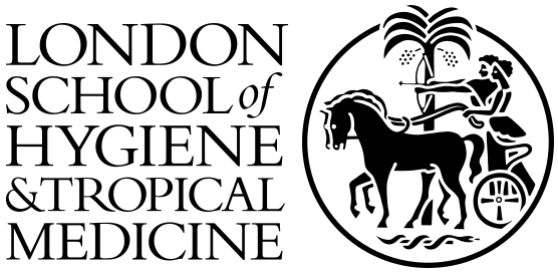


Gender inequality in healthy ageing: a study of the English older population over a decade

BENEDETTA PONGIGLIONE

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of London

March 2017

Department of Medical Statistics
Faculty of Epidemiology and Population Health
LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by Economic and Social Research Council

Declaration

I, Benedetta Pongiglione, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis investigates gender inequalities in healthy ageing among the older English population, using data from the English Longitudinal Study of Ageing. The research aims were achieved by completing the following steps: (i) healthy ageing was intended as advancing to the later stages of the life course without disability; where disability was first theoretically conceptualized and then measured using severity levels that were identified empirically; (ii) gender inequalities in healthy ageing were assessed by studying whether the association between disability and mortality observed over the course of a decade differed between men and women; and (iii) disability and mortality were combined into a summary measure of population health -disability-free life expectancy- in order to estimate how expectancies of healthy life have changed over a decade across the two genders. The work is structured in four papers, denoted Research Paper I-IV.

Research Paper I, a systematic literature review of studies analysing inequalities in health expectancy among the older population, inspired the direction taken by this thesis, as it identified gaps and open questions to be addressed to aid the understanding of the dynamics of healthy ageing. Research Paper II attempted to develop an approach to answer some of these questions. First, a solid and theoretically grounded definition of disability was proposed, based on the WHO's International Classification of Functioning Disability and Health (ICF), and in contrast to the data-dependent (and therefore heterogeneous) measures used in the literature. Then, using this definition, explanations of the gender paradox in health and mortality were attempted by analysing whether the association of disability with mortality differed between women and men over the period for which data were available (2002-2012). In Research Paper III the definition of disability elaborated in Research Paper II was used to foster and advance the debate on the usefulness and relevance of adopting a finer categorization of disability, and discuss why it is important to go beyond a binary classification, and to identify the appropriate number of disability levels

that is most useful for research purposes. Based on these conclusions, the final aim of this thesis was accomplished in Research Paper IV, which studied the trends in disability-free life expectancy in England over the last decade, comparing the changes experienced by men and women at each severity level of disability.

The collective findings of this thesis highlight the importance of defining disability in a consistent and comprehensive way as well as considering different severity levels. This work provides robust empirical evidence for theories of population health change over a decade in the English setting, with gender differences in healthy ageing, and directions of population health changes, found to vary across disability levels.

Acknowledgements

I have been looking forward to writing these acknowledgments. First of all to say thank you to my supervisors, Bianca De Stavola and George Ploubidis. You made my PhD a fantastic journey. Thank you George for the enthusiasm with which you accepted to be my supervisor and that you continued to show throughout all my PhD, for all your advice and support when I had to take important decisions and for encouraging my ideas. Grazie Bianca for taking me under your supervision at the end of my first year and being such an incredible support ever since, for teaching me so much and for understanding so well when an encouragement was especially needed. Thank you both for teaching me that research sometimes is hard but it is still fun!

I am also very much grateful to Hannah Kuper, co-author and member of my advisory committee, from whom I have learnt a lot, especially to spend time to understand the large picture before using numbers to make sense of it.

I want to say thank you to Ross Mac Millan which was the first to encourage me to consider LSHTM as a place for doing a PhD, when I was writing my master thesis. I am extremely thankful to Rosanna Tarricone, who shared with me her PhD experience at the school before I started and instilled lot of enthusiasm and optimism for taking on this challenge. Thank you for your support and curiosity for my research, and for your efforts to find bridges to communicate from different fields and for broadening my perspectives.

