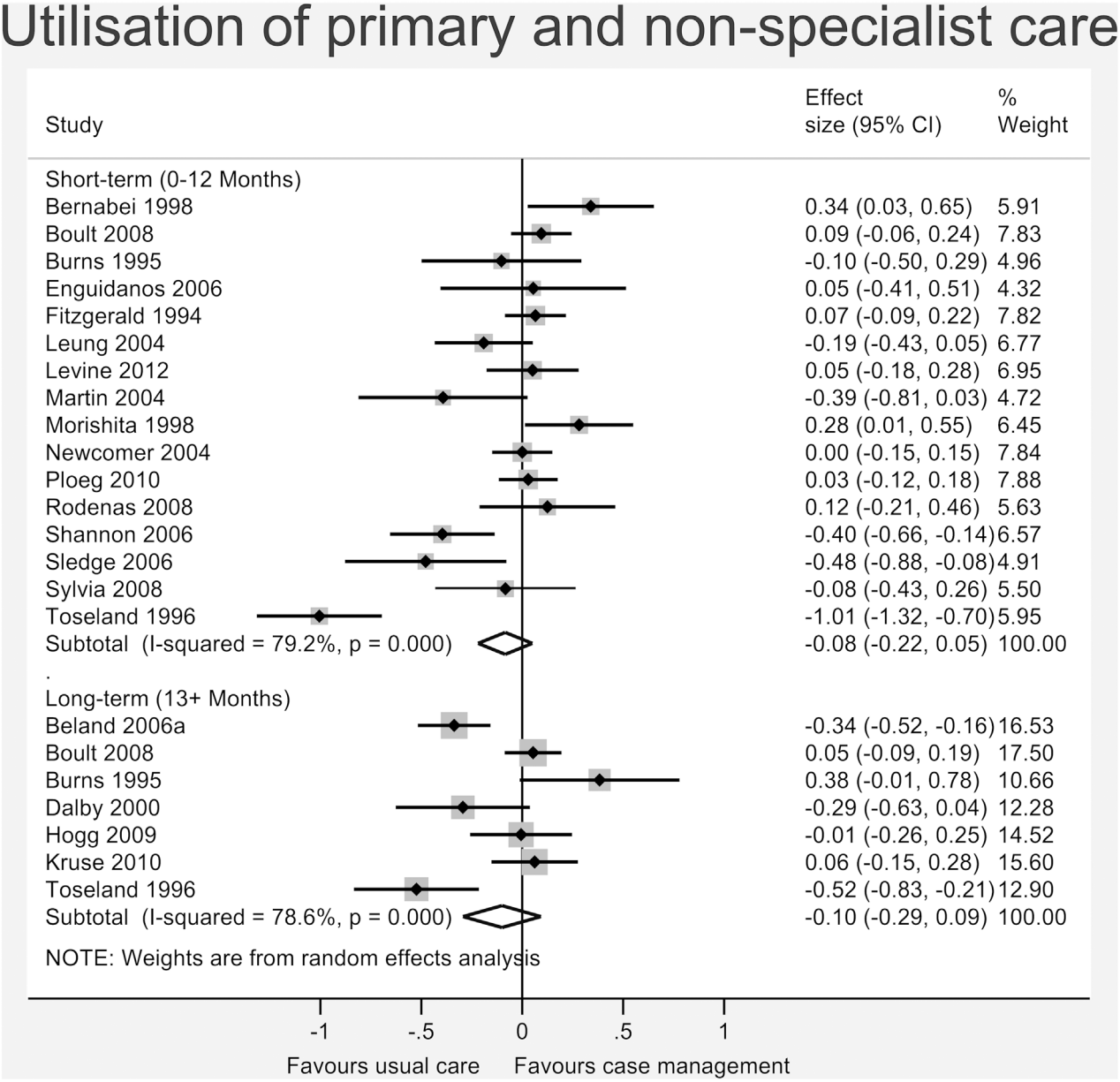


Fig 6. Forrest plot for utilisation of primary and non-specialist care outcome. Effect estimates are the standardised mean difference, where the solid vertical line at 0 indicates no effect. Effect estimates are based on a random-effects model. Each subtotal shows the overall effect estimate for the time-period indicated.

doi:10.1371/journal.pone.0132340.g006

However, results perhaps indicate slightly beneficial effects of delivery of case management by an MDT, with the inclusion of a social worker, and in settings with low strength of primary care. These preliminary findings may merit further investigation. Nevertheless, any significant within-subgroup effects found were extremely small by Cohen's interpretation.

Sensitivity Analysis and Multiple comparisons

Those studies at highest risk of bias reported findings in the short-term (0–12 months) for *utilisation of primary and non-specialist care* and *utilisation of secondary care* [75, 86]. Studies

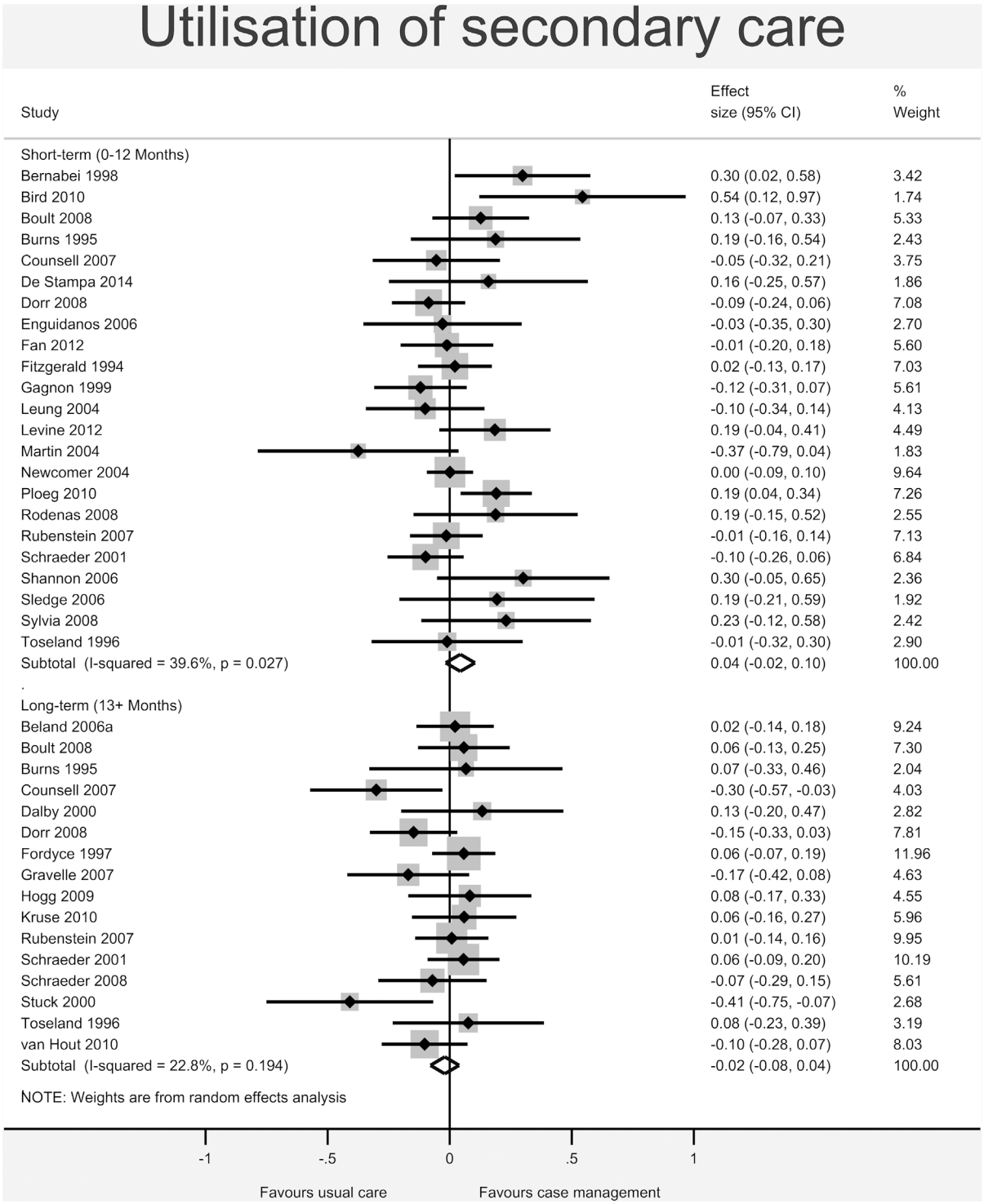


Fig 7. Forrest plot for utilisation of secondary care outcome. Effect estimates are the standardised mean difference, where the solid vertical line at 0 indicates no effect. Effect estimates are based on a random-effects model. Each subtotal shows the overall effect estimate for the time-period indicated.

doi:10.1371/journal.pone.0132340.g007

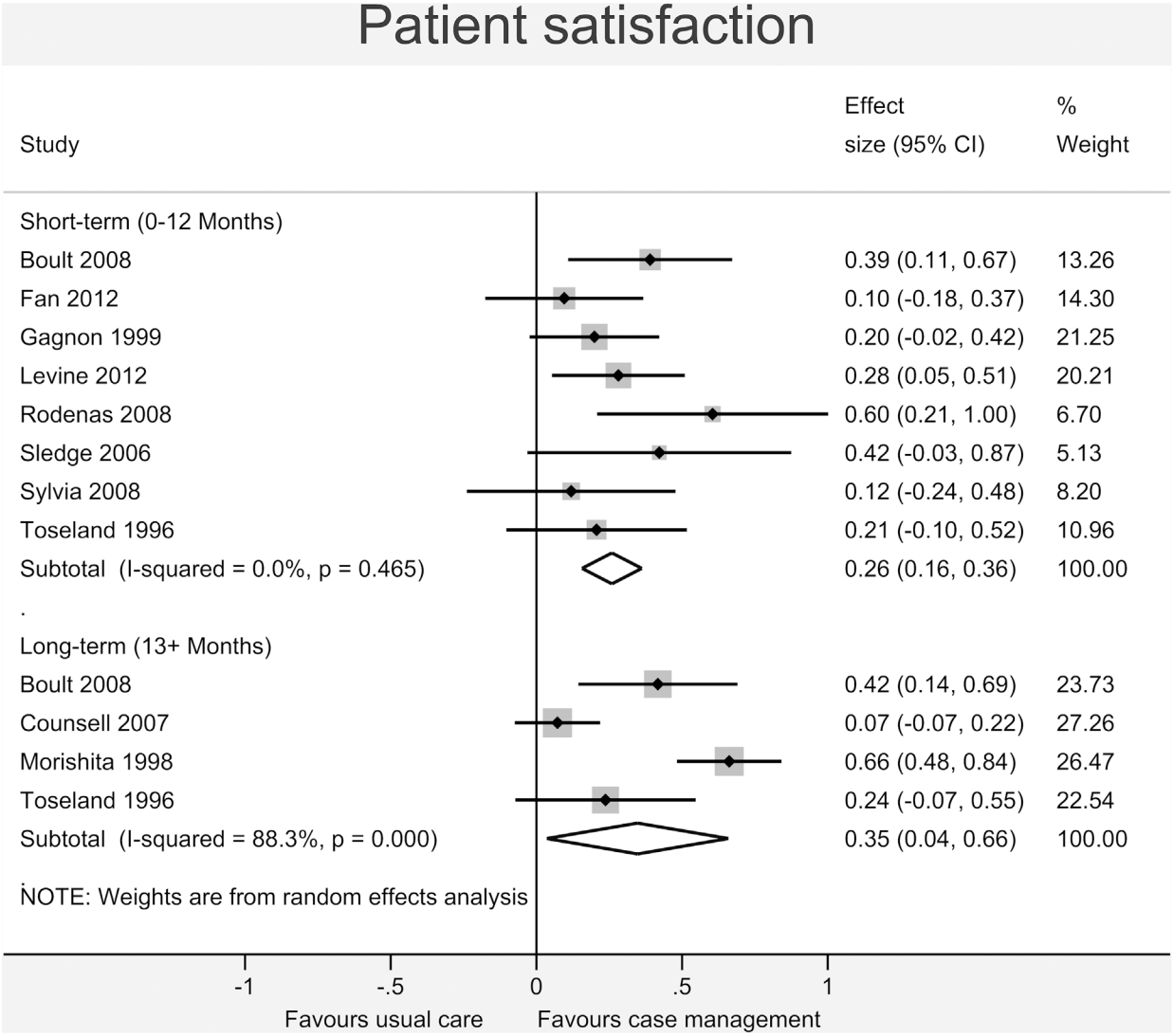


Fig 8. Forrest plot for patient satisfaction outcome. Effect estimates are the standardised mean difference, where the solid vertical line at 0 indicates no effect. Effect estimates are based on a random-effects model. Each subtotal shows the overall effect estimate for the time-period indicated.

doi:10.1371/journal.pone.0132340.g008

using Veteran participants [67, 76, 77, 93, 102], with over 90% males, reported findings in all outcomes and time-periods assessed.

After adjusting for multiple comparisons, excluding these studies showed no significant difference from the results reported above, either for the primary analysis, or between subgroup differences for the subgroup analyses. The results of the sensitivity analysis can be found in S3 Appendix.

After Holm-Bonferroni correction was applied to all results, only two of the statistically significant results held: the finding of a significant effect on *patient satisfaction* in the short-term (0–12 months) in the primary analysis, and the same outcome measure in the sensitivity analysis (excluding studies with Veteran participants).

Table 7. Results of subgroup analyses. No significant differences between subgroups (p<0.05). *Note*: Positive effect size favours case management for all measures.

Outcome (time-period)	Subgroup effect size(number of studies)	
	MDT (21)	**Single** (15)
Mortality (short)	0.20 (0.05 to 0.35)*(6)	0.01 (-0.13 to 0.16)(6)
Mortality (long)	0.04 (-0.06 to 0.14)(6)	0.01 (-0.08 to 0.10)(7)
Self-rated health (short)	0.14 (0.01 to 0.27)*(8)	0.02 (-0.03 to 0.07)(6)
Utilisation primary care (short)	-0.10 (-0.30 to 0.10)(12)	-0.04 (-0.20 to 0.11)(4)
Utilisation secondary care (short)	0.08 (-0.02 to 0.17)(15)	0.01 (-0.06 to 0.09)(8)
Utilisation secondary care (long)	0.02 (-0.04 to 0.09)(9)	-0.08 (-0.18 to 0.03)(7)
	Low PHC score (23)	**Int/high PHC score** (13)
Mortality (short)	0.09 (-0.05 to 0.23)(9)	0.05 (-0.13 to 0.23)(3)
Mortality (long)	0.05 (-0.01 to 0.12)(10)	-0.10 (-0.27 to 0.08)(3)
Self-rated health (short)	0.11 (0.02 to 0.20)*(8)	0.03 (-0.08 to 0.13)(6)
Utilisation primary care (short)	-0.12 (-0.30 to 0.06)(11)	-0.00 (-0.20 to 0.20)(5)
Utilisation secondary care (short)	0.01 (-0.03 to 0.06)(16)	0.08 (-0.10 to 0.26)(7)
Utilisation secondary care (long)	-0.02 (-0.10 to 0.05)(11)	-0.02 (-0.12 to 0.07)(5)
	Clinical Judgement (4)	**Risk modelling** (32)
Mortality (short)	0.10 (0.03 to 0.17)*(2)	0.09 (-0.06 to 0.24)(10)
Mortality (long)	-0.02 (-0.30 to 0.26)(2)	0.02 (-0.05 to 0.09)(11)
Self-rated health (short)	n/a	n/a
Utilisation primary care (short)	n/a	n/a
Utilisation secondary care (short)	-0.06 (-0.18 to 0.06)(3)	0.06 (-0.00 to 0.13)(20)
Utilisation secondary care (long)	-0.01 (-0.15 to 0.14)(3)	-0.02 (-0.09 to 0.04)(13)
	RCT (28)	**Non-RCT** (8)
Mortality (short)	0.07 (-0.07 to 0.22)(9)	0.12 (-0.06 to 0.30)(3)
Mortality (long)	0.03 (-0.05 to 0.10)(10)	-0.00 (-0.18 to 0.17)(3)
Self-rated health (short)	n/a	n/a
Utilisation primary care (short)	n/a	n/a
Utilisation secondary care (short)	0.04 (-0.02 to 0.10)(19)	0.17 (-0.11 to 0.45)(4)
Utilisation secondary care (long)	-0.00 (-0.07 to 0.07)(12)	-0.08 (-0.19 to 0.02)(4)
	Social worker (12)	**No social worker** (24)
Mortality (short)	0.24 (0.10 to 0.37)*(5)	-0.01 (-0.14 to 0.13)(7)
Mortality (long)	0.07 (-0.04 to 0.17)(4)	-0.00 (-0.09 to 0.08)(9)
Self-rated health (short)	0.15 (0.04 to 0.27)*(6)	0.03 (-0.04 to 0.10)(8)
Utilisation primary care (short)	-0.13 (-0.38 to 0.12)(10)	0.03 (-0.05 to 0.10)(6)
Utilisation secondary care (short)	0.10 (0.00 to 0.20)(10)	0.02 (-0.06 to 0.09)(13)
Utilisation secondary care (long)	-0.04 (-0.21 to 0.13)*(4)	-0.02 (-0.08 to 0.05)(12)

* = significant in-subgroup effect (p<0.05)

doi:10.1371/journal.pone.0132340.t007

Discussion

Summary of the key findings

Case management of 'at-risk' patients in primary care has been promoted as a way of reducing health system pressures, and the most recent iteration of the UK GP contract has provided incentives for its delivery. This evidence identified by this review does not provide strong evidence to suggest that case management is an effective way of alleviating pressure on a health system. *Total cost of care*, and *utilisation of secondary care* services do not appear to be

significantly affected by case management. There may be a significant effect on *self-reported health status* with case management. However, the magnitude of the benefit is very modest, does not meet conventional criteria even for a 'small' effect, and was not significant after adjustment for multiple comparisons. Case management does improve *patient satisfaction* when compared to usual care. This is a legitimate outcome for a 'patient-centred' health care system, but is rarely seen as the primary aim of case management interventions.

Strengths and limitations

Strengths of this study include the use of PRISMA guidelines, pre-specification of subgroups, as well as the broad search strategy. Unfortunately, the broad search impaired our ability to double-screen all studies at every stage, although we did double-screen a proportion at every stage, and our inter-rater reliability was consistently good. We did not include grey literature, due to the generally lower quality of this literature [110]. We found no evidence of small study bias in our included studies.

Assessing complex service-level interventions is difficult, and RCTs may be particularly problematic in the context of patients with multimorbidity [111]. We included the range of intervention study types considered by 'The Cochrane Effective Practice and Organisation of Care (EPOC) Group'.

We view the use of meta-analysis as a major strength of this piece of work, which differentiates this review from the narrative syntheses [17, 22–24, 26–28, 30]. Some argue that meta-analysis of complex service-level interventions is inappropriate, because the effects of the intervention are so dependent on context [40], and pooling the results from different contexts is not advisable. However, as shown in the introduction, case management can be defined in terms of a number of common components. In addition, we did try to account for context differences (such as strength of the primary care system), although the precise scope of the term is unclear [112], and a lack of consistent reporting limited what was possible.

Heterogeneity was high for measures of *utilisation of primary and non-specialist care* in both time-periods assessed, and *patient satisfaction* in the long-term. This high level of heterogeneity is expected in analysis of a complex intervention, which is possibly highly dependent on context. On the whole, choosing a random effects model took into account expected heterogeneity arising from comparison of a complex intervention across different settings [113]. Nevertheless, caution must be applied to uncritical interpretation of the pooled effect, due to the level of unexplained variation observed.

When we adjusted for multiple comparisons, only increased *patient satisfaction* in the short-term remained significant. This type of adjustment, while it reduces the risk of false positive findings (type I error), does so at the risk of inflating the number of false negative findings (type II error) [114]. As an intervention with low risk of harm to the patient, we have chosen to present the unadjusted results as the primary analyses, with the results adjusted for multiple comparisons suggesting additional caution in interpretation.

The outcome measures we chose were broadly inclusive. For example, in *self-assessed health status* we included activities of daily living, as well as bed days, and more typical 'health' measures, for instance QALYs. This could be a potential weakness of this study. However, we chose these broad outcome categories attempting to synthesise as much of the relevant data as possible that were reported within the selected studies. Furthermore, these measures were reported as functional outcome measures of health in the individual studies, and were therefore synthesised as such.

Utilisation and cost outcomes have a tendency to be skewed. As expected, the studies we synthesised reporting these outcomes demonstrated significant skew (i.e. the mean is smaller

than twice the standard deviation), indicating that the mean reported is not a good indicator of the centre of the distribution [115]. Future primary studies should make sure these skews are reported, and that the effects of any subsequent log transformation are detailed for more precise synthesis of these outcomes. Furthermore, although costs were detailed in a number of studies, we identified only one cost-effectiveness analysis [82], and one cost-benefit analysis [84].

Interpretation of the results in the context of other studies

It is difficult to directly compare, as most previous reviews on this subject have used narrative synthesis methods [17, 23, 24], or used 'vote counting' to quantify the number of studies with statistically significant results in either direction [22, 26–28, 30]. The majority of existing reviews conclude that despite theoretical benefits, in practice there is only slight evidence of benefits [22, 23], particularly related to patient satisfaction [24, 27], and functional health [26]. The single previous systematic review we identified which employed meta-analysis, additionally included hospital discharge planning interventions (identifying a total of eleven studies, only six of which—three in the primary care setting—were included in meta-analysis), and only used meta-analysis for a single outcome category, 'unplanned hospital admissions', similarly finding no significant effect [25]. Our results are in line with those previous reviews, with the additional benefits of updating the evidence base, quantifying the impacts (emphasising the *small* benefits) across a range of outcome categories, and exploring contextual variations.

Most published reviews focused on implementation have similarly identified inadequate reporting of the methods of applying case management in practice as a major limitation of the literature [17].

Implications for research

Case management has potential to impact on patient safety issues in primary care, such as co-ordination and communication between professionals and levels of the health system [116]. No safety outcomes were identified in the included literature, and primary care patient safety is a notoriously under-researched area [117].

As multimorbid patients are likely at most risk for co-ordination failures, and therefore potentially have most to gain from the integration of care, measures of multimorbidity must be more consistently reported [118], even if this is a simple count of mean number of chronic diseases. Ideally, however, this would give more detailed breakdown by disease type/cluster [119] as a subgroup analysis, enabling further targeting of specific interventions to specific groups of patients who are most likely to benefit [120]. Additionally, there is some evidence that the coexistence of physical and mental health problems could lead to increased management difficulties [121]. Comorbidity of conditions should be better reported in evaluations and explored in further research.

Current evidence comes from a majority of high-income, Western settings. This potential bias requires addressing with evidence from other settings, for example Asia, where the case management approach is currently evolving.

Implications for policy and practice

Given the lack of significant effects across the majority of outcome categories, should case management generally be encouraged or incentivised for the treatment of 'at-risk' patients in primary care? This review would suggest that, as currently delivered, case management should not be regarded as a primary means of reducing overall health service utilisation and that it will not reduce costs or improve health outcomes. While we have shown some statistically

significant benefits, these are not focused on primary outcomes, with the largest overall effect on satisfaction, which did not meet the usual criterion for a 'medium' effect [36]. However, the current results rest on the evidence accumulated from RCTs. There are potential problems associated with this study design in the assessment of complex interventions and conditions [111], although other designs which may be better able to reflect routine delivery of case management (such as controlled before and after designs [80]) have their own problems with internal validity.

Evidence from the subgroup analyses do perhaps point to more effective ways of delivering the intervention, namely: delivery by a MDT as opposed to a single case manager, and the inclusion of a social worker. These findings agree with the wider literature which advocates the use of a multidisciplinary team to successfully manage patients with chronic disease [42], and advocates better integration of health and social care [45]. Case management may be more effective in a system where the strength of primary health care orientation is low. However, these subgroup results should be interpreted with caution, as they are exploratory univariate analyses, which should be investigated further while controlling for potential confounding factors before firm conclusions are drawn. Furthermore, the significance of these effects did not withstand adjustment for multiple comparisons.

Further understanding of factors driving the effectiveness of case management may benefit from on-going evaluation of implementation at the local level. It is important that components of implementation are reported consistently and in detail, so that these can be included in future systematic reviews and effectiveness of individual elements of the intervention can be examined.

Conclusions

Current evidence suggests case management of 'at-risk' patients in primary care is not effective beyond small improvements in patient satisfaction. Case management should not be regarded as a proven technology in the delivery of integrated care, there remains a need for further enhancement and evaluation of its effectiveness, particularly with study designs which better incorporate context, and in lower income settings. More research is needed into more effective methods of delivery (e.g. by an MDT and including a social worker), and implementation (e.g. in a health system with poor primary care orientation), which may additionally improve effectiveness. Even with these improvements, however, case management may never be as effective as it needs to be to deliver major savings through a focus on high risk groups [122]. This highlights the need for a variety of models to deal with system pressures, including integrated care at different levels of the health care system, and with more of focus on the wider population of patients [123].

Supporting Information

S1 Appendix. Full MEDLINE search strategy.
(DOCX)

S2 Appendix. Forest plots for subgroup analyses.
(DOCX)

S3 Appendix. Results of the sensitivity analyses.
(DOCX)

Acknowledgments

Dr Søren Rud Kristensen for his feedback and support during the writing-up phase. Levine et al, for supplying the tables from their paper not available in the online version.

Author Contributions

Conceived and designed the experiments: JS PB KC SCS. Performed the experiments: JS MP RA. Analyzed the data: JS MP RA. Wrote the paper: JS MP RA KC SCS PB.

References

1. Nolte E, McKee M. Caring for people with chronic conditions: a health system perspective: McGraw-Hill International; 2008.
2. Gröne O, Garcia-Barbero M. Integrated care: a position paper of the WHO European office for integrated health care services. International Journal of Integrated Care. 2001; 1.
3. Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications—a discussion paper. Int J Integr Care. 2002; 2:e12. Epub 2006/08/10. PMID: 16896389; PubMed Central PMCID: PMC1480401.
4. Ham C, Walsh N. Making integrated care happen at scale and pace. London: The King's Fund, 2013.
5. Atun R, Aydin S, Chakraborty S, Sümer S, Aran M, Gürol I, et al. Universal health coverage in Turkey: enhancement of equity. The Lancet. 2013; 382(9886):65–99.
6. Ross S, Curry N, Goodwin N. Case Management: What it is and how it can best be implemented: King's Fund; 2011.
7. Powers BW, Chaguturu SK, Ferris TG. Optimizing high-risk care management. Jama. 2015. doi: 10.1001/jama.2014.18171
8. Case Management Society of America. Definition of Case Managagement 2007 [cited 2014 19/11/2014]. Available: http://www.cmsa.org.
9. Goodwin N, Sonola L, Thiel V, Kodner DL. Co-ordinated care for people with complex chronic conditions. London: The King's Fund, 2013.
10. Georghiou T, Steventon A, Billings J, Blunt I, Lewis G, Bardsley M. Predictive risk and health care: an overview. Evidence for better health care [Internet]. 2011. Available: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/Predictive-risk-and-health-care-an-overview_0.pdf.
11. Lewis G, Curry N, Bardsley M. Choosing a predictive risk model: a guide for commissioners in England. London: Nuffield Trust. 2011.
12. Starfield B, Shi L, Macinko J. Contribution of Primary Care to Health Systems and Health. The Milbank quarterly. 2005; 83(3):457–502. doi: 10.1111/j.1468-0009.2005.00409.x PMC2690145. PMID: 16202000
13. Monitor. Closing the NHS funding gap: how to get better value health care for patients 2013.
14. Shi L. The Impact of Primary Care: A Focused Review. Scientifica. 2012; 2012:22. doi: 10.6064/2012/432892
15. van den Akker M, Buntinx F, Metsemakers JFM, Roos S, Knottnerus JA. Multimorbidity in General Practice: Prevalence, Incidence, and Determinants of Co-Occurring Chronic and Recurrent Diseases. Journal of Clinical Epidemiology. 1998; 51(5):367–75. doi: 10.1016/S0895-4356(97)00306-5 PMID: 9619963
16. BMA. General practice contract changes 2014–2015: Unplanned admissions Enhanced Service 2014 [cited 2014 12/11/2014]. Available: http://bma.org.uk/working-for-change/negotiating-for-the-profession/bma-general-practitioners-committee/general-practice-contract/unplanned-admissions-2014.
17. Challis D, Hughes J, Berzins K, Reilly S, Abell J, Stewart K. Self-care and case management in long-term conditions: the effective management of critical interfaces. Report for the National Institute for Health Research Service Delivery and Organisation programme, accessible at http://www.sdo nihr ac uk/files/project/201-final-report pdf. 2010.
18. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012; 380(9836):37–43.

19. Naylor MD, Brooten D, Campbell R, Jacobsen BS, Mezey MD, Pauly MV, et al. Comprehensive discharge planning and home follow-up of hospitalized elders: A randomized clinical trial. Jama. 1999; 281(7):613–20. doi: 10.1001/jama.281.7.613 PMID: 10029122

20. Cochrane Collaboration. LMIC Filters: Norwegian Satellite of the Cochrane Effective Practice and Organisation of Care Group; 2012 [cited 2014 20/4/2014]. Available: https://epocoslo.cochrane.org/lmic-filters.

21. Ovid SP. Expert Searches—Primary Care Filter 2014 [cited 2014 20/4/2014]. Available: http://ovidsp.uk.ovid.com/sp-3.13.1a/ovidweb.cgi.

22. Eklund K, Wilhelmson K. Outcomes of coordinated and integrated interventions targeting frail elderly people: a systematic review of randomised controlled trials. Health & social care in the community. 2009; 17(5):447–58. Epub 2009/02/28. doi: 10.1111/j.1365-2524.2009.00844.x PMID: 19245421.

23. Ferguson JA, Weinberger M. Case management programs in primary care. Journal of General Internal Medicine. 1998; 13(2):123–6. PMID: 9502373

24. Hickam DH, Weiss JW, Guise J-M, Buckley D, Motu'apuaka M, Graham E, et al. Outpatient case management for adults with medical illness and complex care needs. 2013.

25. Huntley AL, Thomas R, Mann M, Huws D, Elwyn G, Paranjothy S, et al. Is case management effective in reducing the risk of unplanned hospital admissions for older people? A systematic review and meta-analysis. Family practice. 2013; 30(3):266–75. doi: 10.1093/fampra/cms081 PMID: 23315222

26. Hutt R, Rosen R, McCauley J. Case-managing Long-term Conditions: What impact does it have in the treatment of older people? London: The King's Fund, 2004.

27. Latour CH, van der Windt DA, de Jonge P, Riphagen II, de Vos R, Huyse FJ, et al. Nurse-led case management for ambulatory complex patients in general health care: a systematic review. Journal of psychosomatic research. 2007; 62(3):385–95. Epub 2007/02/28. doi: 10.1016/j.jpsychores.2006.10.015 PMID: 17324690.

28. Oeseburg B, Wynia K, Middel B, Reijneveld SA. Effects of case management for frail older people or those with chronic illness: a systematic review. Nursing research. 2009; 58(3):201–10. Epub 2009/05/19. doi: 10.1097/NNR.0b013e3181a30941 PMID: 19448524.

29. Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T. Managing patients with multimorbidity: systematic review of interventions in primary care and community settings. BMJ: British Medical Journal. 2012; 345.

30. You E, Dunt D, Doyle C, Hsueh A. Effects of case management in community aged care on client and carer outcomes: a systematic review of randomized trials and comparative observational studies. BMC Health Services Research. 2012; 12(1):395. doi: 10.1186/1472-6963-12-395

31. Starfield B, Shi L. Policy relevant determinants of health: an international perspective. Health policy (Amsterdam, Netherlands). 2002; 60(3):201–18. Epub 2002/04/20. PMID: 11965331.

32. Effective Practice and Organisation of Care (EPOC). Suggested risk of bias criteria for EPOC reviews Oslo: Norwegian Knowledge Centre for the Health Services; 2014 [cited 2014]. Available: https://epoc.cochrane.org/epoc-specific-resources-review-authors.

33. Marshall M, Leatherman S, Mattke S. Selecting indicators for the quality of health promotion, prevention and primary care at the health systems level in OECD countries. 2004.

34. Gebreselassie SG, Gashe FE. A systematic review of effect of prenatal zinc supplementation on birthweight: meta-analysis of 17 randomized controlled trials. Journal of health, population, and nutrition. 2011; 29(2):134. PMID: 21608422

35. Kontopantelis E, Reeves D. MetaEasy: A meta-analysis add-in for Microsoft Excel. J Stat Software. 2009; 30(7).

36. Cohen J. Statistical Power Analysis for the Behavioral Sciences: L. Erlbaum Associates; 1988.

37. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in medicine. 2002; 21(11):1539–58. Epub 2002/07/12. doi: 10.1002/sim.1186 PMID: 12111919.

38. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011; 343:d4002. Epub 2011/07/26. doi: 10.1136/bmj.d4002 PMID: 21784880.

39. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315(7109):629–34. Epub 1997/10/06. PMID: 9310563; PubMed Central PMCID: PMC2127453.

40. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008;337. doi: 10.1136/bmj.a1655

41. Borrill CS, Carletta J, Carter A, Dawson JF, Garrod S, Rees A, et al. The effectiveness of health care teams in the National Health Service: University of Aston in Birmingham; 2000.

42. Wagner EH. The role of patient care teams in chronic disease management. BMJ: British Medical Journal. 2000; 320(7234):569–72. PMC1117605. PMID: 10688568

43. Department of Health Social Services and Public Safety. Case management. A position paper. 2005 [12/02/2015]. Available: http://www.dhsspsni.gov.uk/case_management.pdf.

44. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: a systematic review of the literature. Ageing research reviews. 2011; 10(4):430–9. doi: 10.1016/j.arr.2011.03.003 PMID: 21402176

45. Valentijn PP, Schepman SM, Opheij W, Bruijnzeels MA. Understanding integrated care: a comprehensive conceptual framework based on the integrative functions of primary care. International Journal of Integrated Care. 2013; 13.

46. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. Cochrane Database Syst Rev. 2014; 4: MR000034. Epub 2014/05/02. doi: 10.1002/14651858.MR000034.pub2 PMID: 24782322.

47. Deeks J, Higgins J, Altman D. Chapter 9–Analysing Data and Undertaking Meta-analyses: Cochrane Handbook for Systematic Reviews of Interventions Version 5.1. 0 [updated March 2011]. Cochrane Handbook for Systematic Reviews of Interventions Version. 2011;5(0).

48. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions: Version 4.2.6.2006. Available: http://www.cochrane.org/sites/default/files/uploads/Handbook4.2.6Sep2006.pdf.

49. Lipsey MW, Wilson DB. Practical Meta-Analysis: SAGE Publications; 2001. 264 p.

50. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.; 2013.

51. Harris R, Bradburn M, Deeks J, Harbord R, Altman D, Sterne J. metan: fixed- and random-effects meta-analysis. Stata Journal. 2008; 8(1):3–28.

52. Sterne JAC, Harbord RM. Funnel plots in meta-analysis. Stata Journal. 2004; 4(2):127–41.

53. Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. Stata Journal. 2009; 9(2):197–210.

54. Bender R, Lange S. Adjusting for multiple testing—when and how? Journal of Clinical Epidemiology. 2001; 54(4):343–9. doi: 10.1016/S0895-4356(00)00314-0 PMID: 11297884

55. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics. 1979:65–70.

56. Moher D, Liberati A, Tetzlaff J, Altman DG, The PG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009; 6(7):e1000097. doi: 10.1371/journal.pmed.1000097 PMID: 19621072

57. Beland F, Bergman H, Lebel P, Clarfield AM, Tousignant P, Contandriopoulos AP, et al. A system of integrated care for older persons with disabilities in Canada: results from a randomized controlled trial. The journals of gerontology Series A, Biological sciences and medical sciences. 2006; 61 (4):367–73. Epub 2006/04/14. PMID: 16611703.

58. Beland F, Bergman H, Lebel P, Dallaire L, Fletcher J, Contandriopoulos AP, et al. Integrated services for frail elders (SIPA): a trial of a model for Canada. Canadian journal on aging = La revue canadienne du vieillissement. 2006; 25(1):5–42. Epub 2006/06/14. PMID: 16770746.

59. Bernabei R, Landi F, Gambassi G, Sgadari A, Zuccala G, Mor V, et al. Randomised trial of impact of model of integrated care and case management for older people living in the community1998. 1348 p.

60. Bird S, Noronha M, Sinnott H. An integrated care facilitation model improves quality of life and reduces use of hospital resources by patients with chronic obstructive pulmonary disease and chronic heart failure. Australian journal of primary health. 2010; 16(4):326–33. Epub 2010/12/09. doi: 10.1071/py10007 PMID: 21138701.

61. Boult C, Reider L, Frey K, Leff B, Boyd CM, Wolff JL, et al. Early effects of "Guided Care" on the quality of health care for multimorbid older persons: a cluster-randomized controlled trial. The journals of gerontology Series A, Biological sciences and medical sciences. 2008; 63(3):321–7. PMID: 18375882

62. Leff B, Reider L, Frick KD, Scharfstein DO, Boyd CM, Frey K, et al. Guided care and the cost of complex healthcare: a preliminary report. The American journal of managed care. 2009; 15(8):555–9. Epub 2009/08/13. PMID: 19670959.

63. Boyd CM, Reider L, Frey K, Scharfstein D, Leff B, Wolff J, et al. The effects of guided care on the perceived quality of health care for multi-morbid older persons: 18-month outcomes from a cluster-randomized controlled trial. J Gen Intern Med. 2010; 25(3):235–42. Epub 2009/12/25. doi: 10.1007/s11606-009-1192-5 PMID: 20033622; PubMed Central PMCID: PMCPmc2839336.

64. Boult C, Reider L, Leff B, Frick KD, Boyd CM, Wolff JL, et al. The effect of guided care teams on the use of health services: results from a cluster-randomized controlled trial. Archives of internal medicine. 2011; 171(5):460–6. Epub 2011/03/16. doi: 10.1001/archinternmed.2010.540 PMID: 21403043.

65. Boult C, Leff B, Boyd CM, Wolff JL, Marsteller JA, Frick KD, et al. A matched-pair cluster-randomized trial of guided care for high-risk older patients. J Gen Intern Med. 2013; 28(5):612–21. Epub 2013/01/12. doi: 10.1007/s11606-012-2287-y PMID: 23307395; PubMed Central PMCID: PMCPmc3631081.

66. Boyd ML, Fisher B, Davidson AW, Neilsen CA. Community-based case management for chronically ill older adults. Nursing management. 1996; 27(11):31–2. Epub 1996/11/01. PMID: 8954445.

67. Burns R, Nichols LO, Graney MJ, Cloar FT. Impact of continued geriatric outpatient management on health outcomes of older veterans. Archives of internal medicine. 1995; 155(12):1313–8. Epub 1995/06/26. PMID: 7778963.

68. Burns R, Nichols LO, Martindale-Adams J, Graney MJ. Interdisciplinary geriatric primary care evaluation and management: two-year outcomes. Journal of the American Geriatrics Society. 2000; 48(1):8–13. Epub 2000/01/21. PMID: 10642014.

69. Coburn KD, Marcantonio S, Lazansky R, Keller M, Davis N. Effect of a community-based nursing intervention on mortality in chronically ill older adults: a randomized controlled trial. PLoS medicine. 2012; 9(7):e1001265. doi: 10.1371/journal.pmed.1001265 PMID: 22815653

70. Counsell SR, Callahan CM, Clark DO, Tu W, Buttar AB, Stump TE, et al. Geriatric care management for low-income seniors: a randomized controlled trial. Jama. 2007; 298(22):2623–33. Epub 2007/12/13. doi: 10.1001/jama.298.22.2623 PMID: 18073358.

71. Counsell SR, Callahan CM, Tu W, Stump TE, Arling GW. Cost analysis of the Geriatric Resources for Assessment and Care of Elders care management intervention. Journal of the American Geriatrics Society. 2009; 57(8):1420–6. Epub 2009/08/20. PMID: 19691149; PubMed Central PMCID: PMCPmc3874584.

72. Dalby DM, Sellors JW, Fraser FD, Fraser C, van Ineveld C, Howard M. Effect of preventive home visits by a nurse on the outcomes of frail elderly people in the community: a randomized controlled trial. CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne. 2000; 162(4):497–500. Epub 2000/03/04. PMID: 10701382; PubMed Central PMCID: PMCPmc1231166.

73. de Stampa M, Vedel I, Buyck JF, Lapointe L, Bergman H, Beland F, et al. Impact on hospital admissions of an integrated primary care model for very frail elderly patients. Archives of gerontology and geriatrics. 2014; 58(3):350–5. Epub 2014/02/11. doi: 10.1016/j.archger.2014.01.005 PMID: 24508468.

74. Dorr DA, Wilcox AB, Brunker CP, Burdon RE, Donnelly SM. The effect of technology-supported, multidisease care management on the mortality and hospitalization of seniors. Journal of the American Geriatrics Society. 2008; 56(12):2195–202. Epub 2008/12/20. doi: 10.1111/j.1532-5415.2008.02005.x PMID: 19093919.

75. Enguidanos SM, Jamison PM. Moving from tacit knowledge to evidence-based practice: the Kaiser Permanente community partners study. Home health care services quarterly. 2006; 25(1–2):13–31. Epub 2006/06/29. doi: 10.1300/J027v25n01_02 PMID: 16803736.

76. Fan VS, Gaziano JM, Lew R, Bourbeau J, Adams SG, Leatherman S, et al. A comprehensive care management program to prevent chronic obstructive pulmonary disease hospitalizations: a randomized, controlled trial. Annals of internal medicine. 2012; 156(10):673–83. Epub 2012/05/16. doi: 10.7326/0003-4819-156-10-201205150-00003 PMID: 22586006.

77. Fitzgerald JF, Smith DM, Martin DK, Freedman JA, Katz BP. A case manager intervention to reduce readmissions. Archives of internal medicine. 1994; 154(15):1721–9. Epub 1994/08/08. PMID: 8042889.

78. Fordyce M, Bardole D, Romer L, Soghikian K, Fireman B. Senior Team Assessment and Referral Program—STAR. The Journal of the American Board of Family Practice / American Board of Family Practice. 1997; 10(6):398–406. Epub 1998/01/04. PMID: 9407480.

79. Gagnon AJ, Schein C, McVey L, Bergman H. Randomized controlled trial of nurse case management of frail older people. Journal of the American Geriatrics Society. 1999; 47(9):1118–24. Epub 1999/09/14. PMID: 10484257.

80. Gravelle H, Dusheiko M, Sheaff R, Sargent P, Boaden R, Pickard S, et al. Impact of case management (Evercare) on frail elderly patients: controlled before and after analysis of quantitative outcome data. BMJ: British Medical Journal. 2007; 334(7583):31-. doi: 10.1136/bmj.39020.413310.55 PMC1764106. PMID: 17107984

81. Hogg W, Lemelin J, Dahrouge S, Liddy C, Armstrong CD, Legault F, et al. Randomized controlled trial of anticipatory and preventive multidisciplinary team care: for complex patients in a community-based primary care setting. Canadian family physician Medecin de famille canadien. 2009; 55(12):e76–85. Epub 2009/12/17. PMID: 20008582; PubMed Central PMCID: PMCPmc2793206.

82. Gray D, Armstrong CD, Dahrouge S, Hogg W, Zhang W. Cost-effectiveness of Anticipatory and Preventive multidisciplinary Team Care for complex patients: evidence from a randomized controlled trial. Canadian family physician Medecin de famille canadien. 2010; 56(1):e20–9. Epub 2010/01/22. PMID: 20090057; PubMed Central PMCID: PMCPmc2809192.

83. Kruse RL, Zweig SC, Nikodim B, LeMaster JW, Coberly JS, Colwill JM. Nurse care coordination of older patients in an academic family medicine clinic: 5-year outcomes. Jcom. 2010; 17(5):209–15.

84. Leung AC-t, Liu C-p, Chow NW-s, Chi I. Cost-benefit analysis of a case management project for the community-dwelling frail elderly in Hong Kong. Journal of Applied Gerontology. 2004; 23(1):70–85.

85. Levine S, Steinman BA, Attaway K, Jung T, Enguidanos S. Home care program for patients at high risk of hospitalization. The American journal of managed care. 2012; 18(8):e269–76. Epub 2012/08/30. PMID: 22928795.

86. Martin IR, McNamara D, Sutherland FR, Tilyard MW, Taylor DR. Care plans for acutely deteriorating COPD: a randomized controlled trial. Chronic respiratory disease. 2004; 1(4):191–5. Epub 2005/11/12. PMID: 16281645.

87. Metzelthin SF, van Rossum E, de Witte LP, Ambergen AW, Hobma SO, Sipers W, et al. Effectiveness of interdisciplinary primary care approach to reduce disability in community dwelling frail older people: cluster randomised controlled trial2013.

88. Morishita L, Boult C, Boult L, Smith S, Pacala JT. Satisfaction with outpatient geriatric evaluation and management (GEM). The Gerontologist. 1998; 38(3):303–8. Epub 1998/06/26. PMID: 9640850.

89. Boult C, Boult LB, Morishita L, Dowd B, Kane RL, Urdangarin CF. A randomized clinical trial of outpatient geriatric evaluation and management. Journal of the American Geriatrics Society. 2001; 49(4):351–9. Epub 2001/05/12. PMID: 11347776.

90. Newcomer R, Maravilla V, Faculjak P, Graves MT. Outcomes of preventive case management among high-risk elderly in three medical groups: a randomized clinical trial. Evaluation & the health professions. 2004; 27(4):323–48. Epub 2004/10/20. doi: 10.1177/0163278704270011 PMID: 15492046.

91. Ploeg J, Brazil K, Hutchison B, Kaczorowski J, Dalby DM, Goldsmith CH, et al. Effect of preventive primary care outreach on health related quality of life among older adults at risk of functional decline: randomised controlled trial2010.

92. Ródenas F, Garcés J, Carretero S, Megia MJ. Case management method applied to older adults in the primary care centres in Burjassot (Valencian Region, Spain). Eur J Ageing. 2008; 5(1):57–66. doi: 10.1007/s10433-008-0073-9

93. Rubenstein LZ, Alessi CA, Josephson KR, Trinidad Hoyl M, Harker JO, Pietruszka FM. A randomized trial of a screening, case finding, and referral system for older veterans in primary care. Journal of the American Geriatrics Society. 2007; 55(2):166–74. Epub 2007/02/17. doi: 10.1111/j.1532-5415.2007.01044.x PMID: 17302651.

94. Schraeder C, Shelton P, Sager M. The effects of a collaborative model of primary care on the mortality and hospital use of community-dwelling older adults. The journals of gerontology Series A, Biological sciences and medical sciences. 2001; 56(2):M106–12. Epub 2001/02/24. PMID: 11213273.

95. Schraeder C, Fraser CW, Clark I, Long B, Shelton P, Waldschmidt V, et al. Evaluation of a primary care nurse case management intervention for chronically ill community dwelling older people. Journal of Clinical Nursing. 2008; 17(11c):407–17. doi: 10.1111/j.1365-2702.2008.02578.x

96. Shannon GR, Wilber KH, Allen D. Reductions in costly healthcare service utilization: findings from the Care Advocate Program. Journal of the American Geriatrics Society. 2006; 54(7):1102–7. Epub 2006/07/27. doi: 10.1111/j.1532-5415.2006.00799.x PMID: 16866682.

97. Alkema GE, Wilber KH, Shannon GR, Allen D. Reduced Mortality: The Unexpected Impact of a Telephone-Based Care Management Intervention for Older Adults in Managed Care. Health Services Research. 2007; 42(4):1632–50. doi: 10.1111/j.1475-6773.2006.00668.x PMC1955273. PMID: 17610441

98. Sledge WH, Brown KE, Levine JM, Fiellin DA, Chawarski M, White WD, et al. A randomized trial of primary intensive care to reduce hospital admissions in patients with high utilization of inpatient services. Disease management: DM. 2006; 9(6):328–38. Epub 2006/11/23. doi: 10.1089/dis.2006.9.328 PMID: 17115880.

99. Stuck AE, Minder CE, Peter-Wuest I, Gillmann G, Egli C, Kesselring A, et al. A randomized trial of in-home visits for disability prevention in community-dwelling older people at low and high risk for nursing home admission. Archives of internal medicine. 2000; 160(7):977–86. Epub 2000/04/13. PMID: 10761963.

100. Sylvia ML, Griswold M, Dunbar L, Boyd CM, Park M, Boult C. Guided care: cost and utilization outcomes in a pilot study. Disease management: DM. 2008; 11(1):29–36. Epub 2008/02/19. doi: 10.1089/dis.2008.111723 PMID: 18279112.

101. Boyd CM, Shadmi E, Conwell LJ, Griswold M, Leff B, Brager R, et al. A pilot test of the effect of guided care on the quality of primary care experiences for multimorbid older adults. J Gen Intern Med. 2008;

23(5):536–42. Epub 2008/02/13. doi: 10.1007/s11606-008-0529-9 PMID: 18266045; PubMed Central PMCID: PMCPmc2324149.

102. Toseland RW, O'Donnell JC, Engelhardt JB, Hendler SA, Richie JT, Jue D. Outpatient geriatric evaluation and management. Results of a randomized trial. Medical care. 1996; 34(6):624–40. PMID: 8656727

103. Toseland RW, O'Donnell JC, Engelhardt JB, Richie J, Jue D, Banks SM. Outpatient geriatric evaluation and management: is there an investment effect? The Gerontologist. 1997; 37(3):324–32. PMID: 9203756

104. Engelhardt JB, Toseland RW, O'Donnell JC, Richie JT, Jue D, Banks S. The effectiveness and efficiency of outpatient geriatric evaluation and management. Journal of the American Geriatrics Society. 1996; 44(7):847–56. Epub 1996/07/01. PMID: 8675938.

105. Engelhardt JB, Toseland RW, Gao J, Banks S. Long-Term Effects of Outpatient Geriatric Evaluation and Management on Health Care Utilization, Cost, and Survival. Research on Social Work Practice. 2006; 16(1):20–7. doi: 10.1177/1049731505276047

106. van Hout HPJ, Jansen APD, van Marwijk HWJ, Pronk M, Frijters DF, Nijpels G. Prevention of adverse health trajectories in a vulnerable elderly population through nurse home visits: a randomized controlled trial [ISRCTN05358495]. The journals of gerontology Series A, Biological sciences and medical sciences. 2010; 65(7):734–42. doi: 10.1093/gerona/glq037 PMID: 20457579

107. Macinko J, Starfield B, Shi L. The contribution of primary care systems to health outcomes within Organization for Economic Cooperation and Development (OECD) countries, 1970–1998. Health Serv Res. 2003; 38(3):831–65. Epub 2003/06/26. PMID: 12822915; PubMed Central PMCID: PMCPmc1360919.

108. Grant CC, Forrest CB, Starfield B. Primary care and health reform in New Zealand. The New Zealand medical journal. 1997; 110(1037):35–9. Epub 1997/02/14. PMID: 9066566.

109. Fry J, Horder J. Primary health care in an international context: Nuffield Provincial Hospitals Trust London; 1994.

110. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? Lancet. 2000; 356(9237):1228–31. Epub 2000/11/10. doi: 10.1016/s0140-6736(00)02786-0 PMID: 11072941.

111. Maciejewski ML, Bayliss EA. Approaches to comparative effectiveness research in multimorbid populations. Medical care. 2014; 52 Suppl 3:S23–30. Epub 2014/02/25. doi: 10.1097/mlr.0000000000000060 PMID: 24561754.

112. Bate P, Robert G, Fulop N, Øvretveit J, Dixon-Woods M. Perspectives on context. London, UK: 2014.

113. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986; 7(3):177–88. doi: 10.1016/0197-2456(86)90046-2 PMID: 3802833

114. Gelman A, Hill J, Yajima M. Why we (usually) don't have to worry about multiple comparisons. Journal of Research on Educational Effectiveness. 2012; 5(2):189–211.

115. Altman DG, Bland JM. Statistics Notes: Detecting skewness from summary information1996. 1200 p.

116. Sandars J, Esmail A. The frequency and nature of medical error in primary care: understanding the diversity across studies. Family practice. 2003; 20(3):231–6. Epub 2003/05/10. PMID: 12738689.

117. Panagioti M, Stokes J, Esmail A, Coventry P, Cheraghi-Sohi S, Alam R, et al. Multimorbidity and patient safety incidents in primary care: A systematic review and meta-analysis. PLoS ONE. 2015—Submitted.

118. Kenning C, Coventry PA, Bower P. Self-management interventions in patients with long-term conditions: a structured review of approaches to reporting inclusion, assessment, and outcomes in multimorbidity. 2014. 2014; 4(1):9. Epub 2014-07-24. doi: 10.15256/joc.2014.4.33

119. Islam MM, Valderas JM, Yen L, Dawda P, Jowsey T, McRae IS. Multimorbidity and Comorbidity of Chronic Diseases among the Senior Australians: Prevalence and Patterns. PLoS ONE. 2014; 9(1): e83783. doi: 10.1371/journal.pone.0083783 PMID: 24421905

120. Fried TR, O'Leary J, Towle V, Goldstein MK, Trentelange M, Martin DK. The Effects of Comorbidity on the Benefits and Harms of Treatment for Chronic Disease: A Systematic Review. PLoS ONE. 2014; 9 (11):e112593. doi: 10.1371/journal.pone.0112593 PMID: 25402463

121. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. Lancet. 2007; 370(9590):851–8. Epub 2007/09/11. doi: 10.1016/s0140-6736(07)61415-9 PMID: 17826170.

122. Roland M, Abel G. Reducing emergency admissions: are we on the right track? BMJ. 2012; 345. doi: 10.1136/bmj.e6017

123. Alderwick H, Ham C, Buck D. Population health systems: Going beyond integrated care. London, UK: The King's Fund, 2015.

5.2 Chapter Conclusions

In this chapter, I have shown the limited direct effectiveness of the case management model on the high-risk patients treated across all health-system goals, although patient satisfaction does appear to be consistently improved to a small extent. In addition, subgroup analysis led to the hypothesis that under certain forms of case management (e.g. MDT rather than led by a single healthcare professional, when a social worker is involved, or when the intervention takes place in a health system rated as having weak primary care) the effects of the intervention might be slightly improved. Next, I examine the effectiveness of a single MDT case management intervention in the NHS setting.

6. Integrated care in Central Manchester (MDT Case management)

6.1 Chapter Introduction

This empirical piece builds on the last, addressing gaps identified in the case management literature, e.g. a lack of attention to potential spill-over effects of MDT working, a lack of quasi-experimental studies (here I use a difference-in-differences approach), and those set in contexts outside of the USA. In addition, it evaluates a locally implemented case management model exhibiting what were hypothesised as good practice elements of case management (e.g. delivery by an MDT with involvement of a social worker). The article was published in *BMJ Open* in April 2016.

Any supplementary material referred to in the text is available in Appendix 3.

BMJ Open

Effectiveness of multidisciplinary team case management: difference-in-differences analysis

Jonathan Stokes,[1] Søren Rud Kristensen,[2] Kath Checkland,[3] Peter Bower[1]

To cite: Stokes J, Kristensen SR, Checkland K, *et al*. Effectiveness of multidisciplinary team case management: difference-in-differences analysis. *BMJ Open* 2016;**6**:e010468. doi:10.1136/bmjopen-2015-010468

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2015-010468).

Received 5 November 2015
Revised 28 January 2016
Accepted 9 February 2016

For numbered affiliations see end of article.

Correspondence to
Jonathan Stokes;
jonathan.stokes-3@postgrad.manchester.ac.uk

ABSTRACT

Objectives: To evaluate a multidisciplinary team (MDT) case management intervention, at the individual (direct effects of intervention) and practice levels (potential spillover effects).

Design: Difference-in-differences design with multiple intervention start dates, analysing hospital admissions data. In secondary analyses, we stratified individual-level results by risk score.

Setting: Single clinical commissioning group (CCG) in the UK's National Health Service (NHS).

Participants: At the individual level, we matched 2049 intervention patients using propensity scoring one-to-one with control patients. At the practice level, 30 practices were compared using a natural experiment through staged implementation.

Intervention: Practice Integrated Care Teams (PICTs), using MDT case management of high-risk patients together with a summary record of care versus usual care.

Direct and indirect outcome measures: Primary measures of intervention effects were accident and emergency (A&E) visits; inpatient non-elective stays, 30-day re-admissions; inpatient elective stays; outpatient visits; and admissions for ambulatory care sensitive conditions. Secondary measures included inpatient length of stay; total cost of secondary care services; and patient satisfaction (at the practice level only).

Results: At the individual level, we found slight, clinically trivial increases in inpatient non-elective admissions (+0.01 admissions per patient per month; 95% CI 0.00 to 0.01. Effect size (ES): 0.02) and 30-day re-admissions (+0.00; 0.00 to 0.01. ES: 0.03). We found no indication that highest risk patients benefitted more from the intervention. At the practice level, we found a small decrease in inpatient non-elective admissions (−0.63 admissions per 1000 patients per month; −1.17 to −0.09. ES: −0.24). However, this result did not withstand a robustness check; the estimate may have absorbed some differences in underlying practice trends.

Conclusions: The intervention does not meet its primary aim, and the clinical significance and cost-effectiveness of these small practice-level effects is debatable. There is an ongoing need to develop effective ways to reduce unnecessary attendances in secondary care for the high-risk population.

Strengths and limitations of this study

- This study addresses a number of shortcomings found in related literature from a recent systematic review.
- The difference-in-differences methods can provide a rigorous assessment under certain conditions while evaluating an intervention in a real-world setting.
- Results are analysed and presented at two levels to show direct effects of the intervention, as well as wider spillover effects of integrated care.
- At the practice level, there may be some selection bias due to voluntary recruitment, although we predict this to be minimal based on our robustness checks.
- At the individual level, results may be prone to some bias in favour of control participants due to the ongoing recruitment strategy versus a single time point propensity matching. Again, we predict this to be minimal, as participants and controls were well matched at the first start date.

INTRODUCTION

An ageing population with increasing number of long-term conditions (LTCs) and complex multimorbidity[1 2] has caused policymakers to rethink delivery of care.[3]

There is increasing focus on the benefits of 'integrated care', to enable a more efficient and effective response to LTCs.[3 4] There is no consensus definition of what constitutes 'integrated care',[5] and the concept describes many different changes to the health system that can occur at multiple levels.[6 7 8] Practical implementation examples of integrated care include pooling of funds, joint commissioning, colocation of services, shared clinical records, and at the interface of the health system with the patient (ie, service delivery level) multidisciplinary team (MDT) working and case management.[7 8]

In the UK's National Health Service (NHS), a common model of integrated care is the use of 'multi-disciplinary team (MDT) case management of high-risk patients'.[9 10]

We undertook a systematic review of this model of integrated care and found few effects across a number of relevant outcomes, barring a small effect on patient satisfaction, and short-term changes in self-reported health status.[6]

We also identified gaps in the current literature. In the review, 78% of included studies were randomised controlled trials (RCTs).[6] We suggested a complementary role for rigorous quasi-experiments in routine settings to better balance internal and external validity.[11 12]

The majority of studies also measured only direct (individual-level) effects. MDT case management used to manage a subset of patients could lead to broader changes, such as better 'professional integration' through team working.[13 14] These broader changes could lead to effects on the wider patient population, beyond those patients specifically managed by the MDT (what we call 'spillover effects').

Our contribution to the evidence base for MDT case management thus involved an evaluation of a local integrated care intervention using a robust quasi-experimental study design. We model effects using two distinct analyses: (1) individual-level analysis (to capture direct effects of the intervention) and (2) practice-level analysis (to capture any potential spillover effects).

The intervention

In Central Manchester, the MDT case management is achieved through Practice Integrated Care Teams (PICTs) introduced by the clinical commissioning group (CCG). PICTs conduct case finding, assess the needs of the individual identified, prepare individualised care plans, co-ordinate care and conduct regular review, monitoring and adaptation of the care plan.[15] The aim of the intervention was to reduce unnecessary attendances in secondary care for the high-risk population.[16]

Table 1 gives an overview of the key aspects of the intervention. Compared with our previous systematic review of similar interventions, it is fairly common, where we identified the majority (58%) employing MDT case management (as opposed to a single case manager), and a predictive risk model as the primary method of identifying suitable patients.[6] Less commonly, this intervention took place in a system ranked as delivering 'high' strength of primary care (ie, strength of primary healthcare orientation of the health system as classified by Starfield and Shi[17]—the majority in the review came from a 'low'-strength system, eg, USA: 64%). Additionally, the PICT intervention included involvement of a social worker (33% of studies involved a social worker in our previous review), providing further potential for 'horizontal integration' (ie, integration between separate organisations at the same level of the health system).[3]

Table 1 Overview of intervention

Name of model	Intensity of intervention (patient contacts)	Risk assessment tool (judgement/ threshold/ predictive risk modelling)	MDT composition	Primary location of intervention	24-hour availability of case manager?	Case load per MDT	Training received by case manager/ MDT?	Case management reimbursement method?	IT linkage?
Practice Integrated Care Team (PICT)	Determined case by case —for some weekly, some less frequently	Predictive risk modelling /judgement Combined Predictive Risk Model/clinical judgement	GP, practice nurse, district nurse, social worker and active case manager (with access to other specialists outside of the core team)	GP practice/ telephone/ home visits—location of patient contacts depended on health professional	No—Health professionals available at their core hours. Emergency routes to community services through the night	Stipulated 2% of adult case load per practice	Yes—Action learning sessions, mock MDTs, initial adopters spreading learning	Payment received as part of wider quality process scheme contract	'Graphnet'—holds summary care record accessible by multiple providers (read, but not write access). Care plan can be edited by all

GP, general practitioner; MDT, multidisciplinary team.

METHODS

Our study used a quasi-experimental pre–post design with a suitable control group to examine any change in

outcomes induced by a policy change intervention—an adapted version of difference-in-differences (DD) analysis.[18]

We prepared and analysed data at two distinct levels, each described separately below. Owing to a data governance issue at the CCG, intervention patients could not be identified at the individual level until nearly all practices implemented the intervention (patients were not consented prior to this date, so those joining before could not be included in the analysis—they were also excluded from our control group, so no contamination occurred). Figure 1 summarises the period of analysis for the individual and practice levels, showing the analysis and 'pretrend' period (ie, period prior to any practice/individual joining the intervention group) for each.

With the PICT intervention having no single start date, we adapted our analysis to allow for this staged introduction (using a time fixed effect instead of the usual binary post dummy—see equations in practice level and individual level sections in the online supplementary material appendices).[18 19] The main difference from the standard DD approach is that the intervention and control groups are not static over time, allowing intervention patients/practices to join gradually over the monthly panel datasets, and comparing appropriately at each time point. This method has been used previously in the literature,[19 20] and we have adapted it to suit data at both of our levels of analyses (explained below, and in more detail in the online supplementary material appendices).

We analysed anonymised data held by the CCG, from the 'admitted patient care commissioning dataset', submitted by all providers in England via the Secondary Uses Service (SUS). The dataset included all patient contacts with secondary care services, demographic data, as well as costs calculated through the national payment by results (PbR, together with local prices for local providers where applicable). For the analysis of pseudonymised/anonymised data, no formal ethics process was deemed necessary. The CCG had themselves previously consented the intervention individuals for use of their data for evaluation purposes. For patient satisfaction at the practice level, we used data from the GP Patient Survey (GPPS—see online supplementary material appendices).

Data preparation and analysis

All data preparation and analysis was carried out using STATA (V.13) (StataCorp. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP., 2013). The DD analysis estimate is unbiased only under the key assumption that the average difference between intervention and control units' trends would be the same in the absence of 'treatment' (ie, the PICT intervention).[18] This 'parallel trends' assumption is key to DD analysis and was tested graphically and statistically for each outcome assessed, at each analysis level (see online supplementary material appendices for graphs).

We analysed data distinctly at two levels:

1. Individual level: primary analysis

At the individual level, to obtain parallel pretrends, it was necessary to propensity match intervention patients to controls from within the same CCG (we matched on the characteristics for which the patients were recruited

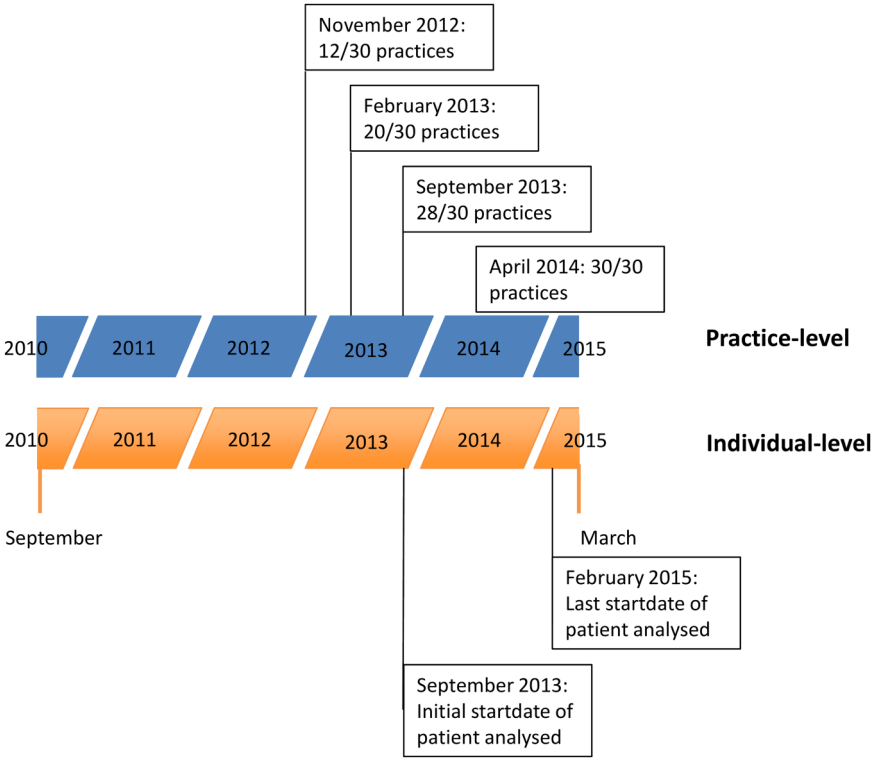


Figure 1 Timeline of analyses highlighting key dates of practices and individual patients included in analysis joining the intervention.

in practice to maximise comparability—see online supplementary material appendices for details). We then analysed 2049 intervention patients versus 2049 matched controls using the best-fitting count model for each outcome.[21] Outcome measures were summed to a count per patient per month over the period September 2010–March 2015 inclusive, to allow a 3-year pretrend period.

In all models, we adjusted for relevant individual covariates from the directed acyclic graph (DAG—see online supplementary material appendices),[22] as well as practice fixed effects (to control for any effects caused by characteristics of a specific practice rather than the intervention itself).[23] We cluster our SEs by practice to deal with concerns of serial correlation.[24] We took the average partial effect of results (for β_2—see online supplementary material appendices for equation) and report these below (ie, the covariates adjusted absolute change in counts per patient per month). We additionally report the effect size (ES; standardised mean difference) as a measure of practical significance of each result.[25] We adopted Cohen's rule of thumb for interpreting ESs, that is, 0.2 indicates a small effect, 0.5 a medium and 0.8 a large effect.[26]

Stratification by risk score

Patients were recruited to the intervention via risk tool score and clinical judgement. To test whether the highest risk patients (according to risk tool score) benefitted more from the PICT intervention than those with lower calculated risks also treated, we generated a 'high-risk' dummy. We reran the individual-level analysis with a difference-in-difference-in-differences (DDD) analysis, using an additional interaction term to determine subgroup effects (see online supplementary material appendices for equation).

2. Practice level: secondary analysis of spillover effects

At the practice level, practices gradually took up the intervention over a period of 18 months. At each time point (updating monthly in our dataset), the time fixed effects compare all intervention practices with all 'controls' (ie, all those practices that have not yet adopted the intervention, even though they will later adopt the intervention).[19 20] Outcomes were summed to a count per 1000 patients per month for each of the practices and analysed over the period September 2010–March 2015 inclusive, to overlap with the individual-level analysis.

We used a linear regression model, adjusting for fixed effects for each practice and time period (monthly—see online supplementary material appendices for equation). We cluster our SEs by practice to deal with concerns of serial correlation.[24]

Outcome measures

Primary outcome measures for both analyses included:

- Inpatient non-elective admissions
- Re-admissions (30 days)
- Inpatient elective admissions
- Accident and emergency (A&E) visits
- Outpatient visits
- Admissions for ambulatory care sensitive conditions (ACSCs, which we used as a measure of patient safety in a health system with universal health coverage—see online supplementary material appendices for details)

Secondary outcome measures included:

- Total cost of secondary care services (£)
- Length of stay (inpatient)
- Patient satisfaction (practice level only: measured through the GPPS—see online supplementary material appendices)
  - General satisfaction
  - LTC-specific satisfaction

Robustness check

At both levels of analysis, we additionally added a robustness check including a practice-specific time trend. This allows intervention and control practices to follow different trends and can help reveal any indication of the observed effect having absorbed any differences in underlying practice time trends.[18]

At the practice level only, due to the voluntary roll-out of the intervention, we attempted to assess the effects of selection bias using a logistic regression model (including % males; % over 65; practice list size; number of general practitioners (GPs) per thousand patients; total Index of Multiple Deprivation (IMD) score 2010; and total % Quality and Outcomes Framework (QOF) achievement score).[27] We additionally reran the practice-level analysis excluding those practices recruited to the intervention in wave 1, assuming these to be the practices at most risk of selection bias if it did indeed occur.[19]

RESULTS

Individual-level analysis

Sample characteristics

A total of 2049 intervention patients were propensity score matched to non-intervention patients from the same CCG. As expected, the differences were small between matched patient baseline characteristics (see table 2).

Table 3 shows the crude absolute differences in mean outcome measures (PICT patients vs matched controls). As for the DD results in the section below, a negative estimate indicates a relative decrease in admissions for PICT patients compared with controls (ie, a negative intervention effect favours the intervention).

DD parallel pretrends

We identified evidence of a significant difference between pretrends for outpatient visits at the individual level. This variable was potentially biased towards a result favouring the PICT intervention over controls.

Table 2 Individual baseline characteristics (before and after matching)

Mean (unless otherwise indicated)	Before matching			After matching		
	PICT (SD)	Controls (SD)	SMD	PICT (SD)	Controls (SD)	SMD
N	2049	93 532		2049	2049	
Male (%)	44.3	47.4		44.3	44.1	
Age	67.2 (17.8)	35.3 (22.2)	−1.44	67.2 (17.8)	65.8 (18.7)	−0.07
IMD 2010	40.2 (14.8)	40.6 (16.0)	0.03	40.2 (14.8)	40.2 (15.8)	0.00
MM count baseline	2.7 (2.1)	0.7 (1.2)	−1.63	2.7 (2.1)	2.4 (2.2)	−0.12
Previous inpatient admissions	1.3 (2.1)	0.3 (1.1)	−0.88	1.3 (2.1)	1.2 (2.2)	−0.05
Previous outpatient visits	7.0 (9.6)	1.9 (4.3)	−1.14	7.0 (9.6)	7.2 (9.8)	0.02
Previous A&E visits	1.4 (2.4)	0.5 (1.2)	−0.73	1.4 (2.4)	1.4 (2.5)	0.00

Previous admissions calculated for period 31 August 2012–1 September 2013 (12 months prior to the first intervention patient start date).
A&E, accident and emergency; IMD, Index of Multiple Deprivation; PICT, Practice Integrated Care Team; SMD, standardised mean difference.

However, we found no statistically significant result favouring either group. All other variables satisfied the parallel trends assumption, with no indication of bias.

DD results

Table 4 shows the DD analysis results at the individual level. After adjustment for age, cumulative multimorbidity, IMD domains (excluding health) and practice and time fixed effects, we found a slight increase in inpatient non-elective admissions (0.0053 per patient per month; 95% CI 0.0004 to 0.0102) and 30-day inpatient readmissions (0.0041; 0.0018 to 0.0064). The ESs (0.02 and 0.03) were small.[26]

Robustness check

All of the estimates withstood the addition of a practice-specific time trend.

Stratification by risk score

We observed no relationship between risk score and time of recruitment into the intervention. Observing the plots of risk score versus total postintervention

Table 3 Average absolute outcomes per patient per month by PICT, preintervention/postintervention

Outcome	Mean (per patient per month)		Difference	Unadjusted intervention effect (difference per patient per month)
	Pre	Post		
Primary outcomes				
Inpatient non-electives				
Controls	0.0362	0.0422	0.006	
PICT	0.0550	0.0704	0.0154	0.0094
Inpatient electives				
Controls	0.0365	0.0369	0.0004	
PICT	0.0438	0.0451	0.0013	0.0009
Outpatient admissions				
Controls	0.5019	0.5611	0.0592	
PICT	0.5701	0.7188	0.1487	0.0895
A&E visits				
Controls	0.0808	0.0805	−0.0003	
PICT	0.1061	0.1217	0.0156	0.0159
ACSCs				
Controls	0.0059	0.0078	0.0019	
PICT	0.0093	0.0124	0.0031	0.0012
Re-admissions (30 days)				
Controls	0.0069	0.0082	0.0013	
PICT	0.0115	0.0191	0.0076	0.0063
Secondary outcomes				
Total cost of secondary care services (£)				
Controls	168.8746	195.1289	26.2543	
PICT	215.0091	276.9591	61.9500	35.6957
Length of stay (days)				
Controls	0.3943	0.4888	0.0945	
PICT	0.5624	0.7903	0.2279	0.1334

A&E, accident and emergency; ACSCs, ambulatory care sensitive conditions; PICT, Practice Integrated Care Team.

Table 4 Individual-level adjusted model results

Outcome	Adjusted* intervention effect (95% CI) (difference per patient per month) Count (NBREG) model	Effect size†
Primary outcomes		
Inpatient non-electives	**0.0053 (0.0004 to 0.0102)‡**	0.02
Inpatient electives	−0.0011 (−0.0092 to 0.0070)	−0.00
Outpatient visits	0.0399 (−0.0068 to 0.0866)	0.03
A&E visits	0.0103 (−0.0001 to 0.0207)	0.03
ACSCs	0.0001 (−0.0017 to 0.0020)	0.00
Re-admissions (30 days)	**0.0041 (0.0018 to 0.0064)‡**	0.03
Secondary outcomes		
Total cost of secondary care services (£)§	8.1687 (−16.0021 to 32.3396)	0.01
Length of stay (days)	0.0528 (−0.1094 to 0.2151)	0.01

N=224 898 observations; 4098 individuals (period September 2010–March 2015).
bold: withstands practice×time robustness check.
*Adjusted for age, cumulative multimorbidity, IMD domains (excluding health), and practice and time fixed effects. Marginal effects on *PICT×Post* reported.
†Standardised mean difference.
‡Significant at p<0.05.
§Zero-inflated negative binomial model based on admission events. A&E, accident and emergency; ACSCs, ambulatory care sensitive conditions; IMD, Index of Multiple Deprivation; NBREG, negative binomial regression.

admissions, however, we see that there does appear to be a relationship between a higher risk score and increased non-elective admissions, A&E visits, total cost of secondary care services, admissions for ACSCs, inpatient length of stay and inpatient 30-day readmissions (see online supplementary material appendices). This implies that the risk score is a good predictor of these future admission types, as expected.

Results of the DDD analysis, however, indicate that those patients with a higher risk score did not benefit more from the intervention, and instead showed statistically significant increased inpatient non-elective admissions (0.0208 per patient per month; 95% CI 0.0083 to 0.0333. ES: 0.09), A&E visits (0.0363; 0.0128 to 0.0598. ES: 0.09) and inpatient length of stay (0.3071; 0.0592 to 0.5549. ES: 0.06—see online supplementary material appendices for full list of DDD estimates) compared with others. Again, the ESs indicate these increases were slight.

Practice-level analysis

Sample characteristics

Table 5 shows the practice characteristics of the intervention and control practices included in the analysis (comparing the practices which joined the intervention in wave 1 with those that joined the intervention at a later date). On average, the practices are very similar, with wave 1 practices with a slightly higher proportion of older patients, and a slightly more even male/female split.

Table 6 shows the crude absolute differences in mean outcome measures (per 1000 patients per month) observed between the wave 1 PICT practices and those practices joining at a later date (shown as 'controls' for illustration purposes), preintervention and postintervention. As for the DD results in the section below, a negative estimate indicates a relative decrease in admissions for PICT practices compared with controls. For satisfaction outcomes, a positive estimate indicates increased satisfaction for the intervention practices compared to usual care.

DD parallel pretrends

We identified no significant differences between pretrends for any outcome at the practice level. These data satisfy the parallel trends assumption, with no indication of bias.

Table 5 Practice characteristics (wave 1 compared to later joining practices)

Mean (unless otherwise indicated)	PICT—wave 1 (SD)	Controls—later joining (SD)	SMD
N	12	18	
Male (proportion of practice)	0.52 (0.04)	0.54 (0.05)	0.04
Over 65 years (proportion of practice)	0.09 (0.03)	0.07 (0.04)	−0.15
IMD 2010	38.5 (10.7)	37.4 (7.7)	−0.12
Practice list size	6022.9 (2656.2)	6879.1 (3503.9)	0.27
GPs per thousand	0.6 (0.1)	0.8 (0.5)	0.29

GPs, general practitioners; IMD, Index of Multiple Deprivation; PICT, Practice Integrated Care Team; SMD, standardised mean difference.

DD results

Table 7 shows the DD analysis results at the practice level. After adjustment for practice and time fixed effects, the difference for inpatient non-elective admissions was significant, with an estimated −0.63 admissions per 1000 patients per month (95% CI −1.17 to −0.09) for PICT practices compared with control practices.

The practical significance, as evidenced by the ES of −0.24, suggests a small effect of PICT on inpatient non-elective admissions at the practice level.

Robustness check

Following our robustness check, including a practice-specific time trend, the estimate for inpatient non-electives was no longer significant: −0.52 (−1.05 to 0.01. ES: −0.20). This may suggest that the intervention effect has absorbed some differences between treated practices due to an underlying practice-specific time trend (which can happen when policies are implemented at different points in time in different units, ie, the practice time trend which was occurring already can drive the results, so once we control for this, the estimated effect is driven towards zero).[18] The ES, however, remained similar to the result reported above.

We were unable to predict wave 1 entry from the characteristics we included in our logistic regression model. Thus, we conclude that selection bias into early adoption, based on these characteristics at least, was minimal. However, this does not preclude the presence of selection bias based on unmeasured characteristics.

When we removed wave 1 practices (assuming these to be at most risk of selection bias, if it did indeed occur), statistical power was reduced (as expected), and the SEs of our estimates were inflated. Subsequently, we found no significant results following this robustness check. The estimate for inpatient non-elective admissions nevertheless remained negative (ie, in favour of the intervention—see online supplementary material appendices for full list of estimates following this robustness check).

DISCUSSION

For direct effects of the intervention, this study finds some statistically significant differences between groups, although effects are very small. The results of our DDD analysis show that even the highest risk patients (as defined by the risk prediction tool) treated did not benefit from the intervention, and in fact admissions for a number of outcomes (inpatient non-electives, A&E visits and inpatient length of stay) increased slightly for these patients.

Additional analysis at the practice level finds indications of potentially small positive spillover effects of integrated working at a higher system level. In particular, we identified a possible reduction in inpatient non-elective admissions (which, however, did not hold up to our robustness check). However, even if these effects are caused by the intervention, which this study cannot prove beyond doubt, the absolute difference observed in the analysis is small.[26] For an average practice of approximately 6000 patients, this would equate to an estimated difference (not an absolute reduction) of −45.6 (95% CI −84.0 to −6.6) inpatient non-elective admissions in a year compared to usual care. If we estimate the average cost of an inpatient non-elective admission to be £1489,[28] this would potentially translate to a £67 898 (95% CI £125 076 to £9827) difference compared to usual care, before accounting for intervention costs. While we did not have data on the precise intervention costs of PICT, the national Directed Enhanced Service (DES), which incentivises similar case management interventions, paid an average-sized practice £5175 for implementing the intervention in 2013/2014.[29] This extra incentive cost of course does not account for actual additional costs of running the intervention, for example, physician time, overheads and opportunity cost of a fairly time-intensive intervention, which would also need to be considered. Additionally, our analysis found no significant effect on total secondary care costs realised during the study period, with a presumable increase in primary care costs to run the intervention (although we did not have data available on primary care costs, so cannot say for certain). Therefore, beyond the cautions we have identified for this potential spillover benefit (ie, absence of a primary effect, and not holding up to robustness checks), cost-effectiveness of the intervention remains questionable.

Comparison of direct and spillover effects

The apparently contradictory findings at the two levels analysed merit specific discussion. First, it is worth highlighting the small proportion of patients managed by the PICT teams directly (a stipulated 2% of each practice's highest risk adult patients). The final pool of intervention patients we analysed (n=2049), therefore, only constitutes 1.04% of the patient population in the 30 practices. The likelihood of the direct effects of the intervention being a driver for practice-level results in terms of numbers treated is therefore negligible.

Second, the patients that were targeted directly by the intervention are by definition the highest risk, and potentially beyond the means of a medical intervention causing significant impact at all. This may be particularly true in the short term, for exacerbation of what are (frequently many) LTCs.[30] Our DDD analysis adds evidence to this effect. Perhaps then, the lower risk patients in the practice would be more likely to benefit from multidisciplinary working.

Additionally, some qualitative work commissioned by the CCG separately reveals that some features of the intervention at the patient level did not occur exactly as planned. For instance, there have been problems with the implementation of the shared summary record through Graphnet, meaning the MDT case management may not have been delivered exactly as planned in every detail (beyond the practice changes introduced by the

Table 6 Average absolute outcomes per 1000 patients per month by PICT (wave 1 compared to practices joining the intervention at a later date), preintervention/postintervention

Outcome	Unadjusted means (per 1000 patients per month)—wave 1 PICT compared to later joining: Pre (before 2012m11)	Post (after 2012m11)	Difference	Unadjusted intervention effect (difference per 1000 patients per month)
Primary outcomes				
Inpatient non-electives				
Controls	6.40	6.61	0.21	
PICT	8.36	8.07	−0.29	−0.50
Inpatient electives				
Controls	6.19	6.66	0.47	
PICT	8.58	8.51	−0.07	−0.54
Outpatient admissions				
Controls	87.64	97.33	9.69	
PICT	116.86	127.11	10.25	0.56
A&E visits				
Controls	25.91	28.64	2.73	
PICT	31.42	34.11	2.69	−0.04
ACSCs				
Controls	0.59	0.66	0.07	
PICT	0.85	0.85	0	−0.07
Re-admissions (30 days)				
Controls	0.87	0.84	−0.03	
PICT	1.22	1.13	−0.09	−0.06
Secondary outcomes				
Total cost of secondary care services (£)				
Controls	29157.28	30530.70	1373.42	
PICT	38923.43	39167.17	243.74	−1129.68
Length of stay (days)				
Controls	60.32	50.97	−9.35	
PICT	77.24	66.21	−11.03	−1.68
Patient satisfaction (general)				
Controls	0.35	0.38	0.03	
PICT	0.38	0.37	−0.01	−0.04
Patient satisfaction (LTC specific)				
Controls	0.13	0.14	0.01	
PICT	0.14	0.16	0.02	0.01

A&E, accident and emergency; ACSCs, ambulatory care sensitive conditions; LTC, long-term condition; PICT, Practice Integrated Care Team.

MDTs in general, and of course the case management those high-risk patients received).[31] So, if the main driver of results was the MDT working, we may plausibly expect these effects to differ by risk group (ie, the general practice being on average at lower risk).

Finally, direct and spillover effects may plausibly act through distinct mechanisms. There are some indications of wider system effects of integrated care in the literature. For example, good team 'climate' (ie, professional integration)[14] has been linked to superior clinical care for a number of LTCs,[32] although evidence of causation is currently lacking.[33] This is one potential mechanism that the MDT spillover effects could act through. Spillover effects, therefore, may not be dependent on the numbers captured by MDTs directly, because they go via the GP and wider care team. If practices 'do' MDT for a few patients, it may influence their care for everyone.

Strengths and weaknesses of the study

Our method of analysis, DD, is a robust method under certain conditions that we tested.[11] We only saw potential bias indicated by non-parallel preintervention trends for a single outcome measure at the individual level (outpatient visits), and we employ robustness checks beyond the primary analysis models. The method allows testing of a complex intervention in routine practice, with potential for greater external validity and generalisability of the findings.[12]

Our results at both levels are plausible. At the individual level, we observed very little differences between the groups, as we would expect from previous literature around this intervention type.[6] At the practice level, the effect we observed was on an outcome (inpatient non-elective admissions) the intervention aimed to affect.[16]

However, our study does suffer from a number of weaknesses. Unfortunately, due to the implementation

Table 7 Practice-level adjusted model results

Outcome	Adjusted* intervention effect (95% CI) (difference per 1000 patients per month) Linear regression model	Effect size†
Primary outcomes		
Inpatient non-electives	−0.63 (−1.17 to −0.09)‡	−0.24
Inpatient electives	0.19 (−0.47 to 0.86)	0.07
Outpatient visits	−2.80 (−9.84 to 4.24)	−0.08
A&E visits	−1.32 (−3.52 to 0.89)	−0.13
ACSCs	−0.04 (−0.15 to 0.06)	−0.08
Readmissions (30 days)	−0.10 (−0.25 to 0.05)	−0.16
Secondary outcomes		
Total cost of secondary care services (£)	−505.73 (−2763.35 to 1751.89)	−0.04
Length of stay (days)	−0.24 (−7.56 to 7.08)	−0.01
Patient satisfaction (general)	−0.03 (−0.09 to 0.02)	−0.24
Patient satisfaction (LTC specific)	0.01 (−0.04 to 0.05)	0.14

N=1650 observations; 30 practices (period November 2010–March 2015).
*Adjusted for practice and time fixed effects with robust SEs.
†Standardised mean difference.
‡Significant at p<0.05.
A&E, accident and emergency; ACSCs, ambulatory care sensitive conditions; LTC, long-term condition.

of the intervention, we were not able to access individual-level data until before the point where nearly all practices implemented the intervention. This is due to an initial problem at the CCG of consenting data use for those individual patients initially included early in the intervention. This limits our ability to ascertain whether the initially recruited patients at each practice were significantly different, or benefited more or less than those recruited later to the intervention. It also prevents direct comparison of the results we saw at the practice level with those at the individual level over exactly the same period of time and limits our ability to look at any longer term effects of the intervention at the individual level. Furthermore, if spillover effects did indeed affect other patients in the practice, then the individual-level effects may be driven towards the null. This is similarly true for the DDD analysis conducted. However, these spillover effects were not strongly indicated at the practice level.

With the intervention so widespread (particularly important for an intervention incentivised nationally), we were extremely careful to choose our comparators (a crucially important step in DD analysis). We chose practices (within the same CCG) for which we knew for definite their intervention status at any time point for the practice-level analysis. Nonetheless, practices volunteered for the intervention, which can potentially introduce some selection bias at the practice level. However, we estimate this possible selection effect to be minimal based on observable practice characteristics. A common limitation of non-experimental studies, however, is we cannot discount differences based on unobservables. Adding practice fixed effects controls for any differences between practices that persist over time, as well as any hospital-level changes during the period that affect all practices.

At the individual level, we matched patients using propensity scores within the CCG achieving the necessary parallel pretrends. However, the intervention patients are selected for their immediate risk, while the control patients were selected based on their matched risk at an earlier date, which may have subsequently subsided (and hence be the reason they were indeed not recruited to the intervention). With 'risk', and so recruitment, defined on time-variant indicators, and so transient over time, there is potential for some bias in favour of the control group for the individual-level results in this analysis. However, with patients well matched at the initial start date, we expect to have minimised this bias.

An important weakness, constrained by the data available to us, is we were not able to analyse outcomes beyond secondary care utilisation and total cost of secondary care. While these utilisation outcomes reflect well the explicit aims of the intervention, they do not allow for a broad representation of the intervention in terms of other important potential outcomes—for example, patient health, quality of life and satisfaction with care. These additional measures could be considered when making commissioning decisions, although they were not the primary stated aim.

Results in relation to other studies

Our recent systematic review and meta-analysis looking at similar interventions likewise showed little effect across relevant health system outcomes for those involved in the intervention directly (ie, non-significant estimated pooled ES of 0.04 for secondary care use in the short term, and −0.02 in the long term).[6] However, the review did show a clear benefit in terms of patient satisfaction for these patients (statistically significant estimated pooled ES of 0.26 in the short term, and 0.35 in the long term). We were unable to replicate this finding

in this study, perhaps due to the data available to us that only allowed us to look at this domain at the practice level, which is likely to be less sensitive. We hypothesised from the results of our review's subgroup analyses that case management by an MDT and involving a social worker may be more effective than other examples also included in the review (eg, single nurse case manager). Results of this subsequent study do not support this previous hypothesis. However, we also suggested that 'low-strength' primary care systems[17] may benefit more from the intervention (where case management may substitute for a strong primary care system). This may explain this deviation from the results of our review, which drew on evidence predominantly from a 'low-strength' primary care country (USA).

Looking at spillover effects from MDT case management was a strength of this paper.[14] Only a few other studies have looked at spillover effects, most notably, evaluation of the Evercare intervention.[34] However, Evercare used only a single case manager, where we might not expect to find large effects, and the study identified no spillover.[34] Analysis of MDT case management in the English 'Integrated Care Pilots' (ICP) likewise looked at direct and spillover effects. Roland *et al* identified an increase in emergency admissions and a decrease in elective admissions and outpatient attendances at the individual level. At the practice level, they identified a slight reduction in outpatient attendances. It is, however, difficult to compare these results directly with this study, with the ICP analysis evaluating six separate sites in combination, each offering slight alterations of MDT case management to different populations. Nevertheless, key differences that stand out include the presence of a social worker in the case management team in this intervention (only two smaller sites in the ICP identified input from a social worker); physical MDT meetings in this intervention rather than 'virtual ward' rounds (as in the ICP sites); and the GP as clinical lead in this intervention, rather than the primarily nurse-led interventions in ICP sites.[35]

Implications for clinicians and policymakers

This study provides further evidence of the limited effectiveness of MDT case management aimed at generally 'at-risk' patients as a tool to reduce care utilisation. MDT case management targeted at high-risk patients importantly does not achieve its primary aim: reducing emergency admissions for those high-risk patients directly managed. Therefore, there may be better alternatives to this intervention, which may be other forms of case management targeted at specific conditions, which have some evidence of beneficial results—for example targeting mental health.[36] Aiming at a very small number of high-risk patients may never alleviate health system pressures alone,[30] and even the potential spillover effects of increased professional integration that may result may not be of sufficient magnitude to achieve the desired effects.

Going beyond the case management model to a more population-based approach may therefore be another avenue to explore, for example, colocation of services, or integrated electronic health records for all patients rather than just a high-risk cohort—interventions further removed from the service delivery level, but which may be regarded as a key foundation for multidisciplinary professional communication and working. We have shown here that this greater professional integration may have scope for improving measurable health system outcomes.

Future research

More work is needed to confirm these initial findings of potentially beneficial spillover effects, particularly qualitative work and process evaluation identifying plausible mechanisms. These did not stand up to our robustness checks in this analysis; however, the indication was always in the direction of favouring the intervention practices with regard to decreasing non-elective admissions at the practice level. Where it is possible, future studies looking at models of integrated care should consider spillover effects.

If commissioning bodies consider evaluation using similar robust, but cost-effective methods in the future, they should be planned from the beginning, where potential bias (discussed above) could be easily avoided. For example, a randomised stepped-wedge design may be an appropriate alternative.[37]

While we improved on previous literature by including a measure of multimorbidity in our study, we only included the most basic of these, a simple count of diseases.[38] Our future research will explore outcomes stratified by different 'types' of multimorbidities, to observe if the intervention can be better targeted for the patients it directly affects, providing a more effective and efficient method of exploiting the potential for wider system effects.

CONCLUSIONS

We show that MDT case management does not fulfil its primary aim, preventing emergency admissions for the high-risk patients it targets. This accords with our previous findings. We show here that the highest risk patients (as identified by the risk tool) receiving the intervention in fact slightly increased admissions in many domains targeted for decrease by the intervention. We do, however, show some indications of beneficial spillover effects of MDT working at the practice level worthy of further exploration. The results highlight the importance of ongoing work on effective ways of avoiding admissions.[36]

Author affiliations
[1]NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
[2]Manchester Centre for Health Economics, University of Manchester, Manchester, UK

[3]NIHR School for Primary Care Research, Centre for Primary Care, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

Twitter Follow Jonathan Stokes at @DukeStokes and Peter Bower at @Bowercpcman

Acknowledgements The authors wish to thank Central Manchester CCG for the data they provided and the time they took to explain to us the details of the intervention.

Contributors JS has had the main responsibility for calculating statistics and writing the paper. JS, SRK, KC and PB designed the analysis plan, and drafted and revised the paper. JS and SRK conducted the analysis.

Funding This work was funded by the National Institute for Health Research Greater Manchester Primary Care Patient Safety Translational Research Centre (NIHR GM PSTRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Office for National Statistics. *An executive summary, 2010-based NPP Reference Volume*. Office for National Statistics, 2012:1–5.
2. Uijen AA, van de Lisdonk EH. Multimorbidity in primary care: prevalence and trend over the last 20 years. *Eur J Gen Pract* 2008;14(Suppl 1):28–32.
3. Nolte E, McKee M. *Caring for people with chronic conditions: a health system perspective*. McGraw-Hill International, 2008.
4. World Health Organization. WHO global strategy on people-centred and integrated health services: interim report. 2015.
5. Armitage GD, Suter E, Oelke ND, *et al*. Health systems integration: state of the evidence, 2009;9:e82.
6. Stokes J, Panagioti M, Alam R, *et al*. Effectiveness of case management for 'at risk' patients in primary care: a systematic review and meta-Analysis. *PLoS ONE* 2015;10:e0132340.
7. Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications--a discussion paper. *Int J Integr Care* 2002;2:e12.
8. Valentijn PP, Vrijhoef HJ, Ruwaard D, *et al*. Towards an international taxonomy of integrated primary care: a Delphi consensus approach. *BMC Fam Pract* 2015;16:64.
9. Department of Health. Health and Social Care Act 2012. http://www.legislation.gov.uk/ukpga/2012/7/contents/enacted
10. Stokes J. CCG implementation of integrated care in the NHS. *BMC Health Serv Res* 2014;14 (2):P119.
11. Maciejewski ML, Bayliss EA. Approaches to comparative effectiveness research in multimorbid populations. *Med care* 2014;52(Suppl 3):S23–30.
12. Steventon A, Grieve R, Bardsley M. 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. *Med Decis Making* 2015;35:1023–36.
13. Bower P, Sibbald B. Do consultation–liaison services change the behavior of primary care providers? A review. *Gen Hosp Psychiatry* 2000;22:84–96.
14. Valentijn PP, Schepman SM, Opheij W, *et al*. Understanding integrated care: a comprehensive conceptual framework based on the integrative functions of primary care. *Int J Integr Care* 2013;13: e010.
15. Ross S, Curry N, Goodwin N. *Case management: what it is and how it can best be implemented*. King's Fund, 2011.
16. Central Manchester Clinical Commissioning Group. Practice Integrated Care Teams Care Planning Toolkit v.2.6, 2013.
17. Starfield B, Shi L. Policy relevant determinants of health: an international perspective. *Health Policy* 2002; 60:201–18.
18. Angrist JD, Pischke JS. *Mostly harmless econometrics: an empiricist's companion*. Princeton university press, 2008.
19. Propper C, Croxson B, Shearer A. Waiting times for hospital admissions: the impact of GP fundholding. *J Health Econ* 2002;21:227–52.
20. Besley T, Burgess R. Can labor regulation hinder economic performance? Evidence from India. *Q J Econ* 2004;119:91–134.
21. Long JS, Freese J. *Regression models for categorical dependent variables using Stata*. 3rd edn. College Station, TX: Stata Press, 2014.
22. Hernan MA, Hernandez-Diaz S, Werler MM, *et al*. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol* 2002;155:176–84.
23. Rice N, Jones A. Multilevel models and health economics. *Health Econ* 1997;6:561–75.
24. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? *Q J Econ* 2004;119:249–75.
25. Lipsey MW, Wilson DB. *Practical meta-analysis*. SAGE Publications, 2001.
26. Cohen J. *Statistical power analysis for the behavioral sciences*: L. Erlbaum Associates, 1988.
27. Baines DL, Whynes DK. Selection bias in GP fundholding. *Health Econ* 1996;5:129–40.
28. Department of Health. Reference costs 2012–13, 2013; https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/261154/nhs_reference_costs_2012-13_acc.pdf
29. NHS England. 2013/14 general medical services (GMS) contract—guidance and audit requirements for new and amended enhanced services: Risk profiling and care management scheme. In: Employers N, editor, 2013.
30. Roland M, Abel G. Reducing emergency admissions: are we on the right track? *BMJ* 2012;345:e6017.
31. Hall Aitken. Central Manchester Practice Integrated Care Team Report. Manchester: Hall Aitken, 2014:1–20.
32. Bower P, Campbell S, Bojke C, *et al*. Team structure, team climate and the quality of care in primary care: an observational study. *Qual Saf Health Care* 2003;12:273–9.
33. Bramwell D, Peckham S, Allen P, *et al*. How can GPs and community health services work more effectively together? *Br J Gen Pract* 2015;65:374–5.
34. Gravelle H, Dusheiko M, Sheaff R, *et al*. Impact of case management (Evercare) on frail elderly patients: controlled before and after analysis of quantitative outcome data. *BMJ* 2007;334:31–31.
35. Roland M, Lewis R, Steventon A, *et al*. Case management for at-risk elderly patients in the English integrated care pilots: observational study of staff and patient experience and secondary care utilisation. *Int J Integr Care* 2012;12:e130.
36. Purdy S. *Avoiding hospital admissions: what does the research evidence say*. King's Fund, 2010.
37. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006;6:54.
38. Huntley AL, Johnson R, Purdy S, *et al*. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med* 2012;10:134–41.

BMJ Open

Effectiveness of multidisciplinary team case management: difference-in-differences analysis

Jonathan Stokes, Søren Rud Kristensen, Kath Checkland and Peter Bower

BMJ Open 2016 6:
doi: 10.1136/bmjopen-2015-010468

6.2 Chapter Conclusions

In this chapter, I have again demonstrated a lack of effectiveness of the case management model regarding the high-risk patients directly treated in terms of the primary aim of the intervention, i.e. decreasing utilisation of secondary care, and so costs of care. However, I have shown that there may be evidence of positive spill-over effects of MDT working (although further work is required, as these did not stand up to the robustness check). Direct and spill-over effects are hypothesised to act through different mechanisms. Next, I examine the effects of the same intervention on subgroups of multimorbid patients.

7. MDT case management and 'type' of multimorbidity

7.1 Chapter Introduction

This empirical piece explores subgroup effects of the case management model evaluated in the last chapter using difference-in-difference-in-differences. In Chapter 1, I introduced the multimorbidity concept as potentially defining those patients most likely to benefit from integrated care, but also a concept that can be implemented in a variety of forms. Here, I look at the dominant approaches to measuring multimorbidity, as identified in a previous systematic review, and examine if modelling multimorbidity in different ways alters the results in terms of effectiveness of case management. The article is currently under review in *BMC Health Services Research*.

Any supplementary material referred to in the text is available in Appendix 4.

BMC Health Services Research

Does the impact of case management vary in different subgroups of multimorbidity? Secondary analysis of a quasi-experiment

--Manuscript Draft--

Manuscript Number:	
Full Title:	Does the impact of case management vary in different subgroups of multimorbidity? Secondary analysis of a quasi-experiment
Article Type:	Research article
Section/Category:	Organization, structure and delivery of healthcare
Funding Information:	National Institute for Health Research — Not applicable
Abstract:	Background Health systems must transition from catering primarily to acute conditions, to meet the increasing burden of chronic disease and multimorbidity. Case management is a popular method of integrating care, seeking to accomplish this goal. However, the intervention has shown limited effectiveness. We explore whether the effects of case management vary in patients with different types of multimorbidity. Methods We extended a previously published quasi-experiment (difference-in-differences analysis) with 2049 propensity matched case management intervention patients, adding an additional interaction term to determine subgroup effects (difference-in-difference-in-differences) by different conceptualisations of multimorbidity: 1) Mental-physical comorbidity versus others; 2) 3+ chronic conditions versus <3; 3) Discordant versus concordant conditions; 4) Cardiovascular/metabolic cluster conditions only versus others; 5) Mental health-associated cluster conditions only versus others; 6) Musculoskeletal disorder cluster conditions only versus others 7) Charlson index >5 versus others. Outcome measures included a variety of secondary care utilisation and cost measures. Results The majority of conceptualisations suggested little to no difference in effect between subgroups. Where results were significant, the vast majority of effect sizes identified in either direction were very small. The trend across the majority of the results appeared to show very slight increases of admissions with treatment for the most complex patients (highest risk). The exceptions to this, patients with a Charlson index >5 may benefit slightly more from case management with decreased ACSC admissions (effect size (ES): -0.06) and inpatient re-admissions (30 days, ES: -0.05), and patients with only cardiovascular/metabolic cluster conditions may benefit slightly more with decreased inpatient non-elective admissions (ES: -0.12). Only the three significant estimates for the musculoskeletal disorder cluster met the minimum requirement for at least a 'small' effect. Two of these estimates in particular were very large. This cluster represented only 0.5% of the total patients analysed, however, so is hugely vulnerable to the effects of outliers, and makes us extremely cautious of interpreting these as 'real' effects. Conclusions Our results indicate no appropriate multimorbidity subgroup at which to target the case management intervention in terms of secondary care utilisation/cost outcomes. The most complex, highest risk patients may legitimately require hospitalisation, and the intensified management may better identify these unmet needs. End of life patients (e.g. Charlson index >5)/those with only conditions particularly amenable to primary care management (e.g. cardiovascular/metabolic cluster conditions) may benefit a very small amount more than others.
Corresponding Author:	Jonathan Stokes, MPH, BSc University of Manchester, NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre Manchester, UNITED KINGDOM

Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Manchester, NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre
Corresponding Author's Secondary Institution:	
First Author:	Jonathan Stokes, MPH, BSc
First Author Secondary Information:	
Order of Authors:	Jonathan Stokes, MPH, BSc
	Søren Rud Kristensen
	Kath Checkland
	Sudeh Cheraghi-Sohi
	Peter Bower
Order of Authors Secondary Information:	
Opposed Reviewers:	

Does the impact of case management vary in different subgroups of multimorbidity? Secondary analysis of a quasi-experiment

Jonathan Stokes. NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. Jonathan.stokes-3@postgrad.manchester.ac.uk

Søren Rud Kristensen. Manchester Centre for Health Economics, University of Manchester, Manchester, UK. soren.kristensen@manchester.ac.uk

Kath Checkland. NIHR School for Primary Care Research, Centre for Primary Care, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. Katherine.H.Checkland@manchester.ac.uk

Sudeh Cheraghi-Sohi. NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. Sudeh.Cheraghi-sohi@manchester.ac.uk

Peter Bower. NIHR Greater Manchester Primary Care Patient Safety Translational Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. peter.bower@manchester.ac.uk

Corresponding author contact information:

Jonathan Stokes
University of Manchester
Centre for Primary Care
Williamson Building, Oxford Road
Manchester, M13 9PL UK
Email: jonathan.stokes-3@postgrad.manchester.ac.uk
Telephone: +44 161 275 0758

Abstract

Background

Health systems must transition from catering primarily to acute conditions, to meet the increasing burden of chronic disease and multimorbidity. Case management is a popular method of integrating care, seeking to accomplish this goal. However, the intervention has shown limited effectiveness. We explore whether the effects of case management vary in patients with different types of multimorbidity.

Methods

We extended a previously published quasi-experiment (difference-in-differences analysis) with 2049 propensity matched case management intervention patients, adding an additional interaction term to determine subgroup effects (difference-in-difference-in-differences) by different conceptualisations of multimorbidity: 1) Mental-physical comorbidity versus others; 2) 3+ chronic conditions versus <3; 3) Discordant versus concordant conditions; 4) Cardiovascular/metabolic cluster conditions only versus others; 5) Mental health-associated cluster conditions only versus others; 6) Musculoskeletal disorder cluster conditions only versus others 7) Charlson index >5 versus others. Outcome measures included a variety of secondary care utilisation and cost measures.

Results

The majority of conceptualisations suggested little to no difference in effect between subgroups. Where results were significant, the vast majority of effect sizes identified in either direction were very small. The trend across the majority of the results appeared to show very slight increases of admissions with treatment for the most complex patients (highest risk). The exceptions to this, patients with a Charlson index >5 may benefit slightly

more from case management with decreased ACSC admissions (effect size (ES): -0.06) and inpatient re-admissions (30 days, ES: -0.05), and patients with only cardiovascular/metabolic cluster conditions may benefit slightly more with decreased inpatient non-elective admissions (ES: -0.12).

Only the three significant estimates for the musculoskeletal disorder cluster met the minimum requirement for at least a 'small' effect. Two of these estimates in particular were very large. This cluster represented only 0.5% of the total patients analysed, however, so is hugely vulnerable to the effects of outliers, and makes us extremely cautious of interpreting these as 'real' effects.

Conclusions

Our results indicate no appropriate multimorbidity subgroup at which to target the case management intervention in terms of secondary care utilisation/cost outcomes. The most complex, highest risk patients may legitimately require hospitalisation, and the intensified management may better identify these unmet needs. End of life patients (e.g. Charlson index >5)/those with only conditions particularly amenable to primary care management (e.g. cardiovascular/metabolic cluster conditions) may benefit a very small amount more than others.

Keywords

Multimorbidity, Case management, Integrated care

Background

It is widely agreed that health systems must transition from catering primarily to acute conditions, to meet the increasing burden of chronic disease and multimorbidity. New care models have been called for to achieve this goal, based in primary care and 'integrated' with wider health and care sectors [1]. In practice, these new interventions and models are based primarily around the concept of 'case management' [2, 3]. Case management involves case finding (identifying 'high-risk' individuals to case manage), individual assessment, care planning and care co-ordination (with regular review, monitoring and adaptation of the care plan) [4]. Many use a multidisciplinary team (MDT) to case manage, which may combine 'professional integration' with co-ordination activity [5].

The logic of the case management model rests on the presence of so-called 'super-utilizers' [6]. This is a small group of 'high-risk' patients (almost exclusively multimorbid [7]) that utilize a disproportionate amount of healthcare resource. For example, the United States Government Accountability Office estimate that the most expensive 5% of Medicaid-only enrolees accounted for 48% of Medicaid-only health spending each year from 2009 to 2011 [8]. The assumption of case management is that by targeting additional and individually tailored primary care at these patients, more costly secondary care admissions (particularly emergency admissions) can be avoided. Thus, overall healthcare spending costs can be reduced, patients can be treated in an environment more satisfactory to them, and in a more holistic and preventative way [9].

However, research suggests that case management does not meet its primary aims for those directly managed, although patient satisfaction does appear to be increased [10, 11].

In practice, MDT case management tends to target those identified as 'high risk' using some sort of algorithm. These tools generate a heterogeneous group of patients, and it may be that there are subgroups for which the direct effects of the intervention are more effective. There are a number of ways of conceptualising multimorbidity, and little evidence as to the advantages and disadvantages of each [12, 13]

These conceptualisations can broadly be categorised into four distinct groups [13]. Within each of these groups, we outline the specific measures we focus on in our analysis with justification for doing so.

1. Simply counting the number of chronic conditions/medications from a pre-specified list.
    a. **Count of diseases –** A simple count of conditions is the most basic conceptualisation of multimorbidity, and the most ubiquitous multimorbidity measure in the current literature [14].
2. Grouping chronic diseases by dyads or triads (i.e. clustering).
    a. **Disease clusters –** As the search for more clinically meaningful conceptualisations of multimorbidity advances, identification of 'non-random clustering' of diseases (i.e. associative multimorbidity) has been a common theme in the literature. A number of plausible mechanisms can explain disease clustering, as outlined by Valderas et al [15]. Furthermore, identifying commonly co-occurring sets of diseases may have important implications for developing clinical guidelines for patients with multiple chronic conditions, for which the current single-disease approach can cause issues [16].

3. Using an index of variable complexity.

    a. **Charlson index –** A number of indices of multimorbidity have emerged, to attempt to better account for the variation in severity of different diseases. No index is more established and prevalent in the literature than the Charlson index, which weights (giving a score between 1 and 6) relevant diseases, and gives a summed score of the weights to the individual [14].

4. Identifying homogeneous groups of people with common diseases and characteristics.

    a. **Mental-physical comorbidities –** These patients are at increased risk for active and precursors to patient safety incidents in primary care [17]. Moreover, a common mental health condition, depression, has been shown to be particularly important in modifying multimorbidity management and outcomes [18].

    b. **Discordant comorbidities –** i.e. co-occurring diseases that are not managed synergistically. Discordant conditions are likely to add to the complexity of clinical treatment/decision-making [19], potentially putting this group at increased risk of management failures. This is opposed to concordant conditions (i.e. co-occurring diseases that are managed synergistically).

To explore whether the effects of case management vary in patients with different types of multimorbidity, we extend an existing analysis [20] using these different definitions of multimorbidity to stratify patients within the sample.

Methods

Detailed methods for the analysis models can be found in our previous publication and the accompanying appendix (see individual-level direct effects model/high-risk subgroup effects model) [20].

Briefly, we evaluated effectiveness of a MDT case management intervention (practice integrated care teams – PICT) in a single Clinical Commissioning Group (CCG) in the UK NHS. We matched 2049 intervention patients one-to-one with controls from the same CCG using propensity scores. We analysed anonymised data held by the CCG relating to hospital admissions and costs. Outcome measures were summed to a count per patient per month over the period September 2010 to March 2015 inclusive, allowing a three-year pre-intervention trend period (patients joined the intervention gradually from September 2013 to February 2015). Primary outcome measures included A&E visits; inpatient non-elective stays, 30-day re-admissions; inpatient elective stays; outpatient visits; and admissions for ambulatory care sensitive conditions. Secondary measures included: inpatient length of stay; total cost of secondary care services.

Previously, we used difference-in-differences (DD) analysis. This method is a quasi-experiment, where the intervention group is compared to a control group (constructed retrospectively, and known to be unaffected by the intervention). The method compares the difference in a measured outcome between intervention and control groups in a pre-intervention period (difference 1), and compares to the difference between the two groups after an intervention is introduced (difference 2), attributing this second difference to the intervention effect. We used a time fixed effect instead of the usual binary post dummy to

account for the gradual intervention joining, comparing appropriately at each time point [21, 22].

Here, we use difference-in-difference-in-differences (DDD) analysis, which adds an additional interaction term to a standard DD approach used in the original study. This allows us to observe subgroup effects (of the different multimorbidity conceptualisations) of the intervention, i.e. the effects of the intervention on a subgroup over and above any baseline DD effect. The models were negative binomial count models (except for total cost of secondary care which was better represented by a zero-inflated negative binomial model based on admission events). We adjusted for age, index of multiple deprivation (IMD) domains (excluding health), practice- and time-fixed effects. We clustered our standard errors by practice to deal with concerns of serial correlation [23]. We report the effect size (standardised mean difference) as a measure of practical significance, calculated from the average partial effect (reported in the Additional file). All data preparation and analysis was carried out using STATA (version 13) [24].

From a list of 20 chronic conditions (see Additional file) represented in the NHS quality and outcomes framework (QOF), we created a number of dummy variables representing different conceptualisations of multimorbidity:

- Count of diseases - a simple cumulative count of the individual's conditions. We created a dummy for 3+ (thought to be a more discriminating definition of multimorbidity than 2+ conditions, better identifying patients with higher needs [25]) versus <3 conditions.

- Disease clusters – We stratified by three common disease clusters identified in the literature [16]. We created a dummy comparing those that only have one or more of the diseases in that cluster versus the rest of the patients in the analysis.
    - Cardiovascular/metabolic cluster: Diabetes, Hypertension, Chronic Heart Disease, Obesity
    - Mental health-associated cluster: Mental health condition (Schizophrenia/Bipolar Disorder/Psychoses/Depression), Hypothyroidism, Dementia, Asthma, Chronic Obstructive Pulmonary Disease (COPD), Rheumatoid Arthritis, Obesity
    - Musculoskeletal disorder cluster: Rheumatoid Arthritis, Osteoporosis, Obesity
- Charlson index - an established measure, with its own set of relevant chronic conditions and weightings, we used the STATA command '*charlson*' to record a Charlson index for each participant [26]. We created a dummy comparing those with a Charlson index >5 (suggested in the literature to be those patients at 'highest risk' of negative outcomes [27]) to all other patients,.
- Mental-physical comorbidities – There were a number of mental health conditions in our disease list (depression, schizophrenia, bipolar disorder, psychoses, and dementia). We created a dummy variable comparing those patients with both a mental and physical conditions to all other patients.
- Discordant comorbidities - we used a list of determined concordant conditions that share a vascular aetiology and common chronic management and treatment goals (coronary heart disease, chronic kidney disease, diabetes, hypertension, heart failure, stroke/transient ischaemic attack, atrial fibrillation, and peripheral

vascular disease), and classified the remainder as discordant conditions [28]. We created a dummy variable comparing patients with discordant conditions to all other patients.

With multiple comparisons, the risk of type I errors is inflated [29]. As a sensitivity analysis, we subsequently correct the results for multiple testing using the Holm-Bonferoni adjustment for multiple comparisons to identify potential false positive results [30]. We report those results that remained significant with the adjusted threshold at the end of the results section.

Results

Each of the dummy variables selected a different proportion of the total patient group analysed (n=4098), as complex/'higher risk' (see Table 1). The proportion of patients selected in total varied by measure. The count (3 or more chronic diseases compared to all others), and discordant comorbidities dummies were least selective, including over half of the total patient group in each case. The specific disease clusters were, as expected, the most selective. The musculoskeletal disorder cluster (rheumatoid arthritis, osteoporosis and obesity), for example, selected only 0.5% of the total patient group in the positive dummy group.