These years would have not been so incredible without all the special persons that I have met and met again in London. Thank you to my petit and bigger comité, Angie, David, Ioana, Donna, Jenny, Sabah, Sol and Sami, for being my family at the school and in London. Thanks in particular to Sabah and Angie for sharing and transmitting the passion for what you do and what you believe and for being such a great friends. I have also been extremely lucky to spend these years in LG20, where I have shared my days with extraordinary fellows. Thanks in particular to Lenka, for explaining me so many things when I was new to the

school, for proofreading my upgrading document and for all your wise advice. Thank you also to Sophie, Laura, Jenny, Chris and Elaine, for the way to get tea together and for making the office such a homey place.

Thank you to all my friends from the Community of Sant'Egidio, for bringing perspective into my life and reminding me what is important and for always being there. I want to say thanks in particular to Emanuela, Carmelo, Silvia, Francesco, Gaia and Irene.

Thank you to all of those I have met in London for the first time despite coming from the same place. Thanks to Elena, Pierpaolo and Cristina for giving me a place to stay when I first arrived and making me feel home, and thanks to Federico for our lunches at Russel square and for all your "if you need anything, just let me know". I am also grateful for the friends that I have met in London again after years, for the good time spent together, the laughter, the chat and friendship, thanks in particular to Francesca and Camilla.

Infine, grazie alla mia famiglia, per supportarmi e sopportarmi, e per credere in me sempre, anche quando io non sono brava a farlo. Per farmi sforzare a non scoprire l'acqua calda e migliorare quello che amo fare, e anche per dirmi ogni tanto di smetterla e andare a fare qualcosa di bello. Grazie ai miei genitori per il loro sostegno incondizionato, che più di tutto mi ha aiutato a raggiungere i miei obiettivi ed essere felice delle scelte che ho fatto. Grazie a mio fratello Filo, per la sua presenza costante, per aver ascoltato quasi tutte le presentazioni che ho fatto, per prendermi sul serio quando serve e prendermi in giro tutto il resto del tempo.

Grazie ai miei amici di sempre, Giorgia, Margherita, Daniele, Caterina, Valentina, Claudia, per farmi sentire a casa come se non fossi mai partita ogni volta che torno e per esserci sempre anche quando siamo distanti. E grazie a chi è arrivato mentre scrivevo le conclusioni di questa tesi, e spero ci sarà anche per i prossimi inizi.

Acronyms and Abbreviations

ADL Activity of Daily Living

AQM Advanced Quantitative Methods

BCS British Cohort Study

BMI Body Mass Index

CAPI Computer-Assisted Personal Interview

CCA Complete Case Analysis

CFAS Cognitive Function and Ageing Studies

DALY Disability-Adjusted Life Years

DFLE Disability-Free Life Expectancy

DLE Disability Life Expectancy

DTSA Discrete-Time Survival Analysis

ELSA English Longitudinal Study of Ageing

ESRC Economic and Social Research Council

EHLEIS European Health and Life Expectancy Information System

FIML Full Information Maximum Likelihood

HLE Healthy Life Expectancy

HLY Healthy Life Years

HRS Health Retirement Study

HSE Health Surveys for England

IADL Instrumental Activity of Daily Living

ICED International Centre for Evidence in Disability

ICF International Classification of Functioning Disability and Health

ICIDH International Classification of Impairments Disabilities and Handicaps

LCA Latent Class Analysis

LMIC Low Middle Income Country

LRT Likelihood Ratio Test

LSHTM London School of Hygiene and Tropical Medicine

MAR Missingness At Random

MCAR Missingness Completely At Random

MCS Millennium Cohort Study

NCDS National Child Development Study

NHSCR National Health Service Central Register

NIA National Institute on Aging

NMAR Non Missingness At Random

NSHD National Survey of Health and Development Cohort

ONS Office for National Statistics

REVES Réseau Espérance de Vie en Santé

RTSG Research Training Support Grant

SC Self-Completion

SEM Structural Equation Model

SHARE Survey of Health, Ageing and Retirement

SMPH Summary Measures of Population Health

SMR Standardised Mortality Ratio

TLE Total Life Expectancy

WG Washington Group

WLSMV Weighted least squares means and variance

YLD Years Lost to Disability

YLL Years of Life Lost

Contents

List of Figures

List of Tables

Part I

Background

Chapter 1

Introduction

Population ageing is *unprecedented*, *pervasive*, *enduring* and *has profound implications* [1].