Table 1: Number (and proportion) of patients positively identified by each dummy variable

	Number of patients positively identified in that dummy measure		
Multimorbidity measure	**PICT (n=2049)**	**Controls (n=2049)**	**Total (n=4098)**
Mental/physical condition	319 (15.6%)	498 (24.3%)	817 (19.9%)
3 + conditions	992 (48.4%)	1165 (56.9%)	2157 (52.6%)
Discordant comorbidities	1030 (50.2%)	1208 (59.0%)	2238 (54.6%)
Cardiovascular/metabolic cluster	300 (14.6%)	209 (10.2%)	509 (12.4%)
Mental health-associated cluster	187 (9.1%)	234 (11.4%)	421 (10.3%)
Musculoskeletal disorder cluster	12 (0.6%)	8 (0.4%)	20 (0.5%)
Charlson >5	398 (19.4%)	347 (16.9%)	745 (18.2%)

PICT = Practice integrated care teams (intervention)

Table 2 shows the correlation between each of the dummy variables used to select patients. The less selective dummy variables (mental/physical conditions; 3 or more conditions; and discordant comorbidities) were well correlated with each other. The remaining more selective measures (the three literature clusters; Charlson index) were far less correlated with any other.

Table 2: Correlation between multimorbidity dummies used

	Mental/ physical condition	3 + conditions	Discordant comorbidities	Cardiovascular/metabolic cluster	Mental health-associated cluster	Musculoskeletal disorder cluster	Charlson >5
Mental/physical condition	1.00						
3 + conditions	0.32	1.00					
Discordant comorbidities	0.45	0.65	1.00				
Cardiovascular/metabolic cluster	-0.19	-0.28	-0.37	1.00			
Mental health-associated cluster	-0.01	-0.30	-0.16	-0.10	1.00		
Musculoskeletal disorder cluster	-0.03	-0.07	-0.08	0.08	0.16	1.00	
Charlson >5	-0.02	0.01	-0.02	-0.17	-0.16	-0.03	1.00

Figure 1 shows a forest plot for each outcome measure, comparing the multimorbidity subgroups within each. The effect sizes represent the impact of the treatment for the group identified as highest risk (i.e. those positively identified as multimorbid according to that conceptualisation) within each dummy. Estimates which lie to the left of the line of no effect (0) favour the intervention for that multimorbid-defined group (i.e. indicate decreased utilisation/cost). Estimates that lie to the right of the line show that intervention treatment led to increased utilisation/cost for that group and outcome measure. Only estimates whose 95% confidence interval does not cross the line of no effect are statistically significant. The

table in the Additional file gives the full list of regression results (i.e. adjusted intervention effect as the difference per patient per month) that these effect sizes were calculated from.

[insert Figure 1]

Figure 1: Forest plot comparing multimorbidity measures by each outcome. Effect size = standardised mean difference. * = statistically significant result (p<0.05). ACSCs = admissions for ambulatory care sensitive conditions

The results from Figure 1 indicate that for the majority of outcomes, the different conceptualisations of multimorbidity have given broadly the same interpretation. For inpatient electives, and total cost of secondary care outcomes, none of the conceptualisations held a statistically significant difference. For outpatient visits (discordant conditions; effect size (ES): 0.08), A&E visits (mental-physical comorbidities; ES: 0.05), and ACSCs (Charlson index >5; ES: -0.06), there was a single (but different) multimorbidity concept which differed significantly. However, all of these differences were to a very small extent (effect size less than the threshold for a 'small' effect, 0.2, in all cases [31]).

Most notably, the musculoskeletal disorder cluster deviated significantly from the other results across a number of outcomes (inpatient non-electives, ES: 3.44; 30-day re-admissions, ES: 2.38; and length of stay, ES: 0.37). In each of these measures, the results suggested that the intervention increased the utilisation of inpatient services for patients with these conditions, with large effect sizes (particularly for inpatient non-electives and 30-day re-admissions).

The significant results for the other multimorbidity measures were far more conservative, but tended to follow this similar trend, suggesting that the highest risk patients increased utilisation following treatment. The exception to this was for the Charlson index, where the effect of case management on those with Charlson >5 led to a larger decrease than those with Charlson <5 for ambulatory care sensitive conditions (ACSCs – effect size (ES): -0.06) and 30-day inpatient re-admissions (ES: -0.05). Those patients with only conditions from the cardiovascular/metabolic cluster also appeared to have slightly fewer inpatient non-elective admissions following treatment (ES: -0.12).

After Holm-Bonferroni correction was applied to all results, only two of the statistically significant results held: the findings of significant increases following treatment of inpatient non-elective admissions and 30-day re-admissions for patients with musculoskeletal disorder cluster conditions.

Discussion

Summary of findings

Do different operationalisations of multimorbidity give different answers in terms of MDT case management effectiveness, and if so, why?

As outlined above, the majority of conceptualisations gave very similar results in relation to effectiveness of the intervention, suggesting little to no difference in effect between subgroups. Where results were significant, the vast majority of effect sizes identified in either direction were very small. Cohen's rule of thumb for interpreting effect sizes is that 0.2 indicates a small effect, 0.5 a medium, and 0.8 a large effect [31]. Only the three significant estimates for the musculoskeletal disorder cluster (rheumatoid arthritis, osteoporosis and obesity) met the minimum requirement for at least a 'small' effect. Two of these estimates in particular were very large (inpatient non-elective admissions ES: 3.44; and 30-day re-admissions ES: 2.38). Musculoskeletal disorders are associated with some of the poorest quality of life, particularly because of the bodily pain and poorer level of physical functioning associated with them [32]. These conditions are difficult to manage leading to a large economic and social burden (with a large proportion of the cost burden due to hospital inpatient admissions) [33]. Perhaps, this difficulty in managing these complex conditions is not alleviated by the MDT, but simply draws attention to unmet need requiring escalation to emergency services? However, interpretation of these large effect sizes in light of the tiny numbers of patients in this group (only 0.5% of the total population analysed) makes us extremely cautious of interpreting these as 'real' effects. These small numbers make this subgroup result hugely vulnerable to the effects of outliers.

What do these results imply for the intervention in practice?

The trend across the majority of the results appeared to show very slight increases of admissions with treatment for the most complex patients (highest risk), again, perhaps indicating identification of unmet need which may plausibly result from an intervention of this type. This would also agree with the sub-group finding in our original analysis paper, where we stratified by risk-tool score, finding that the highest risk patients treated appeared to benefit least [20].

However, an important exception to this finding was for those patients with a Charlson index >5, where those patients may benefit from slightly decreased ACSC admissions (ES: -0.06) and 30-day inpatient re-admissions (ES: -0.05) with the intervention. The Charlson index, unlike the other measures, was developed primarily as a method of predicting mortality [34]. Those with an index greater than or equal to 5 have an 85% chance of 1-year mortality [27]. Perhaps these generally 'end of life' patients are managed differently by the case management team than others? For instance, the treatment plan may be more focused on decreasing the burden of care (e.g. medications – perhaps being less likely to have adverse reactions) at the end of life, or attention may shift to palliative care closer to home resulting in different secondary care use?

Those with conditions only from the cardiovascular/metabolic cluster (diabetes, hypertension, chronic heart disease, obesity), may also have benefited from slightly decreased inpatient non-elective admissions (ES: -0.12). These conditions should all be manageable in primary care [35], perhaps explaining that hospital admissions for these patients may be most susceptible to decrease with increased primary care. However, general ACSC admissions were not significantly affected for this patient group (although the trend was in the same direction; ES: -0.15).

Strengths and weaknesses

As with any subgroup analysis, power to detect subgroup effects will suffer proportionately more than power to detect the overall effect [36]. Therefore, we present this as a very preliminary analysis, results of which may be added to by future studies using similar conceptualisations of multimorbidity. In addition, the interpretation of the subgroup effects should be made in light of the overall effect. These represent the differential effects of the intervention in terms of that subgroup, and may be in the context of an overall null effect, for example.

The strengths and weaknesses identified in our primary analysis publication similarly apply to this analysis [20]. Particularly relevant to this analysis, we previously identified potential spill-over effects of the intervention at the practice-level. If spill-over effects did indeed affect other patients in the practice, then the individual-level effects may be driven towards the null (as the control group was also sampled from implementing practices). This is similarly true for the DDD analysis conducted here. However, these spill-over effects were not strongly indicated at the practice level. Furthermore, if our hypothesis is that these practice-level spill-over effects are as a result of preventative MDT working benefiting lower risk patients predominantly, then the spill-over may not apply to the high-risk matched group analysed in this sample at all. Also important to note, the outpatient visits outcome failed the parallel pre-trends test, and this outcome may therefore be biased in favour of the intervention group.

Although we would have preferred to use chronic conditions data recorded in primary care for our multimorbidity measure, this data was unavailable for this study. However, as

high-risk patients (both the intervention and propensity matched controls), these are probably the most likely to encounter inpatient admissions (and indeed were selected for 'high risk' of admissions). Therefore, these patients should in theory have the most complete recordings at this service level in comparison to the general population. From previous literature though, we can expect our multimorbidity measure to be less sensitive (i.e. be a predictably lower count) because it comes purely from hospitalisation data [37]. Nevertheless, our multimorbidity measure gained strength from being based on a list of 20 chronic conditions, deemed particularly important in the UK's NHS setting, which follows guidance from the multimorbidity literature [25].

Results in relation to other studies

There is little secondary analysis data available on multimorbidity from studies to date [38]. The authors are unaware of any multimorbidity subgroup analysis from evaluations of case management interventions to compare results directly. With the limited power of a single subgroup analysis, there is therefore the need for others to report these type of results, allowing future meta-analysis (or individual patient data meta-analysis, which may be a useful model to adopt [39]) and confirming of initial findings. The single comparison we can make presently is to our original publication, where we stratified results by risk-tool score [20]. As discussed above, the majority of the results obtained here confirmed our findings from the original subgroup analysis, i.e. that generally higher risk, more complex patients, tended towards increased secondary care utilisation with treatment, perhaps indicating identification of unmet need. The exceptions were from our stratification by Charlson index and the cardiovascular/metabolic cluster, which may have been identifying slightly different and distinct populations (those approaching the end of life with the

Charlson index, and those with conditions most amenable to primary care with the cardiovascular/metabolic cluster). This potentially explains the bucking of this trend in these patient groups.

Implications for clinicians and policymakers

Our analysis suggests that there is little to differentiate the effectiveness of the case management intervention by targeting specific multimorbid groups, regardless of conceptualisation. Where we identified subgroups where the intervention may be more beneficial (Charlson index >5 & cardiovascular/metabolic cluster), the differences were extremely small. We hypothesise that these differences may be due to differences in treatment at the end of life/treatment of only conditions that are particularly manageable in primary care. In the majority of cases, however, we expect the intervention to uncover unmet need, particularly for the most complex (highest risk) patients. As stated above, these are currently preliminary findings, and await further testing on other datasets.

Future research

Managing overall healthcare spending and over-utilisation of secondary care remains a vital goal in health systems globally. Further research is needed to evaluate how we may accomplish this goal, particularly for the ever-increasing numbers of multimorbid patients who are inadequately managed and at a high price in our current systems. Case management does not appear to be the singular solution.

These results do not suggest that any of the different ways of conceptualising multimorbidity perform better than any others in understanding the impact of case management. However, in assessing interventions, secondary analyses by multimorbidity subgroup may be a valuable tool to identify specific targetable groups [38], to achieve

maximum cost-effectiveness from an intervention. As we illustrate here, there are numerous ways to operationalise multimorbidity, and perhaps utilising a range of these may hint at subtler indications of intervention effectiveness (or ineffectiveness) which may be explored through yet further analysis (e.g. qualitative methods to explore potential mechanisms of action). When doing so, it is important to adjust for multiple testing in sensitivity analysis, to avoid drawing overly strong conclusions which may be based on false positive results.

Alternatively, other factors than multimorbidity may be of relevance – such as social care needs, or frailty.

Conclusions

Our results indicate no appropriate multimorbidity subgroup at which to target the MDT case management intervention in terms of secondary care utilisation/cost outcomes. The most complex, highest risk patients may legitimately require hospitalisation, and the intensified management may better identify these unmet needs. However, end of life patients/those with only conditions particularly amenable to primary care management may benefit a very small amount more than others. There is an ongoing need to find appropriate ways of addressing health system spending and management of multimorbid patients, however the concept is defined.

424

List of abbreviations

A&E = Accident & emergency

ACSCs = admissions for ambulatory care sensitive conditions

CCG = Clinical Commissioning Group

COPD = Chronic Obstructive Pulmonary Disease

DD = Difference-in-differences

DDD = Difference-in-difference-in-differences

IMD = Index of multiple deprivation

LOS = Length of stay

MDT = Multidisciplinary team

PICT = Practice integrated care teams

QOF = Quality and outcomes framework

Ethics approval and consent to participate

For the analysis of pseudonymised/anonymised data, no formal ethics process was deemed necessary.

Consent for publication

Not applicable

Availability of data and materials

No additional data are available.

Acknowledgements

We wish to thank Central Manchester CCG for the data they provided and the time they took to explain to us the details of the intervention.

Competing interests

The authors declare that we have no competing interests.

Funding

This work was funded by the National Institute for Health Research Greater Manchester Primary Care Patient Safety Translational Research Centre (NIHR GM PSTRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author's contributions

JS, SRK, and PB participated in the research design and project implementation. JS and SRK participated in the data analysis. JS wrote the original text. All authors revised the manuscript, and read and approved the final manuscript.

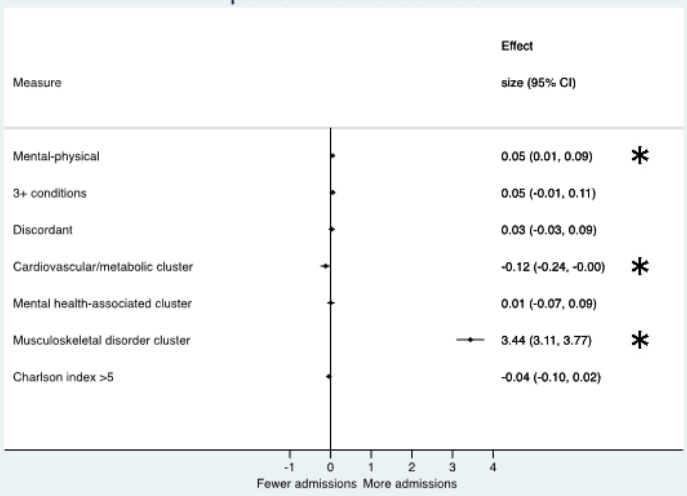
References

1. World Health Organization: **The world health report 2003: shaping the future**: World Health Organization; 2003.
2. Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T: **Interventions for improving outcomes in patients with multimorbidity in primary care and community settings**. *Cochrane Database Syst Rev* 2016, **4**.
3. Stokes J, Checkland K, Kristensen SR, Bower P: **Integrated care: theory to practice**. *Journal of health services research & policy* 2016, **Under review**.
4. Ross S, Curry N, Goodwin N: **Case Management: What it is and how it can best be implemented**: King's Fund; 2011.
5. Valentijn PP, Schepman SM, Opheij W, Bruijnzeels MA: **Understanding integrated care: a comprehensive conceptual framework based on the integrative functions of primary care**. *International Journal of Integrated Care* 2013, **13**.
6. Gawande A: **The hot spotters. Can we lower medical costs by giving the neediest patients better care?** . In: *The New Yorker.* January 24th edn; 2011.
7. Kasteridis P, Street A, Dolman M, Gallier L, Hudson K, Martin J, Wyer I: **The Importance of Multimorbidity in Explaining Utilisation and Costs Across Health and Social Care Settings: Evidence from South Somerset's Symphony Project**. *University of York - Centre for Health Economics Research Papers* 2014(96):1-60.
8. GAO: **MEDICAID: A Small Share of Enrollees Consistently Accounted for a Large Share of Expenditures**. In: *Report to Congressional Requesters.* vol. May 2015: United States Government Accountability Office; 2015: 1-45.
9. Starfield B, Shi L, Macinko J: **Contribution of Primary Care to Health Systems and Health**. *The Milbank quarterly* 2005, **83**(3):457-502.
10. Stokes J, Panagioti M, Alam R, Checkland K, Cheraghi-Sohi S, Bower P: **Effectiveness of Case Management for 'At Risk' Patients in Primary Care: A Systematic Review and Meta-Analysis**. *PLoS ONE* 2015, **10**(7):e0132340.
11. Gravelle H, Dusheiko M, Sheaff R, Sargent P, Boaden R, Pickard S, Parker S, Roland M: **Impact of case management (Evercare) on frail elderly patients: controlled before and after analysis of quantitative outcome data**. *BMJ : British Medical Journal* 2007, **334**(7583):31-31.
12. Brilleman SL, Salisbury C: **Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study**. *Family practice* 2013, **30**(2):172-178.
13. Lefèvre T, d'Ivernois JF, De Andrade V, Crozet C, Lombrail P, Gagnayre R: **What do we mean by multimorbidity? An analysis of the literature on multimorbidity measures, associated factors, and impact on health services organization**. *Revue d'Épidémiologie et de Santé Publique* 2014, **62**(5):305-314.
14. Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C: **Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide**. *Annals of family medicine* 2012, **10**(2):134-141.
15. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M: **Defining comorbidity: implications for understanding health and health services**. *Annals of family medicine* 2009, **7**(4):357-363.

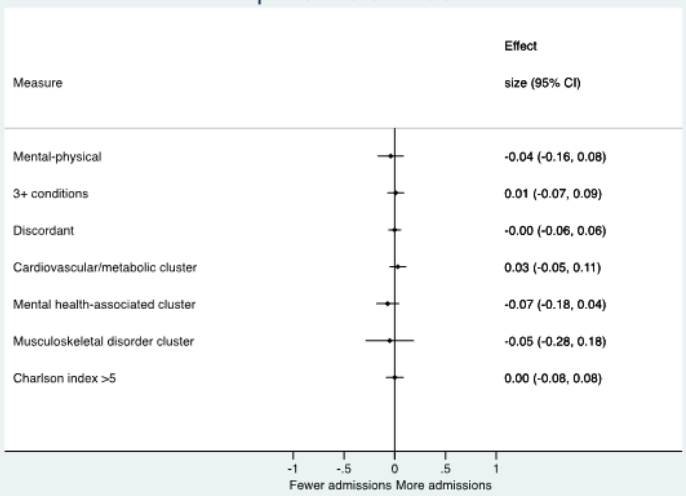
16. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M: **Multimorbidity patterns: a systematic review**. *Journal of Clinical Epidemiology* 2014, **67**(3):254-266.
17. Panagioti M, Stokes J, Esmail A, Coventry P, Cheraghi-Sohi S, Alam R, Bower P: **Multimorbidity and patient safety incidents in primary care: A systematic review and meta-analysis**. *PLoS ONE* 2015 - In Print.
18. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B: **Depression, chronic diseases, and decrements in health: results from the World Health Surveys**. *Lancet* 2007, **370**(9590):851-858.
19. Boyd CM, Fortin M: **Future of multimorbidity research: How should understanding of multimorbidity inform health system design**. *Public Health Rev* 2010, **32**(2):451-474.
20. Stokes J, Kristensen SR, Checkland K, Bower P: **Effectiveness of multidisciplinary team case management: difference-in-differences analysis**. *BMJ Open* 2016, **6**(4):e010468.
21. Angrist JD, Pischke J-S: **Mostly harmless econometrics: An empiricist's companion**: Princeton university press; 2008.
22. Propper C, Croxson B, Shearer A: **Waiting times for hospital admissions: the impact of GP fundholding**. *Journal of health economics* 2002, **21**(2):227-252.
23. Bertrand M, Duflo E, Mullainathan S: **How Much Should We Trust Differences-In-Differences Estimates?** *The Quarterly Journal of Economics* 2004, **119**(1):249-275.
24. StataCorp.: **Stata Statistical Software: Release 13**. College Station, TX: StataCorp LP.; 2013.
25. Fortin M, Stewart M, Poitras M-E, Almirall J, Maddocks H: **A Systematic Review of Prevalence Studies on Multimorbidity: Toward a More Uniform Methodology**. *The Annals of Family Medicine* 2012, **10**(2):142-151.
26. Stagg V: **CHARLSON: Stata module to calculate Charlson index of comorbidity**. *Statistical Software Components* 2006.
27. Charlson ME, Pompei P, Ales KL, MacKenzie CR: **A new method of classifying prognostic comorbidity in longitudinal studies: development and validation**. *J Chronic Dis* 1987, **40**(5):373-383.
28. McLean G, Gunn J, Wyke S, Guthrie B, Watt GC, Blane DN, Mercer SW: **The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study**. *The British journal of general practice : the journal of the Royal College of General Practitioners* 2014, **64**(624):e440-447.
29. Bender R, Lange S: **Adjusting for multiple testing—when and how?** *Journal of Clinical Epidemiology* 2001, **54**(4):343-349.
30. Holm S: **A simple sequentially rejective multiple test procedure**. *Scandinavian journal of statistics* 1979:65-70.
31. Cohen J: **Statistical Power Analysis for the Behavioral Sciences**: L. Erlbaum Associates; 1988.
32. Reginster JY: **The prevalence and burden of arthritis**. *Rheumatology* 2002, **41**(suppl 1):3-6.
33. Yelin E, Callahan LF, The National Arthritis Data Work G: **Special article the economic cost and social and psychological impact of musculoskeletal conditions**. *Arthritis & Rheumatism* 1995, **38**(10):1351-1362.

34. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, Januel J-M, Sundararajan V: **Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries**. *American Journal of Epidemiology* 2011, **173**(6):676-682.
35. Atun R, Gurol–Urganci I, Hone T, Pell L, Stokes J, Habicht T, Lukka K, Raaper E, Habicht J: **Shifting chronic disease management from hospitals to primary care in Estonian health system: analysis of national panel data**. *Journal of Global Health* 2016, **6**(2):020401.
36. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ: **Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test**. *J Clin Epidemiol* 2004, **57**(3):229-236.
37. Tonelli M, Wiebe N, Fortin M, Guthrie B, Hemmelgarn B, James M, Klarenbach S, Lewanczuk R, Manns B, Ronksley P *et al*: **Methods for identifying 30 chronic conditions: application to administrative data**. *BMC Medical Informatics and Decision Making* 2015, **15**(1):31.
38. van den Akker M, Gunn J, Mercer SW, Fortin M, Smith SM: **Secondary analysis of data on comorbidity/multimorbidity: a call for papers**. *2015* 2015, **5**(1):2.
39. Panagioti M, Bower P, Kontopantelis E, Lovell K, Gilbody S, Waheed W, Dickens C, Archer J, Simon G, Ell K *et al*: **Impact of chronic physical conditions on the effectiveness of collaborative care for depression: an individual participant data meta-analysis of 10,962 participants.** *JAMA Psychiatry* 2016(In press).

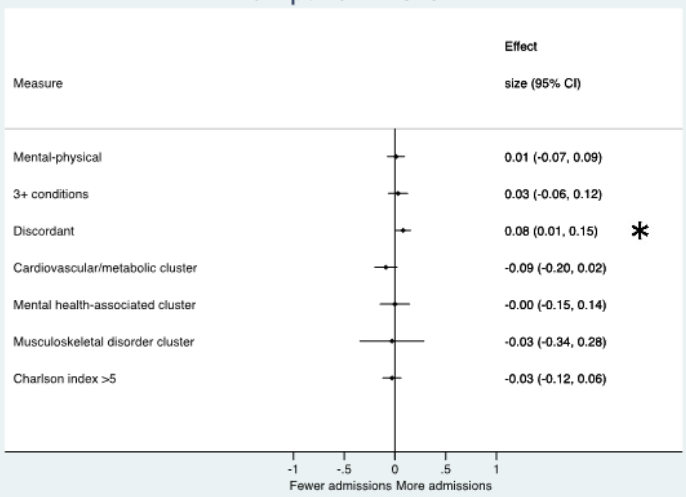
Inpatient non-electives



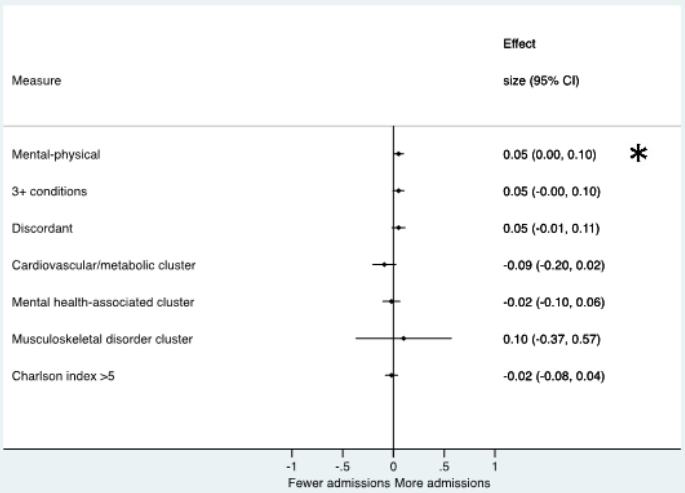
Inpatient electives



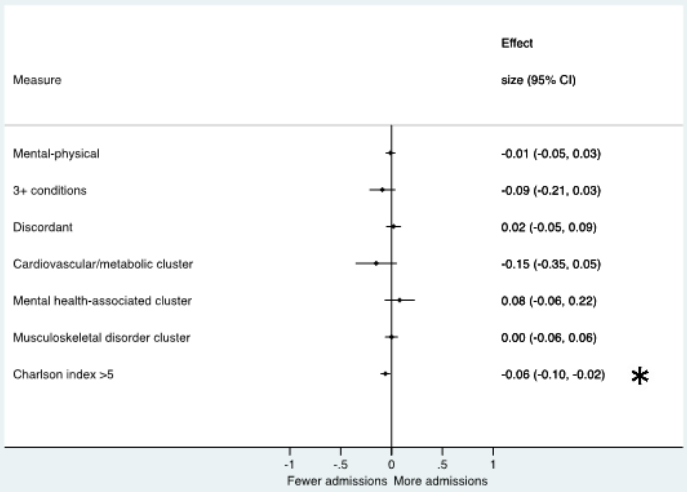
Outpatient visits



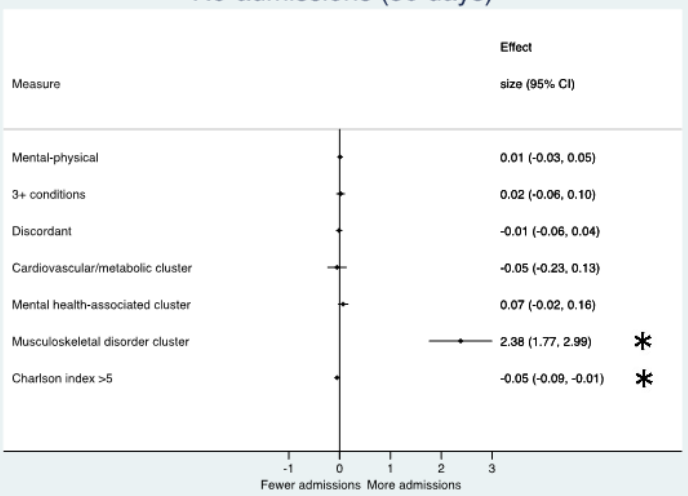
A&E visits



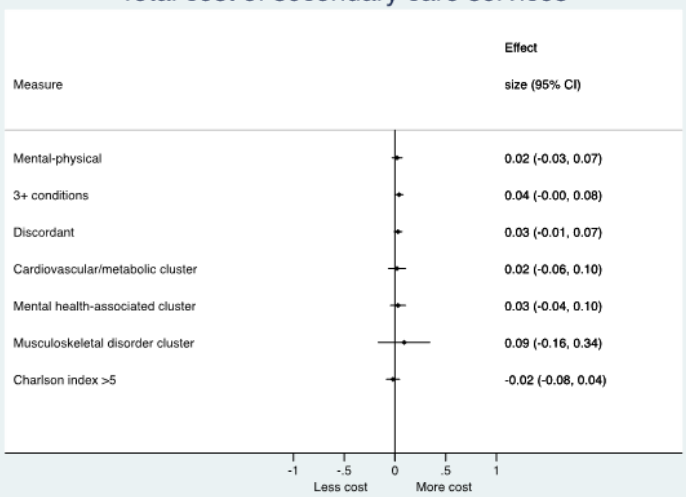
ACSCs



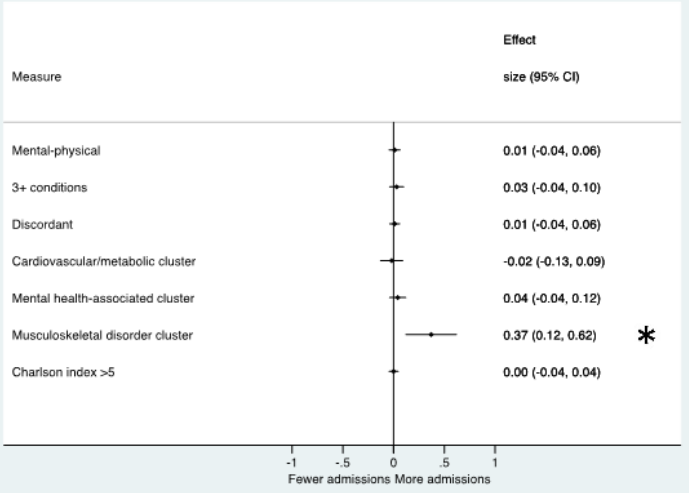
Re-admissions (30 days)



Total cost of secondary care services



Length of stay (days)



7.2 Chapter conclusions

In this chapter, I have shown that there does not appear to be a multimorbidity subgroup which benefits significantly more than others in terms of secondary care utilisation or cost. As was evident in the Chapter 6 subgroup analysis by risk score, the highest-risk patients in fact tend to increase admissions, contrary to the objectives of the intervention. However, from the results here I hypothesise that those patients at the end of life, and/or those with only primary-care sensitive conditions might benefit more than others to a small degree.

8. Discussion

8.1 Chapter Introduction

In this chapter, I summarise the findings of the thesis and discuss its strengths and limitations, findings in the context of the wider literature, theoretical and practical implications, dissemination and impact of the work, and consider further research opportunities in light of this discussion.

8.2 Summary of work

The thesis sought to answer three main research questions, and the work is summarised under these three headings below.

What does 'integrated care' currently look like in practice in the NHS?

In Chapter 2, I illustrated the theoretical confusion regarding the concept of integrated care. There are a multitude of definitions in the literature, and different primary (or multiple simultaneous) goals depending on which is accepted. Unsurprisingly, then, integrated care can take a variety of different forms in practice, and these can take place across all levels of the health system. In Chapter 3, I demonstrated that integrated care has generally been implemented as MDT case management (at the service delivery level) in the NHS to date.

What is the effectiveness of what is currently being carried out as 'integrated care'?

As illustrated in Chapter 2, proponents of integrated care suggest that, if implemented effectively, outcomes would include health benefits, increased patient satisfaction/experience, and cost saving. A systematic review of the international literature, however, showed no clear evidence that case management functions effectively across these domains (Chapter 5), and primary research exploring the implementation of case management in a single CCG (Chapter 6) also demonstrated no

clear benefits. While there is consistent evidence for a small increase in patient satisfaction for those directly treated, health benefits are unlikely in the long term, and there is no evidence of cost saving. What is more, the highest risk patients, appear to respond worse to treatment than those with lower risk scores (i.e. increased secondary-care utilisation, contrary to the aims of the intervention).

There was also no evidence of effectiveness in terms of secondary-care utilisation or cost for those patients directly treated by an MDT. However, MDT case management can theoretically add an additional mechanism of action, through increased professional integration. There were some preliminary indications of small spill-over effects at the practice level, potentially due to this improved professional integration and team working. I hypothesised why this effect may differ between levels. By definition, the average risk of the practice population will be lower than the high-risk group targeted directly, and these patients may be more responsive to preventative measures. Moreover, the mechanism of action for direct and spill-over effects likely differs. The direct effects may act through more holistic care of the high-risk patient individually managed (i.e. a case management effect), whereas the spill-over effect may act through effects of generally better team-working, e.g. quicker referral to allied services such as social care (i.e. a team-working effect).

Are there differential effects of 'integrated care' for different 'types' of multimorbidity?

As highlighted in Chapter 7, there was no multimorbidity subgroup that benefited considerably more than others from the MDT case management intervention evaluated. Again, those more complex (higher risk) patients in the majority of multimorbidity models analysed had slightly increased secondary-care utilisation compared to others, perhaps due to identification of unmet need. However, I hypothesised from the findings that there may be a very small additional benefit of MDT case management for those patients at the end of life (predicted by the Charlson index in the analysis), or with only conditions that are amenable to primary-care management (patients with only cardiovascular/metabolic cluster conditions in the analysis).

8.3 Strengths and limitations

In general, the methods used in this thesis are widely accepted and robust. However, as with all study designs they each have their inherent strengths and weaknesses, and those specific to their implementation in the individual study. The specific strengths and limitations of the individual study findings have been discussed previously (in the discussion section of each of the journal articles, Chapters 5 to 7). In this section, I firstly discuss in turn the strengths and limitations of each of the overarching methods used. Thereafter, I discuss the general strengths and weaknesses of the thesis as a whole.

Methods strengths/weaknesses

Systematic review and meta-analysis

Synthesis of randomised controlled trials (RCTs) evidence by systematic review and meta-analysis is frequently considered the gold standard of causal evidence (Howick et al., 2009). In the synthesis in Chapter 5, the inclusion of RCT evidence was extended with relevant studies implementing strong observational study designs, according to the Cochrane EPOC group, i.e. RCTs, non-randomised controlled trials (nRCTs), controlled before and after studies (CBA), and interrupted time series (ITS). Purists may argue that the causal implications of the study are slightly diminished by these extended inclusion criteria. However, I consider this extension a strength of the study, addressing some of the potential issues with RCT evidence and the population of interest (multimorbid patients) outlined in the methods section (Chapter 4). In addition, subgroup analysis was conducted in the synthesis, to check the RCT findings versus those of other study types to ensure this did not significantly differ.

The approach taken to the systematic review, furthermore, has merit in answering a key question of my doctoral research: What is the effectiveness of the intervention in terms of the policy aims? This is particularly appropriate in the NHS setting, in which the case management approach is incentivised by a national payment mechanism, the Directed Enhanced Services (DES), whereby practices receive additional payment for case management of a specific percentage of their highest risk patients (NHS England, 2013). In this policy context, the case management intervention as a general approach

(with varying practical implementation decided at a local level) is recommended and incentivised. Therefore, synthesising and providing evidence for the approach at this general level is relevant and necessary for evidence-based policy recommendations.

However, while the insights gained from a systematic review might be particularly valued by policy-makers for decision-making, policy-making is both a technical (based on the evidence) and a political (based on other insights and considerations) decision-making process. Therefore, there are a number of other factors contributing factors besides simply 'what works'. One important factor relevant to the review on case management presented in Chapter 5 is timeliness of the research, where some might consider "an 80% right paper before a policy decision is made is worth ten 95% right papers afterwards" (Whitty, 2015). While incentivising the intervention is recent, case management is a long-standing and ingrained policy implementation, as outlined in Chapter 3. Therefore, another important research avenue might be examining the subtleties of delivery of the intervention in the most effective way with rigorous synthesis from the social sciences (Whitty, 2015).

An alternative approach would thus have been realist synthesis, an increasingly popular method of synthesising evidence for this type of intervention. The realist approach differs from the traditional systematic review by focusing on the mechanisms through which interventions work (or do not work) rather than on outcomes alone. They therefore aim to identify the underlying causal mechanisms and assess under which contextual conditions (e.g. delivery of the intervention by whom, under which conditions, to which patient group) they are effective in producing outcomes (Rycroft-Malone et al., 2012). However, the process of realist review is iterative, and requires a team approach with significant expertise in the topic. In addition, it is time-consuming and resource intensive for the team involved, so it can be a more expensive approach than traditional systematic review. Nevertheless, there is arguably a place for the use of this technique in future studies relating to case management and integrated care more generally, particularly if policy-makers choose to pursue the interventions for other political reasons.

Difference-in-differences (DD) analysis

DD analysis is considered one of the most robust forms of controlled observational data evaluation (quasi-experiment). By design, the technique controls for selection on fixed individual observables and unobservables (simulating randomisation), and allows for common time effects that affect both groups of participants (e.g. ageing of the individuals, or effects of other policies that affect all participants) (Meyer, 1995). Used in combination with matching, moreover, potential bias in the data can be further reduced. For instance, the potential bias from regression to the mean from selecting a treatment group based on previous performance on an outcome (e.g. high hospital admissions) is less likely, as this potential bias is more likely to be evenly distributed in both treatment and control groups (i.e. both groups with previously high admissions) (Ryan et al., 2014). In addition, the assumption of common (parallel) pre-trends can be better satisfied by comparing more comparable groups (Angrist and Pischke, 2008). Although this important parallel trends assumption can be tested in the pre-trend period, it is fundamentally untestable in the period after the intervention is introduced (when we must presume that the assumption is likely to hold). Where the assumption is violated, the results of DD may be biased. DD is also based on a 'time invariant composition' assumption which depends on fixed individual characteristics rather than transitory outcome shocks affecting individual participation (Heckman and Leamer, 2007). This relates to a judgement by the researcher(s) of whether treatment and control groups are truly comparable, and therefore cannot be fully tested. Still, in the individual-level analysis in Chapter 6, for example, individuals were matched on a propensity score created from the variables that the intervention patients were selected for by the CCG, so we may expect these groups to be fairly comparable. However, interventions in practice are rarely so simple, and the use of clinical judgement as an additional patient selection tool in our case makes this judgement call more difficult.

With respect to Ryan et al.'s DD checklist (as outlined in Chapter 4), the paper meets all of the criteria listed, with the help of matching and clustered standard errors implemented in the analysis. The only exception is with respect to criterion number seven, "treatment does not 'spill-over' from treatment to comparison group" (Ryan et al., 2014). Through practice-level analysis, we found that the effects of the intervention might indeed spill-over to the wider practice from which the controls were drawn. With

the data available for analysis (where all practices had already implemented the intervention prior to the individuals in the individual-level data we were able to use, due to delays in the CCG's data governance arrangements in their early intervention period), it was not possible to meet this criterion by selecting controls from non-participating practices. However, we have clearly stated this limitation in the discussion in Chapter 6, and, if the hypothesis of the mechanism of spill-over effects primarily acting through lower risk patients holds, then this potential bias would not apply.

An alternative, and arguably more robust, approach to longitudinal observational data analysis is the regression discontinuity (RD) design. Similar to DD, this approach uses a pre-/post-test, but instead assigns a variable with a threshold cut-off point that determines intervention assignment. The idea of RD is that if this threshold is an 'arbitrary' cut-off point (e.g. a specific risk score selected by a policy-maker), then, at the threshold, we would expect participants to be randomly assigned at the point where they marginally achieve or fail to achieve intervention status (Angrist and Pischke, 2008). Because the cut-off point for joining the intervention has this arbitrary nature, the participants at either side of the cut-off point are likely to be very similar, and will exhibit the balanced measured and unmeasured characteristics found if they had been randomised (e.g. in an RCT). While carrying out the DD analysis on the Central Manchester data, I thus planned to carry out an RD analysis (using risk-tool score as the threshold variable) for additional robustness. However, due to data governance regulations at the CCG, I was not given a true risk-tool score for those in the control group, only for intervention participants. When I then attempted to use the generated propensity score as a substitute, there was no common threshold for intervention participation. Observing the spread in risk-tool score for the intervention patients alone, however, the extent to which clinical judgement had played a part in patient selection, which was significant, became apparent. This was confirmed in findings from a qualitative study that the CCG commissioned (Hall Aitken, 2014). Therefore, it is highly likely that even if I had had access to the true risk-tool score for the control patients, I may still have faced the same problem in using an RD approach.

The DD approach was, however, more than adequate for answering the research question: What is the effectiveness of the local MDT case management model? In

addition, the gradual implementation of the case management intervention to practices allowed me to utilise the approach looking both at direct effects (individual level) and potential spill-over effects (practice level) of the MDT implementation. And with this design, furthermore, I was able to address many of the gaps in the literature determined through the systematic review (i.e. a lack of complementary quasi-experimental evidence, located in a setting of high-strength primary care, and observing direct as well as spill-over effects).

Difference-in-difference-in-differences (DDD) subgroup analysis

DDD analysis introduces an additional interaction term to DD, so it can be used to observe subgroup effects, as in a standard regression analysis. DDD analysis builds on DD, and has similar strengths and weaknesses as those discussed above.

There are, however, additional weaknesses to any subgroup analysis when compared to primary analyses in terms of statistical power, as comparing within groups will naturally lead to loss of sample size compared to the higher-level analysis (Brookes et al., 2004). In addition, effects must be interpreted in relation to the findings of the original study. For instance, a relatively beneficial subgroup effect might sit in the context of an overall null effect on the treated population as a whole. Therefore, subgroup analyses should generally be treated with caution, and be hypothesis generating rather than interpreted as implying causation (Oxman and Guyatt, 1992).

Furthermore, although the stratification subgroups were chosen to represent a cross-section of the multimorbidity measurements in use, in line with the results from a systematic review of the literature (Lefèvre et al., 2014), the science is arguably still in development. Thus, different results might be obtained when stratifying using different conceptualisations of multimorbidity. In addition, the programme analysed in Chapters 6 and 7 is only one intervention in one site in England. Therefore, further research is required to enable generalisability of subgroup effects of case management more broadly. However, the Central Manchester programme does represent a common manifestation of the intervention in the NHS currently (as highlighted in Chapter 3), delivered by an MDT and with a social worker involved.

Overall thesis strengths/weaknesses

Overall, the merits of any study should be judged on its ability to meet its objectives. As illustrated above, each of the research methods chosen was suitable for addressing the specific research questions. However, for policy-related research to have any potential for real impact on policy decision-making, there are further criteria to consider. As touched upon above, the policy-making process is frequently observed to be influenced by not only a technical decision (between what works or doesn't work, based on evidence), but also other considerations, including, for example, personal relationships (e.g. trust of the researchers advising), as well as the policy-makers' own judgement of study quality, or their own/the wider political agenda (Whitty, 2015). Furthermore, when interventions are piloted by policy-makers, the piloting process/evaluation of the pilots can be implemented for a variety of reasons, not only to test effectiveness of an intervention for evidence-based policy making (Ettelt et al., 2015). My research was only able to contribute primarily to this single aim (i.e. testing effectiveness) through evaluation, however. There are other aims that might have been higher priority for the policy-makers implementing the Central Manchester intervention, for instance. These might have included: 1. Piloting as a first-step before scaling up; 2. Demonstrating how to successfully implement the intervention; or, 3. Piloting as a learning experience, to overcome implementation barriers and improve processes and outcomes (Ettelt et al., 2015). Therefore, the usefulness of the research to the local policy-makers is partly dependent on my evaluation aim coalescing with theirs.

Nevertheless, testing effectiveness of interventions is a legitimate accomplishment of research more generally, and provides evidence other policy-makers could learn from for their own choice of implementation. Moreover, the aim of testing effectiveness is arguably increasingly important in a cost-constrained system where spending public money on the 'right' interventions will be continuously important, with less money available for improving public services generally (Ettelt et al., 2015). While RCT evidence is generally considered the gold standard evaluation design, and might be considered a particularly persuasive evaluation method for a policy-maker in health policy (Ettelt et al., 2015), practically (particularly for a stand-alone PhD study) it is not always feasible to carry it out. RCTs are also likely to mis-estimate the causal effect of a complex intervention in day-to-day practice (Davies et al., 2015), and frequently (though not always) fail to take adequate account of contextual factors of an intervention (Bonell

et al., 2012). Particularly in the case of an intervention which is extremely unlikely to cause harm, such as case management, and in the context of squeezed public services in the current budgetary environment, it is imperative that researchers produce usable, good quality evidence from the abundant routinely collected data wherever possible (Raine et al., 2016).

It is also important to consider what exactly is being evaluated for effectiveness. In this thesis, the scope has not been to evaluate 'integrated care' more broadly, but instead to: 1. Identify the most common approach being taken to integrate care in England (i.e. MDT case management). This approach is also one which is being incentivised in current policy, and has wider international relevance; 2. Provide meta-analytic evidence about the extent to which it achieves what is claimed; and, 3. Evaluate a specific instance of the policy and test its impact, particularly on multimorbid patients. So, in effect, the thesis can act as evidence for effectiveness of what is currently mainstream integrated care practice (i.e. case management), but not for all of the wider possibilities of integration strategies (outlined in Chapter 3), and it does not compare effectiveness of MDT case management with other integration strategies.

In relation to the complex group of patients who formed the focus of this PhD, multimorbid patients, the general limitations of population-based research methods could also be emphasised. This relates to the conceptual issues of multimorbidity highlighted in the introduction. Particularly when defined by a simple count criterion, this remains a very diverse, heterogeneous group. Thus, attempting to generalise the findings of this population-level analysis to an individual during clinical treatment might prove particularly difficult in this group. There could consequently be scope for development of more personalised results, through further subgroup analysis (and meta-analysis), or perhaps analysis of linked 'big data' sets, where power to observe the stratified effects would be enhanced. Through the DDD analysis, I have attempted to set some groundwork for this first approach, suggesting common measures of the concept by which to stratify results. However, current data availability limits the possibilities of this stratification approach, although this may improve in the future (as discussed further in the multimorbidity part of Section 8.5).

8.4 Findings in the context of the wider literature

There are a number of previous studies that have looked at the effectiveness of case management. The majority of these were captured in the inclusion criteria in the systematic review in Chapter 5, and also discussed in the discussion section of Chapter 6. The most rigorous previous systematic evaluation was from a meta-analysis, which had likewise shown no effect on unplanned hospital admissions (Huntley et al., 2013). This review, however, did not consider effects on satisfaction or health outcomes.

Putting this study in the context of findings from research conducted in the United Kingdom specifically (as highlighted in Chapter 3, a fairly unique context, and case management is an intervention where we can expect context to have an influence on outcomes), the first rigorous assessment of case management was through the Evercare study in the early 2000s.

Evercare was adapted from the US system and involved case management of frail elderly patients, primarily carried out by advanced practice nurses. The evaluation in the BMJ in 2007 also reported no significant effect on rates of emergency admissions, emergency bed days, and mortality (Gravelle et al., 2007). The authors also highlighted the place of case management in NHS policy at the time, through the Community Matron programme. They cautioned policy-makers at the time that "without more radical system redesign this policy is unlikely to reduce hospital admissions" (Gravelle et al., 2007). The Nuffield Trust had originally planned to carry out the evaluation of the Community Matron programme separately (The Nuffield Trust, 2007), although this report does not appear to have ever surfaced in the public domain, so these results cannot be compared with those of this study directly.

More recently, the difference-in-differences evaluation of case management within the Integrated Care Pilots (ICPs) reported a 9% (95% confidence interval, +1% to +16%) increase in emergency admissions, but a 21% (-10% to -32%) reduction in elective admissions and 22% (-16% to -28%) reduction in outpatient attendance for high-risk patients compared to matched controls in the 6-months following introduction of the programme. At the practice-level, they found overall outpatient attendances reduced by 5% (-2% to -8%) two years after the start of the programme. Over half of the reduction in

outpatient admissions was reported to come from a group with cancer (with half of these associated with reduced admissions for chemotherapy). This occured even though there were similar proportions of patients with cancer in the intervention and control group, and cancer care was not a specific focus of the sites (so the authors could not explain the finding). However, the ICP evaluation study combined six separate case management interventions, some of which were aimed at specific conditions only (e.g. COPD – hence the study did not meet the inclusion criteria for the review reported in Chapter 5), so might not be directly comparable. There is some evidence that case management targeted at specific conditions can be more effective than when targeted at generally 'at-risk' patients for reducing admissions (Purdy, 2010). This would also fit with the hypothesis discussed from findings reported in Chapter 7, that those with primary-care sensitive conditions (e.g. COPD, as one such condition) may respond better to the case management intervention.

8.5 Theoretical implications

The PhD study findings have theoretical implications for both the understanding of integrated care and of multimorbidity.

Integrated care

My work suggests that case management, as integrated care, does not reduce secondary-care admissions or cost of care. In the wider literature there is little evidence of this reduction being possible with current primary care and community-based interventions (Smith et al., 2016), and little theoretical justification as to why it should be possible, particularly if targeted only at high-risk groups (Wallace et al., 2016; Roland and Abel, 2012). However, for policy-makers, these outcomes remain a key aim of case management interventions and integrated care more generally. Nevertheless, as illustrated in the systematic review in Chapter 5 (Stokes et al., 2015b), and as is reported of integrated care more widely (Cameron, 2016), patient experience of care (i.e. patient satisfaction) appears to be consistently positively affected. Therefore, while case management and integrated care do not appear to accomplish all of their outcome-oriented goals, especially regarding cost (Cameron, 2016; Nolte and Pitchforth, 2014), they do arguably meet the widely adopted patient-centred outcome-based definition, 'the

feeling of more joined up care' for the patient (Department of Health, 2013), and its single implicit outcome of patient satisfaction.

The majority of current integrated care literature portrays the three outcomes of improved health outcomes, increased cost-effectiveness, and increased patient satisfaction as simultaneously achievable. For example, the WHO's global strategy on people-centred and integrated health services report claims that benefits of these service delivery changes (amongst numerous others) would include, "increased satisfaction with care and better relationships with care providers…reduced overall costs of care per capita…reduced mortality and morbidity from both infectious and non-communicable diseases" (World Health Organization, 2015). The evidence available to date, however, could indicate the naïveté of this assumption. The literature, described below, shows that each of these outcome categories is linked to each of the others. For instance, increased patient satisfaction appears to be attained with improved health outcomes following a healthcare visit (Atkinson and Haran, 2005; Schoenfelder et al., 2011; Bleich et al., 2009; Hardy et al., 1996), but people originally in a better state of health also tend to be more satisfied (i.e. satisfaction may act as both cause and effect) (Stokes et al., 2015a; Moret et al., 2007; Schoenfelder et al., 2011). People in a better state of health, in turn, should theoretically make less use of health resources, but the most satisfied patients (associated with being the healthiest) might also be those with the highest overall healthcare expenditure, and may be those experiencing worse health outcomes, e.g. a higher mortality rate (Fenton et al., 2012). Therefore, it appears that as one of these three overarching outcomes is beneficially influenced, the others might also be influenced (either positively or negatively) by the means of achieving that single outcome (although, clearly, only a single outcome alone can also be *measurably* influenced, as found in the empirical work in this thesis). In sum, there is little evidence that the ideal scenario of not having to pay for improvements in care quality is plausible.

When the focus is primarily on patient satisfaction, there is potentially a conflict between system-level goals (the focus of policy-makers in the current austerity climate, i.e. cost and utilisation reduction) and individual-level goals (making the patient's experience of care coordination better, and possibly benefiting health outcomes by doing so). For example, the Fenton study mentioned above, set in the USA, illustrates the potential dangers of overly 'consumer-focused' healthcare (Fenton et al., 2012). Using

basic economic market theory, this consumer-focus should theoretically strike the correct balance between cost-effectiveness and patient-oriented goals. However, this assumption ignores known market failures in healthcare. Particularly illustrative in this instance is the market failure of information asymmetry (i.e. the expert knows more than others, so can make the more informed decisions) (Morris et al., 2007). While this asymmetry clearly is in favour of the patients in some areas (e.g. experience of disease/ experience of using health services/ treatment burden etc.), it favours the physician/expert in others (e.g. treatment course in many cases/ health system setup or organisation/ cost of care etc.). The patient is therefore not likely to make the most informed choice when it comes to overall health system costs. Focusing too much on 'consumer satisfaction' in a tax-funded health system, then, might conflict with cost outcomes. It is, indeed, difficult to imagine a situation where increased patient satisfaction through patient choice does not involve wrapping more (costly) care around the patient.

If it is not possible to achieve all three overarching outcomes of interest in equal measure simultaneously, there is clearly a need to prioritise what we want integrated care (or any other intervention) to achieve. As I argue above, I consider a primary focus on consumer satisfaction (illustrated by the adopted NHS definition of integrated care) to be wrong in the context of a tax-funded national health service. This is by no means to imply that patient satisfaction is not an important outcome to measure, and to bear in mind while planning health services/systems. However, I would argue that patient satisfaction should be considered as a secondary outcome measure, because patient satisfaction is partially determined by the health outcomes of a healthcare visit. Thus, aiming for better health outcomes without primarily chasing patient satisfaction would be a better goal (i.e. we should still attain satisfaction as a secondary outcome – particularly if care is based in the primary-care setting where this satisfaction is generally increased) (Stokes et al., 2015a; Starfield et al., 2005). Whether this patient satisfaction must necessarily fulfil the patient-centred definition of integrated care (i.e. patient's perception of a more integrated service), or is best measured through other means is a matter for debate.

It is important, then, that we recognise what integrated care may be capable of achieving, agree on prioritised aims (at least within a local context where it is to be

implemented), and establish a common definition of integrated care that adequately reflects these.

Multimorbidity

Similarly, the results of this study have important implications for how we measure/define multimorbidity and which groups we target with interventions. Case management relies upon identifying and targeting the most high-risk patients, usually identified using some kind of risk-stratification tools. These tools tend to target elderly patients and those with many secondary-care admissions (so those already in contact with health services). Elderly patients, however, comprise the minority of those with multimorbidity in absolute terms (Barnett et al., 2012). Therefore, there currently appears to be a major gap in models of care aimed at this wider group in need of better care coordination, including those who have not yet reached the stage where they are recognisable to health services via past utilisation. The slightly lower-risk patients, moreover, may also be those most likely to respond better to prevention-oriented, cost-saving interventions. Such interventions are unlikely to be best delivered via MDTs as currently enabled (although some form of interdisciplinary working will likely play an important part in developing effective future primary-care models in some capacity, potentially tapping in to the proposed benefits we found in our spill-over analysis in Chapter 6).

In the subgroup analysis of the Central Manchester data, I found few differential treatment effects of case management when stratifying by different multimorbidity conceptualisations. However, the few differential effects identified provided a hypothesis for when the intervention might be most effective. When we think of multimorbidity as a complex adaptive system (Sturmberg, 2014), there is unlikely to be one single way in which we should model the relative effects. This implies the need for future studies to agree on standard models of stratification to be used. In Chapter 7, I suggest a starting point for this research with a sample of measures from the four dominant multimorbidity definition types.

There may, however, also be other measures which are ultimately more useful than multimorbidity. For example, the increasingly popular measure of frailty (e.g. the e-frailty index) which may arguably identify patients more amenable to care as it primarily

draws on primary-care records rather than hospital ones (Clegg et al., 2016). It could also be argued, though, that the danger of focusing treatment solely upon those at highest risk (although arguably identified at a slightly earlier stage) and already in contact with health services still exists according to this alternative measure. Alternatively, a combination of measures may be another approach to consider (Koroukian et al., 2016). When choosing an approach, it is ultimately important to bear in mind that while certain measures might be particularly good at predicting certain outcomes, this alone does not mean that targeting the patients identified with added care will necessarily improve these outcomes (i.e. the measures do not necessarily select a group that will respond to treatment). For example, a frail patient approaching the end of life will continue to become increasingly frail (requiring increased healthcare utilisation) whatever interventions are tried.

There is also a theoretical implication with regard to our research methods, and to our guideline-based medicine approach to care. The previous sections of this thesis have questioned the suitability of dominant methods (e.g. RCT design) in relation to multimorbid patient groups. There are ways of improving this study design (e.g. with improved inclusion criteria) to address some of these issues (Coventry et al., 2015; Man et al., 2016; Bonell et al., 2012). But, ultimately multimorbidity draws attention to some of the flaws of population-based research methods when applied to individual patients more generally, since these are particularly illuminated by this extremely heterogeneous group. These insights suggest that a more stratified (personalised) approach might be necessary. But there are important practical considerations that would be necessary to make this personalised approach possible. Arguably, 'big data' with linked and complete health and care datasets would be an important starting point for any stratified approach. However, there are further ethical/public appetite considerations that need to be addressed before use of this tool would be possible. The care.data case study possibly indicates how difficult implementing this could be (Hays and Daker-White, 2015). But there may be some grounds for optimism as early results from recent studies show that when public understanding of data use is enhanced, so too is public acceptability (The University of Manchester, 2016). Likewise, informal evidence from patient and public involvement (PPI) discussion groups, in which I myself have participated, suggests that members of the public tend to be outraged when they discover that their care records are not already linked across care settings.

Moreover, even with complete and linked healthcare data, it is likely that important variables that might help us predict and benefit patients with the correct interventions at the correct time will remain unavailable. For example, missing psychological, social circumstances and lifestyle data will limit the possibilities of a stratified approach. However, our colleagues involved in psychometrics analysis are already exploring exciting opportunities for capturing these missing strata from social media interactions, e.g. data-mining Facebook 'Likes' to accurately predict a variety of attributes, including, "sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender" (Kosinski et al., 2013). The expansion of people contributing to big data sets through wearables offers further opportunity potential. Although there are a number of ethical and governance issues to be discussed and overcome before the full potential of this data could be harnessed effectively (Kosinski et al., 2015), nevertheless, the potential for future public health and health services use is at least plausible.

8.6 Practical implications

There are also a number of practical implications of the findings.

Case management and risk stratification

The findings suggest that commissioners must reassess how/if they use case management and risk stratification, and to what end.

There are a number of options for case management available to commissioners and policy-makers in light of these results:

1) Status quo

Despite the indications of current research, the case management approach may well remain a dominant model of integrated care, as it has been up till now. Section 8.4 illustrates how this approach has remained a policy-focus in light of strong evidence accumulated to date that the intervention does not produce the main aim of reducing emergency admissions, for instance. Thus, policy does not necessarily follow the

evidence. And, it may prove difficult for those decision-makers to go against the face validity of case management (illustrated in Chapter 3) and its perceived ability to accomplish all of their goals simultaneously. But, as discussed above, there is still a gap in the literature regarding the relative effectiveness of various approaches to delivery and their use in different contexts. Future realist synthesis may alleviate this knowledge gap. However, in light of current evidence it is increasingly difficult to justify incentivising case management as a general concept, as the NHS is currently doing. This is especially true at a time when system resources are particularly stretched, and the value of cost-effective spending is increasingly important.

2) **Continue to use case management, but only for what it can achieve**

The evidence summarised in this thesis shows that case management and integrated care more generally does have merit in benefiting a patient's experience of care. This is a legitimate aim of a health system (World Health Organization, 2007), and so decision-makers may be justified in continuing to implement case management to this end. However, as I argue above, it might be important to prioritise outcomes, and patient satisfaction could be considered secondary to others. And again, with cost of care of particular concern at this moment, it could be more difficult to justify the decision of continued use on this basis alone.

3) **Modified/enhanced case management**

With more accurate and standardised reporting of how the case management approach was delivered, to whom, and under what conditions, there is the potential for future syntheses to more accurately explore relatively more effective, modified/enhanced models of case management than was possible on the basis of the current literature (reported in Chapter 5). If policy-makers follow either of the first two options above, this subsequent evaluation reporting and analysis would be a recommendable commensurate action.

4) **Look for alternative approaches/models of care**

If policy-makers continue to emphasise the importance of reducing costs when planning new models of care, alternative approaches might be required. I discuss some

potential options for alternative approaches at the end of this section. Before that, I will briefly discuss one of the tools these models frequently use, risk stratification.

The use of risk stratification tools is widely adopted in the NHS and internationally, including highly successful health system models such as Kaiser Permanente (Chen et al., 2009). However, there is cause to ponder how/if we best use the results obtained from these models in practice. For example, we may wish to consider the levels of risk at which we focus our interventions (and which interventions to focus on) in order to have maximum effect on our aims. While the interventions aimed at higher-risk individuals would be more traditionally considered healthcare interventions (i.e. aimed at the patient who has immediate healthcare needs), there may be a case for broadening this definition (i.e. expanding it to include prevention-based interventions which take the social determinants of health into account (Wilkinson and Marmot, 2003)). For, as detailed above, there is little evidence of effectiveness for interventions aimed at high-risk level patients only (Wallace et al., 2016; Roland and Abel, 2012). However, whether current risk prediction models (based primarily on previous healthcare utilisation data – so identifying those already well known to health services) are able to identify the optimum population at all is possibly unlikely.

The need to consider what the meso/macro environment allows

On the basis of the subgroup result findings of the Chapter 5 systematic review, it was hypothesised that case management might show different effectiveness under different contextual conditions (e.g. case management may be more effective in health systems with low-strength primary care).

This potential differential effect fits with the theoretical understanding of complex interventions introduced previously, in which context is an important factor (Craig et al., 2008). Beyond this, I would argue that certain meso/macro environments also dictate how an intervention can be assimilated and carried out in the first instance. For example, 'extending appointment times' for the treatment of chronic disease and multimorbidity has been suggested as potentially beneficial for addressing the patient holistically, and for improving self-care behaviours, ultimately preventing exacerbation of symptoms and improving patient experience (Man et al., 2016). However, in a health system that rations

its resources through waiting times (as is the case in the NHS), rather than ability to pay (as in the USA), a different intervention would necessarily be implemented in one context than another (never mind with different anticipated effects). For example, the NHS uses 10-minute GP appointment slots to make this resource-rationing system work, and simultaneously attempts to comply with political pressures to meet individual appointment waiting targets. Here, there is little scope to increase appointment times without other delivery adaptations, for example, major increases in workforce or up-skilling other professionals to compensate, or alternatively, group chronic-disease management clinics. Whereas, in a predominantly privately funded system, arguably the price of the increased time could more easily be passed on directly to the insurance company/patient, making it easier to implement with the resources rationed through ability to pay this increased price.

As outlined in Chapter 3, case management has arisen primarily in the USA, where there is historically little recognisable primary care to control costs (e.g. no gatekeeper system). This may make the case management intervention a more important aspect of healthcare delivery in that context, i.e. it may substitute for 'good' primary care. Similarly, the case management intervention probably makes more sense in the payment context of the USA. The insurance-based payers may be less willing to pay before their members are sick, making attempts to contain costs for high-risk patients a priority. We thus see changes towards increased integration and prevention orientation accompanying new payment models, e.g. capitation payments (NHS England, 2014a), and formation of accountable care organisations (Fisher et al., 2007).

In the NHS therefore, prioritising specific aims of the health system could be particularly important due to the nature of funding (i.e. tax-based). In a health system that is based fundamentally on moral values of equity and social justice (Collins and Klein, 1980), does the higher priority of patient choice and experience fit with this value system, or does this originate from models of care adopted from the USA, or market choices available in other aspects of UK life? The separation of funding systems for health (tax-funded, free to all) as opposed to social care (increasingly privately funded, available to those who can afford it) in the UK has further practical implications. This is particularly true in the context in which social care budgets are being cut disproportionately in relation to their healthcare counterparts, as is currently the case

(The King's Fund, 2015b). There are potential problems with the practicalities of truly integrating these systems that are so fundamentally separate at the moral choices level, as indicated in terms of their funding approach. This is particularly true if we choose to target the social determinants of health and prevention of disease, as I would advocate. To do so, we must by definition pay before the illness arises and make future savings by preventing/compressing the treatment costs to a shorter period at the end of life.

Therefore, fundamental value and financing decisions need to be re-assessed within the health and care sectors before the meso/macro environment can adequately accommodate the planned service-delivery changes to take place effectively.

Current ambitions for change in the NHS

The NHS's Five Year Forward View (FYFV) outlines the ambitions for change and new models of care that the NHS in England aspires to. The FYFV outlines a number of new models of care, which have inspired the Vanguards under the same banners (NHS England, 2014b; NHS England, 2016b). These overarching models are outlined in Box 5.

Box 5: Vanguards - new models of care. Source: (NHS England, 2014b; NHS England, 2016b)

1. **Multispecialty Community Providers**

Concentrated on moving specialist care out of hospitals

2. **Integrated Primary and Acute Care Systems**

Concentrated on joining up GP, hospital, community and mental health services

3. **Acute care collaboration**

Concentrated on linking together local hospitals to improve their clinical and financial viability

4. **Urgent and emergency care**

Concentrated on improving the coordination of urgent and emergency care services and reducing the pressure on A&E departments

5. **Enhanced health in care homes**

Concentrated on offering older people better, joined-up health, care and rehabilitation services

This study and the considerations arising from it, as discussed above, relate directly to what we may reasonably expect the FYFV to achieve in practice. The main consideration of the FYFV, as discussed in the introduction, is financing and costs (reducing Demand, improving Efficiency and increasing Funding), i.e. addressing the estimated £30 billion spending gap (NHS England, 2014b).

First, the models of care emphasise the importance placed on integration in order to achieve these goals. However, the evidence from this study, and the wider evidence relating to integrated care (detailed above), shows that this emphasis is misplaced in light of what current models are capable of achieving.

Second, the timeline of five years appears particularly ambitious given that models of care currently in use appear to have achieved little if anything in efficiency savings to date (Nolte and Pitchforth, 2014). I would endorse the ambitions of prevention-oriented demand reduction outlined in the FYFV. However, again, any significant reduction in demand will take a long time (potentially generations of prevented disease) to achieve, as it requires societal shift in terms of lifestyle choices and where patients seek their healthcare. There is, moreover, no indication that the additional funding needed will transpire under a pro-austerity Government, although there are signals from Theresa May that austerity measures might be loosened to some extent in the future.

Last, the document also outlines a vision of a stronger role for the patient in wider healthcare planning, firstly, for disease self-management reasons, but also for strengthening patient choice and direct involvement in “decisions about the future of health and care services” (NHS England, 2014b). However, as I have argued above, these last two involvement types (i.e. excluding for self-management) might be misplaced if the aim of cost saving is the priority.

Indeed, Simon Stevens, Chief Executive of NHS England, has since argued that the ambitions outlined in the FYFV will be unachievable without the commensurate increase in funding of social care services, which has not yet transpired (NHS England, 2016d).

A number of questions, then, arise from examination of the evidence. It also remains to be seen whether the new models of care tested as Vanguards will be able to overcome the national level barriers (e.g. competition agenda and requirement to simultaneously meet nationally set targets while attempting to implement change) and achieve the wider change envisioned by previous government initiatives (e.g. wider organisation/funding integration) (NHS England, 2016c; Erens et al., 2015). This is arguably required for a more population-level approach to integration to occur (i.e. away from the focus only on a small high-risk group). In addition, from examining initial documented plans, MDT case management will once again play a significant part in the Vanguard's service delivery plans (see Appendix 1).

Alternative approaches to integration?

From a pragmatic point of view, we will need to consider alternatives to integration, particularly to deal with the most pressing of health system challenges, the current funding gap in the NHS.

One of the main strategies of policy-makers to close this gap is to reduce secondary-care admissions. In a review by Purdy et al of 'what works for reducing emergency admissions', it was found that prevention, or early-intervention-oriented approaches appeared to be most effective. For example, this included self-management interventions, and continuity of care with a GP. Likewise, the report found that there was no evidence that case management in the community reduced generic admissions (although they found that assertive case management is beneficial for patients with mental health problems). At the time, they also indicated that integrating health and social care, or integrating primary and secondary care might benefit reduced admissions (largely based on initial work at the King's Fund - (Curry and Ham, 2010)) (Purdy, 2010). However, the evidence detailed above now calls this latter claim into question.

Expansion of primary care is frequently floated as a viable option for reducing demand for, and thus costs of, secondary care services. There is, indeed, increasing evidence that primary care offers a cost-effective alternative to secondary care (Atun, 2004; Starfield et al., 2005). There is also some evidence that expansion of access to primary care can (but does not always) lead to reduced secondary-care admissions

(Greater Manchester NIHR CLAHRC, 2015). However, the quality of primary care delivered will clearly be a factor in effectively accomplishing this shift. For example, one study found that "patients cared for by PCPs [primary care practitioners] with better performance on quality metrics may have lower rates of ED [emergency department] and hospital utilization" (Yelibi et al., 2014). Furthermore, if the potential cost savings are to be realised at the system level, secondary-care staff and services would eventually have to be decommissioned (or re-aligned to cheaper primary care) in the long-term, in line with the shift in demand. This would most probably involve a period of increased double-funding in the short-term to prevent any shock to the system and an inability to meet current demand. At the individual level, there may be the opportunity to remove burdensome and unnecessary treatments, reducing resource wastage and potentially dangerous treatment interactions, particularly towards the end of life. However, there are a number of ethical/practical questions regarding how we recognise this 'futility point' of treatment in the first place, and who decides if it is best to step down treatment and when (Søreide and Desserud, 2015).

There may also be a need to up-skill the primary-care workforce to enable it to cope with the increasing service demands placed upon it, and potentially the need for new roles such as community healthcare assistants. These roles are already utilised quite successfully in other health systems – e.g. Brazil/Cuba etc. (Harris and Haines, 2010; Campion and Morrissey 2013), and the healthcare assistants have more time to deliver preventative interventions at a cheaper price per person hour. Or, potentially enhanced roles for professions we already have, e.g. nurses, social workers, etc., may be required. In short, the traditional definition of primary care in the NHS (i.e. GP/nurse care) may need to shift towards what the Alma Ata declaration in 1978 called 'comprehensive primary care' (World Health Organization, 1978). This involves the health sector working more closely with other traditionally non-health sectors (e.g. housing associations, job centres, etc.) to achieve a more preventative, holistic approach.

Historically, the most cost-effective solutions to population health management have always been preventative (through public health interventions) (Owen et al., 2011; Hutton and Haller, 2004). Furthermore, the World Health Organization estimates that if "risk factors were eliminated, at least 80% of all heart disease, stroke and type 2 diabetes…[and] over 40% of cancer would be prevented" (World Health Organization,

2016). There is also evidence that at least a proportion of multimorbidity is associated with preventable risk factors. For example, obesity is estimated to increase the likelihood of multimorbidity 7-fold for males, and nearly 10-fold for females (Jovic et al., 2016). Finally, there is mounting evidence of the ability of better lifestyle factors to compress multimorbidity to a smaller period at the end of life, particularly shown in high-income countries to date, less so in low-income ones (this country-level difference further suggesting the importance of lifestyle factors relating to socioeconomic status – in theory modifiable) (Chatterji et al., 2014). However, as mentioned above, any prevention-driven strategy will likely take a generation to produce any significant results.

In the shorter term, an alternative to the efficiency and demand solutions is of course to address the funding gap more directly. In terms of relative spending, total health expenditure (as a % of GDP) in the UK is below the OECD average (The World Bank, 2016). In addition, per capita healthcare expenditure has been flat for the past few years (in spite of ever-increasing demand, as outlined in the introduction), reduced from a high in 2007, and well below the high-income OECD average (The World Bank, 2016). The Government has committed to some increased spending since the FYFV. However, "this means that between 2009/10 and 2020/21, spending on the NHS in England will rise by nearly £35 billion in cash terms – an increase of 35 per cent. But much of this increase will be swallowed up by rising prices. In fact, around £24 billion will be absorbed by inflation, leaving a real increase of just £11 billion (a 10 per cent rise over eleven years; equivalent to an average annual increase of just 0.9 per cent)" (The King's Fund, 2016). There is the potential to increase funding for the NHS, and as the recent EU referendum campaign has illustrated, public demand for it is also likely to exist.

Reversing 'self-inflicted' demand increases might be another relatively easy option. For example, the current Government's austerity choices have cut public resources to the things that prevent ill-health in the first place. Public health budgets, and budgets for social care services for older people (which are meant to get people out of the expensive hospital setting as quickly as possible and back to the community) have been slashed, along with local council funding (The King's Fund, 2015b). There has also been a rapid increase in recent years of private providers utilising the NHS budget (Lafond, 2015). The majority of this spending is on providers that take a profit slice out of

the health system (The King's Fund, 2015a), thus further exacerbating the spending gap. Moreover, private providers target those parts of the system that are actually profitable and so confound the negative effects of the loss-making parts of the system maintained within the NHS's provision. This increase in private provision is largely due to competition legislation implemented through the 2012 Health & Social Care Act (Department of Health, 2012), and is therefore reversible.

8.7 Dissemination and impact

As discussed above, it is important, if aiming to influence policy decisions, to provide what a policy-maker might consider 'good enough' evidence. In population health decision-making, local data are considered to be the most valued and used by decision-makers, and such data tend to be accessed through personal contacts (Oliver and de Vocht, 2015). In addition, the form in which data is reported to policy-makers appears to make a difference to its use, and it must be accessible (Brownson et al., 2009). Therefore, the additional dissemination activity, such as blogs, newsletter pieces, conference presentations, individual presentation to local CCGs (plus providing them with a short policy brief for circulation), press output (e.g. a pulse article reporting the systematic review - http://www.pulsetoday.co.uk/home/finance-and-practice-life-news/case-management-has-no-effect-on-outcomes-or-costs-find-researchers/20010505.fullarticle) etc., all potentially add strength to this study, by communicating the data in a more accessible manner.

8.8 Further research opportunities

As touched upon throughout this discussion section, there are a number of possibilities for further research. These are brought together and briefly summarised here.

Methodological

There is a need to develop methodological approaches that better characterise effective solutions to the 'multimorbidity problem', and which do not simply define the extent of the problem we face. This may not involve counting conditions, or 'conditions', in the traditional sense at all (e.g. a functional measure may be an alternative). It may be that a much broader and personalised conceptualisation is required. If so, how can we

model this with the available data, or what other data do we need? For example, can we take a leaf out of the psychometrics research agenda and better harness big data effectively? There might also be broader work to be done to address known issues with population-based research methods and their application to individuals. And, there is a need to decide on methods that provide 'good enough' evidence for policy-making in a cost-effective manner. For example, for an intervention where there is no possibility of patient harm, do we really require an expensive RCT to test, and most probably mis-estimate, its causal effects? And, if not, how do we convince policy-makers that this alternative evidence is thorough enough?

Theoretical/ Practical

It is important to test new interventions that may help us prevent and/or compress multimorbidity and its negative effects. With burden of care a particularly important concept for the multimorbid patient, there might be means to withdraw inappropriate care when advisable to do so, simultaneously reducing resource waste. There are further ethical considerations concerning the possibility/acceptability of identifying treatment futility points (points where increasing/ongoing treatment has no beneficial effect for the patient, where the quality of life might alternatively benefit from stepping down care) for individuals, and these should be debated. We might also consider whether there is any over-medicalisation/over-diagnosis of conditions, in order to prevent potential over-inflation of our expectations of multimorbidity, and limit any unnecessary treatments for the individual before they occur. And there is a need to address how we best use the multimorbidity concept to address patient complexity, or if there is a better method/conception by which to stratify groups.

It is necessary for researchers to prioritise aims of integrated care and adopt an accepted definition that addresses these aims adequately. If decision-makers choose to continue to employ case management as their primary method of changing care at the service-delivery level, it is necessary to examine how the model is best employed effectively in different contexts.

As outlined above, there might also be other methods of addressing the significant health-system challenges, present and future. There is the need to consider decommissioning services (and how we do this as well as possible without too much

political controversy) that have a more cost-effective alternative, in order to plan effective *health systems* instead of thinking simply in terms of these individual services (Gray, 2011). There are also moral debates to be had on the best place of the patient in decision-making for the health system, and the implications this has for the meso- and macro-level context. The new models of care emerging should be examined in terms of strategic components and underlying conceptual assumptions so we can work out what exactly is effective and in what context. There is the need also to evaluate the effectiveness of these models in terms of the outcomes that they seek to address.

8.9 Chapter Conclusions

In this chapter, I have discussed the thesis findings in the context of the broader literature, suggested alternative policy options to achieve health system aims, and highlighted areas in need of further research.

9. Thesis Conclusions

The current challenges facing health systems have increased the requirement and urgency of adapting health systems and services to better meet the needs of the 21st century. Integrated care has been suggested as a means of accomplishing better health outcomes for those with long-term conditions and multimorbidity, increasing their satisfaction with services, while reducing costs. The current implementation of the concept predominantly through MDT case management, however, has not proven able to meet these multiple aims. The current financial climate might mean that other means of achieving prioritised aims are required in the short-term, with *comprehensive* primary care and population health strategies employed to better prevent/compress the negative effects of lifestyle-associated conditions in the longer-term.

References

Angrist, J. D. & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*: Princeton university press.

Armitage, G. D., Suter, E., Oelke, N. D. & Adair, C. E. (2009). Health systems integration: state of the evidence. *International Journal of Integrated Care,* 9(2)**,** 1-11.

Atkinson, S. & Haran, D. (2005). Individual and district scale determinants of users' satisfaction with primary health care in developing countries. *Social Science & Medicine,* 60(3)**,** 501-513.

Atun, R. (2004). What are the advantages and disadvantages of restructuring a health care system to be more focused on primary care services. Available: http://www.euro.who.int/__data/assets/pdf_file/0004/74704/E82997.pdf [Accessed 01/10/2013].

Atun, R. (2012). Health systems, systems thinking and innovation. *Health Policy and Planning,* 27(suppl 4)**,** iv4-iv8.

Atun, R., Aydin, S., Chakraborty, S., Sümer, S., Aran, M., Gürol, I., Nazlioilu, S., Özgülcü, e., Aydoian, Ü., Ayar, B., Dilmen, U. u. & Akda, R. (2013). Universal health coverage in Turkey: enhancement of equity. *The Lancet,* 382(9886)**,** 65-99.

Atun, R. & Menabde, N. (2008). Health systems and systems thinking. *Health systems and the challenge of communicable diseases: experiences from Europe and Latin America. Berkshire (United Kingdom): Open University Press, McGraw Hill Education***,** 121-140.

Bähler, C., Huber, C. A., Brüngger, B. & Reich, O. (2015). Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. *BMC Health Services Research,* 15(1)**,** 1-12.

Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S. & Guthrie, B. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet,* 380(9836)**,** 37-43.

Bayliss, E. A., Edwards, A. E., Steiner, J. F. & Main, D. S. (2008). Processes of care desired by elderly patients with multimorbidities. *Family Practice,* 25(4)**,** 287-293.

Bell, D. (1976). The coming of the post-industrial society. *The Educational Forum,* 40(4)**,** 574-579.

Berwick, D. M., Nolan, T. W. & Whittington, J. (2008). The triple aim: care, health, and cost. *Health Aff (Millwood),* 27(3)**,** 759-769.

Bevan, G., Karanikolos, M., Exley, J., Nolte, E., Connolly, S. & Mays, N. (2014). The four health systems of the United Kingdom: how do they compare? Available: http://www.nuffieldtrust.org.uk/sites/files/nuffield/140411_four_countries_health_systems_full_report.pdf [Accessed 01/02/2015].

Black, N. (1996). Why we need observational studies to evaluate the effectiveness of health care. *BMJ,* 312(7040)**,** 1215.

Bleich, S. N., Ozaltin, E. & Murray, C. K. (2009). How does satisfaction with the health-care system relate to patient experience? *Bulletin of the World Health Organization,* 87(4)**,** 271-278.

BMA (2014). What is social care, and how can health services better integrate with it? : British Medical Association.

Bodenheimer, T., Wagner, E. H. & Grumbach, K. (2002a). Improving primary care for patients with chronic illness. *JAMA,* 288(14)**,** 1775-1779.

Bodenheimer, T., Wagner, E. H. & Grumbach, K. (2002b). Improving primary care for patients with chronic illness: The chronic care model, part 2. *JAMA,* 288(15)**,** 1909-1914.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T. & Moore, L. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine,* 75(12)**,** 2299-2306.

Boult, C., Karm, L. & Groves, C. (2008). Improving Chronic Care: The "Guided Care" Model. *The Permanente Journal,* 12(1)**,** 50-54.

Bouvard, V., Loomis, D., Guyton, K. Z., Grosse, Y., Ghissassi, F. E., Benbrahim-Tallaa, L., Guha, N., Mattock, H. & Straif, K. (2015). Carcinogenicity of consumption of red and processed meat. *The Lancet Oncology,* 16(6)**,** 1599-1600.

Bower, P., Campbell, S., Bojke, C. & Sibbald, B. (2003). Team structure, team climate and the quality of care in primary care: an observational study. *Quality & safety in health care,* 12(4)**,** 273-279.

Bower, P., Hann, M., Rick, J., Rowe, K., Burt, J., Roland, M., Protheroe, J., Richardson, G. & Reeves, D. (2013). Multimorbidity and delivery of care for long-term conditions in the English National Health Service: baseline data from a cohort study. *J Health Serv Res Policy,* 18(2 Suppl)**,** 29-37.

Boyd, C. M., Darer, J., Boult, C., Fried, L. P., Boult, L. & Wu, A. W. (2005). Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA,* 294(6)**,** 716-724.

Boyd, C. M. & Fortin, M. (2010). Future of multimorbidity research: How should understanding of multimorbidity inform health system design. *Public Health Rev,* 32(2)**,** 451-474.

Boyle, S. (2011). United Kingdom (England): Health system review. *Health Systems in Transition,* 13(1)**,** 1-486.

Bray, G. A., Nielsen, S. J. & Popkin, B. M. (2004). Consumption of high-fructose corn syrup in beverages may play a role in the epidemic of obesity. *The American Journal of Clinical Nutrition,* 79(4)**,** 537-543.

Brilleman, S. L., Gravelle, H., Hollinghurst, S., Purdy, S., Salisbury, C. & Windmeijer, F. (2014). Keep it simple? Predicting primary health care costs with clinical morbidity measures. *Journal of Health Economics,* 35**,** 109-122.

Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A. & Peters, T. J. (2004). Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol,* 57(3)**,** 229-36.

Brownson, R. C., Chriqui, J. F. & Stamatakis, K. A. (2009). Understanding Evidence-Based Public Health Policy. *American Journal of Public Health,* 99(9)**,** 1576-1583.

Bunker, J. P. (2001). The role of medical care in contributing to health improvements within societies. *International Journal of Epidemiology,* 30(6)**,** 1260-1263.

Cameron, A. (2016). What have we learnt about joint working between health and social care? *Public Money & Management,* 36(1)**,** 7-14.

Campion , E. W. & Morrissey , S. (2013). A Different Model — Medical Care in Cuba. *New England Journal of Medicine,* 368(4)**,** 297-299.

Chatterji, S., Byles, J., Cutler, D., Seeman, T. & Verdes, E. (2014). Health, functioning, and disability in older adults - present status and future implications. *The Lancet,* 385(9967)**,** 563-575.

Chechulin, Y., Nazerian, A., Rais, S. & Malikov, K. (2014). Predicting Patients with High Risk of Becoming High-Cost Healthcare Users in Ontario (Canada). *Healthcare Policy,* 9(3)**,** 68-79.

Checkland, K., Allen, P., Coleman, A., Segar, J., McDermott, I., Harrison, S., Petsoulas, C. & Peckham, S. (2013). Accountable to whom, for what? An exploration of the early development of Clinical Commissioning Groups in the English NHS. *BMJ Open,* 3(12)**,** e003769.

Chen, C., Garrido, T., Chock, D., Okawa, G. & Liang, L. (2009). The Kaiser Permanente Electronic Health Record: transforming and streamlining modalities of care. *Health Aff (Millwood),* 28(2)**,** 323-333.

Clegg, A., Bates, C., Young, J., Ryan, R., Nichols, L., Ann Teale, E., Mohammed, M. A., Parry, J. & Marshall, T. (2016). Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age and Ageing,* 45(3)**,** 353-360.

Coleman, A., Segar, J., Checkland, K., McDermott, I., Harrison, S. & Peckham, S. (2015). Leadership for health commissioning in the new NHS: Exploring the early development of clinical commissioning groups in England. *Journal of Health Organization and Management,* 29(1)**,** 75-91.

Collins, E. & Klein, R. (1980). Equity and the NHS: self-reported morbidity, access, and primary care. *BMJ,* 281(6248)**,** 1111-1115.

Coventry, P., Lovell, K., Dickens, C., Bower, P., Chew-Graham, C., McElvenny, D., Hann, M., Cherrington, A., Garrett, C., Gibbons, C. J., Baguley, C., Roughley, K., Adeyemi, I., Reeves, D., Waheed, W. & Gask, L. (2015). Integrated primary care for patients with mental and physical multimorbidity: cluster randomised controlled trial of collaborative care for patients with depression comorbid with diabetes or cardiovascular disease. *BMJ,* 350**,** h638.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ,* 337**,** a1655.

Cross, M. (2006). Will Connecting for Health deliver its promises? *BMJ,* 332(7541)**,** 599-601.

Curry, N. & Ham, C. (2010). Clinical and Service Integration: The route to improved outcomes. Available: https://www.kingsfund.org.uk/sites/files/kf/Clinical-and-service-integration-Natasha-Curry-Chris-Ham-22-November-2010.pdf [Accessed 25/10/2013].

Darzi, A. & Howitt, P. (2012). Integrated care cannot be designed in Whitehall. *International Journal of Integrated Care,* 12(2)**,** e43.

Davies, A., Ariti, C., Geoghiou, T. & Bardsley, M. (2015). Evaluation of complex health and care interventions using retrospective matched control methods. 1-28. Available: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/evaluation_report_final_0.pdf [Accessed 03/01/2016].

Davis, K., Stremikis, K., Schoen, C. & Squires, D. (2014). Mirror, mirror on the wall, 2014 update: how the US health care system compares internationally. Available: http://www.commonwealthfund.org/~/media/files/publications/fund-report/2014/jun/1755_davis_mirror_mirror_2014.pdf [Accessed 01/02/2015].

Department of Health (2009). NHS Constitution. London, UK: Department of Health.

Department of Health (2012). Health and Social Care Act 2012. London, UK: Department of Health.

Department of Health. (2013). Integrated Care and Support: Our Shared Commitment. 1-46. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198748/DEFINITIVE_FINAL_VERSION_Integrated_Care_and_Support_-_Our_Shared_Commitment_2013-05-13.pdf [Accessed 01/09/2013].

Department of Health. (2014a). Adult social care in England: overview. 1-62. Available: https://www.nao.org.uk/wp-content/uploads/2015/03/Adult-social-care-in-England-overview.pdf [Accessed 09/10/2015].

Department of Health. (2014b). Comorbidities: A framework of principles for system-wide action. 1-7. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/307143/Comorbidities_framework.pdf [Accessed 05/05/2015].

Diederichs, C., Berger, K. & Bartels, D. B. (2011). The measurement of multiple chronic diseases--a systematic review on existing multimorbidity indices. *J Gerontol A Biol Sci Med Sci,* 66(3)**,** 301-311.

Dimick, J. B. & Ryan, A. M. (2014). Methods for evaluating changes in health care policy: The difference-in-differences approach. *JAMA,* 312(22)**,** 2401-2402.

Drennan, V. & Goodman, C. (2004). Nurse-led case management for older people with long-term conditions. *Br J Community Nurs,* 9(12)**,** 527-533.

Drennan, V., Goodman, C., Manthorpe, J., Davies, S., Scott, C., Gage, H. & Iliffe, S. (2011). Establishing new nursing roles: a case study of the English community matron initiative. *J Clin Nurs,* 20(19-20)**,** 2948-2957.

Erens, B., Wistow, G., Mounier-Jack, S., Douglas, N., Jones, L., Manacorda, T. & Mays, N. (2015). Early evaluation of the Integrated Care and Support Pioneers Programme. Available: http://www.piru.ac.uk/assets/files/Early_evaluation_of_IC_Pioneers_Final_Report.pdf [Accessed 10/01/2016].

Ettelt, S., Mays, N. & Allen, P. (2015). The Multiple Purposes of Policy Piloting and Their Consequences: Three Examples from National Health and Social Care Policy in England. *Journal of Social Policy,* 44(02)**,** 319-337.

Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases,* 23(7)**,** 455-468.

Fenton, J. J., Jerant, A. F., Bertakis, K. D. & Franks, P. (2012). The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch Intern Med,* 172(5)**,** 405-411.

Fisher, E. S., Staiger, D. O., Bynum, J. P. W. & Gottlieb, D. J. (2007). Creating Accountable Care Organizations: The Extended Hospital Medical Staff. *Health Affairs,* 26(1)**,** w44-w57.

Fortin, M., Dubois, M.-F., Hudon, C., Soubhi, H. & Almirall, J. (2007a). Multimorbidity and quality of life: a closer look. *Health and Quality of Life Outcomes,* 5**,** 52.

Fortin, M., Soubhi, H., Hudon, C., Bayliss, E. A. & van den Akker, M. (2007b). Multimorbidity's many challenges. *BMJ,* 334(7602)**,** 1016.

Fortin, M., Stewart, M., Poitras, M.-E., Almirall, J. & Maddocks, H. (2012). A Systematic Review of Prevalence Studies on Multimorbidity: Toward a More Uniform Methodology. *The Annals of Family Medicine,* 10(2)**,** 142-151.

Fraccaro, P., Arguello Casteleiro, M., Ainsworth, J. & Buchan, I. (2015). Adoption of Clinical Decision Support in Multimorbidity: A Systematic Review. *JMIR Medical Informatics,* 3(1)**,** e4.

Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine,* 303(3)**,** 130-135.

Fulop, N., Allen, P., Clarke, A. & Black, N. (2001). *Studying the organisation and delivery of health services: research methods*. London, UK: Psychology Press.

GAO. (2015). MEDICAID: A Small Share of Enrollees Consistently Accounted for a Large Share of Expenditures. 1-45. Available: http://www.gao.gov/assets/680/670112.pdf [Accessed 01/06/2016].

Gawande, A. (2011). The hot spotters. Can we lower medical costs by giving the neediest patients better care? . *The New Yorker* [Online], (January 24, 2011). Available: http://www.newyorker.com/magazine/2011/01/24/the-hot-spotters [Accessed 01/02/2016].

Glynn, L. G., Valderas, J. M., Healy, P., Burke, E., Newell, J., Gillespie, P. & Murphy, A. W. (2011). The prevalence of multimorbidity in primary care and its effect on health care utilization and cost. *Family Practice,* 28(5)**,** 516-523.

Goodwin, N., Smith, J., Davies, A., Perry, C., Rosen, R., Dixon, A., Dixon, J. & Ham, C. (2012). Integrated care for patients and populations: Improving outcomes by working together. Available: http://www.kingsfund.org.uk/sites/files/kf/integrated-care-patients-populations-paper-nuffield-trust-kings-fund-january-2012.pdf [Accessed 09/09/2013].

Government Statistical Service. (2015). The 2011 Rural-Urban Classification for Local Authority Districts in England. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427117/RUCLAD_leaflet_May2015.pdf [Accessed 05/04/2016].

Gravelle, H., Dusheiko, M., Sheaff, R., Sargent, P., Boaden, R., Pickard, S., Parker, S. & Roland, M. (2007). Impact of case management (Evercare) on frail elderly patients: controlled before and after analysis of quantitative outcome data. *BMJ,* 334(7583)**,** 31.

Gray, J. A. M. (2011). *How to build healthcare systems*. Oxford, UK: Offox Press.

Greater Manchester NIHR CLAHRC. (2015). NHS Greater Manchester Primary Care Demonstrator Evaluation. Available: http://clahrc-gm.nihr.ac.uk/wp-content/uploads/PCDE-final-report-full-final.pdf [Accessed 01/07/2015].

Gröne, O. & Garcia-Barbero, M. (2001). Integrated care: a position paper of the WHO European office for integrated health care services. *International Journal of Integrated Care,* 1**,** e21.

Guyatt, G. H., Sackett, D. L., Sinclair, J. C. & et al. (1995). Users' guides to the medical literature: IX. A method for grading health care recommendations. *JAMA,* 274(22)**,** 1800-1804.

Hall Aitken (2014). Central Manchester Practice Integrated CareTeam Report. Manchester, UK: Hall Aitken.

Hardy, G. E., West, M. A. & Hill, F. (1996). Components and predictors of patient satisfaction. *British Journal of Health Psychology,* 1(1)**,** 65-85.

Harris, M. & Haines, A. (2010). Brazil's Family Health Programme. *BMJ,* 341.

Hays, R. & Daker-White, G. (2015). The care.data consensus? A qualitative analysis of opinions expressed on Twitter. *BMC Public Health,* 15(1)**,** 1-13.

Heckman, J. J. & Leamer, E. (2007). *Handbook of econometrics* (Vol. 6): Elsevier.

Holgate, S., Grigg, J., Agius, R., Ashton, J. R., Cullinan, P., Exley, K., Fishwick, D., Fuller, G., Gokani, N. & Griffiths, C. (2016). Every breath we take: The lifelong impact of air pollution, Report of a working party. Available: https://www.rcplondon.ac.uk/file/2912/download?token=5pFurNnk [Accessed 01/09/2016].

Houweling, T. A., Kunst, A. E., Looman, C. W. & Mackenbach, J. P. (2005). Determinants of under-5 mortality among the poor and the rich: a cross-national analysis of 43 developing countries. *International Journal of Epidemiology,* 34(6)**,** 1257-1265.

Howick, J., Glasziou, P. & Aronson, J. K. (2009). The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? *Journal of the Royal Society of Medicine,* 102(5)**,** 186-194.

Hubert, H. B., Feinleib, M., McNamara, P. M. & Castelli, W. P. (1983). Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation,* 67(5)**,** 968-977.

Huntley, A. L., Thomas, R., Mann, M., Huws, D., Elwyn, G., Paranjothy, S. & Purdy, S. (2013). Is case management effective in reducing the risk of unplanned hospital admissions for older people? A systematic review and meta-analysis. *Family Practice,* 30(3)**,** 266-275.

Hutton, G. & Haller, L. (2004). Evaluation of the costs and benefits of water and sanitation improvements at the global level. Available: http://www.who.int/water_sanitation_health/wsh0404.pdf [Accessed 09/09/2013].

Iacobucci, G. (2014). A third of NHS contracts awarded since health act have gone to private sector, BMJ investigation shows. *BMJ,* 349.

Jenkins, S. P. (1995). Accounting for inequality trends: decomposition analyses for the UK, 1971-86. *Economica***,** 29-63.

Johnson, T. L., Rinehart, D. J., Durfee, J., Brewer, D., Batal, H., Blum, J., Oronce, C. I., Melinkovich, P. & Gabow, P. (2015). For Many Patients Who Use Large Amounts Of Health Care Services, The Need Is Intense Yet Temporary. *Health Affairs,* 34(8)**,** 1312-1319.

Jovic, D., Marinkovic, J. & Vukovic, D. (2016). Association between body mass index and prevalence of multimorbidity: a cross-sectional study. *Public Health,* 139**,** 103-111.

Kasteridis, P., Street, A., Dolman, M., Gallier, L., Hudson, K., Martin, J. & Wyer, I. (2015). Who would most benefit from improved integrated care? Implementing an analytical strategy in South Somerset. *International Journal of Integrated Care,* 15**,** e001.

Keller, K., Corbett, J. & Nichols, D. (1998). Repetitive strain injury in computer keyboard users: Pathomechanics and treatment principles in individual and group intervention. *Journal of Hand Therapy,* 11(1)**,** 9-26.

Kersbergen, A. L. (1996). Case management: A rich history of coordinating care to control costs. *Nursing Outlook,* 44(4)**,** 169-172.

Kodner, D. L. & Spreeuwenberg, C. (2002). Integrated care: meaning, logic, applications, and implications--a discussion paper. *Int J Integr Care,* 2**,** e12.

Koroukian, S. M., Schiltz, N., Warner, D. F., Sun, J., Bakaki, P. M., Smyth, K. A., Stange, K. C. & Given, C. W. (2016). Combinations of Chronic Conditions, Functional Limitations, and Geriatric Syndromes that Predict Health Outcomes. *J Gen Intern Med,* 31(6)**,** 630-637.

Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V. & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *Am Psychol,* 70(6)**,** 543-556.

Kosinski, M., Stillwell, D. & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences,* 110(15)**,** 5802-5805.

Lafond, S. (2015). Current NHS spending in England. Available: http://www.health.org.uk/sites/default/files/FundingOverview_CurrentNHSSpendingInEngland.pdf [Accessed 02/02/2016].

Lawn, J. E., Rohde, J., Rifkin, S., Were, M., Paul, V. K. & Chopra, M. (2008). Alma-Ata 30 years on: revolutionary, relevant, and time to revitalise. *The Lancet,* 372(9642)**,** 917-927.

Lefèvre, T., d'Ivernois, J. F., De Andrade, V., Crozet, C., Lombrail, P. & Gagnayre, R. (2014). What do we mean by multimorbidity? An analysis of the literature on multimorbidity measures, associated factors, and impact on health services organization. *Revue d'Épidémiologie et de Santé Publique,* 62(5)**,** 305-314.

Leutz, W. N. (1999). Five laws for integrating medical and social services: lessons from the United States and the United Kingdom. *Milbank Q,* 77(1)**,** 77-110, iv-v.

Lewis, G., Curry, N. & Bardsley, M. (2011). Choosing a predictive risk model: a guide for commissioners in England. Available: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/choosing_predictive_risk_model_guide_for_commissioners_nov11.pdf [Accessed 02/12/2013].

Lipsitz, L. A. (2012). Understanding Health Care as a Complex System: The Foundation for Unintended Consequences. *JAMA,* 308(3)**,** 243-244.

Lloyd, J. & Wait, S. (2006). Integrated care: a guide for policymakers. Available: www.ilcuk.org.uk/images/uploads/publication-pdfs/pdf_pdf_7.pdf [Accessed 02/10/2013].

Man, M.-S., Chaplin, K., Mann, C., Bower, P., Brookes, S., Fitzpatrick, B., Guthrie, B., Shaw, A., Hollinghurst, S., Mercer, S., Rafi, I., Thorn, J. & Salisbury, C. (2016). Improving the management of multimorbidity in general practice: protocol of a cluster randomised controlled trial (The 3D Study). *BMJ Open,* 6(4)**,** e011261.

Marengoni, A., Angleman, S., Melis, R., Mangialasche, F., Karp, A., Garmen, A., Meinow, B. & Fratiglioni, L. (2011). Aging with multimorbidity: a systematic review of the literature. *Ageing research reviews,* 10(4)**,** 430-439.

McKee, M. (1993). Routine data: a resource for clinical audit? *Quality in Health Care,* 2(2)**,** 104.

McKee, M., Basu, S. & Stuckler, D. (2012a). Health systems, health and wealth: the argument for investment applies now more than ever. *Social Science & Medicine,* 74(5)**,** 684-687.

McKee, M., Karanikolos, M., Belcher, P. & Stuckler, D. (2012b). Austerity: a failed experiment on the people of Europe. *Clin Med,* 12(4)**,** 346-350.

McKeigue, P., Shah, B. & Marmot, M. (1991). Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians. *The Lancet,* 337(8738)**,** 382-386.

McLaughlin, J. A. & Jordan, G. B. (1999). Logic models: a tool for telling your programs performance story. *Evaluation and program planning,* 22(1)**,** 65-72.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics,* 13(2)**,** 151-161.

Moffat, K. & Mercer, S. (2015). Challenges of managing people with multimorbidity in today's healthcare systems. *BMC Family Practice,* 16(1)**,** 129.

Monitor. (2013). Closing the NHS funding gap: how to get better value health care for patients 1-23. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/284044/ClosingTheGap091013.pdf [Accessed 05/04/2015].

Moret, L., Nguyen, J. M., Volteau, C., Falissard, B., Lombrail, P. & Gasquet, I. (2007). Evidence of a non-linear influence of patient age on satisfaction with hospital care. *Int J Qual Health Care,* 19(6)**,** 382-389.

Morris, S., Devlin, N. & Parkin, D. (2007). *Economic analysis in health care*. West Sussex, UK: John Wiley & Sons.

Newton, J. N., Briggs, A. D. M., Murray, C. J. L., Dicker, D., Foreman, K. J., Wang, H., Naghavi, M., Forouzanfar, M. H., Ohno, S. L., Barber, R. M., Vos, T., Stanaway, J. D., Schmidt, J. C., Hughes, A. J., Fay, D. F. J., Ecob, R., Gresser, C., McKee, M., Rutter, H., Abubakar, I., Ali, R., Anderson, H. R., Banerjee, A., Bennett, D. A., Bernabé, E., Bhui, K. S., Biryukov, S. M., Bourne, R. R., Brayne, C. E. G., Bruce, N. G., Brugha, T. S., Burch, M., Capewell, S., Casey, D., Chowdhury, R., Coates, M. M., Cooper, C., Critchley, J. A., Dargan, P. I., Dherani, M. K., Elliott, P., Ezzati, M., Fenton, K. A., Fraser, M. S., Fürst, T., Greaves, F., Green, M. A., Gunnell, D. J., Hannigan, B. M., Hay, R. J., Hay, S. I., Hemingway, H., Larson, H. J., Looker, K. J., Lunevicius, R., Lyons, R. A., Marcenes, W., Mason-Jones, A. J., Matthews, F. E., Moller, H., Murdoch, M. E., Newton, C. R., Pearce, N., Piel, F. B., Pope, D., Rahimi, K., Rodriguez, A., Scarborough, P., Schumacher, A. E., Shiue, I., Smeeth, L., Tedstone, A., Valabhji, J., Williams, H. C., Wolfe, C. D. A., Woolf, A. D. & Davis, A. C. J. (2015). Changes in health in England, with analysis by English regions and areas of deprivation, 1990-2013;2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet,* 386(10010)**,** 2257-2274.

NHS Choices. (2015). The history of the NHS in England - The early years 1948-1959. Available: http://www.nhs.uk/NHSEngland/thenhs/nhshistory/Pages/NHShistory1948.aspx [Accessed 06/05/2016].

NHS England. (2013). 2013/14 general medical services (GMS) contract - guidance and audit requirements for new and amended enhanced services: Risk profiling and care management scheme. Available:

http://www.nhsemployers.org/~/media/Employers/Publications/2013-14-GMS-contract-Guidance-audit-requirements.pdf [Accessed 01/02/2015].
NHS England. (2014a). Capitation: a potential new payment model to enable integrated care. 1-44. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/445731/LPE_Capitation.pdf [Accessed 09/06/2015].
NHS England. (2014b). Five Year Forward View. 1-41. Available: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf [Accessed 01/11/2015].
NHS England. (2016a). Better Care Fund. 2016. Available: https://www.england.nhs.uk/ourwork/part-rel/transformation-fund/bcf-plan/ [Accessed 08/08/2016].
NHS England. (2016b). New care models - vanguard sites. 2016. Available: https://www.england.nhs.uk/ourwork/futurenhs/new-care-models/ [Accessed 18/02/2016].
NHS England. (2016c). People helping people: Year two of the pioneer programme. 1-45. Available: https://www.england.nhs.uk/pioneers/wp-content/uploads/sites/30/2016/01/pioneer-programme-year2-report.pdf [Accessed 10/08/2016].
NHS England. (2016d). Simon Stevens speech to NHS Confederation Conference 2016. 2016. Available: https://www.england.nhs.uk/2016/06/simon-stevens-confed-speech/ [Accessed 30/06/2016].
NICE. (2016a). Multimorbidity: clinical assessment and management (Draft for consultation). 1-450. Available: https://www.nice.org.uk/guidance/NG56/documents/draft-guideline [Accessed 01/04/2016].
NICE. (2016b). Multimorbidity: clinical assessment and management NICE guideline. 1-23. Available: https://www.nice.org.uk/guidance/ng56/resources/multimorbidity-clinical-assessment-and-management-1837516654789 [Accessed 22/09/2016].
Noble, H. & Lewis, R. (2008). Assessing palliative care needs in end-stage kidney disease. *Nursing Times,* 104(23)**,** 26-27.
Nolte, E. & McKee, M. (2008). Caring for people with chronic conditions: a health system perspective. Available: http://www.euro.who.int/__data/assets/pdf_file/0006/96468/E91878.pdf [Accessed 01/10/2013].
Nolte, E. & Pitchforth, E. (2014). What is the evidence on the economic impacts of integrated care. Available: http://www.euro.who.int/__data/assets/pdf_file/0019/251434/What-is-the-evidence-on-the-economic-impacts-of-integrated-care.pdf [Accessed 01/05/2015].

Oliver, K. A. & de Vocht, F. (2015). Defining 'evidence' in public health: a survey of policymakers' uses and preferences. *The European Journal of Public Health*, 1-6.

Omran, A. R. (1971). The epidemiologic transition: a theory of the epidemiology of population change. *The Milbank Memorial Fund Quarterly,* 49(4), 509-538.

ONS. (2014). Population estimates. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates [Accessed 06/05/2016].

ONS. (2015). 2011 Census analysis: Ethnicity and religion of the non-UK born population in England and Wales: 2011. Available: http://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/2011censusanalysisethnicityandreligionofthenonukbornpopulationinenglandandwales/2015-06-18 [Accessed 06/05/2016].

Orueta, J. F., Garcia-Alvarez, A., Garcia-Goni, M., Paolucci, F. & Nuno-Solinis, R. (2014). Prevalence and costs of multimorbidity by deprivation levels in the basque country: a population based study using health administrative databases. *PLoS ONE,* 9(2), e89787.

Osborne, G. (2014). We need a Northern powerhouse. *Museum of Science and Industry, Manchester* [Online]. Available: https://www.gov.uk/government/speeches/chancellor-we-need-a-northern-powerhouse [Accessed 25/06/2016].

Ouwens, M., Wollersheim, H., Hermens, R., Hulscher, M. & Grol, R. (2005). Integrated care programmes for chronically ill patients: a review of systematic reviews. *Int J Qual Health Care,* 17(2), 141-146.

Owen, L., Morgan, A., Fischer, A., Ellis, S., Hoy, A. & Kelly, M. P. (2011). The cost-effectiveness of public health interventions. *Journal of Public Health*, 1-9.

Oxman, A. D. & Guyatt, G. H. (1992). A consumer's guide to subgroup analyses. *Annals of internal medicine,* 116(1), 78-84.

Panagioti, M., Stokes, J., Esmail, A., Coventry, P., Cheraghi-Sohi, S., Alam, R. & Bower, P. (2015). Multimorbidity and Patient Safety Incidents in Primary Care: A Systematic Review and Meta-Analysis. *PLoS ONE,* 10(8), e0135947.

Parliamentary Office of Science and Technology. (2016). POSTNOTE - Electronic Health Records. Available: http://researchbriefings.files.parliament.uk/documents/POST-PN-0519/POST-PN-0519.pdf [Accessed 01/09/2016].

Piantadosi, S., Byar, D. P. & Green, S. B. (1988). The ecological fallacy. *American journal of epidemiology,* 127(5), 893-904.

Powell Davies, G., Williams, A. M., Larsen, K., Perkins, D., Roland, M. & Harris, M. F. (2008). Coordinating primary health care: an analysis of the outcomes of a systematic review. *Med J Aust,* 188(8 Suppl), S65-68.

Prazeres, F. & Santiago, L. (2015). Prevalence of multimorbidity in the adult population attending primary care in Portugal: a cross-sectional study. *BMJ Open,* 5(9), e009287.

Purdy, S. (2010). Avoiding hospital admissions: what does the research evidence say. Available: http://www.kingsfund.org.uk/sites/files/kf/Avoiding-Hospital-Admissions-Sarah-Purdy-December2010.pdf [Accessed 05/05/2014].

Raine, R., Fitzpatrick, R., Barratt, H., Bevan, G., Black, N., Boaden, R., Bower, P., Campbell, M., Denis, J. L., Devers, K., Dixon-Woods, M., Fallowfield, L., Forder, J., Foy, R., Freemantle, N., Fulop, N. J., Gibbons, E., Gillies, C., Goulding, L., Grieve, R., Grimshaw, J., Howarth, E., Lilford, R. J., McDonald, R., Moore, G., Moore, L., Newhouse, R., O'Cathain, A., Or, Z., Papoutsi, C., Prady, S., Rycroft-Malone, J., Sekhon, J., Turner, S., Watson, S. I. & Zwarenstein, M. (2016). Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. Available: http://www.rds-eastmidlands.nihr.ac.uk/resources/FullReport-hsdr04160.pdf [Accessed 01/10/2016].

RAND Europe & Ernst & Young LLP. (2012). National Evaluation of the Department of Health's Integrated Care Pilots. 1-129. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215103/dh_133127.pdf [Accessed 01/10/2013].

Richardson, W. S. & Doster, L. M. (2014). Comorbidity and multimorbidity need to be placed in the context of a framework of risk, responsiveness, and vulnerability. *J Clin Epidemiol,* 67(3)**,** 244-246.

Roberts, C. & Torgerson, D. J. (1999). Baseline imbalance in randomised controlled trials. *BMJ,* 319(7203)**,** 185-185.

Roberts, E. & Mays, N. (1998). Can primary care and community-based models of emergency care substitute for the hospital accident and emergency (A & E) department? *Health Policy,* 44(3)**,** 191-214.

Roland, M. & Abel, G. (2012). Reducing emergency admissions: are we on the right track? *BMJ,* 345.

Rose, G. (1981). Strategy of prevention: lessons from cardiovascular disease. *British Medical Journal (Clinical research ed.),* 282(6279)**,** 1847-1851.

Ross, S., Curry, N. & Goodwin, N. (2011). Case Management: What it is and how it can best be implemented. Available: http://www.kingsfund.org.uk/sites/files/kf/Case-Management-paper-The-Kings-Fund-Paper-November-2011_0.pdf [Accessed 05/11/2013].

Rothstein, H. R., Sutton, A. J. & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*: John Wiley & Sons.

Ryan, A. M., Burgess, J. F., Jr. & Dimick, J. B. (2014). Why We Should Not Be Indifferent to Specification Choices for Difference-in-Differences. *Health Serv Res,* 50(4)**,** 1211-1235.

Rycroft-Malone, J., McCormack, B., Hutchinson, A. M., DeCorby, K., Bucknall, T. K., Kent, B., Schultz, A., Snelgrove-Clarke, E., Stetler, C. B., Titler, M., Wallin, L. & Wilson, V. (2012). Realist synthesis: illustrating the method for implementation research. *Implementation Science,* 7(1)**,** 1-10.

Salisbury, C., Johnson, L., Purdy, S., Valderas, J. M. & Montgomery, A. A. (2011). Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *The British Journal of General Practice,* 61(582)**,** e12.

Schoenfelder, T., Klewer, J. & Kugler, J. (2011). Determinants of patient satisfaction: a study among 39 hospitals in an in-patient setting in Germany. *International Journal for Quality in Health Care,* 23(5)**,** 503-509.

Shahnavaz, H. (1987). Workplace injuries in the developing countries. *Ergonomics,* 30(2)**,** 397-404.

Sigfrid, L. A., Turner, C., Crook, D. & Ray, S. (2006). Using the UK primary care Quality and Outcomes Framework to audit health care equity: preliminary data on diabetes management. *Journal of Public Health,* 28(3)**,** 221-225.

Smith, S. M., Soubhi, H., Fortin, M., Hudon, C. & O'Dowd, T. (2012). Managing patients with multimorbidity: systematic review of interventions in primary care and community settings. *BMJ,* 345.

Smith, S. M., Soubhi, H., Fortin, M., Hudon, C. & O'Dowd, T. (2016). Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. *Cochrane Database Syst Rev,* 4.

Søreide, K. & Desserud, K. F. (2015). Emergency surgery in the elderly: the balance between function, frailty, fatality and futility. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine,* 23**,** 10.

Sorenson, C., Drummond, M. & Bhuiyan Khan, B. (2013). Medical technology as a key driver of rising health expenditure: disentangling the relationship. *ClinicoEconomics and Outcomes Research: CEOR,* 5**,** 223-234.

Starfield, B. (1994). Is primary care essential? *The Lancet,* 344(8930)**,** 1129-1133.

Starfield, B. & Shi, L. (2002). Policy relevant determinants of health: an international perspective. *Health Policy,* 60(3)**,** 201-218.

Starfield, B., Shi, L. & Macinko, J. (2005). Contribution of Primary Care to Health Systems and Health. *Milbank Q,* 83(3)**,** 457-502.

Stokes, J. (2014). CCG implementation of integrated care in the NHS. *BMC Health Services Research,* 14(Suppl 2)**,** P119.

Stokes, J., Checkland, K. & Kristensen, S. R. (2016). Integrated care: theory to practice. *Journal of Health Services Research & Policy,* 21(4)**,** 282-285.

Stokes, J., Gurol-Urganci, I., Hone, T. & Atun, R. (2015a). Effect of health system reforms in Turkey on user satisfaction. *Journal of Global Health,* 5(2)**,** 020403.

Stokes, J., Panagioti, M., Alam, R., Checkland, K., Cheraghi-Sohi, S. & Bower, P. (2015b). Effectiveness of Case Management for 'At Risk' Patients in Primary Care: A Systematic Review and Meta-Analysis. *PLoS ONE,* 10(7)**,** e0132340.

Sturmberg, J. P. (2014). Multimorbidity and chronic disease: an emergent perspective. *Journal of Evaluation in Clinical Practice,* 20(4)**,** 508-512.

The King's Fund. (2012). Future trends - Demography. Available: http://www.kingsfund.org.uk/time-to-think-differently/trends/demography [Accessed 06/05/2016].