Population ageing is the phenomenon observed when the median age of a population increases due to rising life expectancy and/or declining fertility rates. Worldwide population ageing over the past decades has mainly been driven by an increase in life expectancy, and the most significant trend now affecting longevity is the decline in mortality among the elderly [2].

The pace and magnitude of population ageing are *unprecedented* and this phenomenon is expected to grow even faster in the next century. The number of people aged over 60 years in the world population today is almost 1 in 10 and by 2050 it is forecasted to be 1 in 5; the number of people aged 80 years or over, the 'oldest-old', is increasing even faster than the number of older persons overall. Projections indicate that in 2050 the oldest-old will triple in number since 2015, when there were 125 million [3]. Forecast for England and Wales recently produced [4] have shown that national life expectancy in 2030 is expected to reach 85.7 (95 % credible interval 84.2-87.4) years for men and 87.6 (86.7-88.9) years for women, reducing the female advantage, and that most of the gains in longevity will be in those older than 65 years of age.

Population ageing is *pervasive*, it is a global phenomenon, affecting every man and woman worldwide. Countries, however, are at very different stages of the process, and the pace of change differs greatly. In high-income countries the phenomenon has begun few decades ago, while low-middle income settings have started the process later and will have

less time to adjust. [5].

Population ageing is *enduring* in the sense that, since the 1950s, the proportion of older persons has been rising steadily, passing from 8 percent in 1950 to over 12 percent in 2015, and is expected to reach over 21 percent in 2050 [3]. There is general agreement over the fact that the maximum life span -the longest number of years a human being has lived- has increased spectacularly in the past century. Scholarly opinion diverges, however, as to whether these increases will continue or whether human longevity is approaching its limit [6]. Two main views have emerged with regard to this matter. The supporters of the "limited-lifespan paradigm" maintain that the human body is biologically not designed for extended survival and life expectancy will reach its limit [7, 8, 9]. The other view, known as the "mortality-reduction" paradigm, does not see an end in the rising trends in life expectancy and maintains that mortality at older ages will continue to decline [10, 11, 12, 13].

Population ageing has *profound implications* for society. Ageing is an individual experience, as well as a collective phenomenon. As an individual experience, being an 'older person' is shaped by events and experiences throughout lifetime, including for example education, work and retirement, and health and lifestyle choices. Ageing, however, is a process that is also intimately intertwined with other people, institutions, and structures. It implies facing new health needs and its impact on a progressively larger proportion of the population will lead to their requiring assistance from other individuals and society, likely for long periods of time. The financial burden on the health care system is one of the main challenges currently faced by governments in many countries. Another consequence of ageing relates to the pension system and retirement age. How long people will live after retirement, and for how long they will be able to function autonomously, is crucial to the withstanding of pension systems and for setting sustainable retirement ages. This is currently one of the most intense points of debate. The greying of populations also poses social challenges on how to promote successful ageing, and how to maintain high quality of life for individuals as they age, as well as for those around them.

Theories of population health change

Whether the years of life gained through increased longevity are spent in good or poor health is a question which has been asked since the rise in life expectancy has shown to be consistent and continuous, starting from the 1960s [14]. The answers have mainly crystallized around three alternative theories: *expansion of morbidity*, *compression of morbidity* and *dynamic equilibrium*. They are graphically represented in figure 1.1, borrowed from chapter 3 of The Health of Population book. [15, p. 73]

The **expansion** of morbidity hypothesis was put forward by Gruenberg [16] in 1977 and Kramer [17] in 1980. According to this theory, increases in life expectancy are driven mainly by improvements in medical care and secondary prevention strategies that avert fatal outcome from degenerative diseases, whilst the epidemiology of these conditions remains more or less the same. Mortality rates decline because people survive chronic diseases, but in turn they live a longer part of their life with the condition, i.e. morbidity expands together with longevity. The implicit assumption of this hypothesis is that advancements in modern medicine reduce fatal incidence of degenerative diseases, but the progress of pathologies will remain resistant to and unaffected by improvements in public health.