The King's Fund. (2015a). Is the NHS being privatised? Available: http://www.kingsfund.org.uk/projects/verdict/nhs-being-privatised [Accessed 30/06/2016].
The King's Fund. (2015b). Spending on social care for older people. Available: http://www.kingsfund.org.uk/projects/nhs-in-a-nutshell/spending-social-care-older-people [Accessed 30/07/2016].
The King's Fund. (2016). The NHS budget and how it has changed. Available: http://www.kingsfund.org.uk/projects/nhs-in-a-nutshell/nhs-budget [Accessed 30/06/2016].
The Nuffield Trust. (2007). Evaluation of community matrons on hospital use. Available: http://www.nuffieldtrust.org.uk/node/463 [Accessed 10/08/2016].
The University of Manchester. (2016). Citizens' jury says "share my data" following the UK's first ever health records 'trial'. Available: http://www.manchester.ac.uk/discover/news/citizens-jury-says-share-my-data-following-the-uks-first-ever-health-records-trial [Accessed 30/07/2016].
The World Bank. (2016). Data - United Kingdom. Available: http://data.worldbank.org/country/united-kingdom [Accessed 06/05/2016].
Toynbee, P. (2007). NHS: the Blair years. *BMJ,* 334(7602)**,** 1030-1031.
United Nations. (2015). World Population Prospects: The 2015 Revision. Available: http://esa.un.org/unpd/wpp/publications/files/key_findings_wpp_2015.pdf [Accessed 01/02/2016].
Valderas, J. M., Starfield, B., Sibbald, B., Salisbury, C. & Roland, M. (2009). Defining comorbidity: implications for understanding health and health services. *Ann Fam Med,* 7(4)**,** 357-63.
Valentijn, P. P., Boesveld, I. C., van der Klauw, D. M., Ruwaard, D., Struijs, J. N., Molema, J. J. W., Bruijnzeels, M. A. & Vrijhoef, H. J. (2015). Towards a taxonomy for integrated care: a mixed-methods study. *International Journal of Integrated Care,* 15**,** 1-18.
Valentijn, P. P., Schepman, S. M., Opheij, W. & Bruijnzeels, M. A. (2013). Understanding integrated care: a comprehensive conceptual framework based on the integrative functions of primary care. *International Journal of Integrated Care,* 13**,** 1-12.
van den Akker, M., Buntinx, F. & Knottnerus, J. A. (1996). Comorbidity or multimorbidity: what's in a name? A review of literature. *European Journal of General Practice,* 2(2)**,** 65-70.
van den Akker, M., Buntinx, F., Metsemakers, J. F. M., Roos, S. & Knottnerus, J. A. (1998). Multimorbidity in General Practice: Prevalence, Incidence, and Determinants of Co-Occurring Chronic and Recurrent Diseases. *Journal of Clinical Epidemiology,* 51(5)**,** 367-375.
van den Akker, M., Gunn, J., Mercer, S. W., Fortin, M. & Smith, S. M. (2015). Secondary analysis of data on comorbidity/multimorbidity: a call for papers. *2015,* 5(1)**,** 2.

Wagner, E. H. (1998). Chronic disease management: what will it take to improve care for chronic illness? *Effective clinical practice,* 1(1)**,** 2-4.
Wagner, E. H. (2000). The role of patient care teams in chronic disease management. *BMJ,* 320(7234)**,** 569-572.
Wallace, E., Salisbury, C., Guthrie, B., Lewis, C., Fahey, T. & Smith, S. M. (2015). Managing patients with multimorbidity in primary care. *BMJ,* 350.
Wallace, E., Smith, S. M., Fahey, T. & Roland, M. (2016). Reducing emergency admissions through community based interventions. *BMJ,* 352.
Weiss, C. O., Varadhan, R., Puhan, M. A., Vickers, A., Bandeen-Roche, K., Boyd, C. M. & Kent, D. M. (2014). Multimorbidity and evidence generation. *J Gen Intern Med,* 29(4)**,** 653-660.
Whitty, C. J. (2015). What makes an academic paper useful for health policy? *BMC Medicine,* 13(1)**,** 1.
Wilkinson, R. G. & Marmot, M. G. (2003). *Social determinants of health: the solid facts*. Copenhagen, Denmark: World Health Organization.
Willadsen, T. G., Bebe, A., Koster-Rasmussen, R., Jarbol, D. E., Guassora, A. D., Waldorff, F. B., Reventlow, S. & Olivarius, N. F. (2016). The role of diseases, risk factors and symptoms in the definition of multimorbidity - a systematic review. *Scand J Prim Health Care***,** 1-10.
WK Kellogg Foundation. (2004). Logic Model Development Guide. Available: https://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide [Accessed 31/03/2016].
Wooldridge, J. (2007). What's new in econometrics? Imbens/Wooldridge lecture notes; summer institute 2007, lecture 10: Difference-in-differences estimation. Available: http://www.nber.org/WNE/lect_10_diffindiffs.pdf [Accessed 01/03/2015].
World Health Organization. (1948). Constitution of the World Health Organization. Available: http://www.who.int/governance/eb/who_constitution_en.pdf [Accessed 01/11/2013].
World Health Organization. (1978). Declaration of Alma-Ata. Available: http://www.who.int/publications/almaata_declaration_en.pdf [Accessed 01/09/2013].
World Health Organization. (2007). Everybody's business -strengthening health systems to improve health outcomes: WHO's framework for action. Available: http://www.who.int/healthsystems/strategy/everybodys_business.pdf [Accessed 01/05/2014].
World Health Organization. (2010). The World Health Report - Health Systems Financing: The Path to Universal Coverage. Available: http://apps.who.int/iris/bitstream/10665/44371/1/9789241564021_eng.pdf [Accessed 01/10/2013].
World Health Organization. (2015). WHO global strategy on people-centred and integrated health services: interim report. Available:

http://apps.who.int/iris/bitstream/10665/155002/1/WHO_HIS_SDS_2015.6_eng.pdf [Accessed 01/02/2016].

World Health Organization. (2016). Overview - Preventing chronic diseases: a vital investment. Available: http://www.who.int/chp/chronic_disease_report/part1/en/index11.html [Accessed 01/03/2016].

Yelibi, C., Ward, C. E., Chang, Y., Ashburner, J. M., Hong, C. S. & Atlas, S. J. (2014). Do primary care physicians who deliver better quality of care also use less health care resources? *37th Annual Meeting of the Society of General Internal Medicine.* Springer New York.

Yoon, J., Zulman, D., Scott, J. Y. & Maciejewski, M. L. (2014). Costs associated with multimorbidity among VA patients. *Med Care,* 52 Suppl 3**,** S31-36.

Zulman, D., Asch, S., Martins, S., Kerr, E., Hoffman, B. & Goldstein, M. (2013). Quality of Care for Patients with Multiple Chronic Conditions: The Role of Comorbidity Interrelatedness. *J Gen Intern Med.,* 29(3)**,** 529-537.

Appendices

Appendix 1: Primary focus of integrated care in the NHS (England)

Site	Governance & Organisation	Financing	Resource management	Service Delivery	Primary focus:	Quote:	Notes
CCGs:							
Harrow	•"Working closely" with Council (joint working) •GP Network development	•Pooled budgets (looking to develop - Integration Transformation Fund)	•Integrated IT (looking to develop) •Multi-agency safety Hub developed to co-locate social workers, police, education and health •Healthcare hubs to move some care out of hospital (looking to develop) •Risk stratification	•Case management, 'virtual wards' (GP-led MDT supporting self-management - focused on high admission diabetes/frail elderly) •Primary care mental health team •Intermediate care service (facilitated discharge) •Community outpatient pathways •Specialist palliative care pathway •Care homes CM	MDT case management	"Integrated care focuses on services for patient with ongoing care needs, proactively supporting individuals to maintain fitness and independence across health and social care through active case management and support networks." ('NHS Harrow Commissioning Intentions 13/14' - http://www.harrowccg.nhs.uk/media/7039/harrow_commissioning_intentions_2013-14.pdf)	Early work very centered around case management. Moving towards 'whole systems' integrated care model, at least in terms of planning. Part of NW London Integrated Care Pilot
Hardwick	•Joint commissioning (commited to strengthening)		•Specialist community centers/housing for dementia •Handyvan/trusted trader scheme •Frail Elderly Unit •Risk stratification	•Mental health pathways (developing) •Self-care support •Supporting carers •MDT case management of high-risk patients, 'virtual ward' •End of life care (developing) •Befriending schemes •Home care and enhanced out of hours	MDT care management (virtual ward)	"Hardwick CCG aims to continue to increase the number of patients who can be treated at home or in primary care through the virtual ward model...Having successfully rolled out this model during 2013 the CCG is working with patients and partner organisations to further develop the model, areas of work include dementia care, links to the voluntary and community sector and the launch of the falls partnership vehicle in November 2013" ('NHS HARDWICK CCG GOVERNING BODY MEETING, Tuesday 26 November 2013' - http://www.derbyshire.nhs.uk/ccg/GB-Papers-nov-V3.pdf)	Early work centered on virtual ward case management. Expanding to dementia care, and strengthening links with voluntary sector etc.
Airedale, Wharfdale And Craven	•Joint commissioning (ambition to implement)	•Devolved health and social care budget (working towards) •new tariff and funding model to incentivise care around person needs (ambition)	•IT and estates (in development - "back office systems which will be critical to enabling integration") •Risk stratification	•Telehealth/monitoring •Health and social care teams •Home care •MDT case management of high-risk •Collaborative mental health services (ambition)	MDT case management	"Integrated care involves identifying people with complex and high risk needs and helping them to remedy issues before they become urgent and lead to unplanned hospital or care home admissions." (In Touch, News from the Bradford, Airedale, Wharfedale and Craven CCGs -http://www.airedalewharfedalecravenccg.nhs.uk/wp-content/uploads/2013/07/InTouch-August-2013.pdf)	Early focus on case management as central to integrated care. Ambitions detailed towards 'system transformation, with devolved budgets etc. Also a PIONEER site
Guildford And Waverley	•'Whole system partner group' (plan to implement to monitor and manage changes)		•Risk stratification	•MDT case management teams, virtual ward (frail elderly) •Palliative care pathways •Telehealth/care for frail elderly (when appropriate) •home-based monitoring and support •discharge assessments	MDT care management (virtual ward)	"11.1.2. Integrated Care Through the Local transformation Board, which involves all provider partners we are working together to create an integrate model of care that is patient centred rather than Provider centred • Develop case management "Priority care" Model of care and multi-disciplinary community based teams • Introduction of Frail Elderly pathway including Dementia and EOLC developments" ('NHS Guildford and Waverley Clinical Commissioning Group Integrated Commissioning Plan 2012/13 – 2015/16' - http://www.guildfordandwaverleyccg.nhs.uk/website/X09413/files/L1A-GU-1_1_w-cover_sheet_TEMP.pdf)	Focus on high-risk case management, particularly frail elderly
Central Manchester	•Joint commissioning boards •Identified contracting and performance management environment as key enabler	•identified supportive financial framework as key enabler of integration	•Identified IT setup as key enabler •Graphnet for case managed patients (not fully utilised/rolled out yet) •Risk stratification	•'Manchester Integrated Care Gateway' (triage of referrals) •PICTs MDT case management hospital teams •Falls services •LTC pathways	MDT case management	"• Integrating Care This section describes the key features of our approach and sets out our priorities, including progress in developing Integrated care teams for high risk patients." ('Central Manchester CCG Commissioning Plan' - http://www.manchester.nhs.uk/document_uploads/Commissioning/Central%20Manchester%20CCG%20Commissioning%20plan%20final.pdf)	Focus on MDT case management for now. DevoManc and Healthier Together (PIONEER programme) are both focused on wider system transformation e.g. pooled budget etc.
Lincolnshire West	•Joint commissioning		•section 75 framework enabling easy transfer of resources between partners •Risk stratification	•Frailty pathway •integrated community teams - MDT case management (targeting elderly)	MDT case management	"As a CCG we are continually working to improve the quality of care for our patients. Some of our achievements to date include: • Development of an Integrated Frailty Pathway • Introduction of integrated community teams" ('Lincolshire West CCG, Our healthcare plan 2013/14' - http://www.lincolnshirewestccg.nhs.uk/document-library/cat_view/104-key-documents/131-public-information)	Primarily elderly focus
Bromley	•Joint commissioning	•incentive schemes for prescribing/ case management of complex and vulnerable in own home	•Section 256 agreement for joint resource development with Council •integrated IT (working towards) •Risk stratification	•MDT case management (ProMISE - focused on frail elderly high-risk) •Self-care, telehealth, comprehensive falls prevention service (piloting) •Intermediate care services (working towards stepping up) •End of life care	MDT case management	"Our key work programme for Long Term Conditions is our ProMISE (Proactive Management and Integrated Services for the Elderly) programme." ('Bromley CCG Integrated Plan 2012-15' - http://www.bromleyccg.nhs.uk/NewsPublications/Policies/Policies/Bromley%20CCG%20Integrated%20Plan%202012-15.pdf) "3.2.2 Integrated Care Establishes clearly defined teams of professionals within specified localities to improve communication and learning between the various professional disciplines. This will lead to more streamlined, better co-ordinated services." ('ProMISE Business Case' - http://www.bromleyccg.nhs.uk/about/ourboard/Papers/Enc%2005%20-%20proMISE%20Business%20Case.pdf)	MDT focused, and case management within that. Talk of moving towards 'whole system integration' in the future, learning from NW London specifically
Chiltern	•Joint commissioning		•shared records (working towards) •111 service in combination with social care (developing) •Risk stratification	•self-care support •falls reduction project •telehealth/ telecare	MDT case management	"develop increased integration between health, including mental health, and social care services, building on the agreement to work together "as if a single organisation". Practical schemes include the introduction of "multi-agency groups" in GP Practices" ('NHS CHILTERN CLINICAL COMMISSIONING GROUP OPERATIONAL PLAN 2013/14' - http://www.chilternccg.nhs.uk/practices/bpcc/Operational-Plan-VF1-120413---Final-2.pdf)	Case management focused with desire to improve 'integrated IT' especially
North East Essex	•Joint commissioning and contracting	•Pooled health and social care budget	•Integrated IT (planned to set up and implement by Sept 2013) •Integrated Pathway Hub •Integrated urgent care centre •single point of referral engine •Risk stratification	•primary care mental health services •Frailty pathway •Virtual wards - MDT case management	Pooled budget/ MDT case management	"Integrated care: 3.32 The integration of the provision of services, including the pooling of budgets to reflect local need, should be an explicit consideration in local area planning." ('Everyone Counts Planning for Patients 2013/14' - http://www.neessexccg.nhs.uk/library_uploads/files/everyonecounts-planning.pdf)	Pooled budgets considered key component enabling delivery of integrated care to patient, primarily through case management

West London (K&C & Qpp)	• Integrated Primary Care Mental Health Service •Joint commissioning (for certain services e.g. elderly)	•Pooled budget (for certain services e.g. elderly)	•Provider Network Hubs (being developed for co-location) •Integrated IT with patient access (planned) •Risk stratification	risk) •Rapid Response and Community Independence Service for urgent responses •Wellbeing promotion (working with voluntary sector - ambition to axpand approach) •Intermediate care teams (discharge planning) •Telehealth (planned) •Medicines support	MDT case management	""Putting Patients First is our approach to delivering integrated care for patients with long-term or complex needs who are likely to be frequent users of health and social care services and at risk of emergency admission to hospital." ('An introduction to your local Clinical Commissioning Group (CCG)' - http://www.westlondonccg.nhs.uk/media/6385/wlccg_prospectus.pdf)	Initially focused on MDT case management and reducing admissions of high-risk, but again broadening to try and move towards 'whole system'
Lancashire North	•Joint commissioning			•Transitional Care Pathway •Integrated education teams	Joint commissioning/planning	"5.2.12. Act with a view to promoting integration of both health services with other health services and health services with health-related and social care services where the Group considers that this would improve the quality of services or reduce inequalities" ('NHS LANCASHIRE NORTH CLINICAL COMMISSIONING GROUP Constitution version 2.4' - http://www.lancashirenorthccg.nhs.uk/download/corporate-documents/LNCCG%20Constitution%20v2.4-55883.pdf)	Early focus on working with other services to plan
Bristol	•Locality meetings •LTC/Frail Elderly workstreams		•Investment in out of hospital care/7-day working (planned) •Investing in 'real-time information systems (planned) •Risk stratification	•MDT case management •Co-ordinating nurse leader in practice •Self-care •Geriatricians in community setting •Primary care input to emergency departments •Re-ablement & rehab services	MDT case management	"Our goal is to deliver a practice-based approach to case management that integrates across the health system." ('Bristol CCG Full Integrated Plan' - http://www.bristolccg.nhs.uk/media/4571/full_integrated_plan.pdf)	Case management focused
Eastern Cheshire	•Working together across sectors •Joint commissioning		•Care co-ordination hub supporting case management •Single point of referral •Integrated IT & shared records •Risk stratification	•MDT case management •self-management support •Assistive technology •Shared decision making •Telehealth	MDT case management/Whole system	"It will create a shared view of the population's health and social care needs and will drive improvements by identifying those most at risk and vulnerable. This will be supported by a care co-ordination service providing a central point of information and access, the sharing of information and the use of new technologies to monitor some health conditions remotely." ('Inspiring better health and wellbeing 2013/14 Prospectus' - http://www.easterncheshireccg.nhs.uk/downloads/publications/Strategies/Eastern%20Cheshire%20Prospectus%202013%20-%20FINAL2.pdf)	Case management as focal point - mentioned learning from Kaiser Permanente. Additional ' whole system' integration enablers e.g. IT
East Riding Of Yorkshire	•Joint planning •Integrated discharge steering group		•Integrated primary care centre •24/7 nursing care for dementia (developing) •Risk stratification	•Pathways for ACSCs •Integrated hospital team •Re-ablement schemes •MDT case management •MM clinics (devloping) •self-care support •Dementia model •Telehealth •Homecare •End of life care •Systemised management of LTCs	Self-care/MDT case manageme	"Delivery of an overall 10% reduction in non-elective admissions for patients with a long term condition by 2014. Delivered by •Risk stratification and multi-disciplinary team working will be rolled out to all practices across Hull and the East Riding •Systemised approach to management of patients with Long Term Conditions and other specific conditions agreed and implemented •Increased access to effective pulmonary rahabilitation by end 2012/13" ('OPERATING PLAN 2012 – 17, version 5.1' - http://www.eastridingofyorkshireccg.nhs.uk/data/uploads/publications/eryccg-operating-plan-2012-2017.pdf)	Focus on case management as integration, and self-management as preventer of need for case management
Bradford Districts	•Joint commissioning	•Grants for primary care, social care, voluntary sector to incentivise integrated working practices	•Risk-stratification tool •Integrated IT (for case management - developing) •Telehealth hub	•MDT case management •Virtual ward from hospitals (discharge focus) •Self-care support •Medication review •Telemedicine	MDT case management	"Bradford Districts and City CCGs where the integrated case management approach is live, and where agreed approaches are currently in development." ('Integrated Care Newsletter May 2013' - http://www.bradforddistrictsccg.nhs.uk/wp-content/uploads/2013/05/Integrated-care-newsletter-May-2013.pdf)	CCG has defined integrated care in terms of patient experience and says that in practice this is case management
Leeds North	•Joint planning •Joint commissioning (working towards)		•Risk stratification •Shared records (pilot - plus online access for patient)	•MDT case management •Home care •Integrated urgent care pathways •Telecare/medicine •Systematic self-management support	MDT case management	"Integrated provision of health and social care: Bringing together health and social care professionals to provide integrated care. Integrated services will support those with long term conditions or with a need for on-going care through personalised tailored care" ('Leeds North Clinical Commissioning Group Clear and Credible Plan 2013/14 – 2015/16' - http://www.leedsnorthccg.nhs.uk/Downloads/About%20us/Clear%20and%20Credible%20Plan.pdf)	MDT case management focused. View to add shared IT records etc to support, and self-care to prevent
Newbury And District	•Joint commissioning		•Risk stratification •Shared record (developing)	•(MDT) Case management •Telecare •Self-care •Disease management	(MDT) case management	"Community based care co-ordinators will use a risk stratification tool to identify patients who would most benefit from individualised support and care planning. This may involve a multi-disciplinary team (including Local Authority social care.) High risk patients will have a proactive care plan. Less people will be admitted as an emergency associated with disease progression." ('Newbury & District Clinical Commissioning Group, COMMISSIONING PLAN 2013/14' - http://www.newburyanddistrictccg.nhs.uk/images/publications/PDFs/NDCCG_Commissioning_Plan_2013-14_FINAL_v1.3.pdf)	Case management (which may go towards MDT, but starting GP/case manager-led) and joint commissioning focus
South Devon And Torbay			•Pairing care homes with GP practices and community teams	•MDT case management (virtual ward - input from hospital) •Home care •Mental health integration in primary care •Community nursing •Self-care	MDT care management (virtual ward)	"In 2012 we won both the NHS Alliance Acorn Award and the NAPC Vision Award for integrated care, for our work on 'virtual wards.' These focus on those people who have long-term health conditions and are most at risk of needing to be admitted to hospital if their condition worsens. The virtual wards aim to make sure they get the treatment and care they need so that they can stay well enough to remain in their own home." ('About Us, South Devon and Torbay CCG 2013' - http://southdevonandtorbayccg.nhs.uk/index.php/about-us/op/185-south-devon-and-torbay-ccg-2013)	Virtual ward pioneers. Case management with strong vertical integration

Bradford City		•Trasformational investment fund for pump priming new models	•Shared case management records •Risk stratification	•MDT case management •Self-care support •Virtual ward for LTCs/elderly by hospital staff (intermediate care)	MDT case management	"Working in an integrated way helps organisations to support people as individuals, rather than focusing on only one part of the person at a time. It involves identifying people with the most complex and high risk needs and helping them to remedy issues before they become urgent and lead to unplanned hospital admissions." ('Bradford City CCG Prospectus 2013' - http://www.bradfordcityccg.nhs.uk/wp-content/uploads/2013/05/NHS-Bfd-City-CCG-prospectus.pdf)	Focus on the DES for risk stratification and MDT case management
Central London (Westminster)	•Co-design •Joint commissioning •Dermatology re-tender •Aggregated 'networks of care' based on GP lists	•Integration Transformation Fund, single pooled budget health and social (planned)	•Respiratory services/gynaecology/district nursing redesign •New wheelchair service •Sharing of homelessness data •Sared reecords (starting process)	•MDT case management •Home care •Telehealth •Discharge teams	MDT case management	"Description of approach A network of integrated services will be delivered by a multi-disciplinary team who will work in an together to ensure the patient pathway is seamless, reduces duplication of assessment and ensures the correct outcomes are achieved." ('Central London CCG Commissioning Intentions, version 3.0' - http://www.centrallondonccg.nhs.uk/media/9276/nhscl_ccg_commissioning_intentions_v3_0_final_iteration_dec_13.pdf)	MDT case management focus, moving towards 'whole system integration'
Dorset				•MDT case management	(MDT) case management	"Key priorities for delivery in 2013/14 are: • implementation of integrated teams- expansion of district nursing, intensive case managers" ('NHS Dorset Clinical Commissioning Group Annual delivery plan 2013/14' - http://www.dorsetccg.nhs.uk/Downloads/aboutus/annual%20delivery%20plan%20FINAL%20low%20res%20version.pdf)	Early stage of implementing. Appears to focus on going from single to MDT case management. Little details though
ICPs:							
16 sites	•Formation of single primary and secondary care Trusts •Joint Commissioning	•Pooled budgets	•Shared case management records •Risk stratification •Co-location •Care Hub	•MDT case management •End of life care •Hospital-at-home •Virtual wards •Dementia/mental-health specific services •Rapid response teams •Falls prevention •Self-management support •Prevention services	(MDT) case management	"Most sites adopted some approach that identified populations – the risk in question varied between sites, with the commonest being people at risk of emergency hospital admission. The chosen interventions varied, but a common feature was use of an integrated or multidisciplinary team, with implementation strategies varying from regular meetings between different professionals involved with same patients, to a single, multiprofessional team working within the same building. The virtual ward (a forum in which a patient who is not present is discussed, often by a number of professionals from different specialities) was implemented in a number of sites. Methods by which patients were identified for admission, processes for operating virtual wards and the level of intensity of additional patient care varied between sites. Most commonly, virtual wards were maintained through a key worker or case manager who visited patients and reported back to other clinicians involved in the patient's care through multidisciplinary team (MDT) meetings. Other variations of case management were also used, including one in which a primary professional was assigned to coordinate care for a patient or a group of patients across organisations." ('National Evaluation of the Department of Health's Integrated Care Pilots FINAL REPORT: FULL VERSION' - https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215103/dh_133127.pdf)	Some mix of models, but majority focused on MDT case management. Evaluation report details a number barriers met when sites had originally planned to implement wider organisational/structural integration
PIONEERS:							
25 sites	•GP Networks •Joint commissioning/planning •Involving voluntary sector	•Removing financial disincentives	•Single points of referral •Community hubs (co-location) •Community care hubs (for high-risk patients) •Integrated discharge arrangements •Risk stratification •Shared records •Developing blended workforce roles	•MDT case management •Virtual wards •Self-care support •Patient education (prevention) •Incorporating mental health and dementia services •Telehealth	Whole system integration'/MDT case management	"3.1.1.2 Multi-disciplinary working As described in the pioneer's individual profiles, all are developing new ways for practitioners to work together around the individual. Although exact team configurations vary according to local need, typical features of an integrated care model include the use of joint assessment and care planning, with accountability resting with a lead professional. Single points of referral or access are common too, often maximising the use of technology. Cornwall, for example, developed a 'community line' which is manned by volunteers and provides a trusted link between local multi-disciplinary teams, key workers and local resources. South Devon and Torbay is also developing a single point of access with a unified call centre, services directory, e-hub and website." "The importance of prevention is now well understood. The pioneers' care models aim to increase people's ability to manage their own health, for example through increasing patients' access to information, encouraging them to make healthier choices or involving them in decisions about their own care, which research shows increases wellbeing and satisfaction." ('Integrated Care and Support Pioneer Programme Annual Report 2014' - http://www.local.gov.uk/documents/10180/6927502/Integrated+Care+Pioneer+Programme+Annual+Report+2014/76d562c3-4f7d-4169-91bc-69f7a9be481c) Year 2 report: "Many of the highlights reported by pioneers have taken 18 months to come to fruition, sometimes longer, as sites continue to grapple with the balance of short-term operational requirements and longerterm sustainable transformation." ('People helping people Year two of the pioneer programme' - https://www.england.nhs.uk/pioneers/wp-content/uploads/sites/30/2016/01/pioneer-programme-year2-report.pdf)	Still, predominant shared model is MDT working, and case management. But, more recognition and emphasis given to the barriers and facilitators to integration e.g. payment, governance, IT, workforce transformation etc. Many of the PIONEER sites are building other aspects into their care models therefore, e.g. larger GP networks, co-location in hubs, prevention focus with patient education, self-management support etc. Criteria for selection this time included 'whole system integration'
Vanguards:							MDT case management still key part of most. Interestingly, some moving down to less 'at risk' patients, e.g. Erewash/Sunderland MDT prevention team

Integrated primary and acute care systems (9 sites)	•Formation of single health & social care organisation	•Capitation-based contracts •Single pooled budgets	•Shared records •Primary care hubs •Co-location •7-day Primary Care •Rationalised hospital sector	•MDT case management •Intermediate care teams •Crisis response teams •Specialty triage •Health coaching •Telehealth •Discharge planning •Self-care support	•Co-location/MDT case management	"Integrated primary and acute care systems – joining up GP, hospital, community and mental health services" ('Integrated primary and acute care systems vanguard sites' - https://www.england.nhs.uk/ourwork/futurenhs/new-care-models/primary-acute-sites/)	9 sites. Wirral Partners: MDT case management focused on elderly at risk of falls - keeping them at home. Mid Nottinghamshire: Move to capitated contract for all providers, integrated ED, locality based MDT case management, intermediate care teams, crisis response teams, specialty triage for referral, shared IT. South Somerset: single budget, three hubs with co-location of primary and secondary care (for MDT case management of MM patients), health coaching for less complex patients, remote monitoring, shared access to records online. Nothumberland: Primary care hubs, open 7 days, co-ordinated discharge, shared IT. Salford: Formation of single Integrated Care Organisation, MDT case management, working with voluntary sector (socially isolated people as target), single point of contact e.g. health coaching, information and support. Morecambe Bay: Single budget for health and care, smaller hospital service. North East Hampshire and Farnham: Shared resources and skills across health and social care, MDT case management (respiratory and cardiac focus). Harrogate: community hubs, MDT case management, 24/7 access to support. Isle of Wight: Telehealth, single point of access, self-management and monitoring
Multispecialty community provider (14 sites)	•Primary care partnership organisations	•Capitated outcomes-based contracting	•Shared records •24-hour Urgent care centre	•MDT case management •MDT prevention teams •Extended access to GP (7-days) •Self-care support •Discharge planning •Regular check-up appointments •Longer consultations for most complex •Home care	•Extended hours GP/MDT case management	"Multispecialty community providers – moving specialist care out of hospitals into the community" ('Multispecialty community providers vanguard sites' - https://www.england.nhs.uk/ourwork/futurenhs/new-care-models/community-sites/)	14 sites. Calderdale: single point of access, MDT case management. Erewash: MDT prevention team (focus on those who do not require hospital services), extending access to GP services, shared records for most vulnerable patients. Fylde Coast: MDT case management (frail and elderly), shared IT record. Birmingham & Sandwell: Care co-ordinator (case manager), extended primary care services including outpatient and diagnostic services. West Wakefield: MDT case management (team of care navigators), shared IT records, 7 day service. Sunderland: MDT care and prevention team, self-care. Dudley: MDT case management (frail/LTCs), 24-hour urgent care centre (single point of access). Whitstable: Trained care workers (working with MDT) case management, technology enabling. Stockport: Single point of access for urgent care (integrated with community teams), MDT case management (complex/end of life). Tower Hamlets: MDT case management. Southern Hampshire: self-management encouragement, discharge planning, regular check-up appointments, shared IT record. West Cheshire: 3 programmes focusing on different groups, children, MDT case management (LTCs, frail/complex). Northamptonshire: urgent care model, ambulatory care service, LTC management service, GP-led MDT case management (complex care), outpatient and planned care services, MDT primary care, most vulnerable 7% - longer in-depth consultations with enhanced continuity of care-, home care. Southern Nottinghamshire: primary care partnership organisation taking contractual responsibility for health, quality and costs of care with capitated outcomes-based resource allocation contract, MDT case management.
Devolution:							
DevoManc	•Single shared services creation •Joint commissioning	•Pooled health & social care budget (£6 billion estimated) •Contracting and payment design	•Shared records •Rationalised hospital services •Navigation hubs	•MDT case management •Population health prevention •Self-care support •Mental health service integration	•Pooled budget/ MDT case management	"Key features will be targeted case management of the population most in need delivered by upskilled multi-disciplinary teams, together with streamlined discharge planning in order to reduce the demand placed on acute hospitals." ('Taking charge of our health and social care in reater Manchester. The Plan FINAL DRAFT, version 11.3' - http://www.gmhealthandsocialcaredevo.org.uk/assets/GM-Strategic-Plan-Final.pdf)	Main focus is more macro/meso level changes. 1. Pooled budget, but also devolving power with these finances. Attempt to make it easier to overcome these national barriers to integration at scale? 5 areas of focus identified in strategic plan: '1 Radical upgrade in population health prevention (self-management, pharmacy and prescribing, plus tackle burden of CVD and diabetes). 2 Transforming community based care & support (care closer to home, integrated primary, acute, community, mental health and social care. "Key features will be targeted case management of the population most in need delivered by upskilled multi-disciplinary teams, together with streamlined discharge planning in order to reduce the demand placed on acute hospitals."). 3 Standardising acute & specialist care (single shared services creation, IT standardised). 4 Standardising clinical support and back office services (coordination centres to help patients navigate services). 5 Enabling better care ("creation of innovative organisational forms, new ways of commissioning, contracting and payment design and standardised information management and technology to incentivise ways of working across GM, so that our ambitious aims can be realised").'

Appendix 2: Chapter 5 Supplementary files

S1 APPENDIX: FULL MEDLINE SEARCH STRATEGY

	Keywords
1	Family Practice/
2	Primary Health Care/
3	Physicians, Family/
4	Community Health Services/
5	Community Dentistry/
6	Community Health Nursing/
7	Community Mental Health Services/
8	Community Pharmacy Services/
9	Home Care Services/
10	Community Mental Health Centers/
11	family pract$.tw.
12	general practice$.tw.
13	community based.tw.
14	community care.tw.
15	family medicine.tw.
16	family physician$.tw.
17	primary care.tw.
18	(primary health care or primary healthcare).tw.
19	family doctor$.tw.
20	primary medical care.tw.
21	general physician$.tw.
22	general practitioner$.tw.
23	primary care practitioner$.tw.
24	(community adj (health or healthcare or health care)).tw.
25	primary healthcare team$.tw.

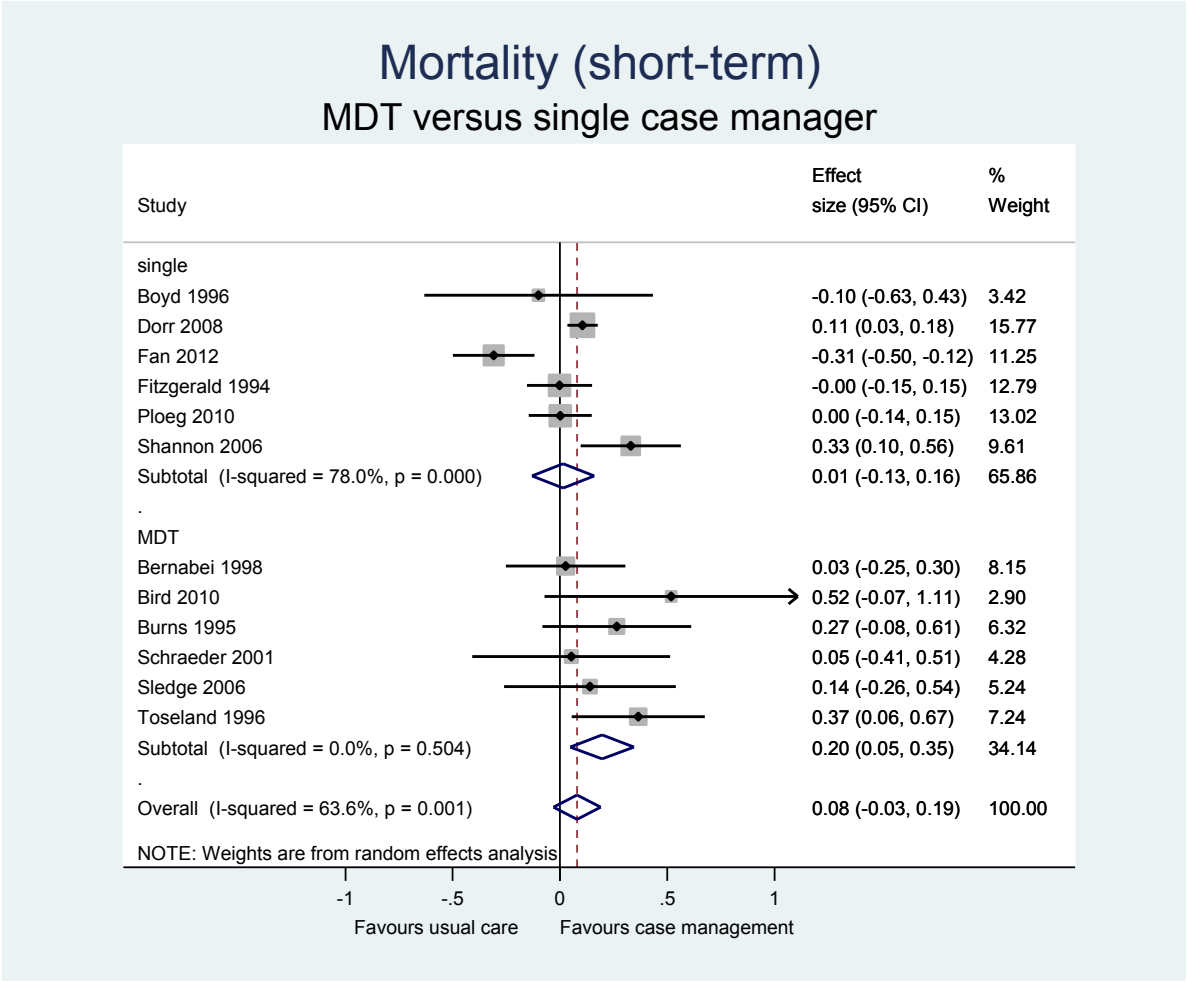
26	primary health care team$.tw.
27	primary medical care team$.tw.
28	practice nurse$.tw.
29	practice manager$.tw.
30	(gpsi or gpwsi).tw.
31	(practitioner$ adj3 special interest$).tw.
32	(primary care or primary health care or general practice or family practice or family medicine).nw.
33	or/1-32
34	case management.tw.
35	care management.tw.
36	care co?ordination.tw.
37	collaborat* care.tw.
38	collaborat* practice.tw.
39	virtual ward*.tw.
40	care integrat*.tw.
41	care team?.tw.
42	co?ordinated care.tw.
43	multidisciplin* care.tw.
44	interdisciplin* care.tw.
45	multidisciplin* team?.tw.
46	interdisciplin* team?.tw.
47	multidisciplin* management.tw.
48	interdisciplin* management.tw.
49	(integrated adj4 care).tw.
50	exp Patient Care Team/
51	care plan*.tw.
52	practice team?.tw.

53	exp Patient Care Planning/
54	exp "Delivery of Health Care, Integrated"/
55	or/34-54
56	randomized controlled trial.pt.
57	controlled clinical trial.pt.
58	multicenter study.pt.
59	(randomis* or randomiz* or randomly allocat* or random allocat*).ti,ab.
60	groups.ab.
61	(trial or multicenter or multi center or multicentre or multi centre).ti.
62	(intervention* or controlled or control group or compare or compared or (before adj5 after) or (pre adj5 post) or pretest or pre test or posttest or post test or quasiexperiment* or quasi experiment* or evaluat* or effect or impact or time series or time point? or repeated measur*).ti,ab.
63	or/56-62
64	exp Animals/
65	Humans/
66	64 not (64 and 65)
67	review.pt.
68	meta analysis.pt.
69	news.pt.
70	comment.pt.
71	editorial.pt.
72	cochrane database of systematic reviews.jn.
73	comment on.cm.
74	(systematic review or literature review).ti.
75	or/66-74
76	63 not 75
77	33 and 55 and 76

S2 APPENDIX: FOREST PLOTS FOR SUBGROUP ANALYSES*

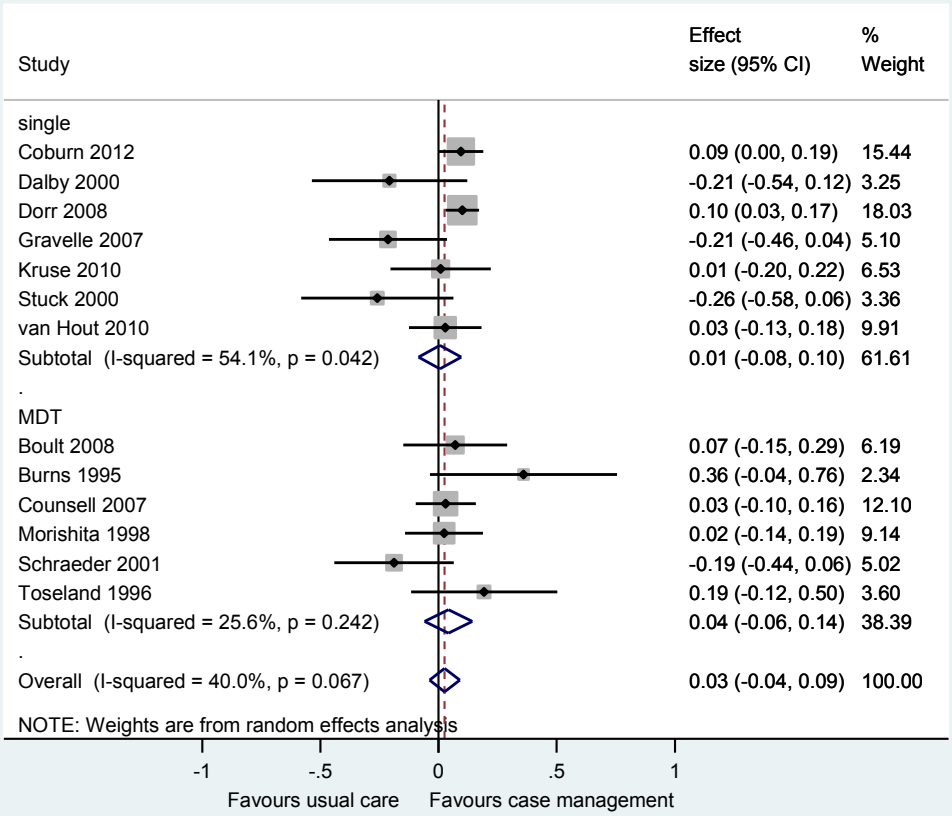
*** No significant results remain following Holm-Bonferroni correction for multiple comparisons**

MULTIDISCIPLINARY TEAM VERSUS SINGLE CASE MANAGER



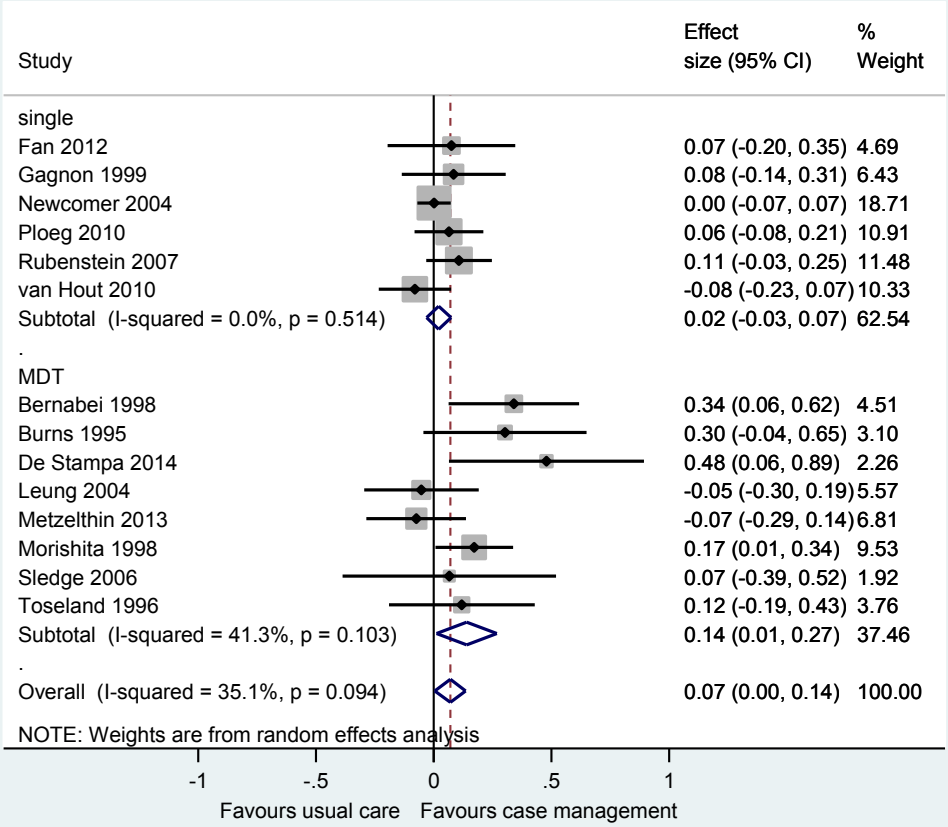
Mortality (long-term)

MDT versus single case manager



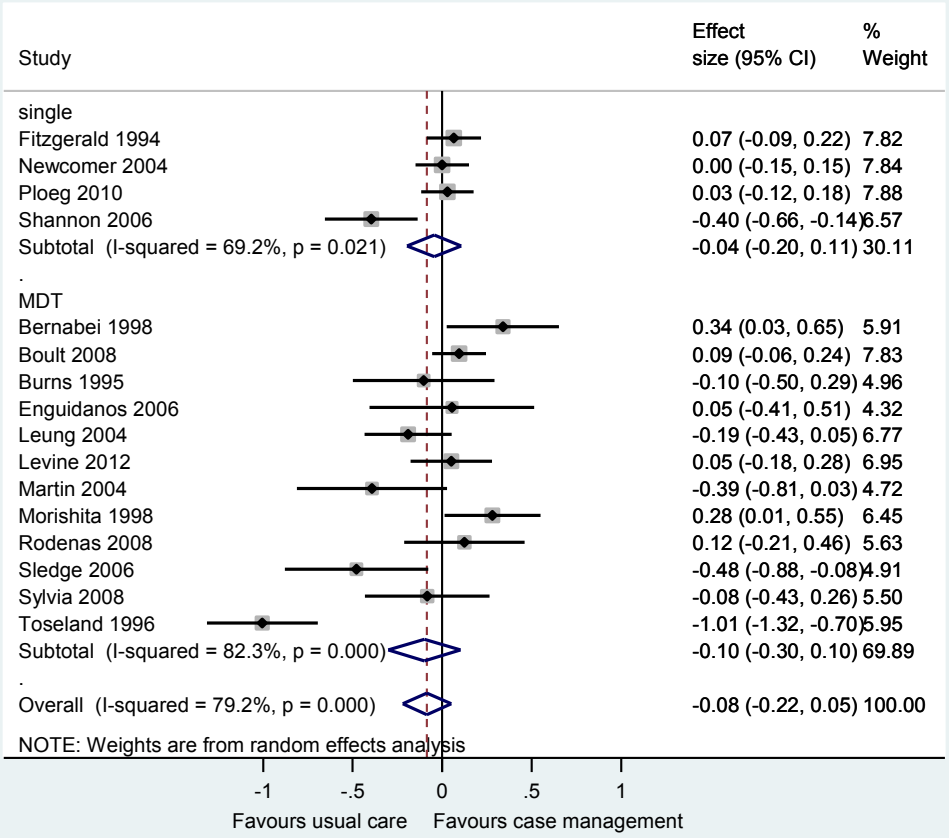
Self-rated health (short-term)

MDT versus single case manager



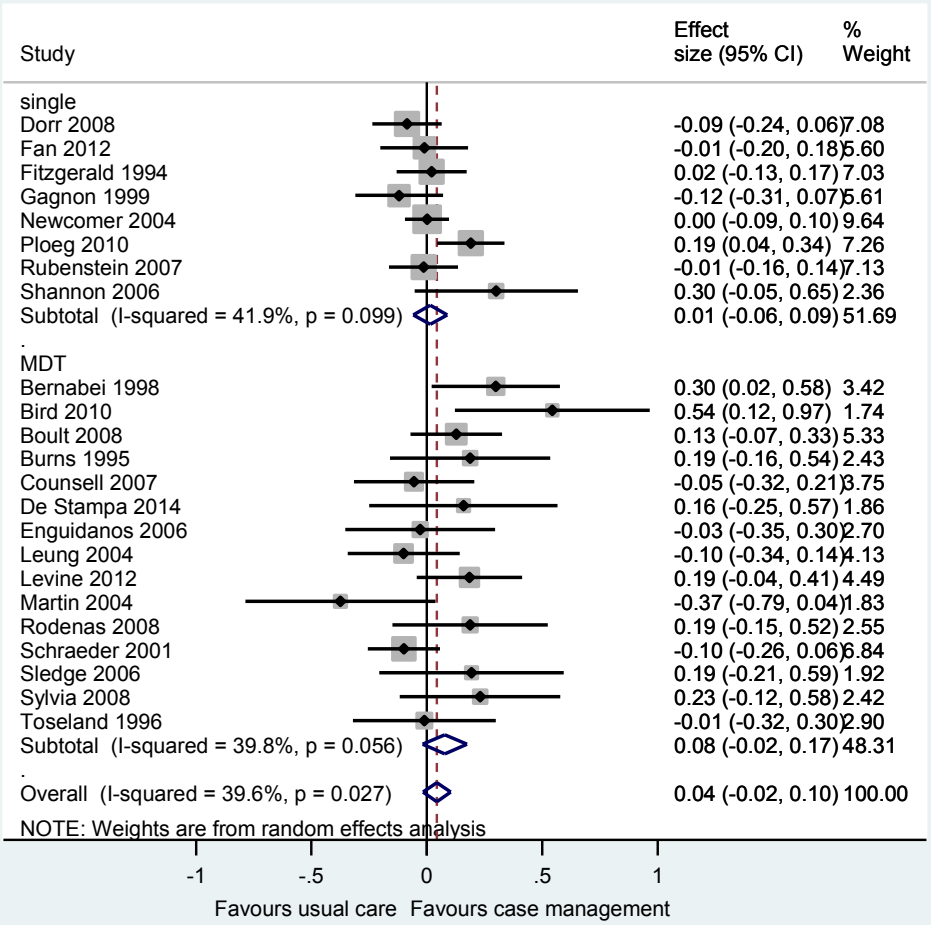
Utilisation of primary care (short-term)

MDT versus single case manager



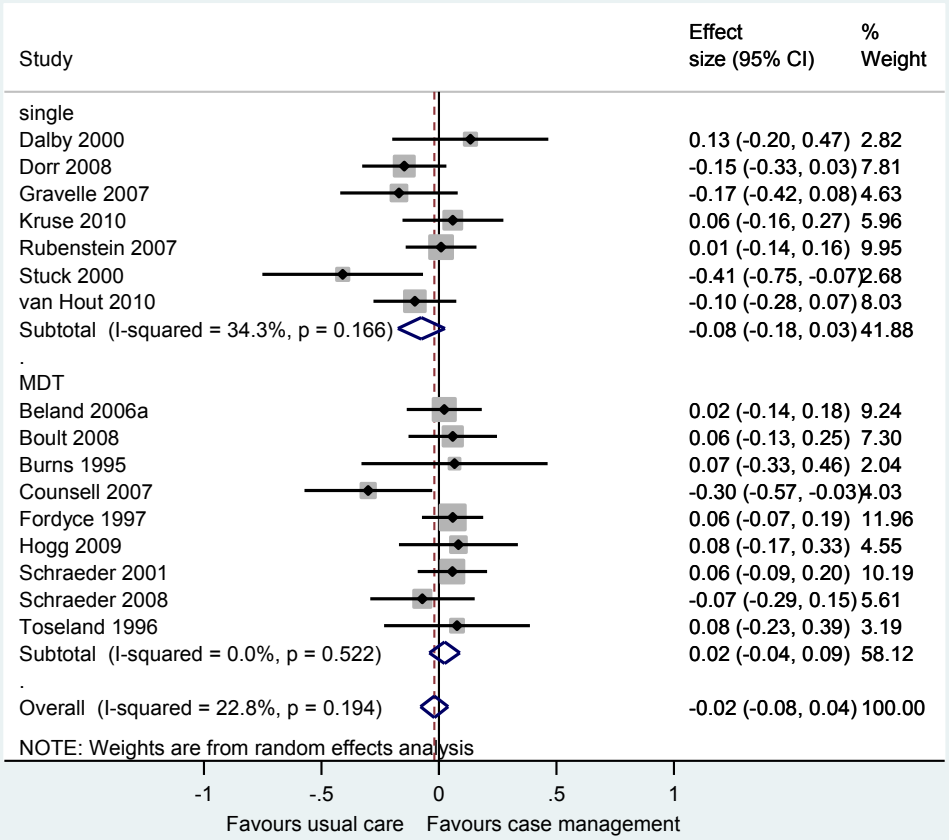
Utilisation of secondary care (short-term)

MDT versus single case manager

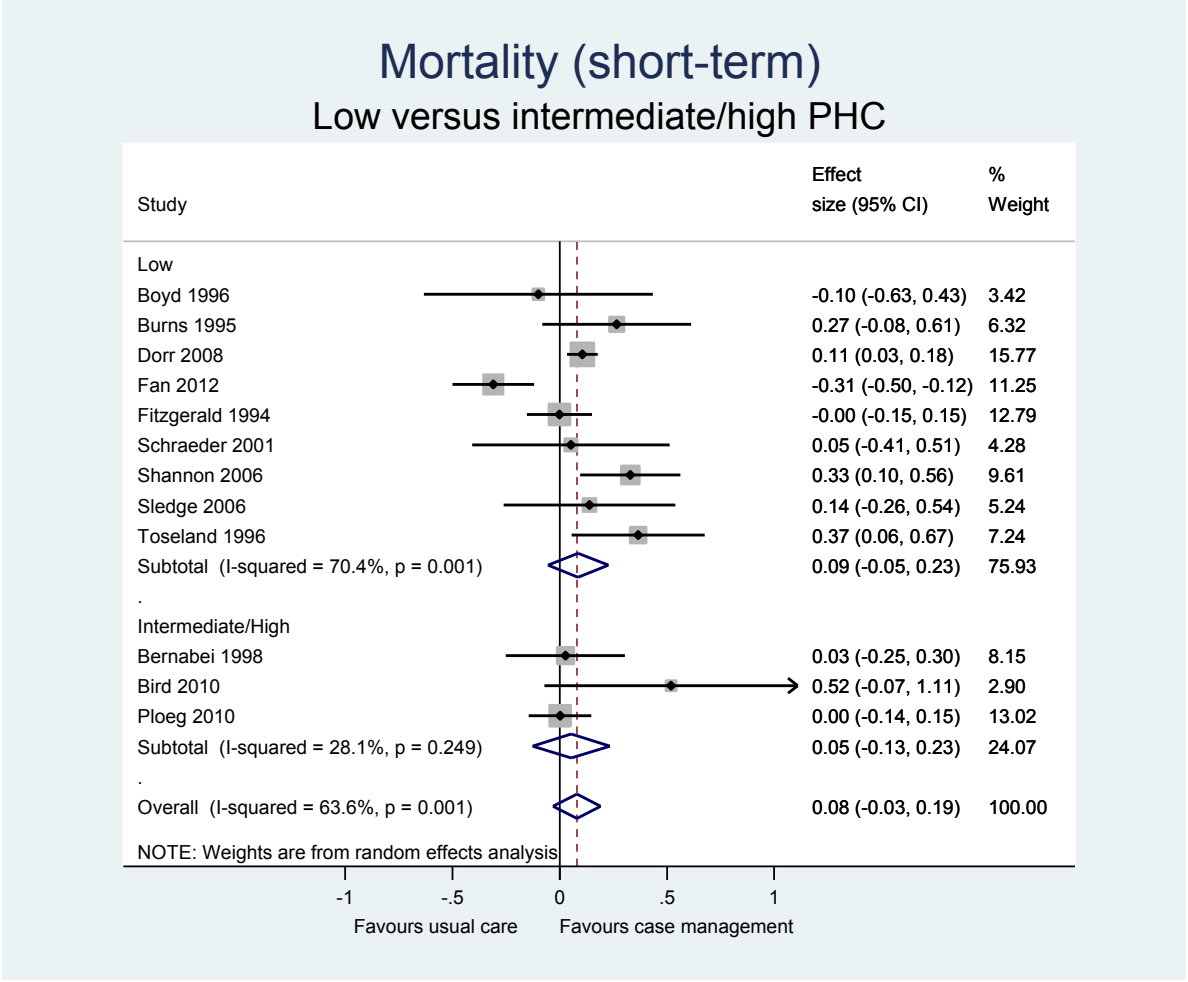


Utilisation of secondary care (long-term)

MDT versus single case manager

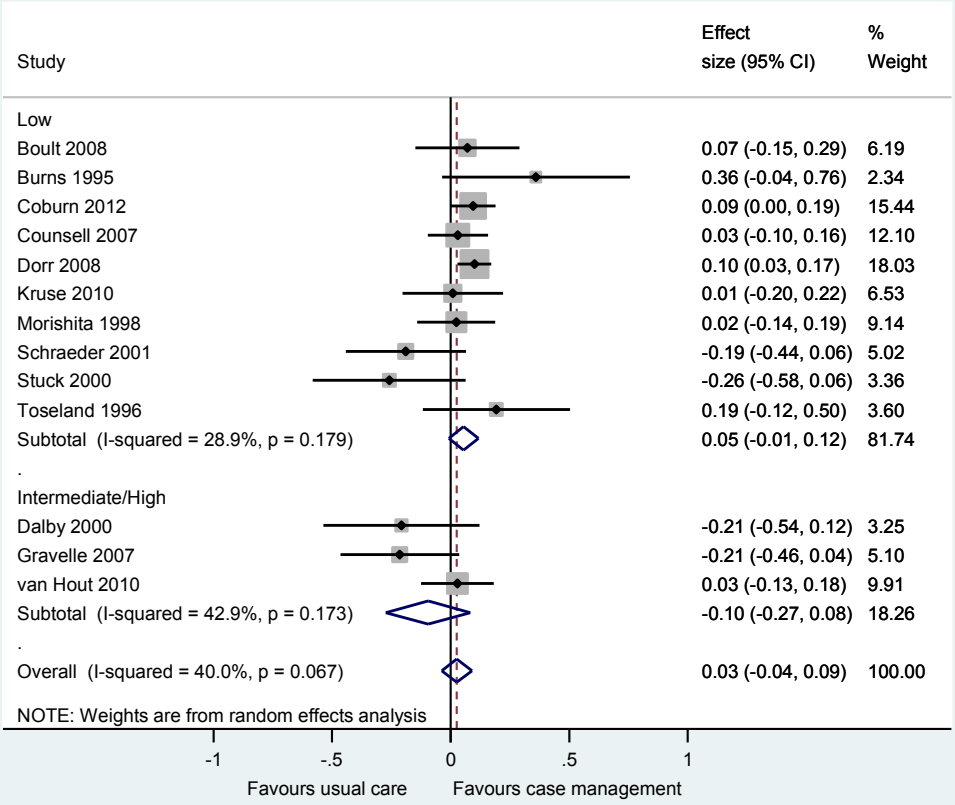


STRENGTH OF PRIMARY CARE ORIENTATION



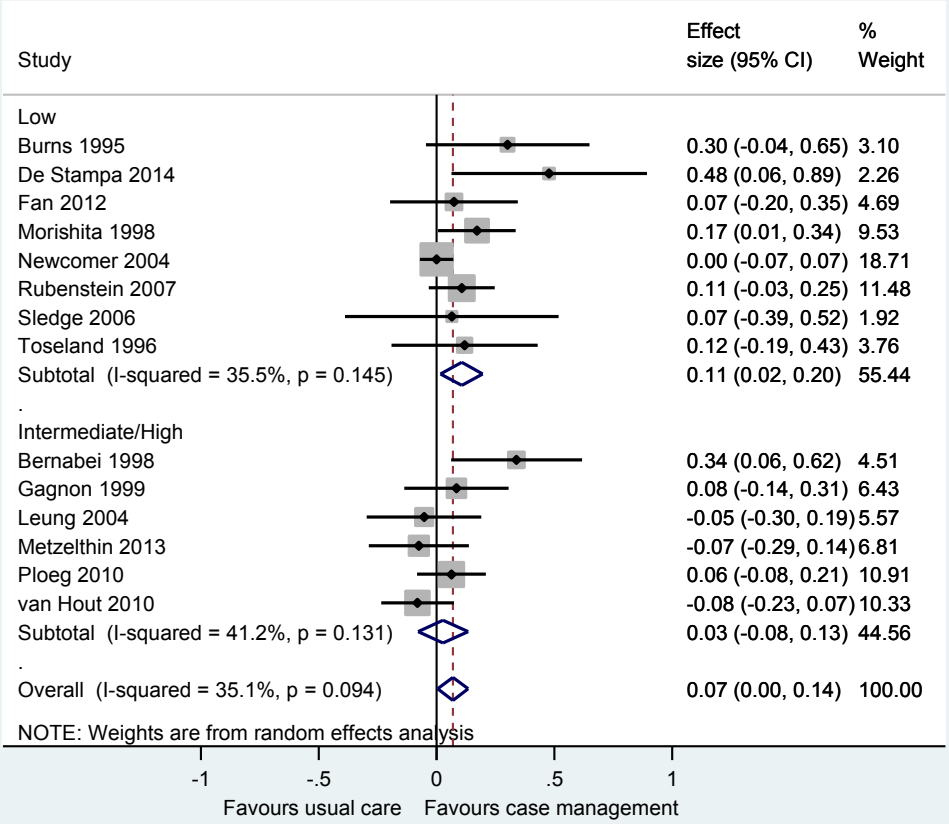
Mortality (long-term)