The **compression** of morbidity hypothesis was first proposed by Fries, in 1980 [18]. Opposite to the expansion theory, the hypothesis of compression of morbidity maintains that the causes that have led to decreased mortality would also be linked to a lower incidence of chronic diseases and delays in onset of chronic diseases and disability. The higher age at onset of conditions implies that the time lived with disease would be compressed into a shorter period at the end of life, if the expansion is greater than the increase in longevity. One of the underlying assumptions in the compression hypothesis proposed by Fries is that life expectancy has a limiting biological maximum [19], and delaying onset of diseases or disability certainly leads to reducing the period of life spent with the conditions. Such an assumption is not exempt from objections, especially considering current mortality trends. It is however possible for compression of morbidity to occur if Healthy Life Expectancy (HLE) increases faster than life expectancy, even if life expectancy continues to grow.

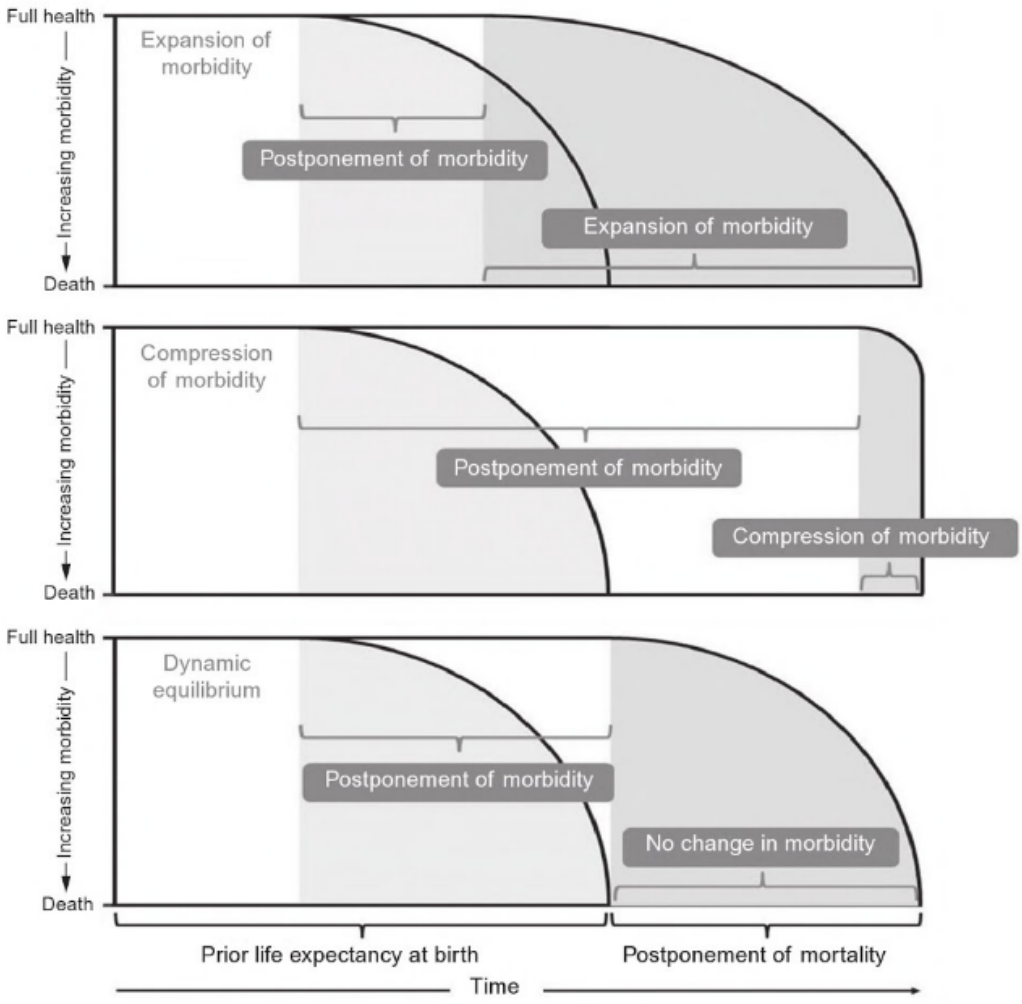
The hypothesis of **dynamic equilibrium** of morbidity lies in between expansion and compression theories. It was first advanced by Manton in 1982 [20], which offered an

alternative view of the processes behind the postponement of death from chronic disease. According to Manton, the severity and progression of chronic diseases have changed, with progression of disease halted at earlier stages, and the population being affected by disease but less so from its consequences, such as disability and imminent death. The slowdown in the rate of progression of disease would lead to (i) an increase in overall prevalence due mostly to increases in the prevalence of mild and less disabling disease states, (ii) largely stable rates of severe disease.