Low versus intermediate/high PHC



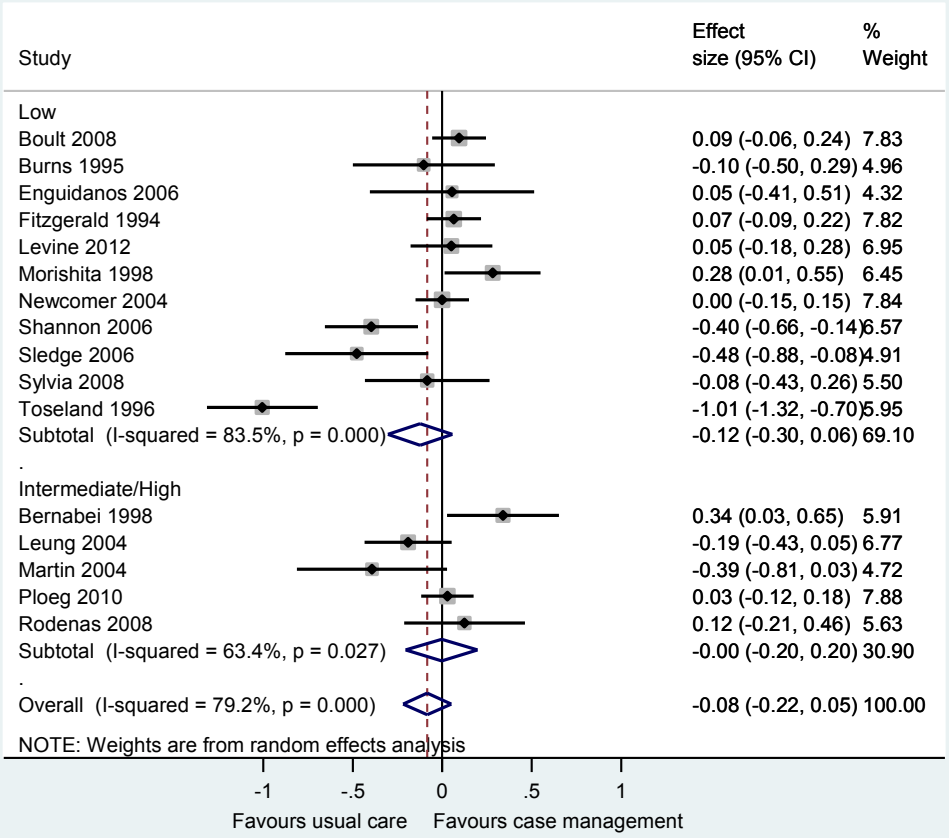
Self-rated health (short-term)

Low versus intermediate/high PHC



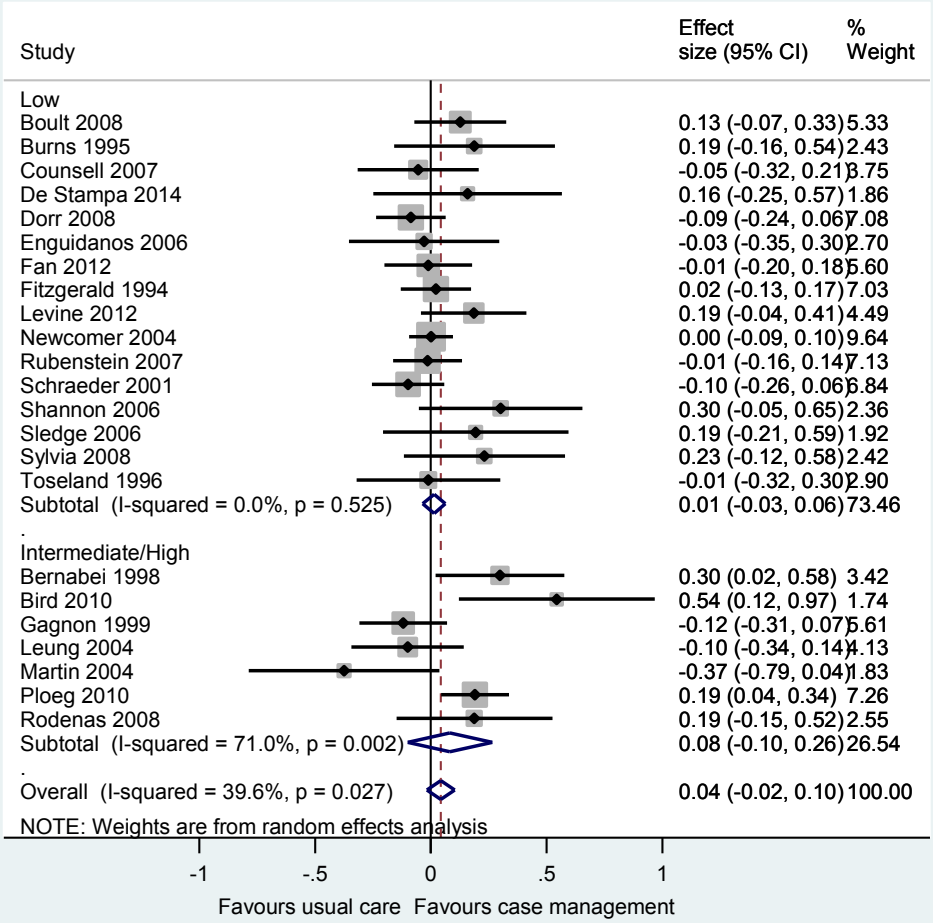
Utilisation of primary care (short-term)

Low versus intermediate/high PHC



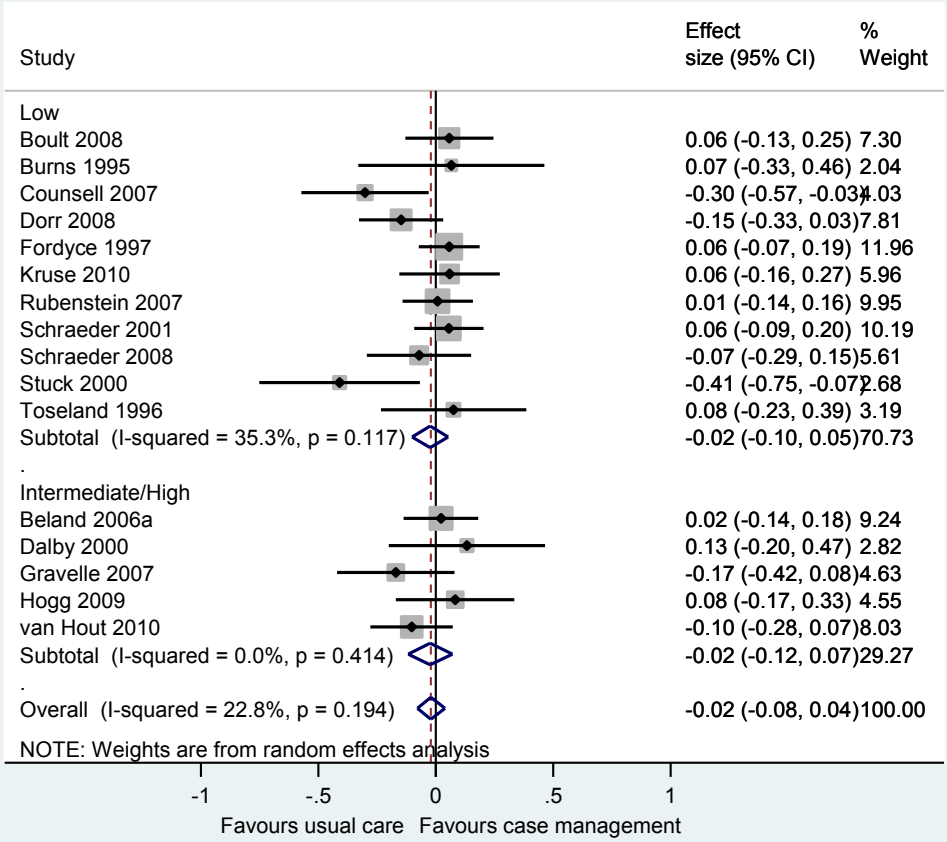
Utilisation of secondary care (short-term)

Low versus intermediate/high PHC

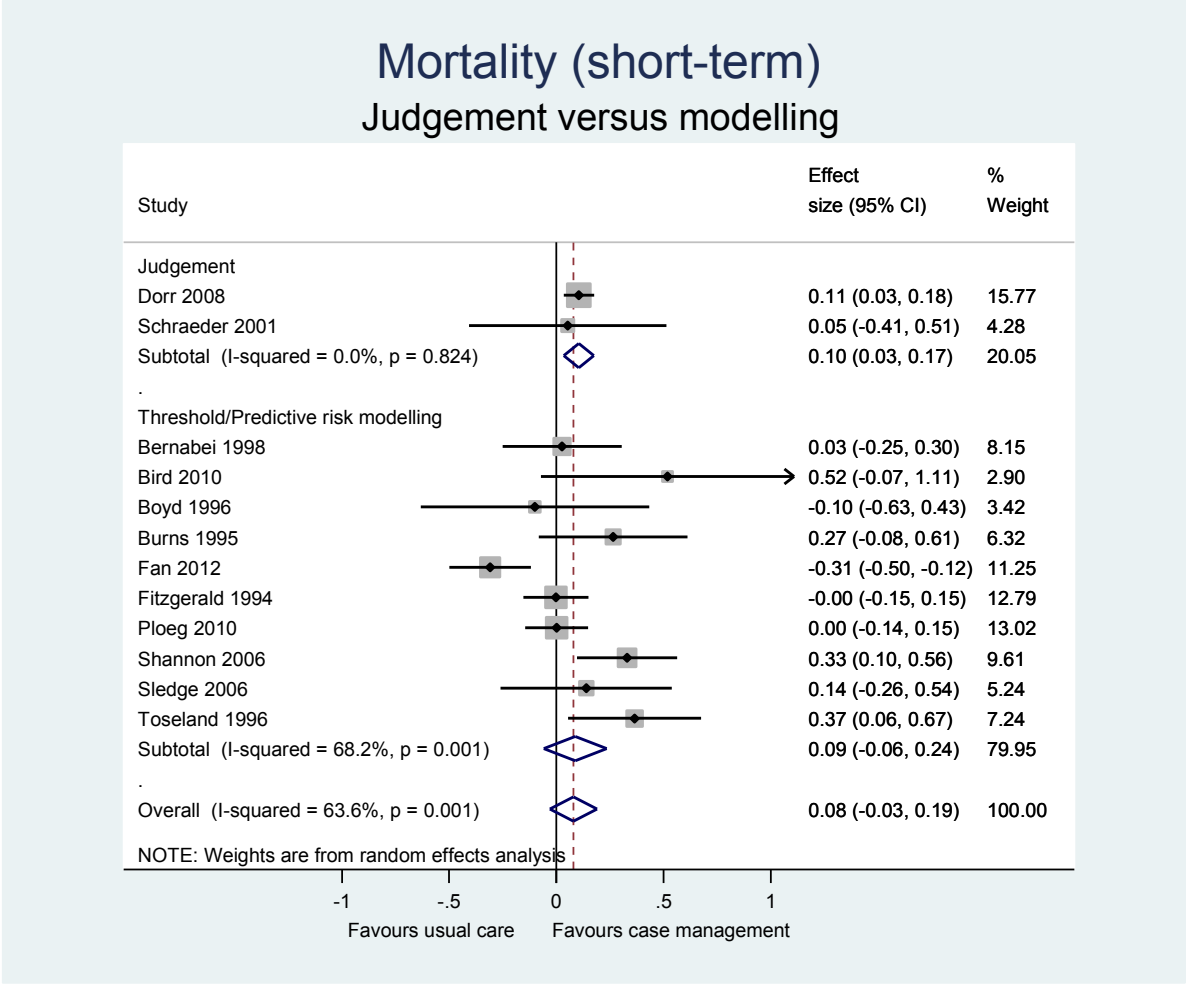


Utilisation of secondary care (long-term)

Low versus intermediate/high PHC

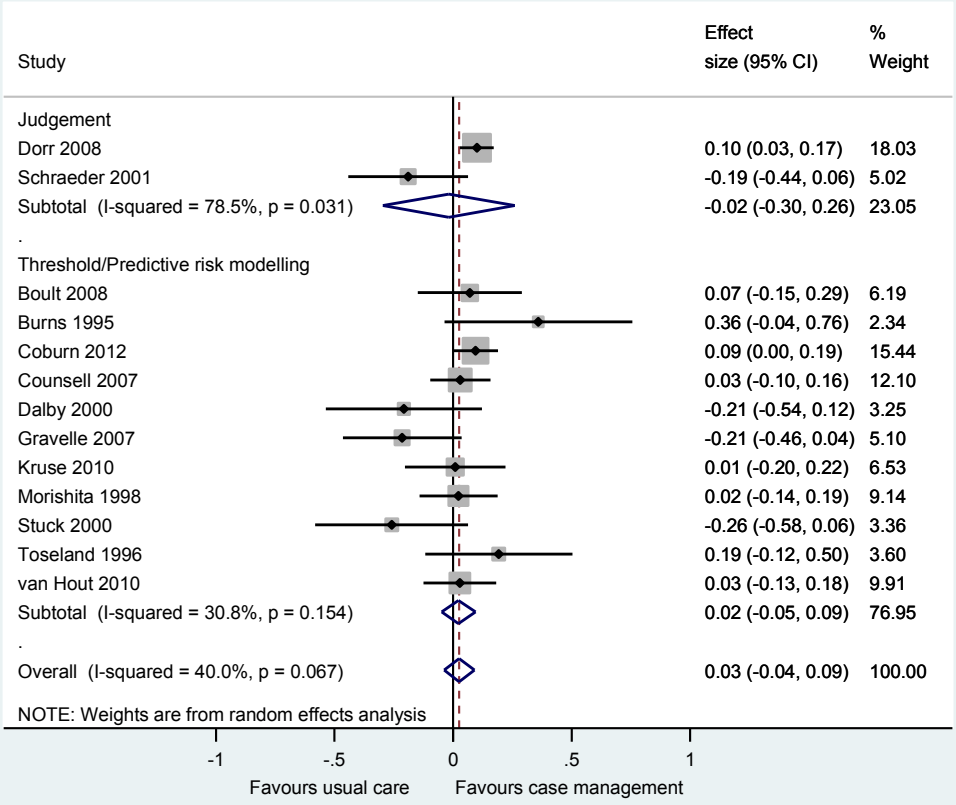


TYPE OF RISK TOOL USED



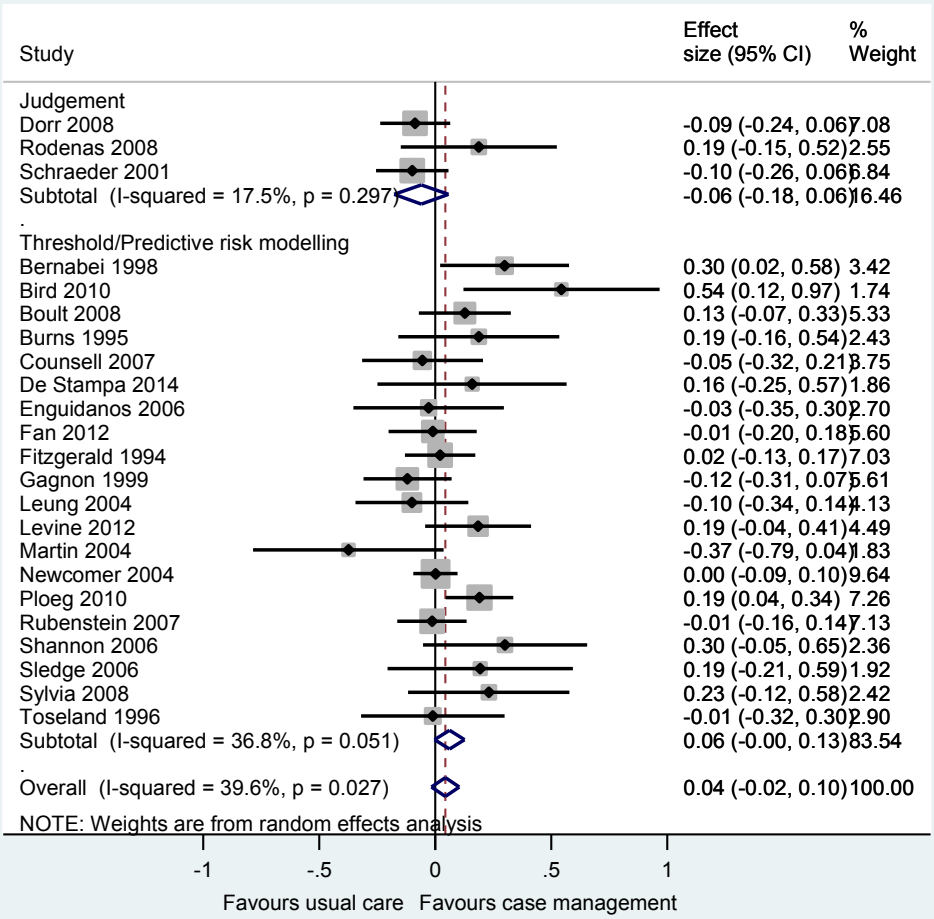
Mortality (long-term)

Judgement versus modelling



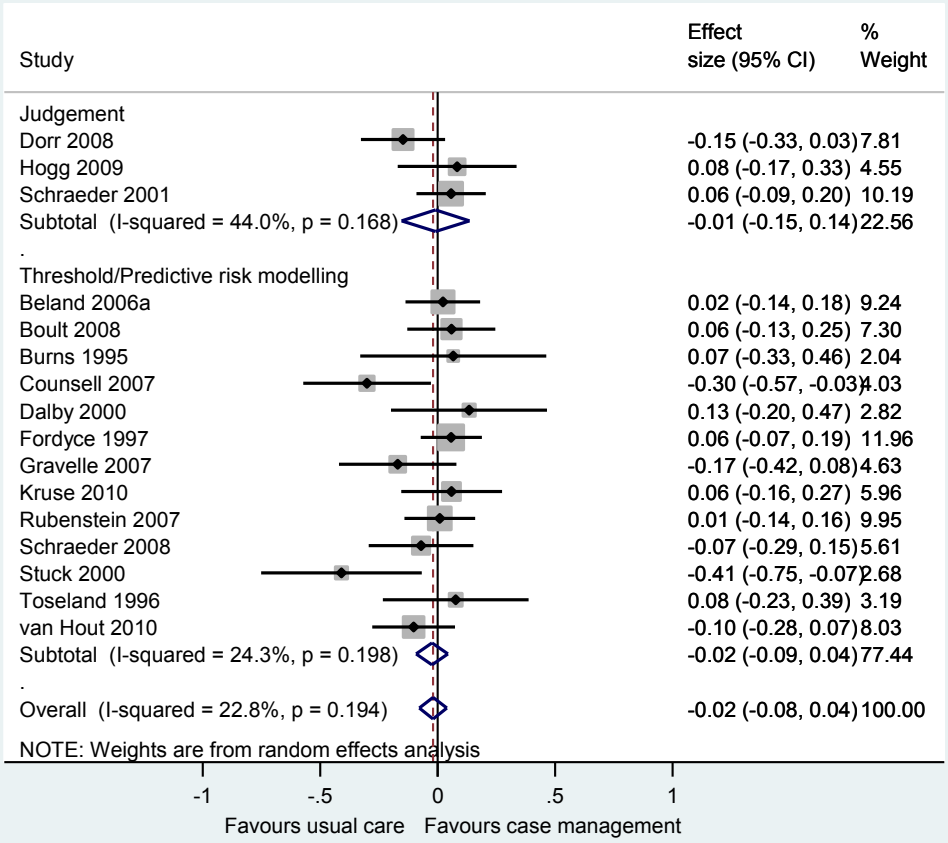
Utilisation of secondary care (short-term)

Judgement versus modelling

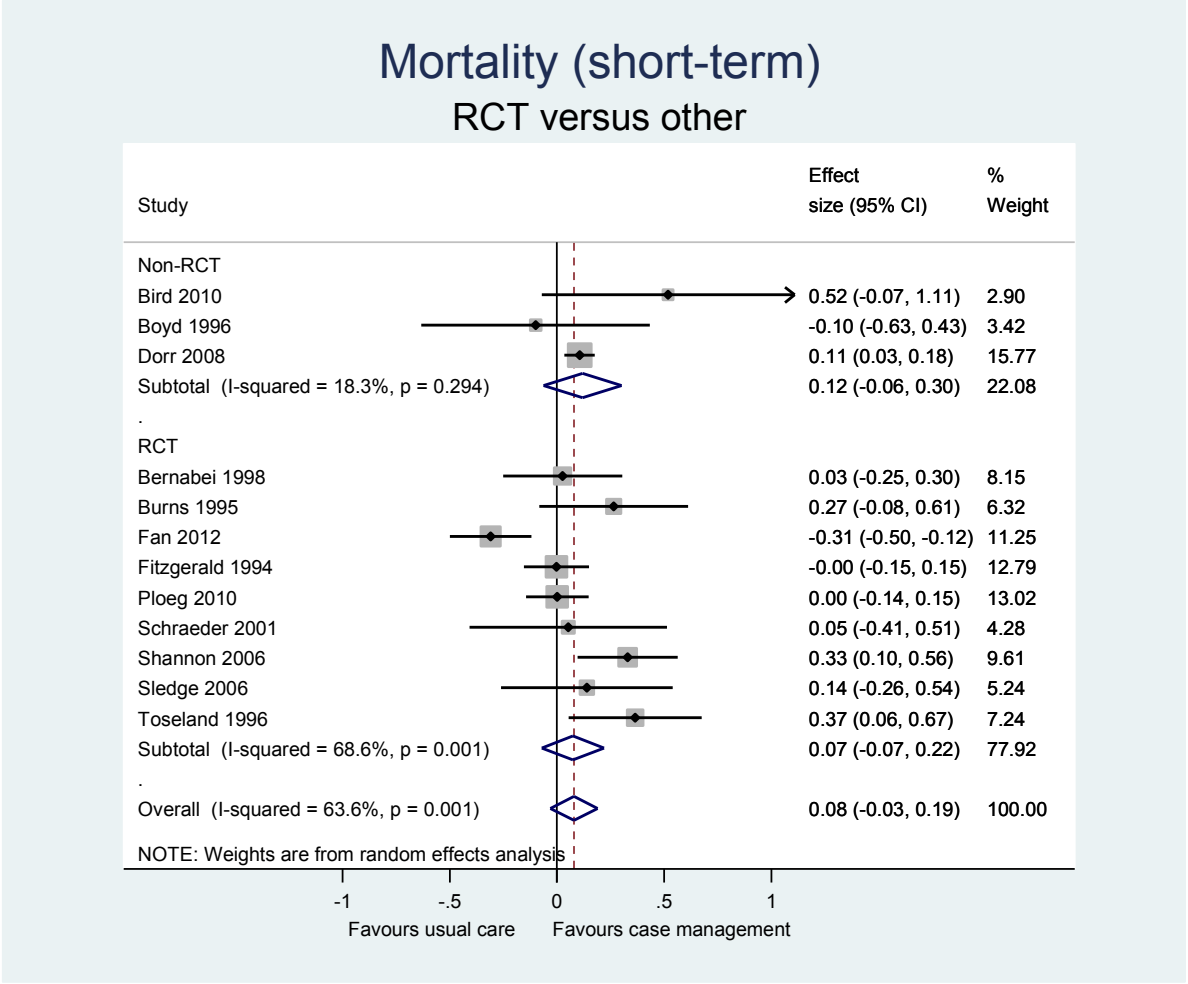


Utilisation of secondary care (long-term)

Judgement versus modelling

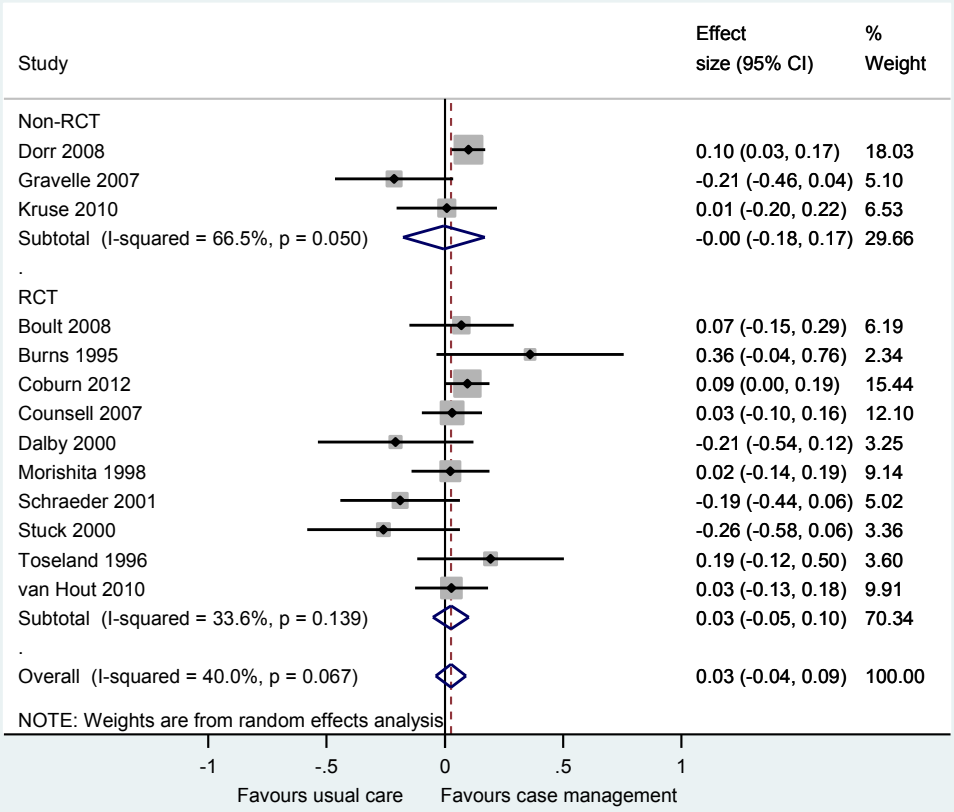


RCT VERSUS NON-RCT



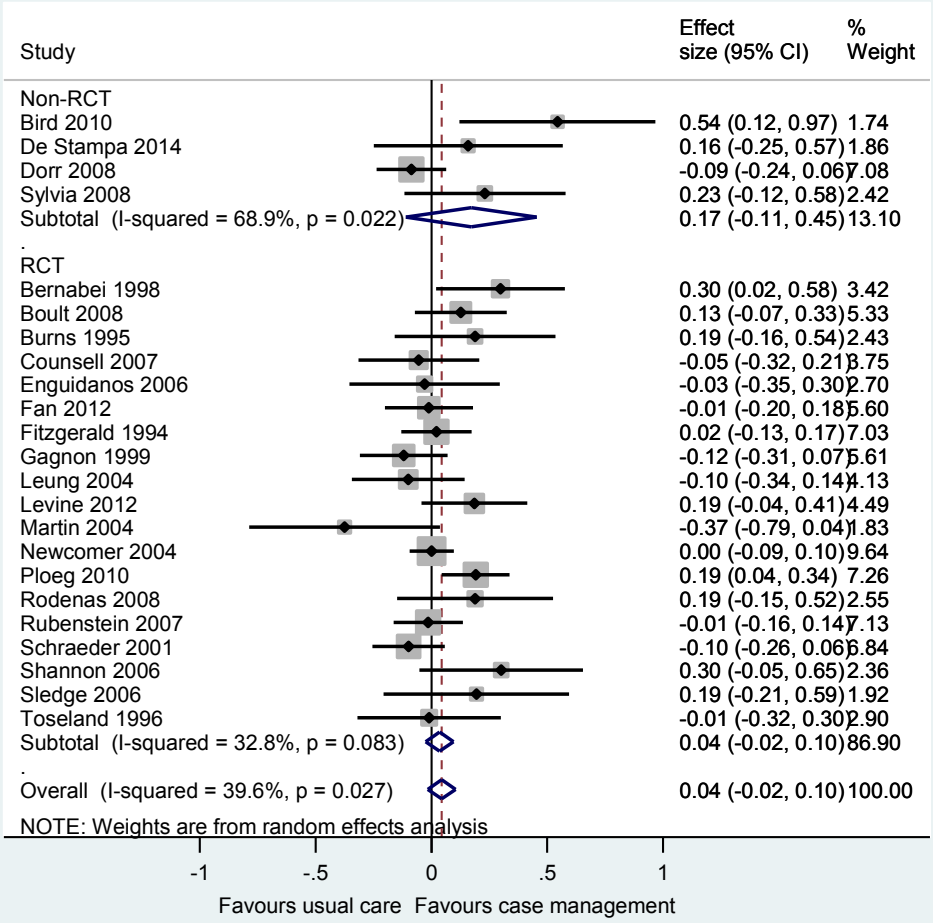
Mortality (long-term)

RCT versus other



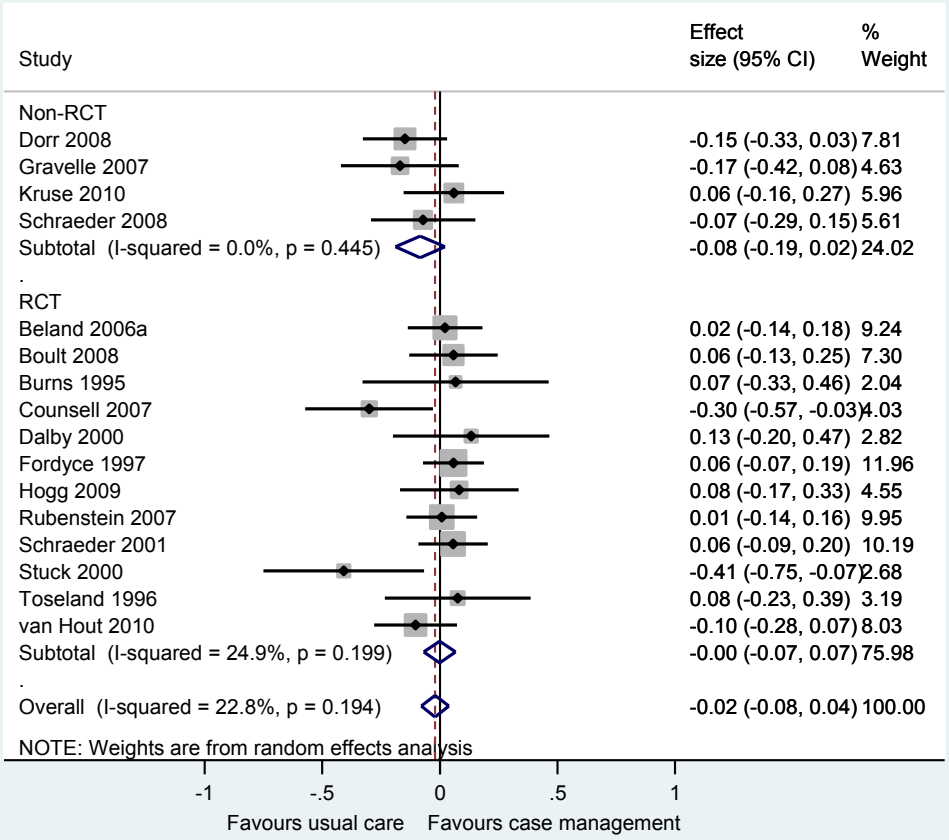
Utilisation of secondary care (short-term)

RCT versus other

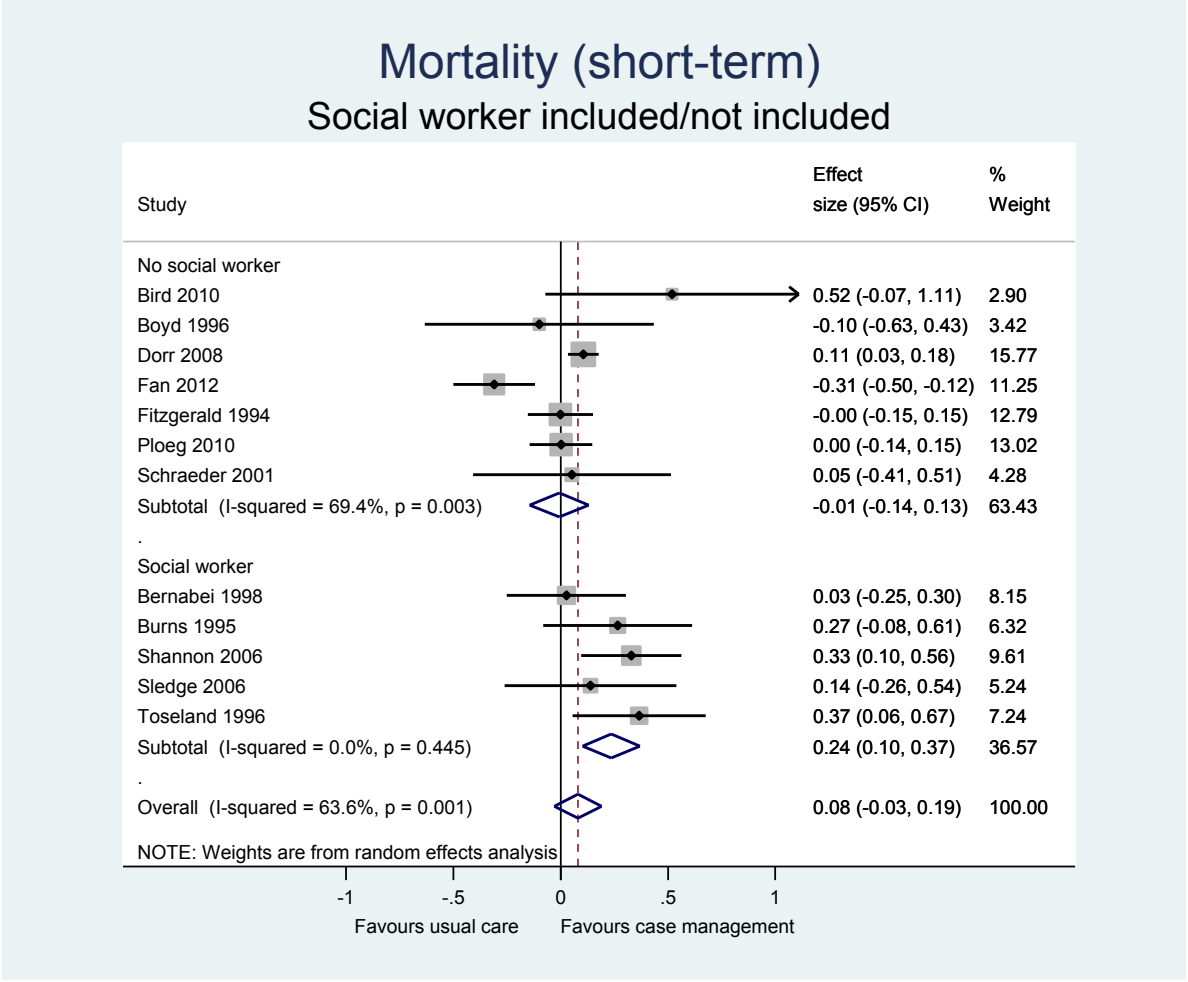


Utilisation of secondary care (long-term)

RCT versus other

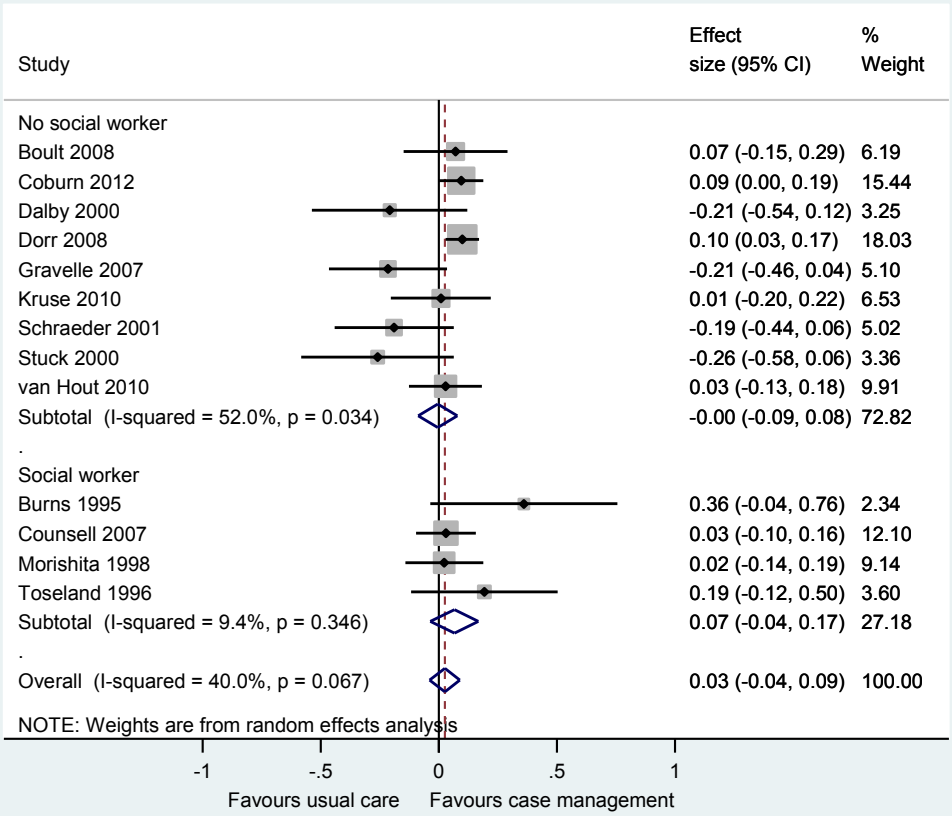


INCLUSION OF A SOCIAL WORKER IN CASE MANAGEMENT



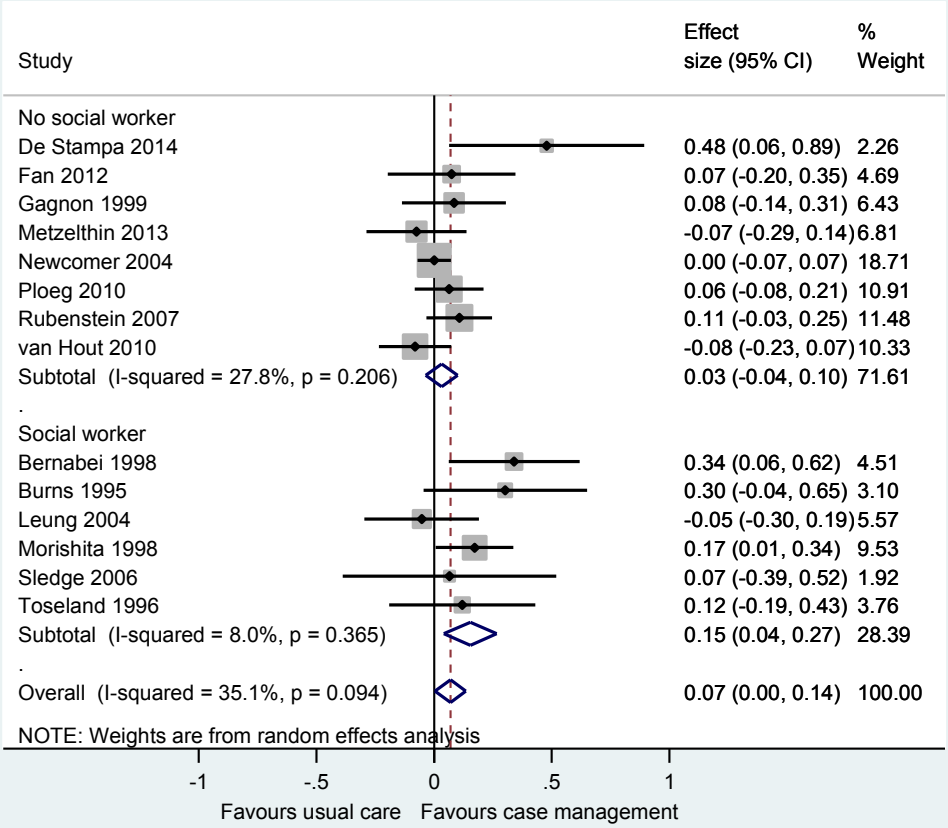
Mortality (long-term)

Social worker included/not included



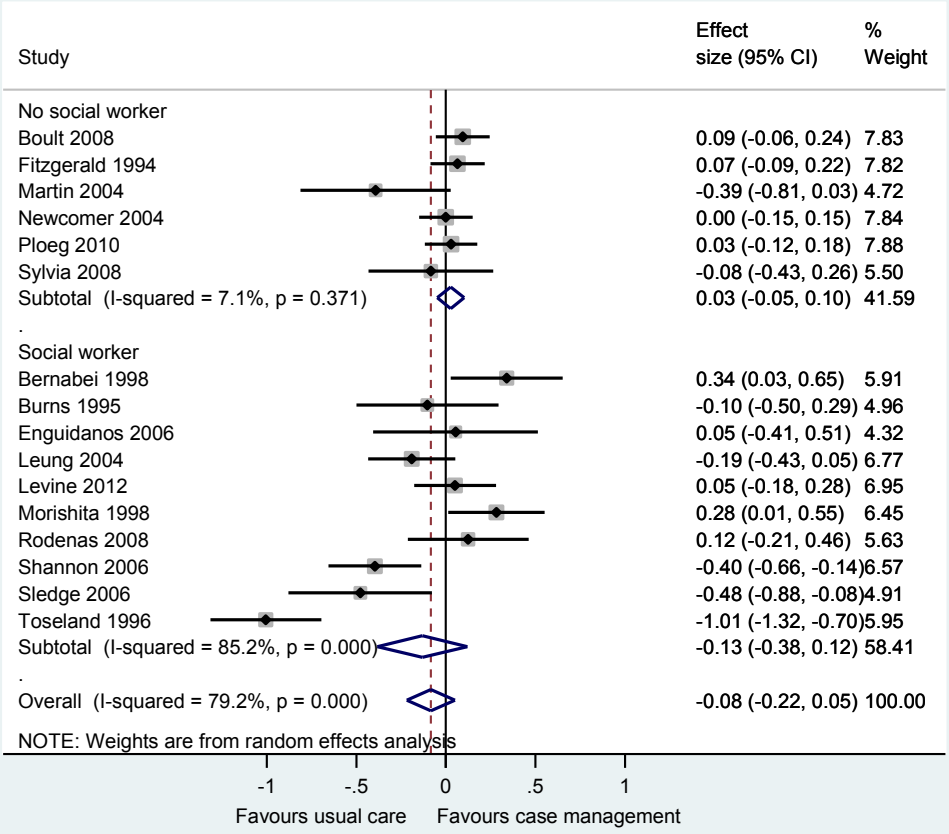
Self-rated health (short-term)

Social worker included/not included



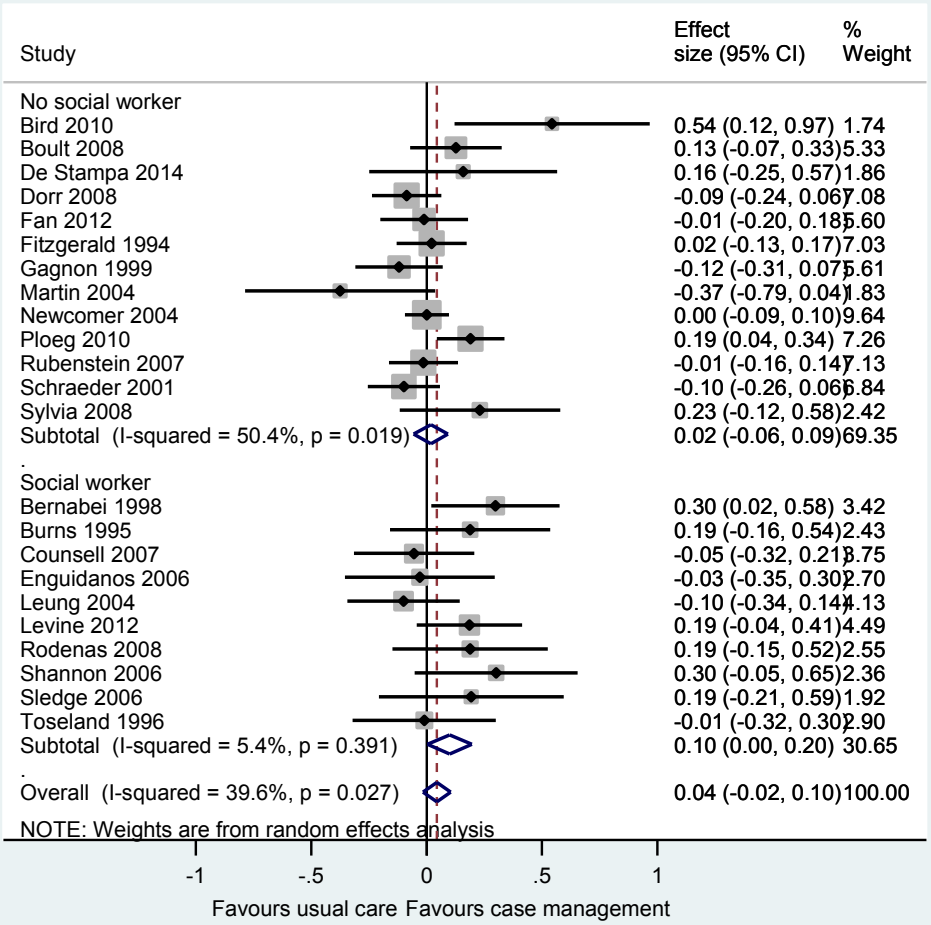
Utilisation of primary care (short-term)

Social worker included/not included



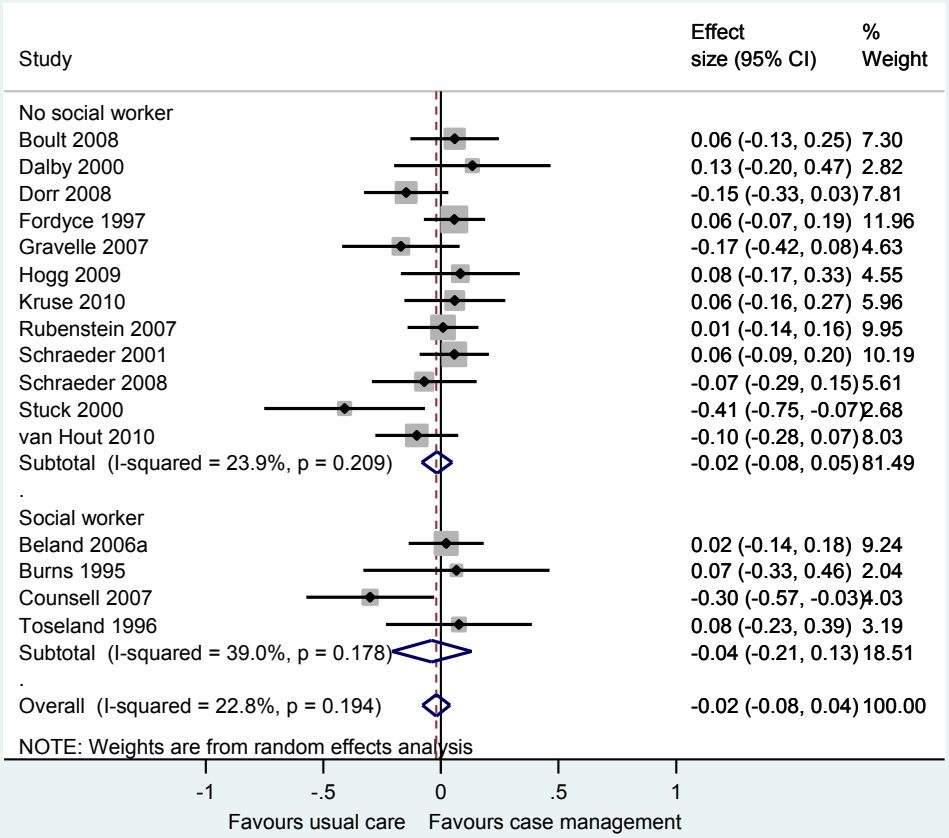
Utilisation of secondary care (short-term)

Social worker included/not included



Utilisation of secondary care (long-term)

Social worker included/not included

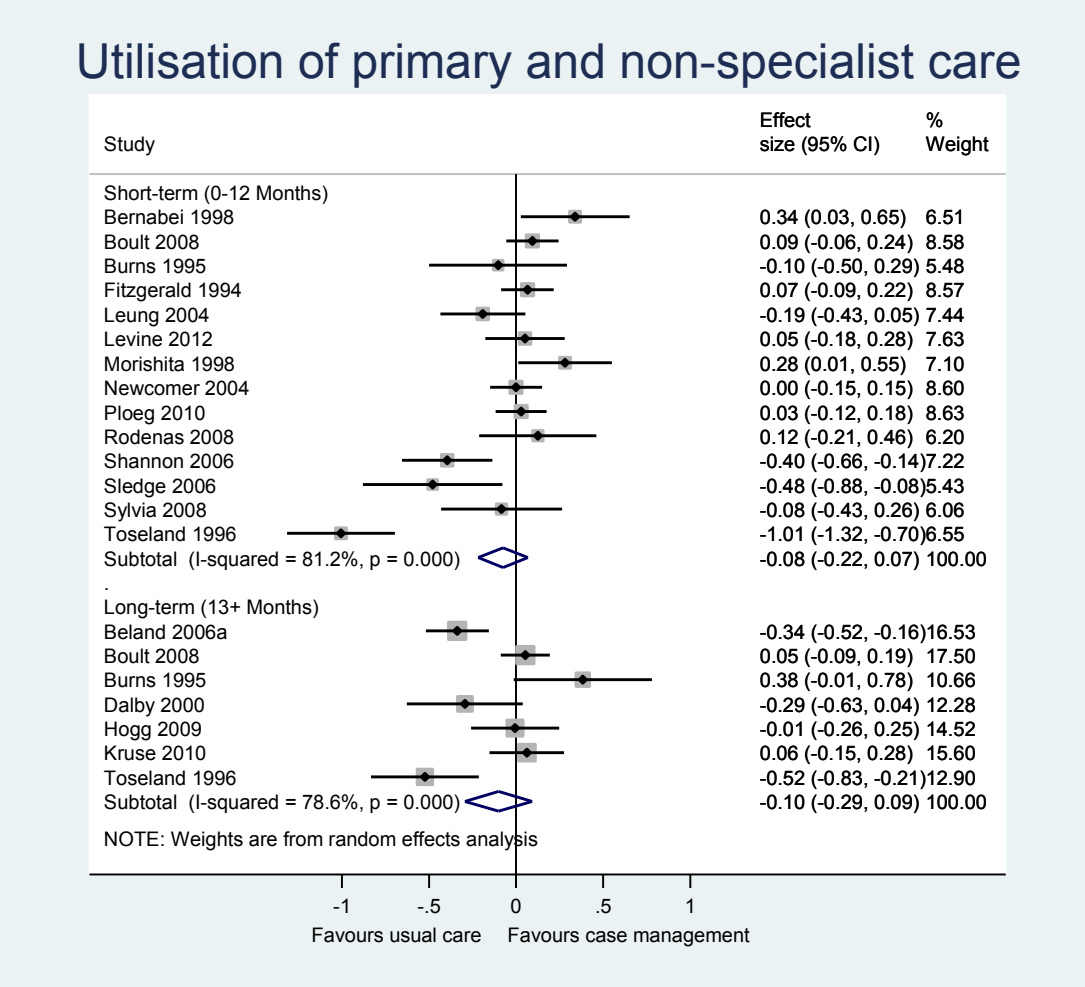


S3 Appendix:

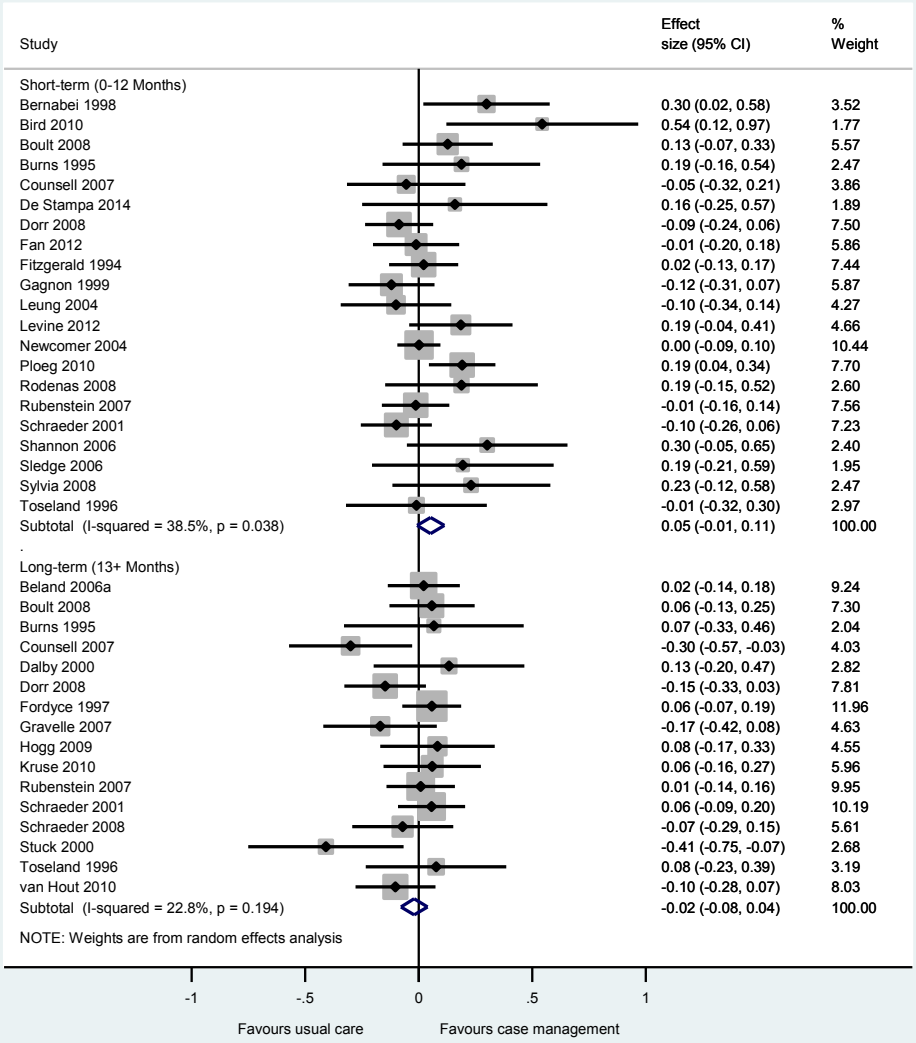
Results of the sensitivity analysis excluding studies at high risk of bias*

*** No significant results remain following Holm-Bonferroni correction for multiple comparisons**

Primary analysis – sensitivity analysis



Utilisation of secondary care



SECONDARY ANALYSIS – SENSITIVITY ANALYSIS

Outcome (time-period)	**Subgroup effect size (number of studies)**	
	MDT (19)	**Single (15)**
Utilisation primary care (short)	-0.09 (-0.31 to 0.14) (10)	-0.04 (-0.20 to 0.11) (4)
Utilisation secondary care (short)	0.10 (0.01 to 0.20)* (13)	0.01 (-0.06 to 0.09) (8)
	Low PHC score (22)	**Int/high PHC score (12)**
Utilisation primary care (short)	-0.14 (-0.33 to 0.05) (10)	0.05 (-0.14 to 0.25) (4)
Utilisation secondary care (short)	0.01 (-0.03 to 0.06) (15)	0.13 (-0.05 to 0.31) (6)
	Clinical Judgement (4)	**Risk modelling (30)**
Utilisation primary care (short)	n/a	n/a
Utilisation secondary care (short)	-0.06 (-0.18 to 0.06) (3)	0.07 (0.01 to 0.14)* (18)
	RCT (26)	**Non-RCT (8)**
Utilisation primary care (short)	n/a	n/a
Utilisation secondary care (short)	0.04 (-0.02 to 0.10) (17)	0.17 (-0.11 to 0.45) (4)
	Social worker (11)	**No social worker (23)**
Utilisation primary care (short)	-0.15 (-0.42 to 0.12) (9)	0.04 (-0.03 to 0.11) (5)
Utilisation secondary care (short)	0.11 (0.01 to 0.22)* (9)	0.03 (-0.04 to 0.10) (12)

No significant difference between subgroups ($p<0.05$)

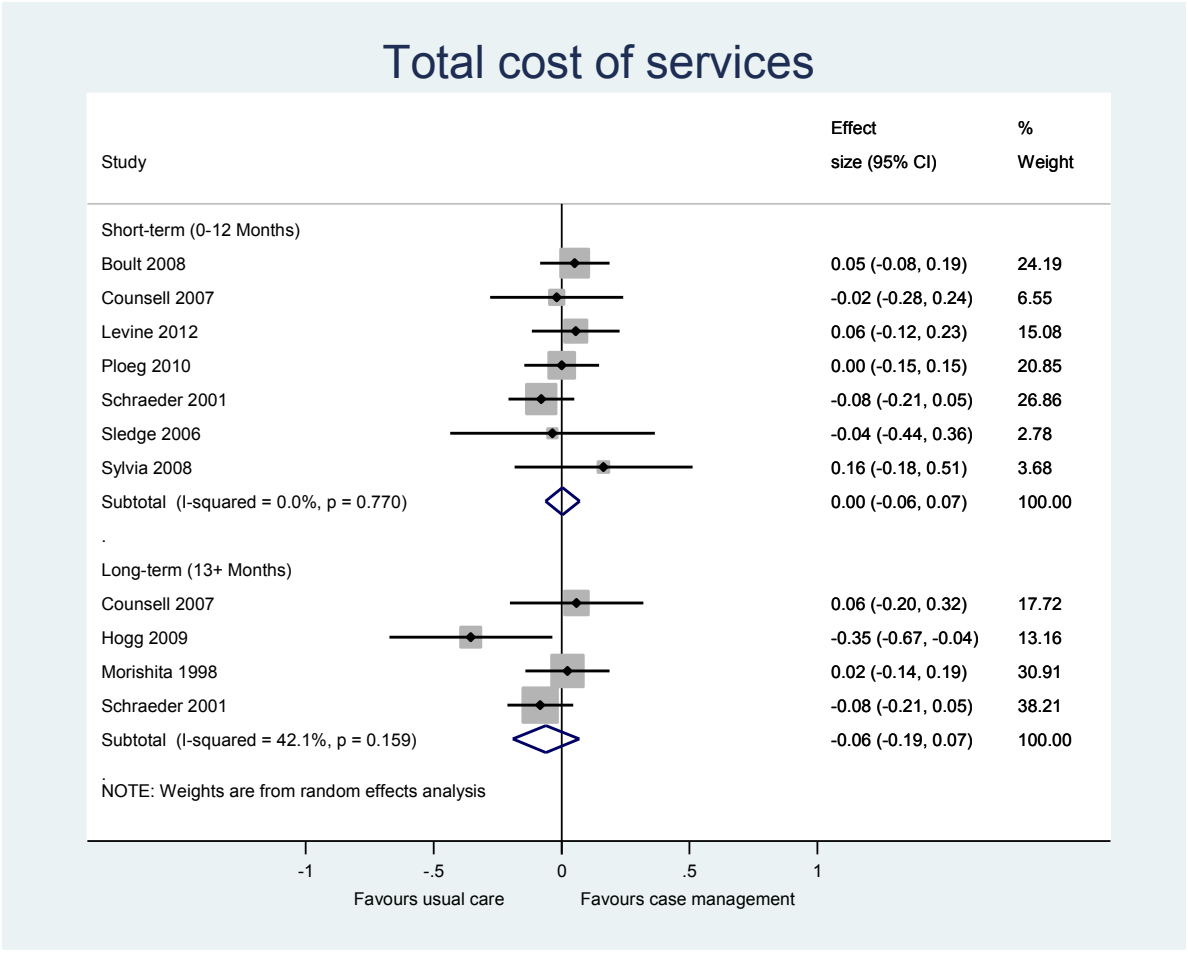
* = significant in-subgroup effect ($p<0.05$)

Note: Positive effect size favours case management for all measures

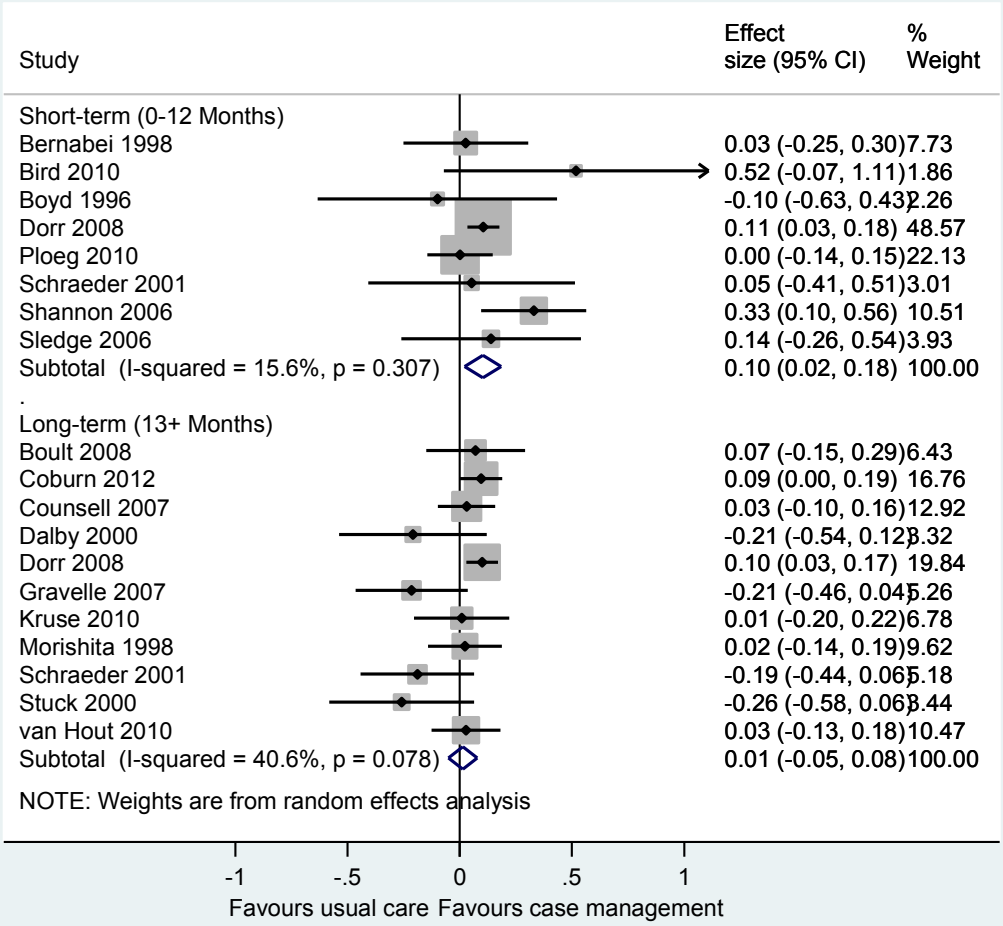
RESULTS OF THE SENSITIVITY ANALYSIS EXCLUDING STUDIES CONDUCTED IN VETERAN'S SETTINGS (OVER 90% MALE POPULATION)*

*** No significant results remain following Holm-Bonferroni correction for multiple comparisons**

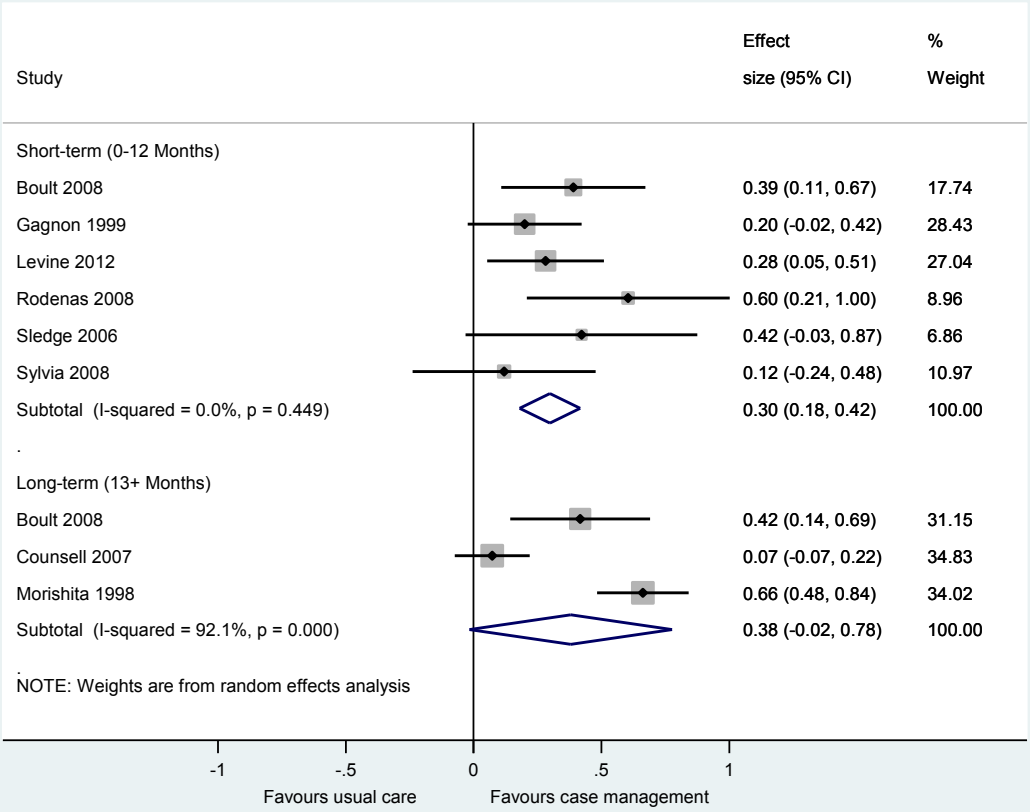
PRIMARY ANALYSIS – SENSITIVITY ANALYSIS



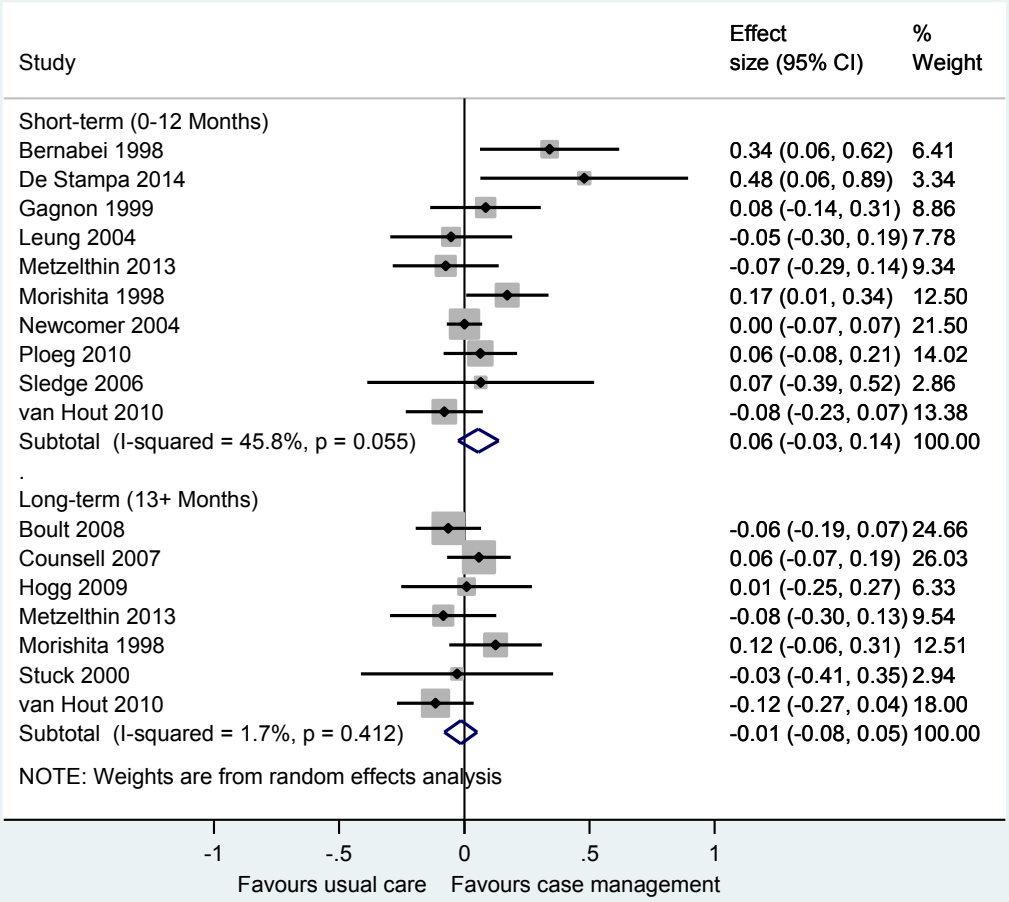
Mortality


Study
Effect size (95% CI)
% Weight
Short-term (0-12 Months)
Bernabei 1998 0.03 (-0.25, 0.30) 7.73
Bird 2010 0.52 (-0.07, 1.11) 1.86
Boyd 1996 -0.10 (-0.63, 0.43) 2.26
Dorr 2008 0.11 (0.03, 0.18) 48.57
Ploeg 2010 0.00 (-0.14, 0.15) 22.13
Schraeder 2001 0.05 (-0.41, 0.51) 3.01
Shannon 2006 0.33 (0.10, 0.56) 10.51
Sledge 2006 0.14 (-0.26, 0.54) 3.93
Subtotal (I-squared = 15.6%, p = 0.307) 0.10 (0.02, 0.18) 100.00
.
Long-term (13+ Months)
Boult 2008 0.07 (-0.15, 0.29) 6.43
Coburn 2012 0.09 (0.00, 0.19) 16.76
Counsell 2007 0.03 (-0.10, 0.16) 12.92
Dalby 2000 -0.21 (-0.54, 0.12) 3.32
Dorr 2008 0.10 (0.03, 0.17) 19.84
Gravelle 2007 -0.21 (-0.46, 0.04) 5.26
Kruse 2010 0.01 (-0.20, 0.22) 6.78
Morishita 1998 0.02 (-0.14, 0.19) 9.62
Schraeder 2001 -0.19 (-0.44, 0.06) 5.18
Stuck 2000 -0.26 (-0.58, 0.06) 3.44
van Hout 2010 0.03 (-0.13, 0.18) 10.47
Subtotal (I-squared = 40.6%, p = 0.078) 0.01 (-0.05, 0.08) 100.00
NOTE: Weights are from random effects analysis
-1
-.5
0
.5
1
Favours usual care
Favours case management

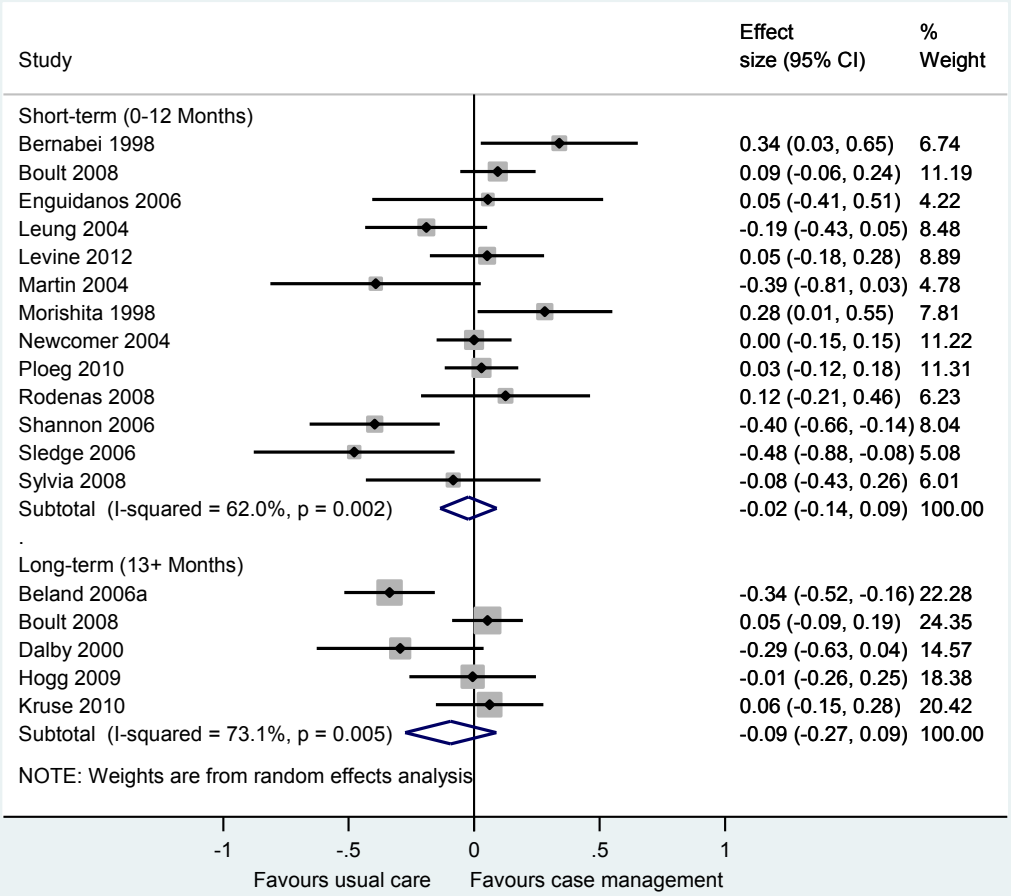
Patient satisfaction



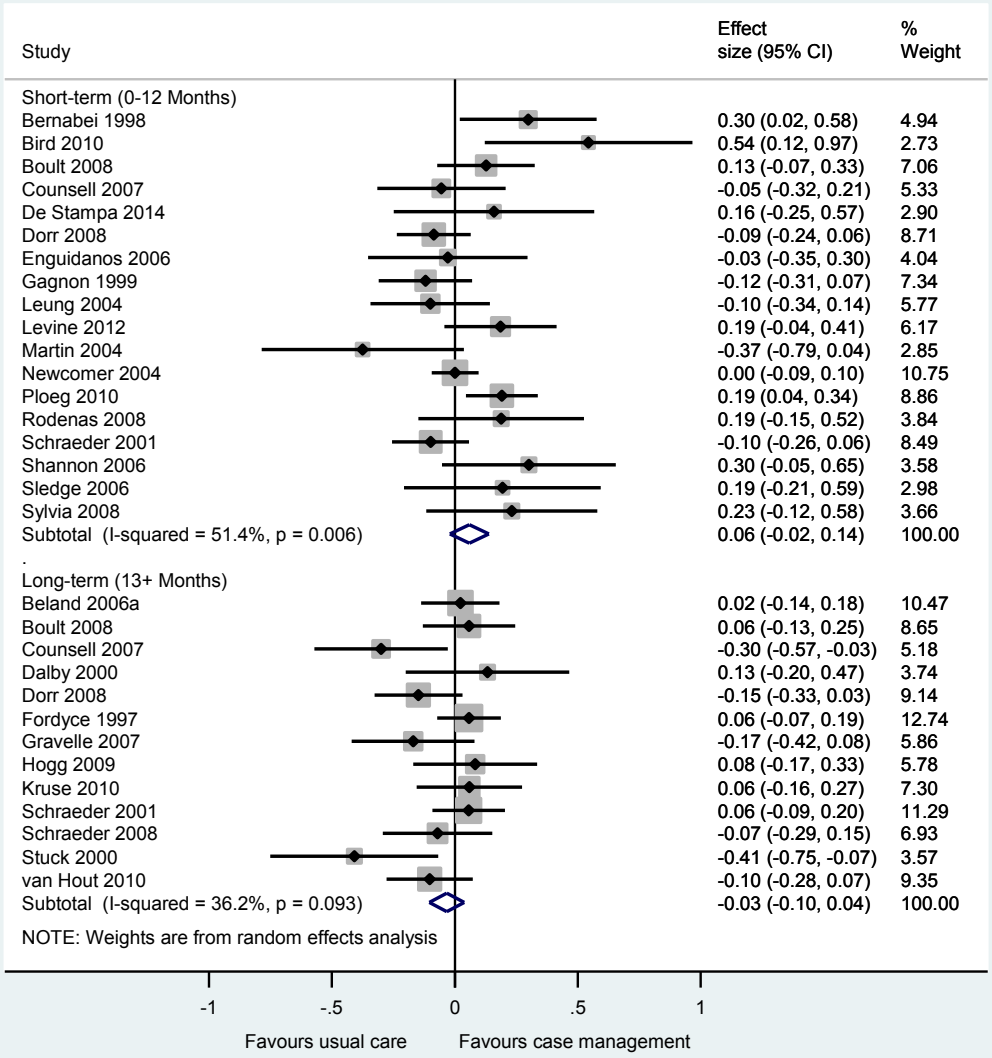
Self-assessed health status



Utilisation of primary and non-specialist care



Utilisation of secondary care



SECONDARY ANALYSIS – SENSITIVITY ANALYSIS

Outcome (time-period)	**Subgroup effect size (number of studies)**	
	MDT (19)	**Single (12)**
Mortality (short)	n/a	n/a
Mortality (long)	0.01 (-0.08 to 0.10) (4)	0.01 (-0.08 to 0.10) (7)
Self-rated health (short)	0.13 (-0.03 to 0.29) (6)	0.00 (-0.05 to 0.06) (4)
Utilisation primary care (short)	0.01 (-0.14 to 0.15) (10)	-0.09 (-0.30 to 0.12) (3)
Utilisation secondary care (short)	0.08 (-0.03 to 0.19) (13)	0.03 (-0.10 to 0.15) (5)
Utilisation secondary care (long)	0.02 (-0.06 to 0.09) (7)	-0.10 (-0.22 to 0.0) (6)
	Low PHC score (18)	**Int/high PHC score (13)**
Mortality (short)	n/a	n/a
Mortality (long)	0.04 (-0.02 to 0.11) (8)	-0.10 (-0.27 to 0.08) (3)
Self-rated health (short)	0.12 (-0.05 to 0.29) (4)	0.03 (-0.08 to 0.13) (6)
Utilisation primary care (short)	-0.04 (-0.19 to 0.11) (8)	-0.00 (-0.20 to 0.20) (5)
Utilisation secondary care (short)	0.03 (-0.05 to 0.10) (11)	0.08 (-0.10 to 0.26) (7)
Utilisation secondary care (long)	-0.05 (-0.15 to 0.05) (8)	-0.02 (-0.12 to 0.07) (5)
	Clinical Judgement (4)	**Risk modelling (27)**
Mortality (short)	n/a	n/a
Mortality (long)	-0.02 (-0.30 to 0.26) (2)	0.01 (-0.06 to 0.08) (9)
Self-rated health (short)	n/a	n/a
Utilisation primary care (short)	n/a	n/a
Utilisation secondary care (short)	-0.06 (-0.18 to 0.06) (3)	0.09 (-0.00 to 0.18) (15)
Utilisation secondary care (long)	-0.01 (-0.15 to 0.14) (3)	-0.04 (-0.13 to 0.04) (10)
	RCT (23)	**Non-RCT (8)**
Mortality (short)	n/a	n/a
Mortality (long)	0.01 (-0.06 to 0.08) (8)	-0.00 (-0.18 to 0.17) (3)
Self-rated health (short)	n/a	n/a
Utilisation primary care (short)	n/a	n/a
Utilisation secondary care (short)	0.05 (-0.04 to 0.13) (14)	0.17 (-0.11 to 0.45) (4)
Utilisation secondary care (long)	-0.02 (-0.10 to 0.07) (9)	-0.08 (-0.19 to 0.02) (4)
	Social worker (10)	**No social worker (21)**
Mortality (short)	n/a	n/a
Mortality (long)	0.03 (-0.07 to 0.13) (2)	-0.00 (-0.09 to 0.08) (9)
Self-rated health (short)	0.14 (-0.02 to 0.30) (4)	0.02 (-0.07 to 0.10) (6)
Utilisation primary care (short)	-0.02 (-0.23 to 0.18) (8)	0.01 (-0.09 to 0.11) (5)
Utilisation secondary care (short)	0.11 (-0.01 to 0.22) (8)	0.03 (-0.07 to 0.13) (10)
Utilisation secondary care (long)	-0.12 (-0.43 to 0.19) (2)	-0.02 (-0.10 to 0.05) (11)

No significant difference between subgroups ($p<0.05$)

* = significant in-subgroup effect ($p<0.05$)

Note: Positive effect size favours case management for all measures

Appendix 3: Chapter 6 Supplementary files

Appendices

Individual-level methods extension

The parallel trends assumption was not fulfilled when we compared the PICT patients to all possible control patients in the CCG. We therefore had to match intervention patients with potential controls from within the CCG using one-to-one propensity score. We matched using characteristics for which the intervention patients were selected in implementation (based on age; sex; index of multiple deprivation (IMD) 2010; total multimorbidity count previous to the first available start date of our intervention patients; previous inpatient, outpatient and A&E attendance in the previous year before the first intervention patient start date).[1]

As recommended,[2] prior to matching, we first imputed missing data for IMD based on the other complete variables, using multivariate normal multiple imputation (we used the STATA command 'mi impute' and used the average of

10 imputations).[3 2] We used STATA's 'pscore' and 'psmatch2' commands for propensity score matching.[4 5]

When we identified any control patients who had died before their matched start date, we eliminated these controls from the matching dataset (as well as all those intervention and control patients already matched adequately) and returned to this reduced matching pool to repeat the propensity match for the remaining intervention patients. This process was repeated twice, until all intervention patients were adequately matched to a living patient at their start date (round 1, n=1982 intervention patients correctly matched; round 2, n=62; round 3, n=5; total, n=2049).

The fact that nearly all practices were already implementing the intervention prior to individual-level analysis means the risk score of the control patients is likely to be slightly lower than the risk score of the intervention patients (i.e. we would expect that the highest risk patients in each practice would have already been signed up to the intervention). However, the DD analysis technique we use does not require control and intervention groups to be exactly the same, only that they are comparable based on parallel pre-trends of each of the outcomes assessed.

Outcome measures were summed to a count per patient per month over the period September 2010 to March 2015 inclusive, to allow a 3-year pre-trend period. All summed co-variates were once again sourced from the master

dataset, together with the cumulative multimorbidity count (described in the Appendices). Missing observations were filled as detailed in the Appendices.

Analysis models were chosen based on the results of the stata 'countfit' command, which compares count models based on the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and Vuong's closeness test. For the majority of outcomes, the best fitting model was the negative binomial model. The exception to this was for *total cost of secondary care services*, where we used a zero-inflated negative binomial model (inflating based on admission events). In each case, all three comparison tests agreed to the count model chosen. The equation used was:

$$y_{ijt} = \beta_1 PICT_{ijt} + \beta_2 PICT_{ijt} \times Post_{ijt} + \delta_t + \boldsymbol{xk}_{it} + \alpha_j + \varepsilon_{ijt}$$

Where:

y_{ijt} = outcome of person i in practice j in time t

$PICT_{ijt}$= dummy for treatment status

$PICTxPost_{ijt}$ = a dummy variable which equals one for treatment units in the post-treatment period, and is otherwise zero

δ_t= time fixed-effects

$\boldsymbol{xk}_{it}$ = *individual covariates* (age, cumulative multimorbidity count, IMD 2010 domains (excluding health domain))

α_j = constant with absorbed practice fixed-effects

ε_{ijt} = random error

Stratification by risk score

We were interested to observe whether those treated patients with a higher risk score, were indeed those at highest risk of future emergency admissions, and to test whether those at higher risk benefited more from the intervention.

While we were only able to access the actual risk scores of the intervention patients and not the controls, due to information governance rules, we assessed the correlation of the actual risk score of the intervention patients with the propensity score we used to match the intervention patients to controls. We found these to be moderately,[6] statistically significantly correlated ($r=0.38$; $p<0.0001$), and therefore assigned the same actual risk score of the intervention patients to their one-to-one propensity matched control for this sub-analysis.

We first looked descriptively at the association of risk score and post-intervention outcomes by calculating a single 'post-intervention' admissions value for each of the secondary care utilisation measures, and plotted this graphically against the risk score. To ensure the highest risk patients were not also those recruited first to the intervention (thus artificially inflating our simple measure for this sub-analysis), we additionally plotted each intervention patient start date together with the risk score to determine any time-varying relationship that might exist.

We assigned those intervention patients and matched controls with a risk score more than the 75th percentile to the high-risk group, and conducted the DDD

analysis. The extended equation for this extra interaction-term becomes, where the average partial effect of results for β_7 is reported:

$$y_{ijt} = \beta_1 PICT_{ijt} + \beta_2 Post_{ijt} + \beta_3 RISK_i + \beta_4 PICT_{ijt} \times Post_{ijt} + \beta_5 PICT_{ijt} \times RISK_i + \beta_6 Post_{ijt} \times RISK_i + \beta_7 PICT_{ijt} \times Post_{ijt} \times RISK_i + \delta_t + \boldsymbol{xk}_{it} + \alpha_j + \varepsilon_{ijt}$$

Where:

y_{ijt} = outcome of person i in practice j in time t

$PICT_{ijt}$= dummy for treatment status

$RISK_i$ = dummy for high risk score according to risk tool

$Post_{ijt}$ = dummy for the post-treatment period

δ_t= time fixed-effects

$\boldsymbol{xk}_{it}$ = *individual covariates* (age, cumulative multimorbidity count, IMD 2010 domains (excluding health domain))

α_j = constant with absorbed practice fixed-effects

ε_{ijt} = random error

Practice-level methods extension

The practice integrated care teams were initially introduced as an option within a mandatory (contracted and paid) Quality Process scheme in the CCG, therefore practices initially volunteered for PICT. Due to the staggered start dates of practices signing up to the PICT intervention (see **Error! Reference source not found.** in main paper), a natural experiment emerged at the practice-level. We exploited this rollout with data collapsed to the practice-level for 30 practices in

the CCG, for which we had detailed start date information. Outcomes were summed to a count per 1000 patients per month for each of the practices, and analysed over the period September 2010 to March 2015 inclusive, to overlap with the individual-level analysis.

The practice fixed-effects in our analysis model act to remove any observed and unobserved, time-invariant confounders at the practice level (e.g. should PICT practices be relatively more focussed on reducing secondary care attendance than non-PICT practices), while the time fixed-effects act to remove any fluctuations in outcome due to the general trend over time. We report the results of β_1 in the main text, together with the corresponding effect size. The equation used was:

$$y_{jt} = \beta_1 PICT_{jt} \times Post_{jt} + \delta_t + \alpha_j + \varepsilon_{jt}$$

Where:

y_{jt} = outcome of practice j in time t

$PICT_{jt}xPost_{jt}$ = a dummy variable which equals one for treatment units in the post-treatment period, and is otherwise zero

δ_t = time fixed effects

α_j = constant with absorbed practice fixed-effects

ε_{jt} = random error

At the practice-level only, due to the voluntary roll-out of the intervention, we attempted to assess the effects of selection bias using a logistic regression model. We attempted to predict the wave of entry to the intervention using the practice

characteristics recorded (including: % males; % over 65; list size; number of GPs per thousand patients; total IMD score 2010; and total % QOF achievement score).[7] As a further robustness check, we additionally re-ran the practice-level analysis excluding those practices recruited to the intervention in wave 1, assuming these to be the practices at most risk of selection bias if it did indeed occur.[8]

Multimorbidity measure

For the individual-level analysis, a multimorbidity measure was prepared from the previously recorded (period June 2006 to March 2015) inpatient admissions, for 20 chronic conditions recorded in the Quality and Outcomes Framework (QOF) (see below for list of conditions and icd-10 codes).[9-11] This measure was recorded cumulatively by month over the dataset, with a binary indicator for each condition 'switched on' by the recording in the inpatient record, subsequently staying 'on', with addition of any further diagnoses recorded at a later date. The cumulative monthly total was used in the analysis.

Although we would have preferred to use chronic conditions data recorded in primary care for our multimorbidity measure, this data was unavailable for this study. However, as high-risk patients (both the intervention and propensity matched controls), these are probably the most likely to encounter inpatient admissions (and indeed were selected for 'high risk' of admissions), so should in theory have the most complete recordings at this service level in comparison to the general population. From previous literature though, we can expect our multimorbidity measure to be less sensitive (i.e. predictably lower count)

because it comes purely from hospitalisation data.[12] Nevertheless, our multimorbidity measure gained strength from its cumulative nature, accounting for changes over the life-course. Furthermore, we calculated the count based on a list of 20 chronic conditions, deemed particularly important in the UK's NHS setting, which follows guidance from the multimorbidity literature.[13]

1. Asthma
J45-J47

2. Atrial fibrillation
I48

3. Cancer
C00–C14, C15–C26, C30–C39,
C40–C41, C43–C44, C45–C49,
C50, C51–C58, C60–C63,
C64–C68, C69–C72, C73–C75,
C81–C96, C76–C80, C97,
D00–D09, D37–D48 (Koller et al 2014)

4. Chronic kidney disease
N18

5. Coronary heart disease
I20-I25

6. COPD
J40-J44

7. Dementia
F00-F03

8. Depression
F32-F33

9. Diabetes mellitus
E10-E14

10. Epilepsy
G40-G41

11. Heart failure
I50

12. Hypertension
I10-I15

13. Hypothyroidism
E00-E03

14. Learning disability
F80-F89

15. Mental health (schizophrenia, bipolar affective disorder and other psychoses)
F20-F29, F31, F34-F39

16. Obesity
E66

17. Osteoporosis
M80-M82

18. Atherosclerosis/Peripheral arterial oclusive disease (PAOD)
I65–I66, I67.2, I70, I73.9

19. Rheumatoid arthritis
M05-M06

20. Cerebral ischemia/chronic stroke
I60–I64, I69, G45

Missing data

Any missing Lower Layer Super Output Area (LSOA) codes (n= 295) were updated with the dominant LSOA where patients live, according to their GP practice code.[14] IMD domains (excluding the health domain) were matched to the LSOA codes.

Missing panel data (where the patient had no secondary care service use in a given month) was filled with zero observations using the STATA command 'tsfill' to make a balanced panel dataset.[15] Any filled observations following inpatient

mortality were deleted for that given patient, as well as any observations giving a negative age (for the youngest patients included in the intervention/control).

All other data used was complete in the dataset.

Ambulatory Care Sensitive Conditions (ACSCs)

Ambulatory Care Sensitive Conditions (ACSCs) are specified conditions, which should be managed adequately at the primary care level and prevented from worsening to the extent that secondary care is necessary. They are deemed an indication of access to and effectiveness of primary care in a health system.[16 17] When these set conditions are not managed effectively (particularly in a national health system which should provide universal access, as is the case in the UK), a harm is befalling the patient and a safety incident can be said to have occurred.[18]

We used the NHS's definition of an ACSC.[19] We included the following icd-10 codes, when they were coded as the primary diagnosis, for an emergency admission only.

Vaccine preventable
B18.0 Chronic viral hepatitis B with delta-agent
B18.1 Chronic viral hepatitis B without delta-agent
Asthma
J45 Asthma
J46X Status asthmaticus
Congestive heart failure
I11.0 Hypertensive heart disease with (congestive) heart failure
I50 Heart failure
J81X Pulmonary oedema
I13.0 Hypertensive heart and renal disease with (congestive) heart failure
Diabetes

E10 Insulin-dependent diabetes mellitus
E11 Non-insulin-dependent diabetes mellitus
E12 Malnutrition-related diabetes mellitus
E13 Other specified diabetes mellitus
E14 Unspecified diabetes mellitus
Chronic obstructive pulmonary disease
J20 Acute bronchitis
J41 Simple and mucopurulent chronic bronchitis
J42X Unspecified chronic bronchitis
J43 Emphysema
J44 Other chronic obstructive pulmonary disease
J47X Bronchiectasis
Angina
I20 Angina pectoris
I25 Chronic ischaemic heart disease
Iron deficiency anaemia
D50.1 Sideropenic dysphagia
D50.8 Other iron deficiency anaemias
D50.9 Iron deficiency anaemia, unspecified
D51 Vitamin B12 deficiency anaemia
D52 Folate deficiency anaemia
Hypertension
I10X Essential (primary) hypertension
I11.9 Hypertensive heart disease without (congestive) heart failure
Convulsions and epilepsy
G40 Epilepsy
G41 Status epilepticus
Dementia
F00 Dementia in alzheimers
F01 Vascular dementia
F02 Dementia in other diseases
F03 Unspecified dementia
Atrial fibrillation and flutter
I48X Atrial fibrillation and flutter

Patient satisfaction

Patient satisfaction was assessed at the practice-level using data from the GP Patient Survey (GPPS). The GPPS was provided by Ipsos MORI at the individual-level.[20] The satisfaction measure was analysed using two questions from the GPPS:

1. Overall, how would you describe your experience of your GP surgery? (general satisfaction)

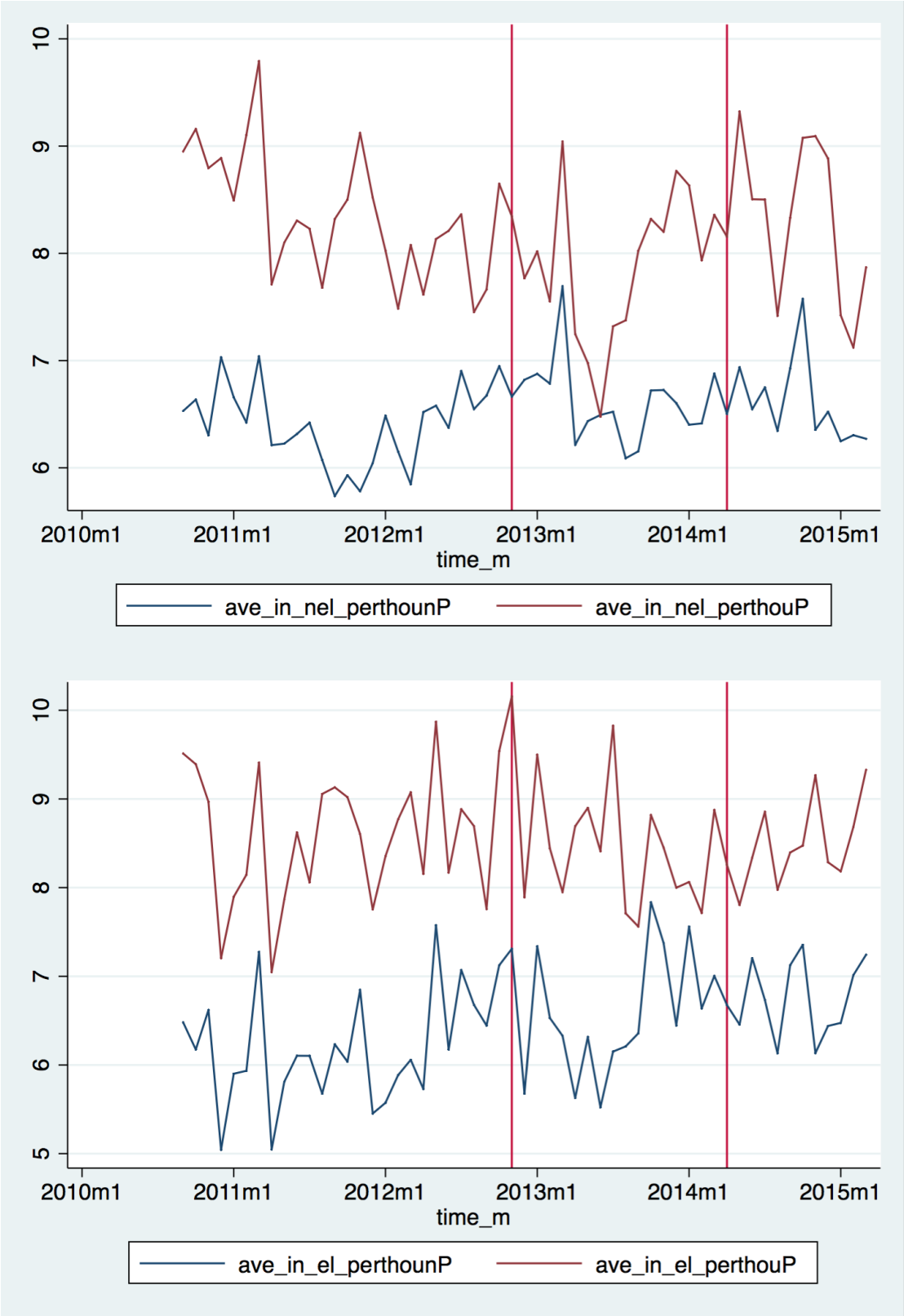
2. In the last 6 months, have you had enough support from local services or organisations to help you to manage your long-term health condition(s) (LTC)? (LTC-specific satisfaction)

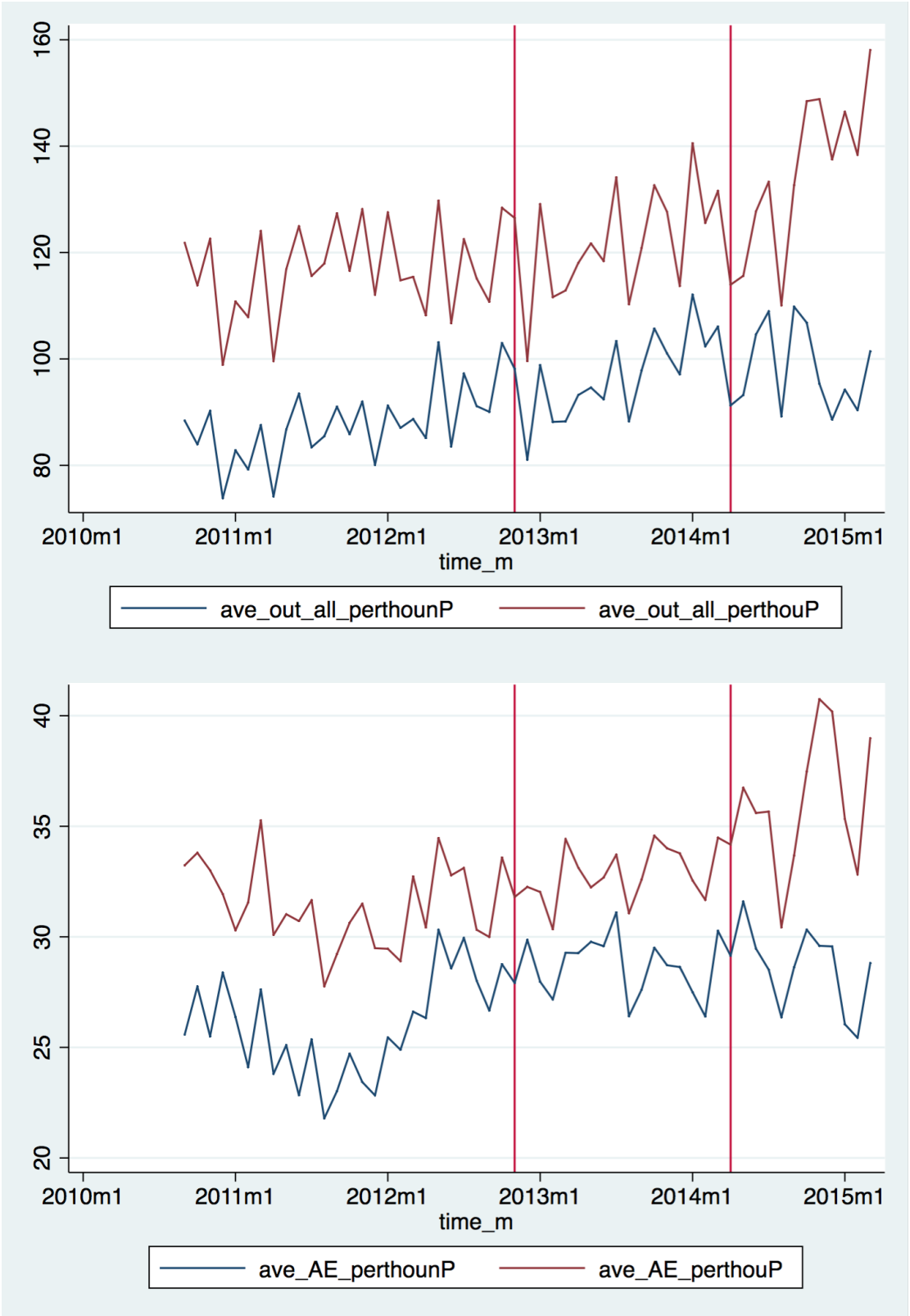
Responses were coded to a binary variable (1 for 'Very good'/'Yes, definitely', respectively; 0 for any less positive response, with missing responses coded as missing) for each patient within each of the 30 practices. This data was collapsed (summed) to the practice level (with weights used to collapse to representative samples of practice populations), along with a variable coded as '1' for each observation, which was consequently used to weight the outcomes by number of responses received for each practice, so that no practice was over-represented in the final analysis. The 'pre-period' was three survey waves, comprising the period from July 2011 to September 2012; and the 'post-period' for all practices was two survey waves, comprising the period from January 2013 to September 2013.

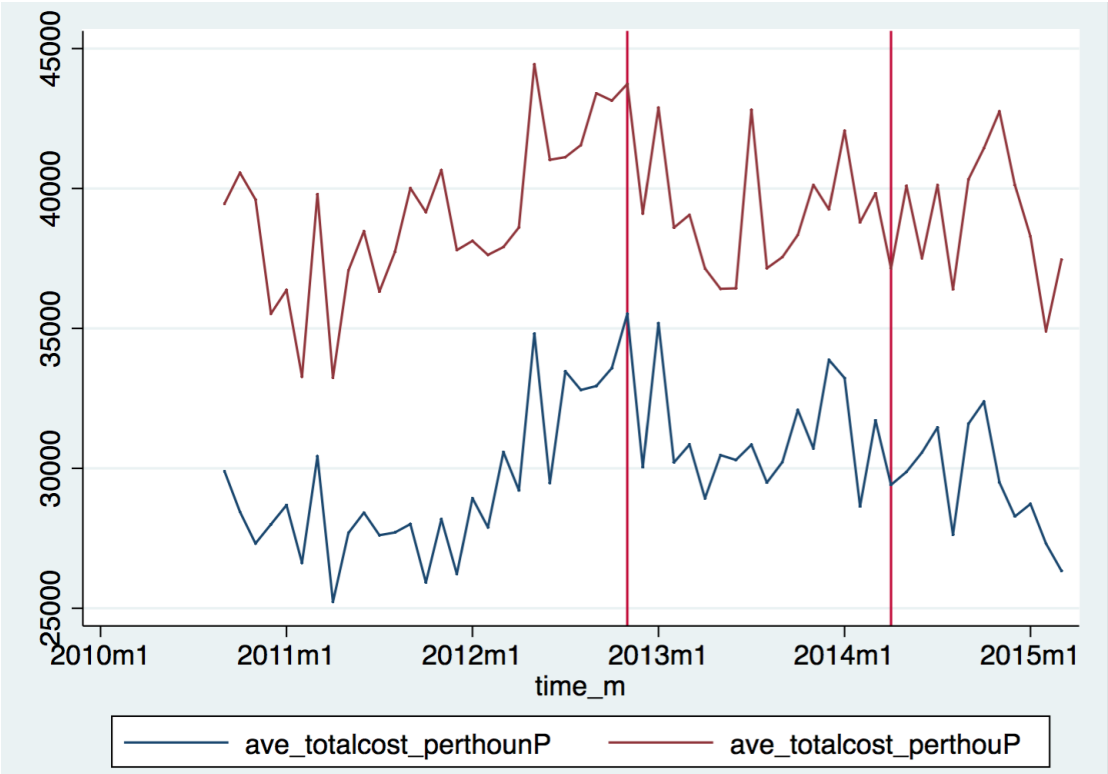
Parallel Trends graphs

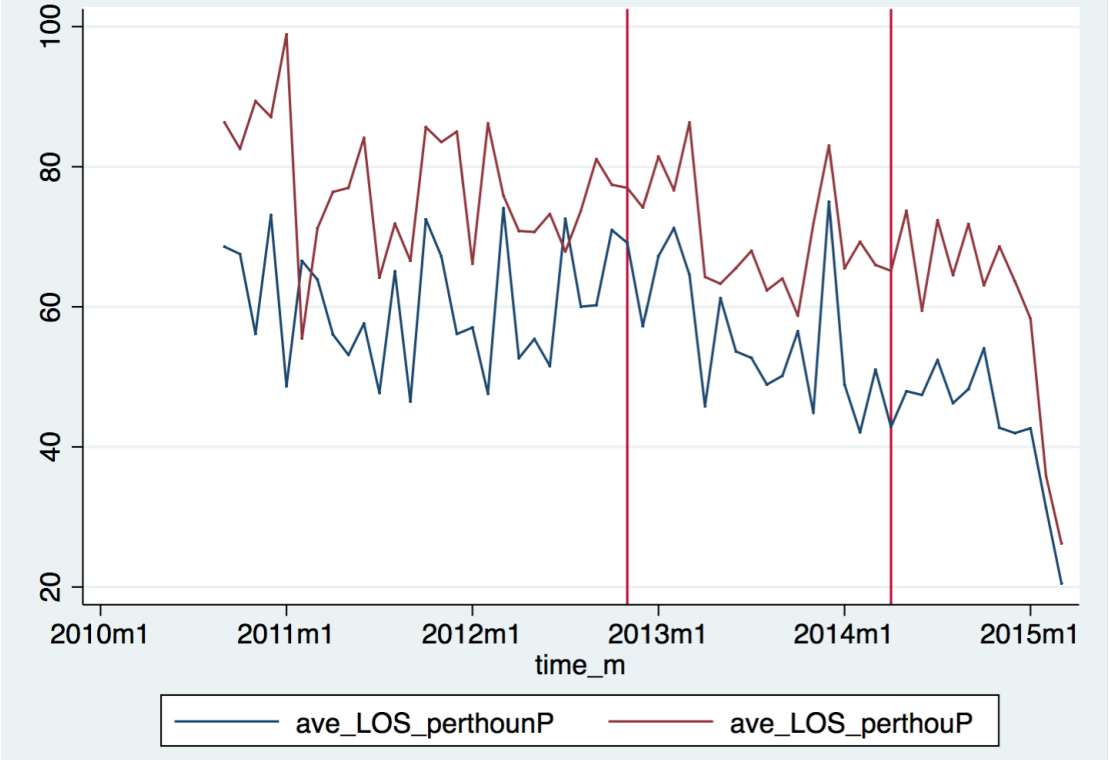
Pre-trends were additionally tested statistically using a continuous time linear time trend interacted with the treatment dummy, as well as multiple time dummies interacted with treatment dummy and using an f-test to assess overall significance, both methods using the pre-intervention data only. The results of these tests are reported in the main paper.