Empirical evidence has not yet clearly pointed at any of these theories as the one to be preferred and the debate is still lively and open. For example, the hypothesis of expansion of disability was supported for the US from 1964 to 1974 in a paper examining the trend in the health of Americans, where overall life expectancy has increased over this decade, with almost all of this increase being years of disability [21, 22]. The compression of morbidity hypothesis was supported in some studies investigating trends in physical functioning and activity limitations in the US during the 1980s, 1990s and the early twentieth century [23, 24]. Other studies set in the US showed declines in the prevalence of Activity of Daily Living (ADL) and Instrumental Activity of Daily Living (IADL)) [25, 26, 27, 28, 29]. Compression was also observed in Austria [30] and Spain [31]. The dynamic equilibrium hypothesis was supported in the study of Graham and colleagues in New Zealand [32]. A study set in France [33] showed that patterns differed depending on the disability measure underlying the health indicator.

The lack of predominance of any of the aforementioned theories and the coexistence of theories worldwide [34] and in particular in the UK, as described below in section 1.2, along with the need of having a clear understanding on the current and future development of the health state of longevity have motivated the development of the work of this thesis. This thesis seeks to contribute to the debate on population health changes and to offer new evidence for the English setting. Therefore, the three theories of population health change are central references in this thesis, and are mentioned repeatedly through chapters while they are, albeit succinctly, described again in Research Papers I, III and IV.

Figure 1.1: Schematic representation of survival curves for the three main competing hypotheses of population health change.



Source: The Health of Populations: Beyond Medicine [15, p. 73]

1.1 Aims and objectives

This thesis investigates gender inequalities in healthy ageing among the older English population, over the past decade. The general research aim is achieved through a number of consequent steps: (i) healthy ageing is intended as advancing to the later stages of the life course without disability; where disability is first theoretically conceptualised, then measured accordingly, and finally severity levels are identified empirically; (ii) gender inequalities in healthy ageing are assessed by studying whether the association between disability and mortality observed over the course of a decade differ between men and women; and (iii) each level of disability is combined with mortality into a Summary Measures of Population Health (SMPH) -Disability-Free Life Expectancy (DFLE)- in order to estimate how expectancies of healthy life have changed over a decade in England across the two genders.

The thesis is structured in four papers, denoted Research Paper I-IV, and the research aims are achieved by targeting a number of linked objectives. Each research paper addresses one or more of the objectives. They are summarised in table 1.1, with the corresponding Research Paper and page of the thesis listed next to them.