Practice-level average trends (wave 1 practices versus later joining practices)



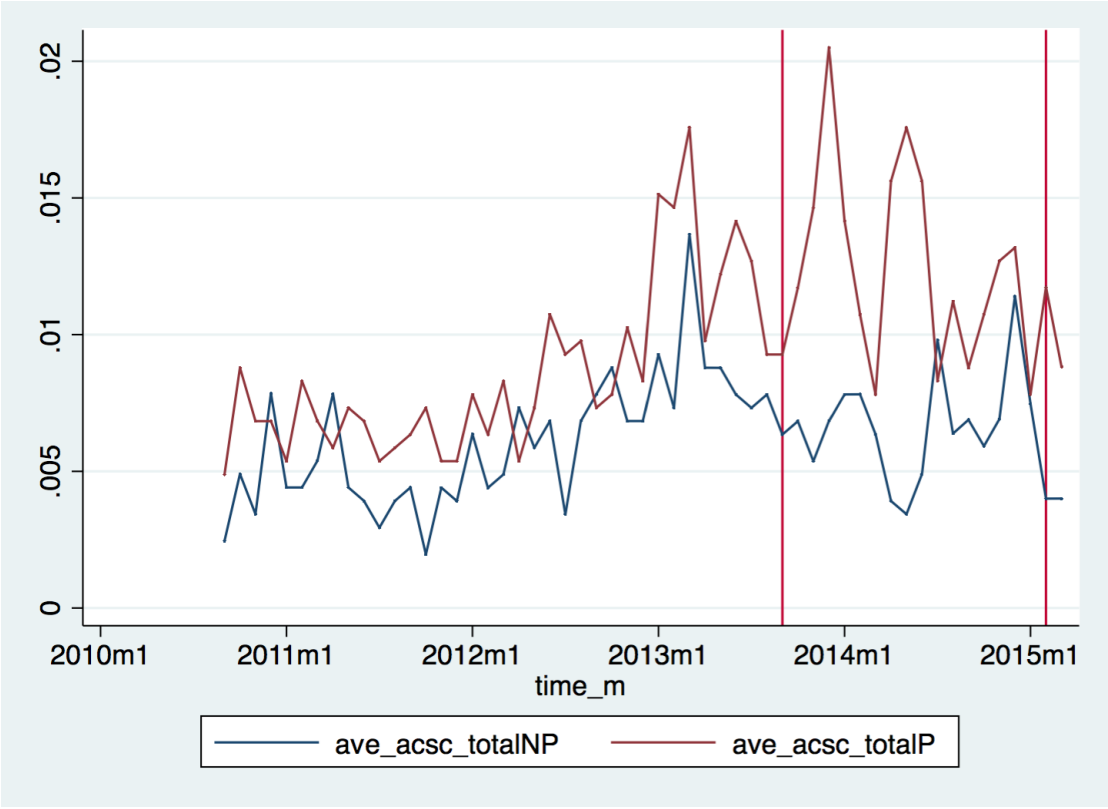

160
140
120
100
80
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_out_all_perthounP
ave_out_all_perthouP
40
35
30
25
20
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_AE_perthounP
ave_AE_perthouP

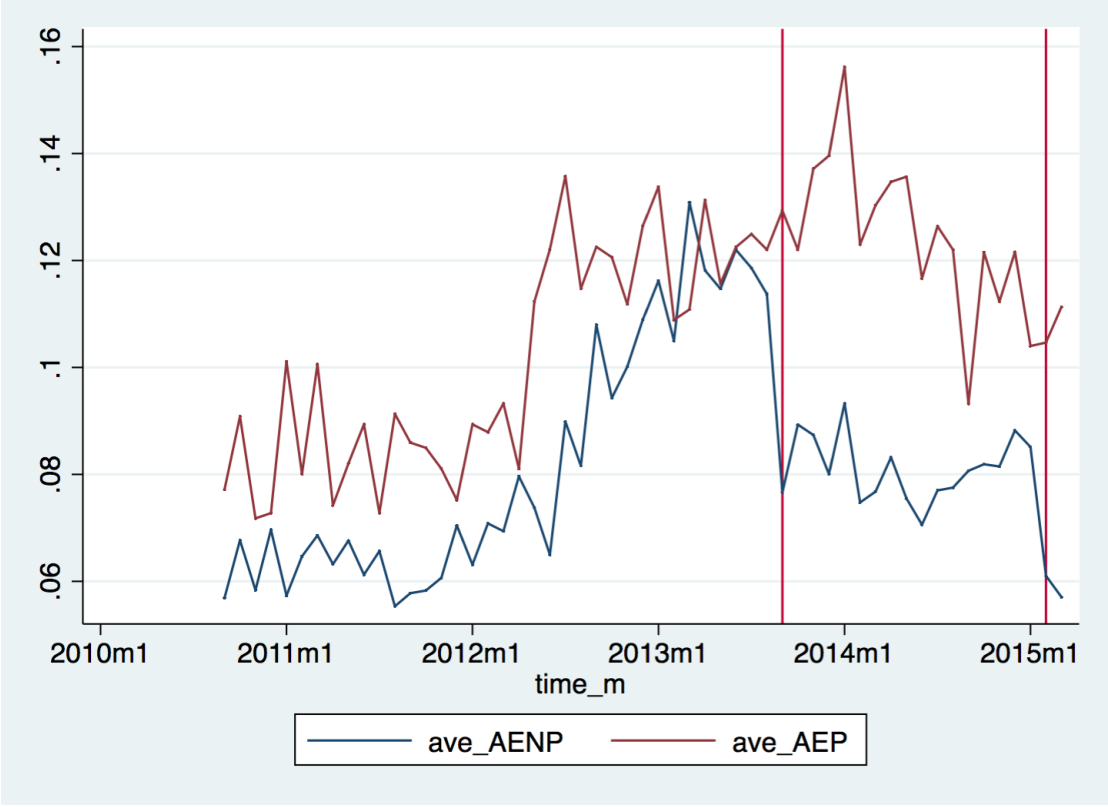

45000
40000
35000
30000
25000
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_totalcost_perthounP
ave_totalcost_perthouP

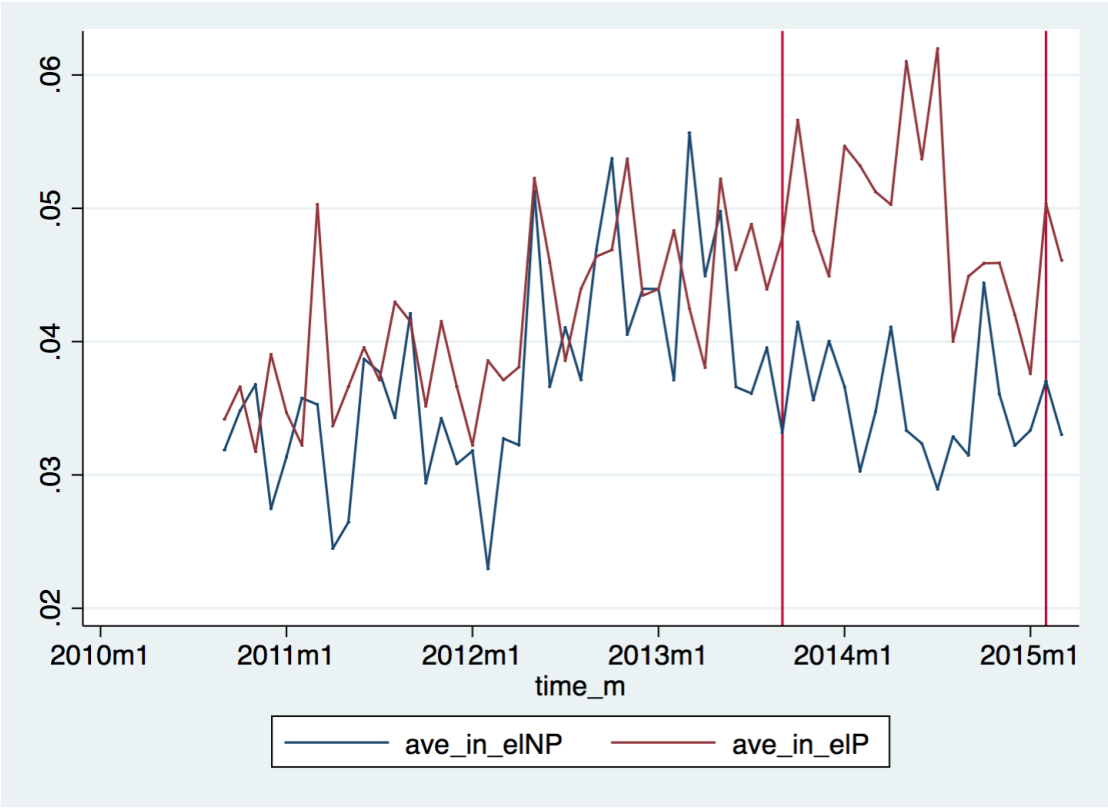

100
80
60
40
20
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_LOS_perthounP
ave_LOS_perthouP

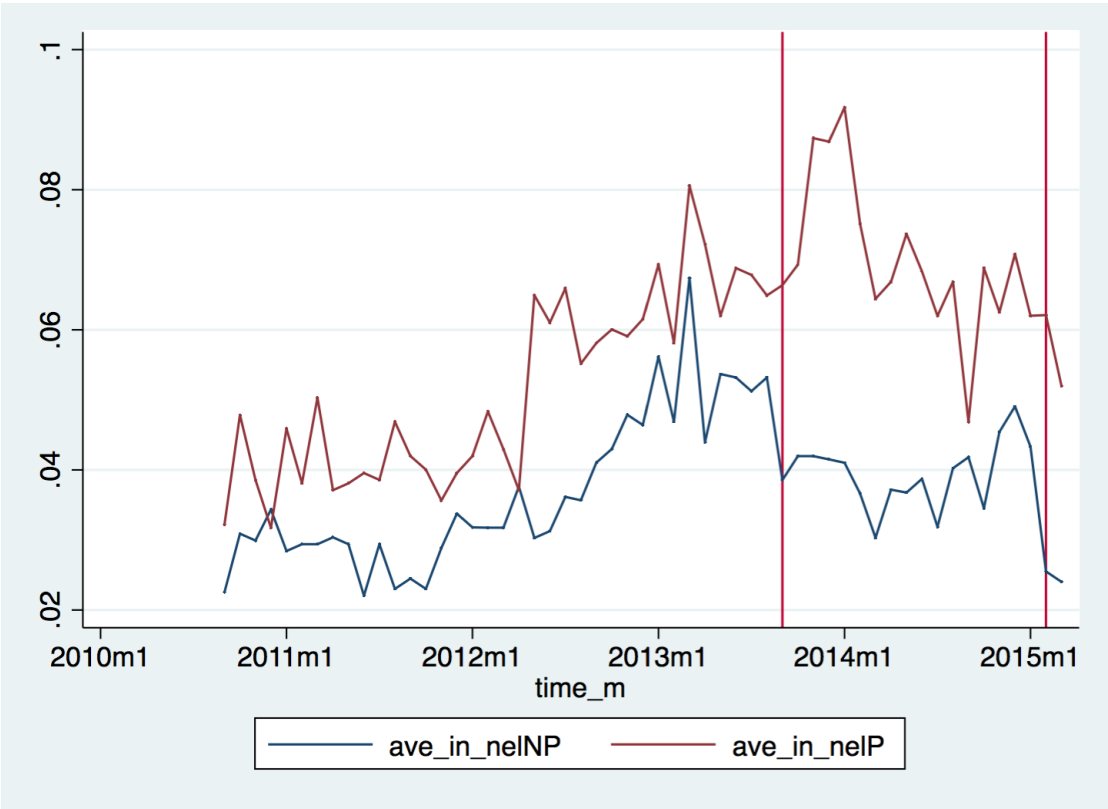

.4
.6
.8
1
1.2
1.4
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_acsc_perthounP
ave_acsc_perthouP
0
.5
1
1.5
2
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_readmit_perthounP
ave_readmit_perthouP

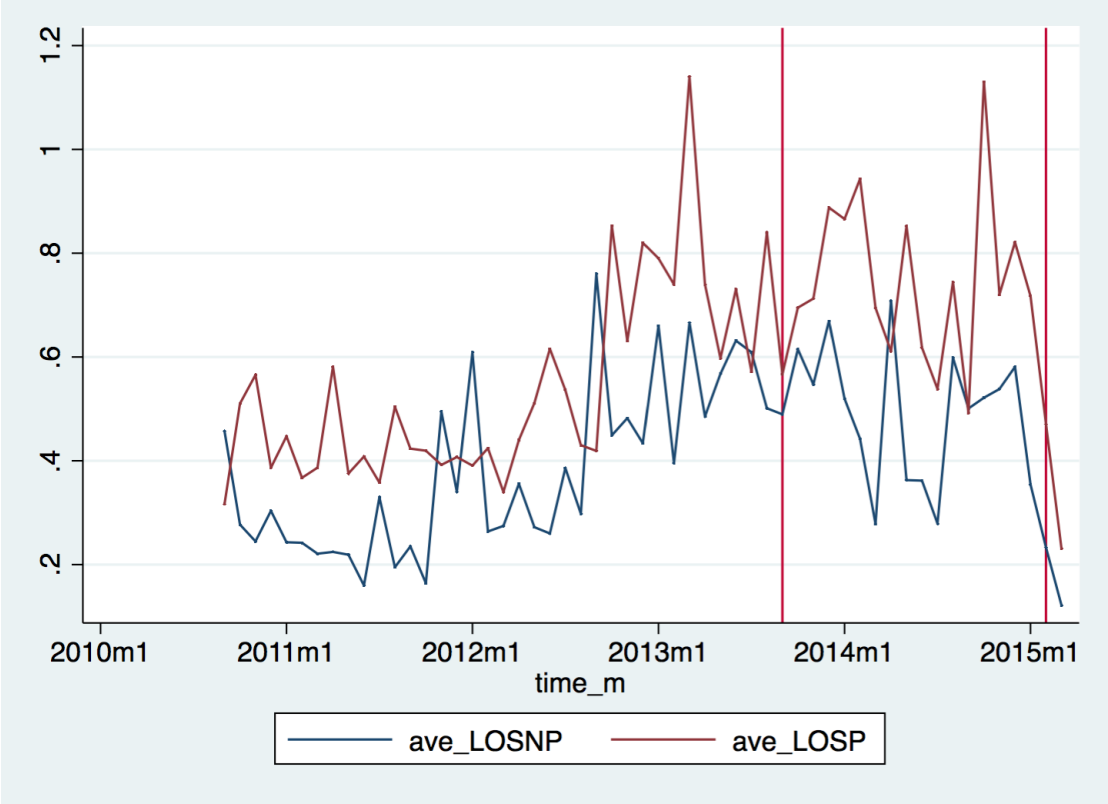
Individual-level average trends

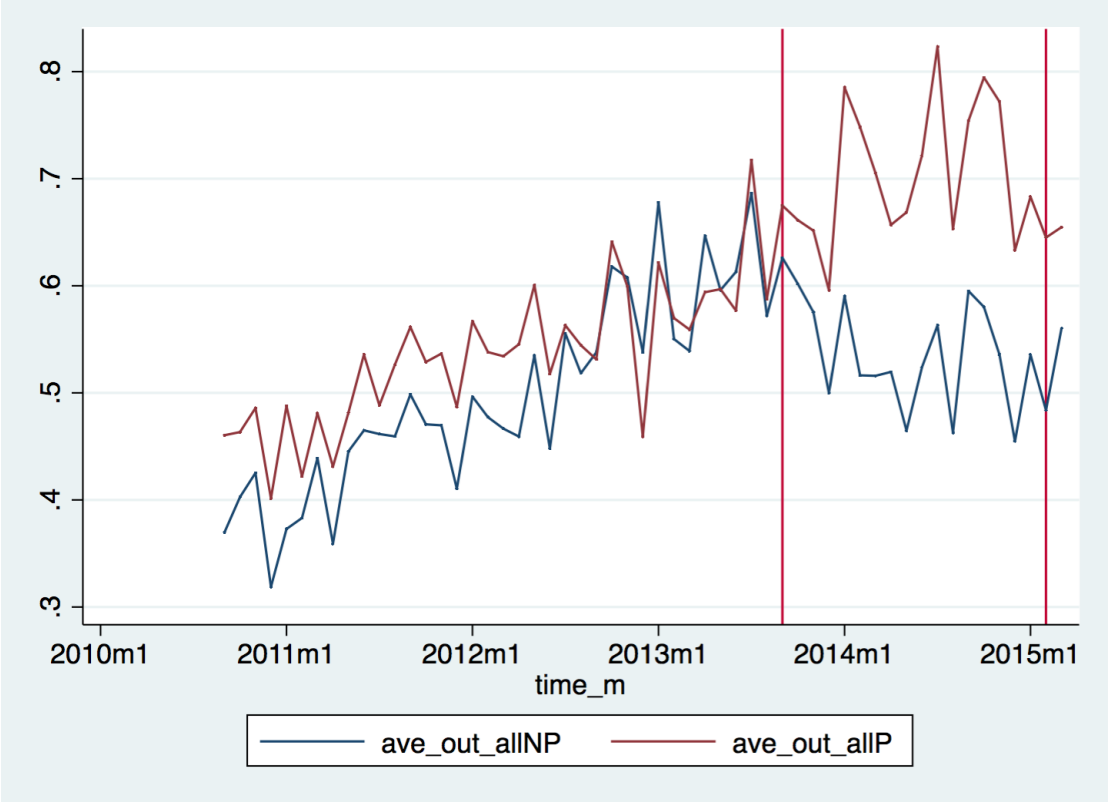


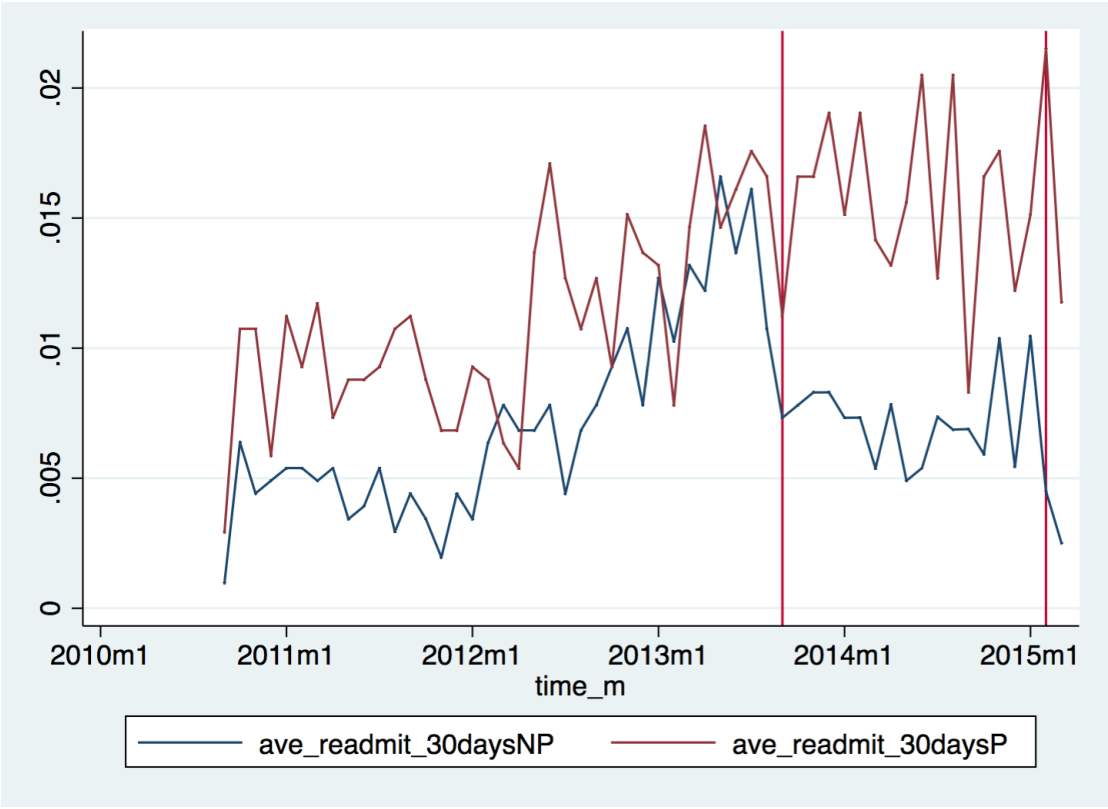


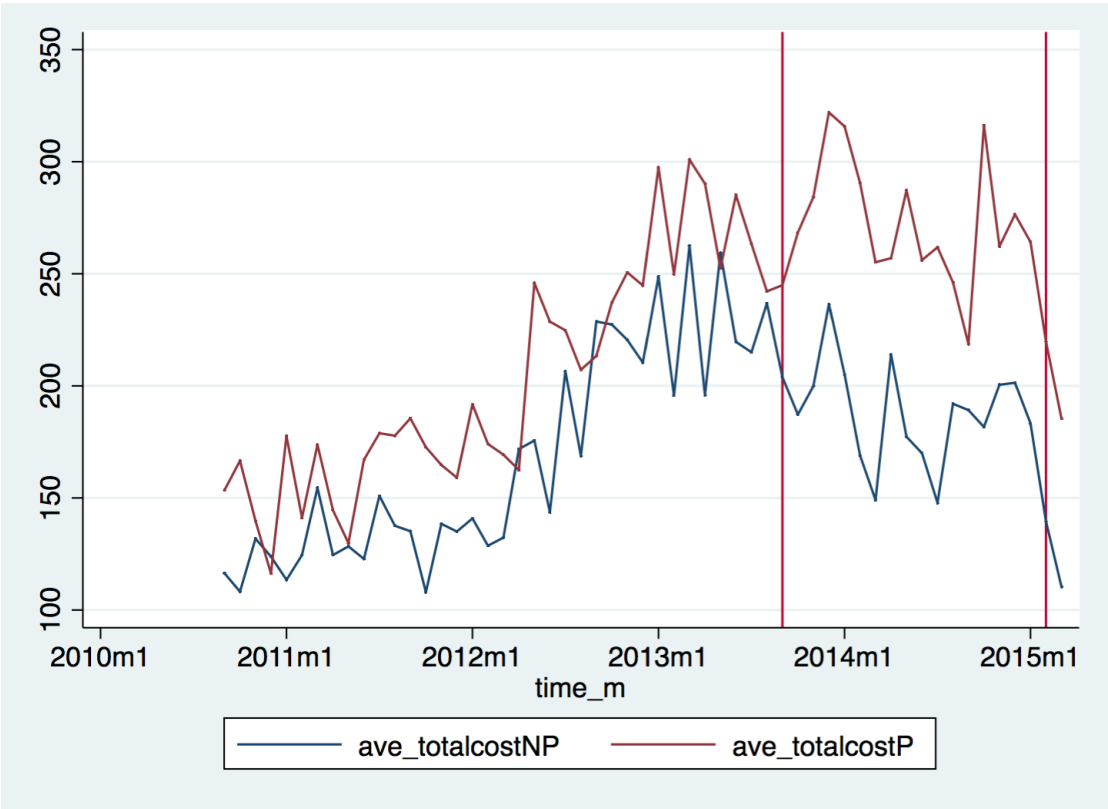

.02
.03
.04
.05
.06
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_in_elNP
ave_in_elP


.02
.04
.06
.08
.1
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_in_nelNP
ave_in_nelP


1.2
1
.8
.6
.4
.2
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_LOSNP
ave_LOSP


.8
.7
.6
.5
.4
.3
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_out_allNP
ave_out_allP


0
.005
.01
.015
.02
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_readmit_30daysNP
ave_readmit_30daysP


100
150
200
250
300
350
2010m1
2011m1
2012m1
2013m1
2014m1
2015m1
time_m
ave_totalcostNP
ave_totalcostP

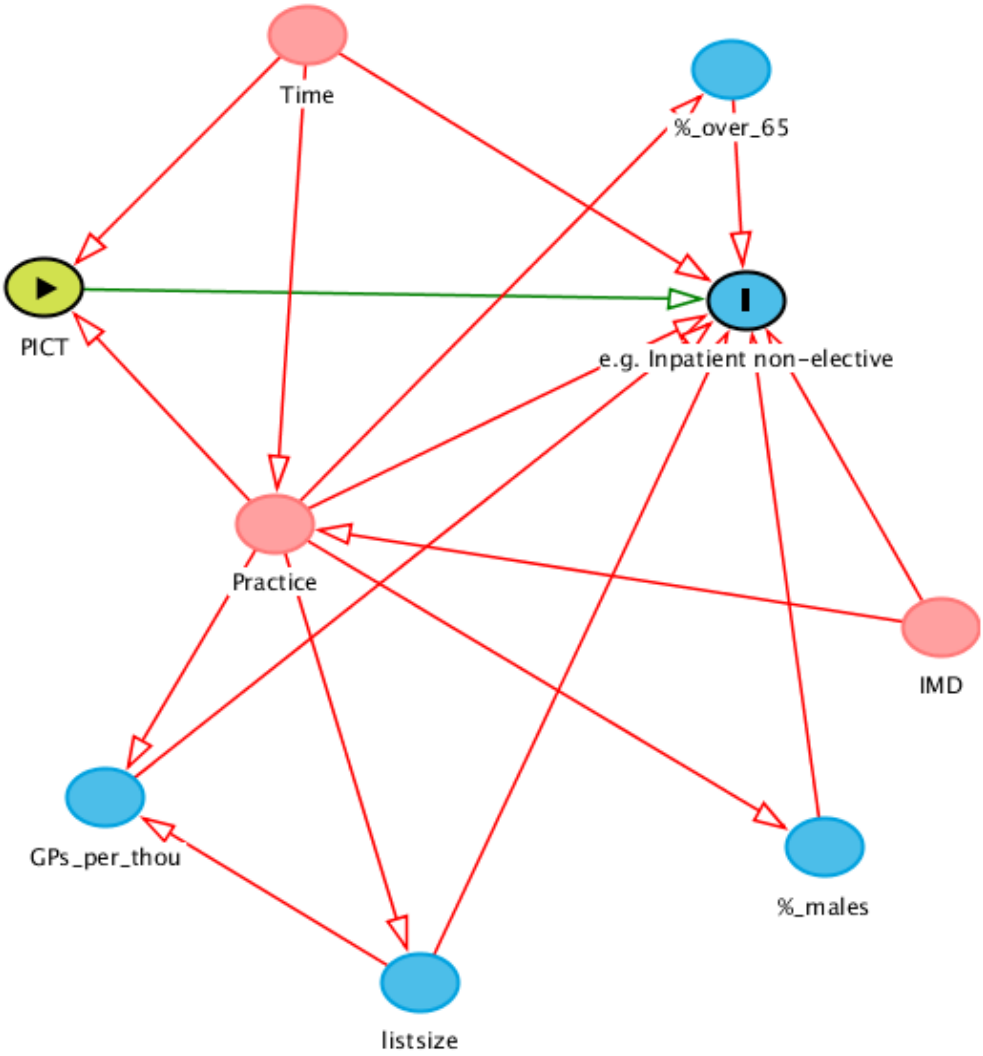
Directed Acyclic Graphs (DAGs)

Co-variates

We adjusted for co-variates in each model, chosen based on causal diagrams, known as directed acyclic graphs (DAGs – see Appendices).[21] DAGs are graphical descriptions that require us to set down clearly our assumptions about causal relationships, and can be used to select the minimal sufficient adjustment for determining the causal relationship of interest (in this case the effect of the PICT intervention on each of the outcome measures listed above), while ensuring minimum bias.[21]

The DAGs were created using free open-access software available at http://www.dagitty.net/dags.html#.

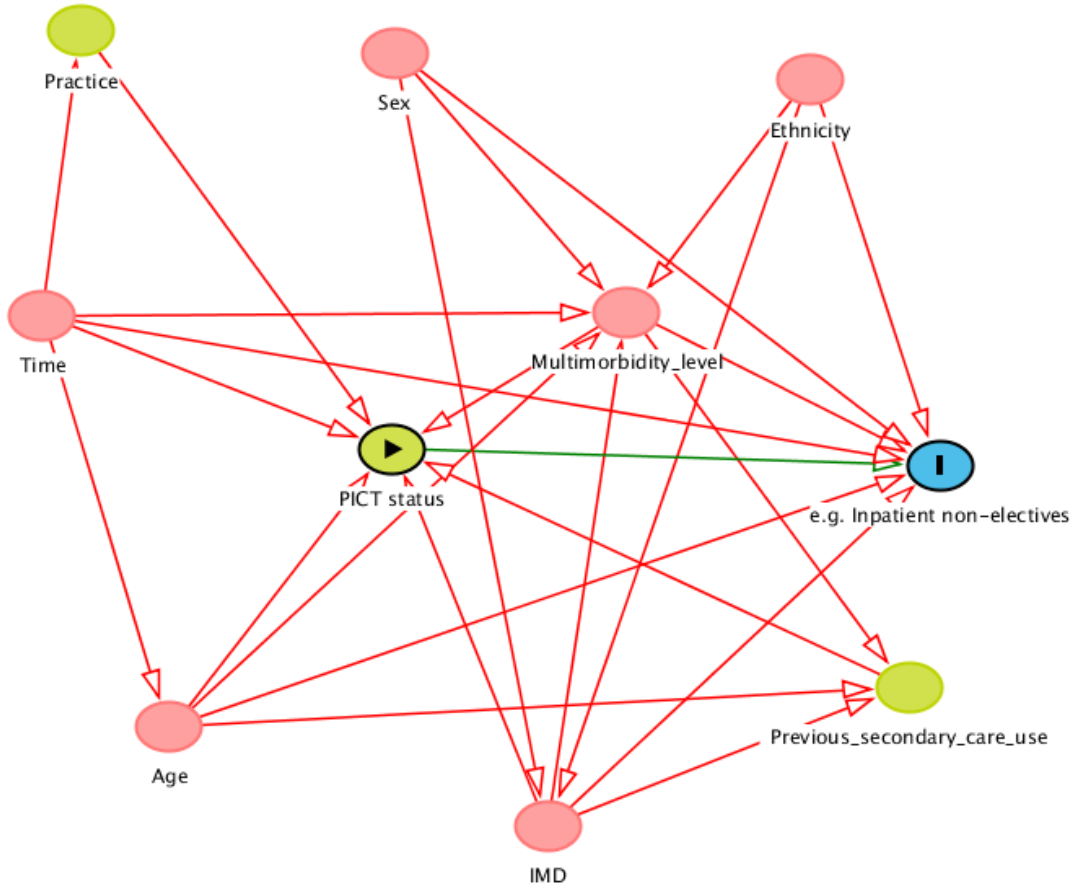
Practice-level



Minimal sufficient adjustment sets for estimating the total effect of PICT on e.g. Inpatient non-elective:

- Practice, Time

Individual-level



Minimal sufficient adjustment sets for estimating the total effect of PICT status on e.g. Inpatient non-electives:

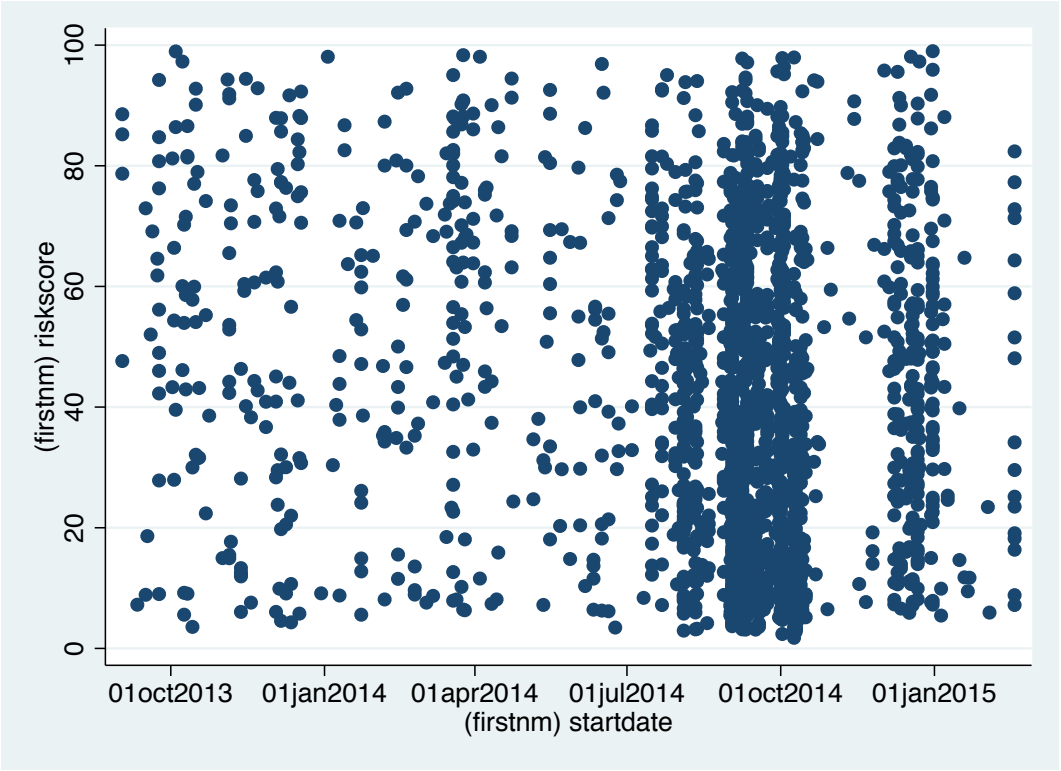
- Age, IMD, Multimorbidity_level, Time

Practice-level robustness check estimates (after removal of wave 1 practices)

Outcome	Adjusted+ Intervention effect (95% CI)	Effect size*
	Linear regression model	
Primary outcomes:		
Inpatient non-electives	-0.23 (-1.02 to 0.56)	-0.09
Inpatient electives	0.04 (-0.56 to 0.64)	0.01
Outpatient admissions	0.99 (-6.17 to 8.15)	0.03
A&E visits	0.35 (-2.35 to 3.06)	0.03
ACSCs	-0.05 (-0.20 to 0.10)	-0.10
Re-admissions (30 days)	0.01 (-0.22 to 0.22)	0.02
Secondary outcomes:		
Total cost of 2º care services	610.36 (-2434.52 to 3655.25)	0.05
Length of stay (days)	4.03 (-4.95 to 13.02)	0.10
Patient satisfaction (general)	-0.03 (-0.10 to 0.05)	-0.24
Patient satisfaction (LTC-specific)	0.01 (-0.05 to 0.07)	0.14

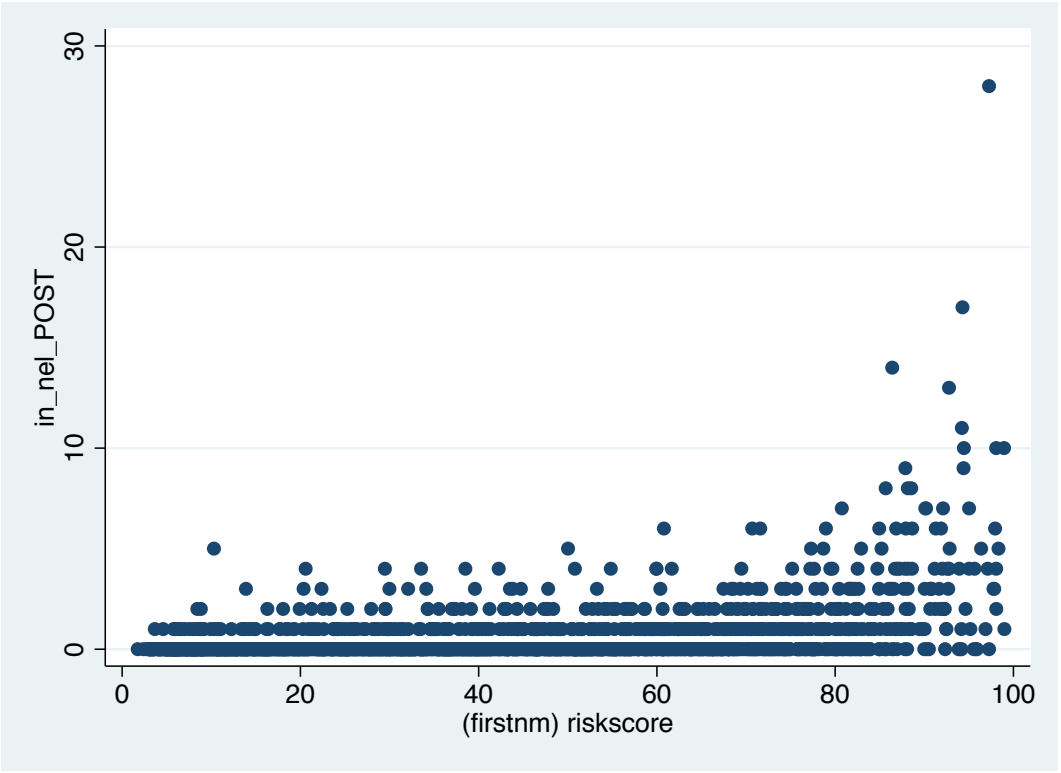
+: adjusted for practice and time fixed-effects with robust standard errors
*: standardised mean difference
#: significant at $p<0.05$
n= 990 observations; 18 practices (period November 2010 to March 2015)

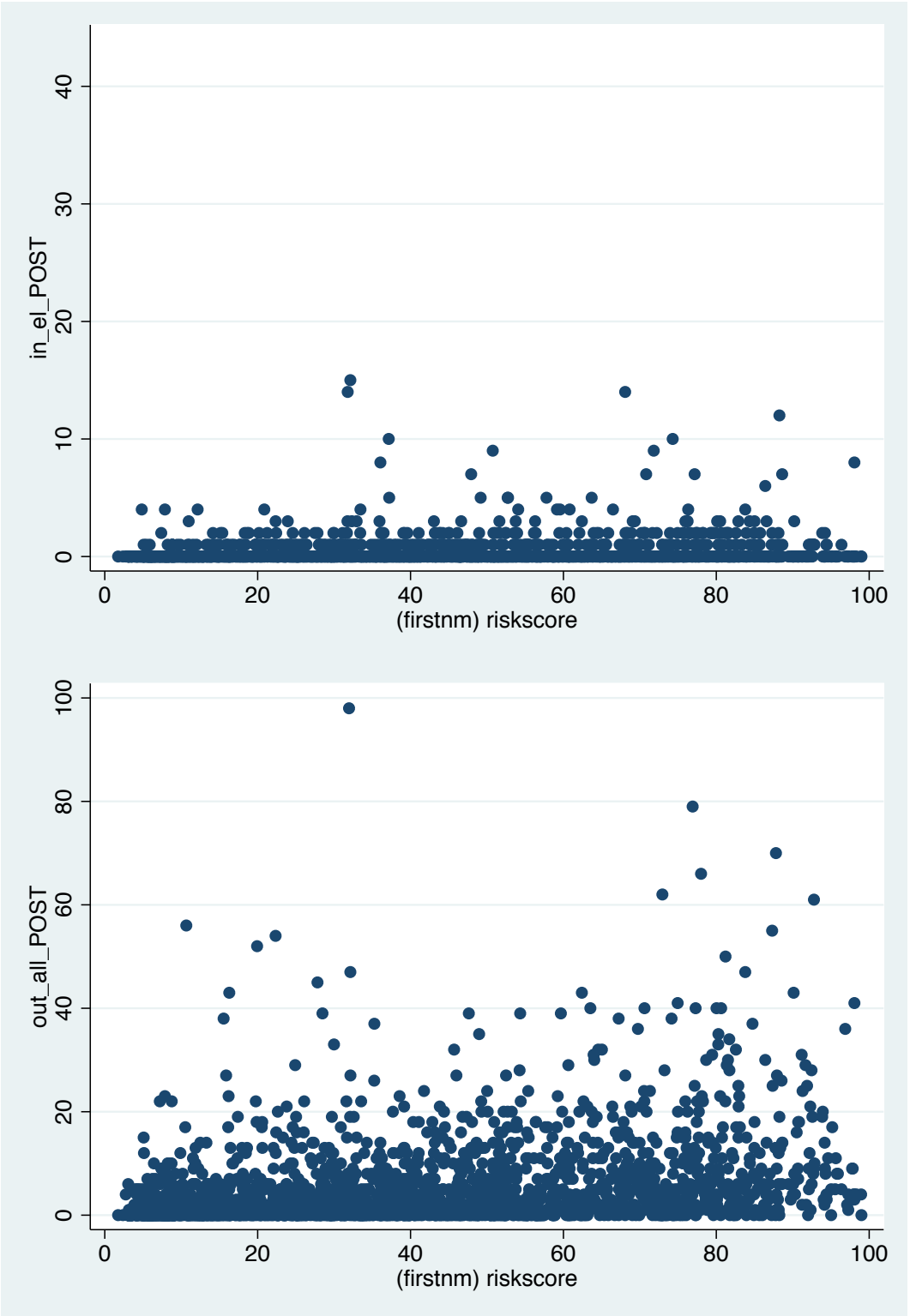
Risk score stratification results (individual-level)

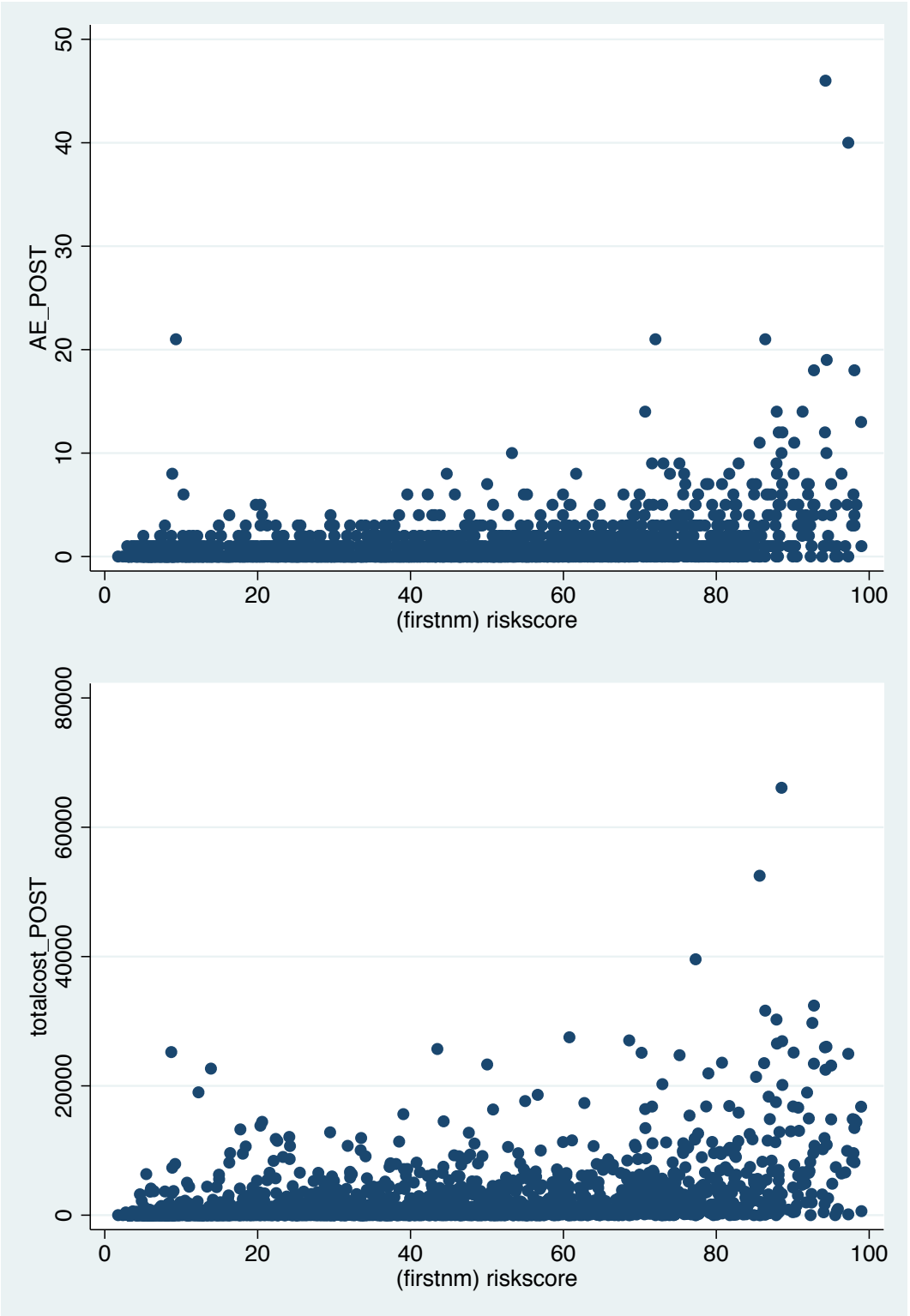


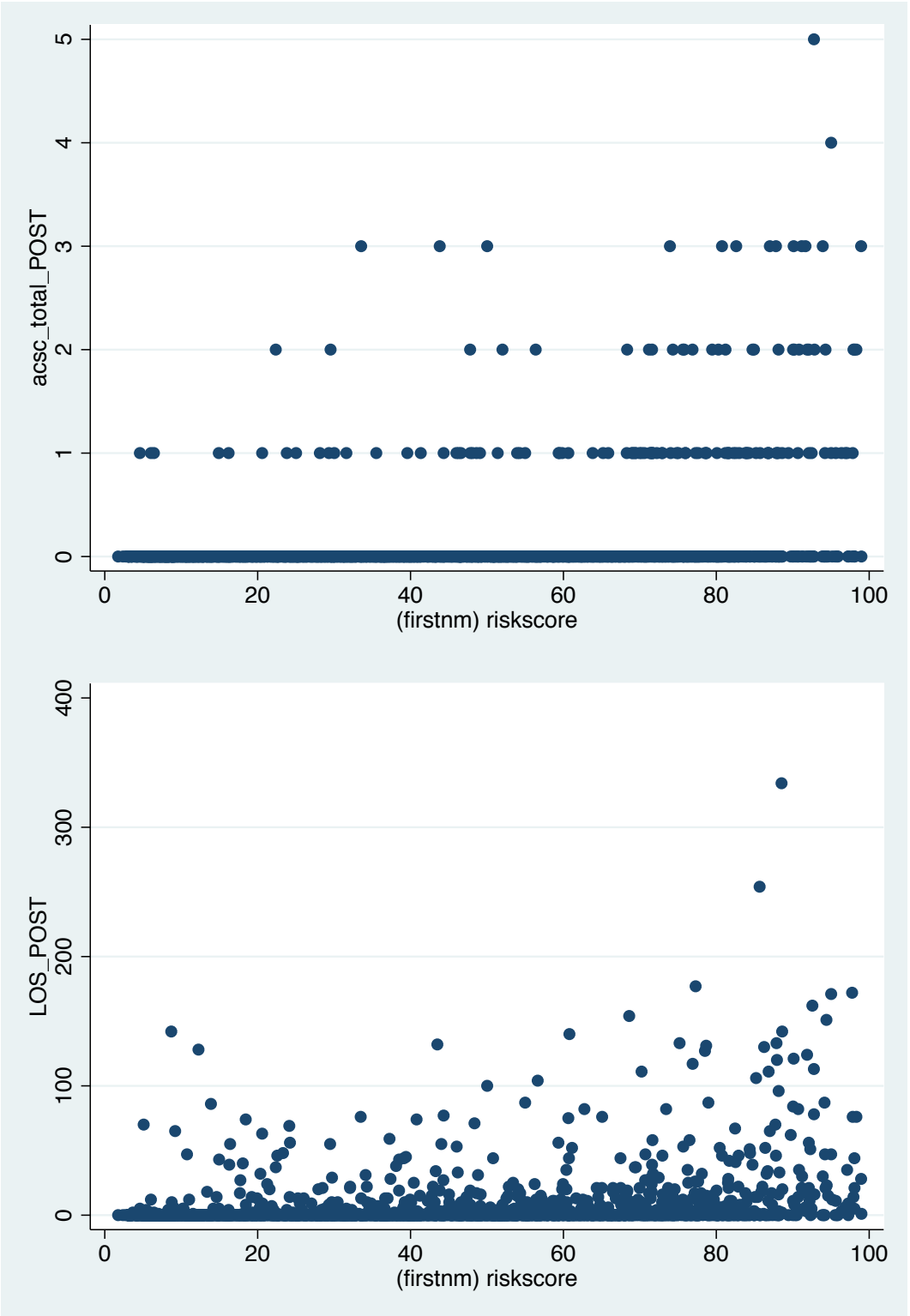
No indication of relationship between risk score and time recruited to intervention.

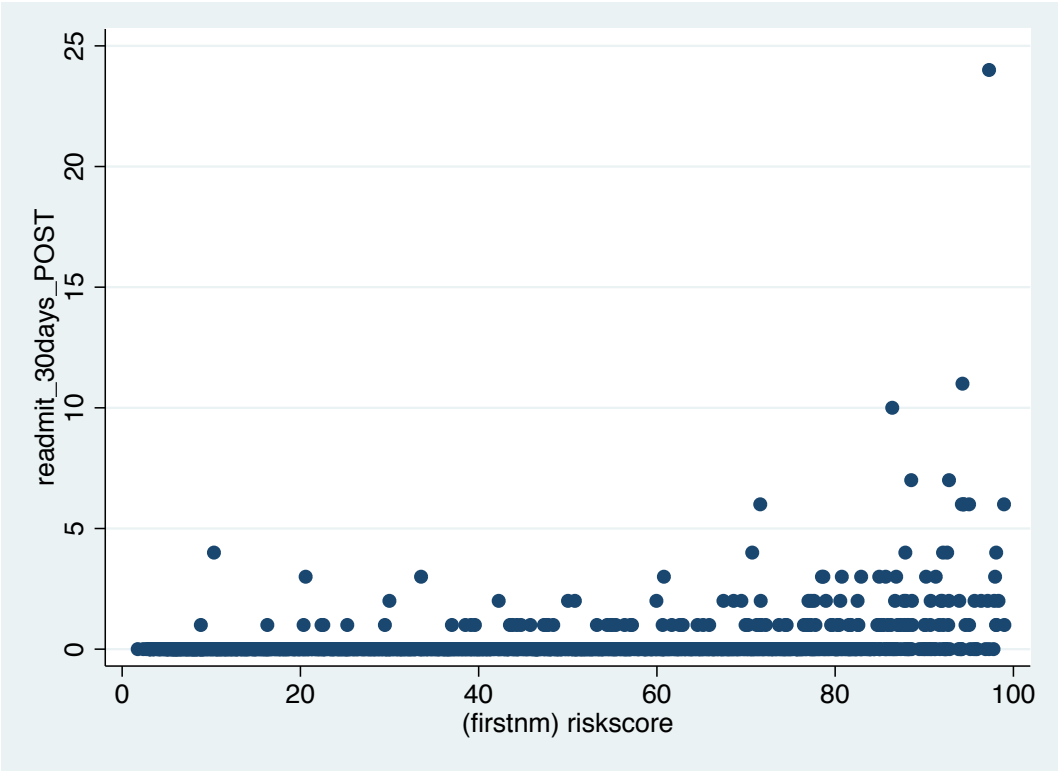
How well does actual risk score predict number of POST-intervention admissions?




in_el_POST
0
10
20
30
40
0
20
40
60
80
100
(firstnm) riskscore
out_all_POST
0
20
40
60
80
100
0
20
40
60
80
100
(firstnm) riskscore


AE_POST
0
10
20
30
40
50
0
20
40
60
80
100
(firstnm) riskscore
totalcost_POST
0
20000
40000
60000
80000
0
20
40
60
80
100
(firstnm) riskscore


acsc_total_POST
0
1
2
3
4
5
0
20
40
60
80
100
(firstnm) riskscore
LOS_POST
0
100
200
300
400
0
20
40
60
80
100
(firstnm) riskscore


readmit_30days_POST
0
5
10
15
20
25
0
20
40
60
80
100
(firstnm) riskscore

Risk Score DDD Estimates

Outcome	**Adjusted[+] intervention effect (95% CI) (difference per patient per month)**		**Effect size[*]**
	Count (nbreg) model		
Primary outcomes:			
Inpatient non-electives	0.0208 (0.0083 to 0.0333)	#	0.09
Inpatient electives	-0.0009 (-0.0174 to 0.0156)		-0.00
Outpatient admissions	0.0943 (-0.0042 to 0.1927)		0.08
A&E visits	0.0363 (0.0128 to 0.0598)	#	0.09
ACSCs	0.0020 (-0.0029 to 0.0069)		0.02
Re-admissions (30 days)	0.0059 (-0.0004 to 0.0123)		0.05
Secondary outcomes:			
Total cost of 2[o] care services (£)[~]	19.2166 (-20.5574 to 58.9907)		0.02
Length of stay (days)	0.3071 (0.0592 to 0.5549)	#	0.06

+: Adjusted for: age, cumulative multimorbidity, imd domains (excluding health), practice- and time- fixed-effects. Marginal effects on *PICT x Post* reported.
*: Standardised mean difference
#: significant at $p<0.05$
nbreg: negative binomial regression
~: zero-inflated negative binomial models based on admission events
n= 224,898 observations; 4098 individuals (period September 2010 to March 2015)

References

1. Wennberg D, Siegel M, Darin B, Filipova N, Russell R, Kenney L, et al. Combined predictive model: final report and technical documentation. *London: Department of Health, The Kings Fund, NYU, Health Dialogue* 2006.
2. Hill J. Reducing bias in treatment effect estimation in observational studies suffering from missing data. 2004.
3. StataCorp. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP., 2013.
4. Leuven E, Sianesi B. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components* 2014.
5. Becker SO, Ichino A. Estimation of average treatment effects based on propensity scores. *The stata journal* 2002;2(4):358-77.
6. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*: L. Erlbaum Associates, 1988.
7. Baines DL, Whynes DK. Selection bias in GP fundholding. *Health economics* 1996;5(2):129-40.
8. Propper C, Croxson B, Shearer A. Waiting times for hospital admissions: the impact of GP fundholding. *Journal of health economics* 2002;21(2):227-52.
9. HSCIC. Quality and Outcomes Framework (QOF) - 2013-14, 2014.
10. Koller D, Schon G, Schafer I, Glaeske G, van den Bussche H, Hansen H. Multimorbidity and long-term care dependency--a five-year follow-up. *BMC geriatrics* 2014;14:70.
11. World Health Organization. International statistical classification of diseases and health related problems (The) ICD-10. World Health Organization, 2004.
12. Tonelli M, Wiebe N, Fortin M, Guthrie B, Hemmelgarn B, James M, et al. Methods for identifying 30 chronic conditions: application to administrative data. *BMC Medical Informatics and Decision Making* 2015;15(1):31.
13. Fortin M, Stewart M, Poitras M-E, Almirall J, Maddocks H. A Systematic Review of Prevalence Studies on Multimorbidity: Toward a More Uniform Methodology. *The Annals of Family Medicine* 2012;10(2):142-51.
14. HSCIC. Numbers of Patients Registered at a GP Practice - January 2014, 2014.
15. StataCorp L. *Stata Time-series: Reference Manual*: StataCorp LP, 2007.
16. Ansari Z, Laditka JN, Laditka SB. Access to health care and hospitalization for ambulatory care sensitive conditions. *Medical care research and review : MCRR* 2006;63(6):719-41.
17. Caminal J, Starfield B, Sanchez E, Casanova C, Morales M. The role of primary care in preventing ambulatory care sensitive conditions. *European journal of public health* 2004;14(3):246-51.
18. Parchman ML, Culler S. Primary care physicians and avoidable hospitalizations. *The Journal of family practice* 1994;39(2):123-8.
19. NHS England. Everyone Counts: Planning for Patients 2013/14 Technical Definitions: NHS Commissioning Board, 2012:1-103.
20. Ipsos MORI. GP Patient Survey – Technical Annex 2013-2014 annual report: Social Research Institute, 2014.

21. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol* 2002;155(2):176-84.

Appendix 4: Chapter 7 Supplementary files

ADDITIONAL FILE

MULTIMORBIDITY MEASURES

The multimorbidity measures (except Charlson index which has its own pre-specified list of condition codes [1]) were prepared from the previously recorded (period June 2006 to March 2015) inpatient admissions, for 20 chronic conditions recorded in the NHS Quality and Outcomes Framework (QOF) (see below for list of conditions and icd-10 codes) [2-4].

1. Asthma

J45-J47

2. Atrial fibrillation

I48

3. Cancer

C00–C14, C15–C26, C30–C39,

C40–C41, C43–C44, C45–C49,

C50, C51–C58, C60–C63,

C64–C68, C69–C72, C73–C75,

C81–C96, C76–C80, C97,

D00–D09, D37–D48 (Koller et al 2014)

4. Chronic kidney disease

N18

5. Coronary heart disease

I20-I25

6. COPD

J40-J44

7. Dementia

F00-F03

8. Depression

F32-F33

9. Diabetes mellitus

E10-E14

10. Epilepsy

G40-G41

11. Heart failure

I50

12. Hypertension

I10-I15

13. Hypothyroidism

E00-E03

14. Learning disability

F80-F89

15. Mental health (schizophrenia, bipolar affective disorder and other psychoses)

F20-F29, F31, F34-F39

16. Obesity

E66

17. Osteoporosis

M80-M82

18. Atherosclerosis/Peripheral arterial oclusive disease (PAOD)

I65–I66, I67.2, I70, I73.9

19. Rheumatoid arthritis

M05-M06

20. Cerebral ischemia/chronic stroke

I60–I64, I69, G45

RESULTS TABLE

Outcome	Mental-physical	3 or more conditions	Discordant versus other
	Count (nbreg) model. Adjusted[+] intervention effect (95% CI) (difference per patient per month)		
Primary outcomes:			
Inpatient non-electives	0.0122 (0.0018 to 0.0228)#	0.0129 (-0.0027 to 0.0286)	0.0078 (-0.0075 to 0.0232)
Inpatient electives	-0.0109 (-0.0409 to 0.0191)	0.0023 (-0.0166 to 0.0212)	-0.0008 (-0.0156 to 0.0139)
Outpatient visits	0.0169 (-0.0820 to 0.1158)	0.0393 (-0.0717 to 0.1503)	0.0900 (0.0040 to 0.1760)#
A&E visits	0.0200 (0.0009 to 0.0391)#	0.0186 (-0.0031 to 0.0402)	0.0187 (-0.0058 to 0.0432)
ACSCs	-0.0006 (-0.0050 to 0.0038)	-0.0081 (-0.0199 to 0.0036)	0.0019 (-0.0046 to 0.0084)
Re-admissions (30 days)	0.0006 (-0.0042 to 0.0053)	0.0024 (-0.0065 to 0.0113)	-0.0011 (-0.0074 to 0.0053)
Secondary outcomes:			
Total cost of 2[o] care services (£)~	13.8866 (-35.6720 to 63.4452)	36.1449 (-2.2763 to 74.5661)	24.9248 (-9.3942 to 59.2437)
Length of stay (days)	0.0612 (-0.2395 to 0.3619)	0.1814 (-0.1933 to 0.5560)	0.0496 (-0.2348 to 0.3340)

Outcome	**Lit Cluster 1**	**Lit Cluster 2**	**Lit Cluster 3**
	Count (nbreg) model. Adjusted⁺ intervention effect (95% CI) (difference per patient per month)		
Primary outcomes:			
Inpatient non-electives	-0.0284 (-0.0565 to -0.0003)#	0.0014 (-0.0190 to 0.0219)	0.8171 (0.7376 to 0.8966)#
Inpatient electives	0.0071 (-0.0121 to 0.0263)	-0.0162 (-0.0431 to 0.0106)	-0.0109 (-0.0670 to 0.0452)
Outpatient visits	-0.1067 (-0.2334 to 0.0200)	-0.0043 (-0.1715 to 0.1628)	-0.0397 (-0.3970 to 0.3175)
A&E visits	-0.0353 (-0.0796 to 0.0090)	-0.0092 (-0.0431 to 0.0248)	0.0387 (-0.1392 to 0.2166)
ACSCs	-0.0138 (-0.0317 to 0.0042)	0.0075 (-0.0060 to 0.0210)	0.0001 (-0.0054 to 0.0057)
Re-admissions (30 days)	-0.0062 (-0.0277 to 0.0152)	0.0078 (-0.0034 to 0.0190)	0.1805 (0.1624 to 0.1986)#
Secondary outcomes:			
Total cost of 2º care services (£)~	14.4987 (-60.8355 to 89.8328)	23.1385 (-46.6572 to 92.9341)	83.3950 (-146.0357 to 312.8258)
Length of stay (days)	-0.0993 (-0.7057 to 0.5072)	0.2234 (-0.1919 to 0.6387)	1.9649 (0.2253 to 3.7046)#

Outcome	**Charlson Index**
	Count (nbreg) model. Adjusted[+] intervention effect (95% CI) (difference per patient per month)
Primary outcomes:	
Inpatient non-electives	-0.0085 (-0.0227 to 0.0056)
Inpatient electives	0.0011 (-0.0197 to 0.0219)
Outpatient visits	-0.0340 (-0.1398 to 0.0718)
A&E visits	-0.0066 (-0.0303 to 0.0172)
ACSCs	-0.0059 (-0.0099 to -0.0019)#
Re-admissions (30 days)	-0.0058 (-0.0106 to -0.0010)#
Secondary outcomes:	
Total cost of 2º care services (£)~	-15.0060 (-73.2848 to 43.2728)
Length of stay (days)	0.0081 (-0.2412 to 0.2573)

+: adjusted for: age, imd domains (excluding health), practice- and time- fixed-effects. Marginal effects on *PICT x Post* reported.
*: standardised mean difference
#: significant at p<0.05
nbreg: negative binomial regression
~: zero-inflated negative binomial model based on admission events
n= 224,898 observations; 4098 individuals (period September 2010 to March 2015)

REFERENCES

1. Stagg, V., *CHARLSON: Stata module to calculate Charlson index of comorbidity.* Statistical Software Components, 2006.
2. HSCIC. *Quality and Outcomes Framework (QOF) - 2013-14*. 2014 25/08/2015]; Available from: http://www.hscic.gov.uk/catalogue/PUB15751.
3. Koller, D., et al., *Multimorbidity and long-term care dependency--a five-year follow-up.* BMC Geriatr, 2014. **14**: p. 70.
4. World Health Organization, *International statistical classification of diseases and health related problems (The) ICD-10*. 2004, World Health Organization.