1.2 Research context

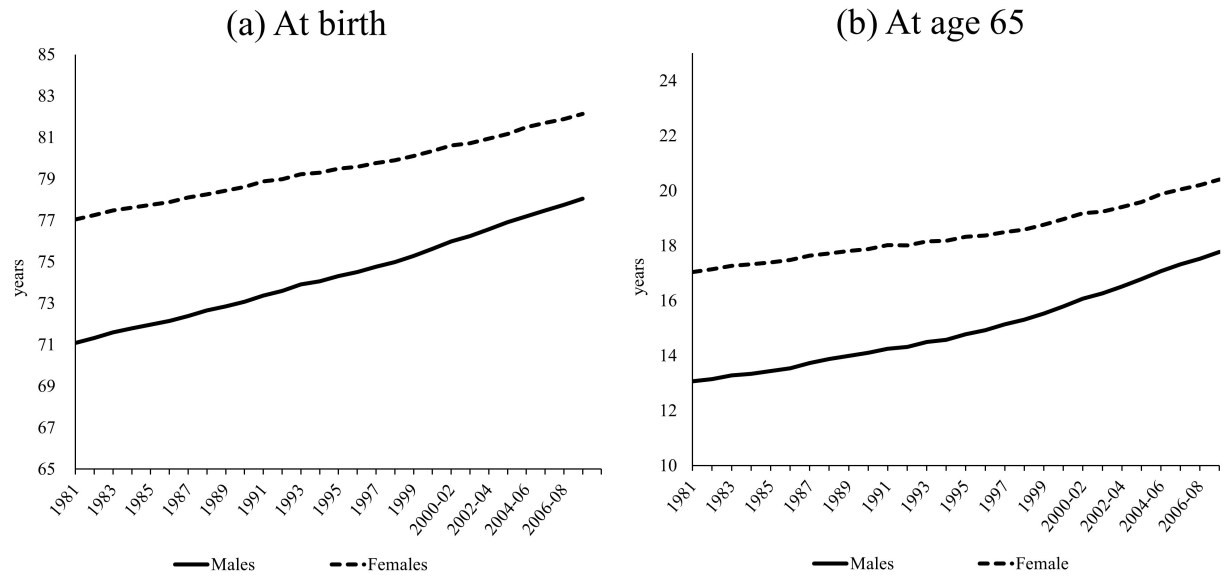
This thesis uses data from the English Longitudinal Study of Ageing (ELSA) (described in details in Chapter Four) which is set in England. With regards to the aims of the study, England represents an interesting case because, as in all developed countries, over the past decades its population has been experiencing a steady increase in Total Life Expectancy (TLE). Life tables, based on the population estimates and deaths provided by the Office for National Statistics (ONS) show that over the past three decades TLE at birth has continuously increased, rising from 71.1 years in 1981 to 78 years in 2007/09 (+9.8%) for men and from 77 years to 82.1 years for women (+6.6%). This increase in TLE has been more dramatic in the older population: from 1981 to 2007/09, TLE of men aged 65 has risen from 13.1 years to 17.8 years (+35%) and from 17 to 20.4 years for women (+19.8%) (Figure 1.2).

Time-series of life expectancy and health expectancy have been available since the

Table 1.1: Research objectives and Research Papers where objectives were addressed.

n	Objectives	Research Paper	Page
1	To systematically review the literature that investigates socioeconomic and demographic inequalities in health and mortality using summary measures of population health.	Research Paper I	16
2	To theoretically conceptualize and empirically measure disability.	Research Paper II	69
3	To explore gender inequalities in the effect of disability measured at baseline on mortality observed over a decade.	Research Paper II	69
4	To identify an optimal number of severity levels of disability based on empirical evidence.	Research Paper III	103
5	To combine each level of disability with mortality in a single indicator of population health, DFLE, and use it to estimate how expectancies of healthy life have changed over a decade across the two genders and across different levels of disability.	Research Paper IV	126
6	To explore possible demographic and behavioural factors explaining changes in DFLE observed over a decade.	Research Paper IV	126

Figure 1.2: Trends in life expectancy from 1981 to 2009, by gender



Source: Office for National Statistics

1980s for the UK. However, no clear pattern has been found. Some research seems to support the pessimistic view of expansion of morbidity. According to the study of Bebbington [35], in England and Wales since 1976 there has been a steady improvement in life expectancy without disability, but the rate of improvement has not been more than the rate of increase in TLE, hence most of the increase in TLE was with years with chronic disability. However, an improvement, both in TLE and in DFLE, was found at older ages. Another study monitoring HLE from 1980 to 1996 in Great Britain found that both TLE and HLE increased, but the latter did not increase by as much as the former [36]. As a result, both men and women, both at birth and at age 65, were expected to live most of the gained years of life in poor health or with a limiting long-standing illness. A similar work [37], investigating the same period but only in England and Wales, found that HLE at age 65 had increased at a similar rate as TLE, with the result that the proportion of remaining years of life free of disability was almost stable in this period. Similarly, Bone et al. [38] showed that life expectancy without severe disability between 1976 and 1994, roughly progressed in parallel with TLE in males at 65, which means that the number of years lived with severe disability stagnated and even showed a decrease. More recently, a review of the UK government's Foresight Future of an Ageing Population project [39], using

the latest routine national data made available by the ONS, has shown that increases in DFLE and HLE in the UK are not keeping pace with gains in TLE at later ages, and real health improvements are experienced by the younger rather than the older population. Finally, a study set in England [40], using health measures available in the Cognitive Function and Ageing Studies (CFAS), has produced various health expectancies and shown evidence supporting an absolute compression of cognitive impairment, a relative compression of self-perceived health, and dynamic equilibrium of disability, over a period of twenty years, from 1991 to 2011. The increase in number of years with any disability was found to be higher for years with mild disability than for years with moderate or severe disability. To ease the understanding of the evolution of life expectancy and health expectancy over the past three decades, studies are summarised in table 1.2 and sorted by period of investigation. However, the comparison and combination of findings are limited due to the fact that the studies referred to different age groups and disability was measured in different ways.

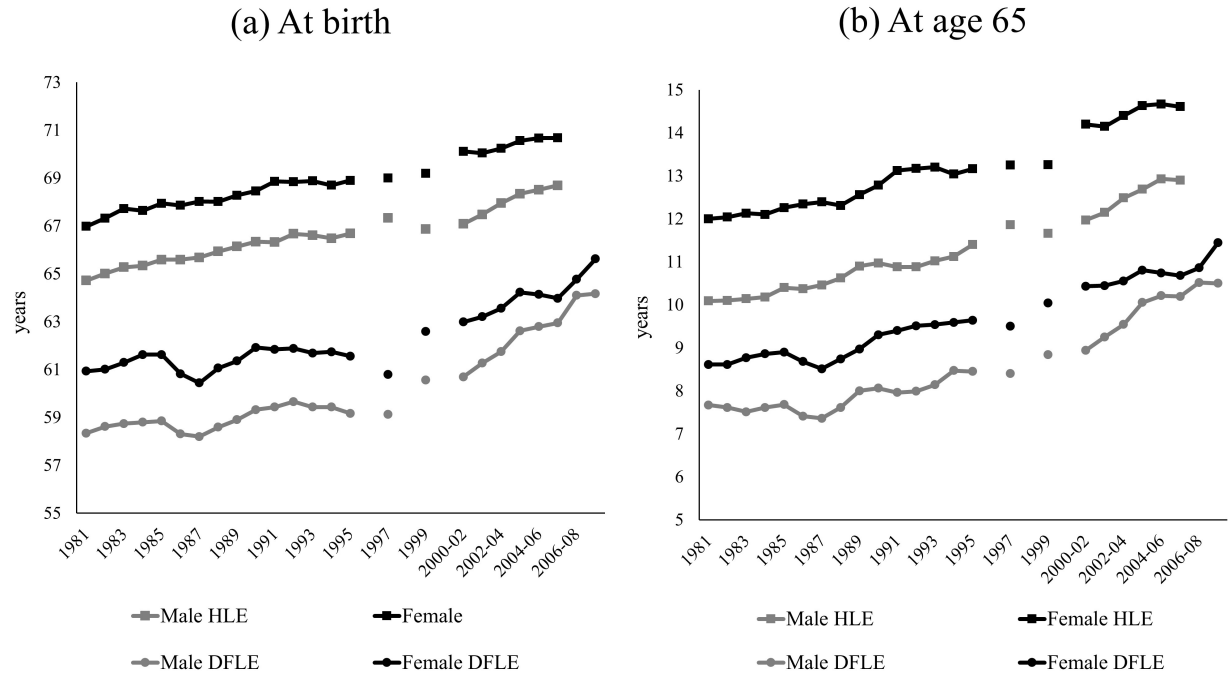
Lastly, for explanatory purposes, estimates of HLE in good self-rated health and DFLE (defined in terms of long-standing illness) produced by the ONS are illustrated for the period 1981-2008 (figure 1.3). Over the past three decades, both indicators have increased over time (i.e. more years of life in good health and without disability). This was observed at birth and at age 65 years, among both women and men. On the other hand, the trend of DFLE and HLE in proportional terms with TLE from 1980 to 2009 were rather fluctuating (figure 1.4). In the last decade, in particular, the proportions of DFLE have slightly increased (lines with round marker), suggesting a reduction of the burden of disability. But this was not the case for HLE, neither at birth nor at age 65.

As suggested, whether morbidity is expanding or compressing remains unclear, and in part this seems to be dependent on how health and disability are measured.

1.3 Structure of the thesis

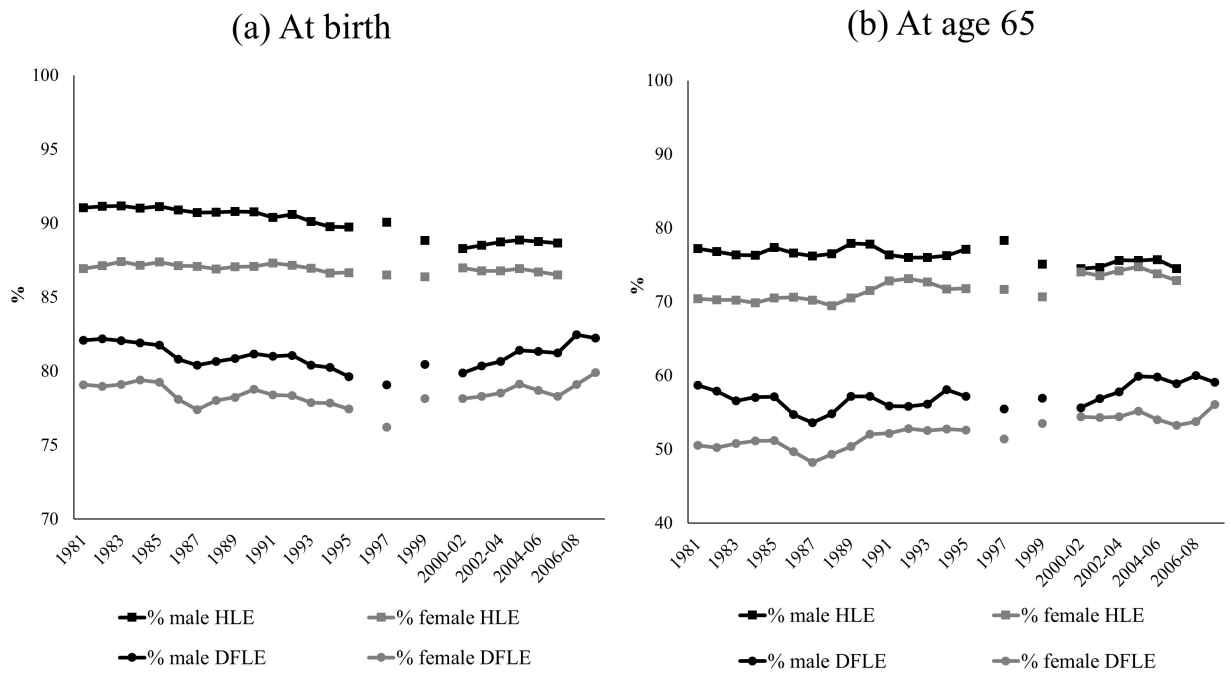
This thesis is presented in a series of research papers. It includes four published or submitted for review academic papers (Research Papers I-IV), and four additional chapters. Each research paper is preceded by a preamble, which outlines the rationale for the study, contextualises the research within the thesis, and links the article to the findings and ma-

Figure 1.3: Trends in HLE and DFLE from 1981 to 2009, by gender



Source: Office for National Statistics

Figure 1.4: Trends in proportional HLE and DFLE from 1980 to 2009, by gender



Source: Office for National Statistics

Table 1.2: Collection of studies on trends in health expectancy

Study	Period	Where	Type of HE	Age group	Conclusions
Bebbington [35](1988)	1976, 1981, 1985	England and Wales	LLI	All ages	Expansion
Bone et al. [38](1995)	1976-1994	UK	Handicap	Birth and 65 years	Dynamic equilibrium
Bebbington and Darton [37](1996)	1980-94	England and Wales	LLI, ADL	All ages	Dynamic equilibrium
Kelly et al. [36](2000)	1980-96	Great Britain	SRH and LLI	Birth and 65 years	Expansion
Jagger et al. [40](2016)	1991-2011	England	ADL and IADL	65 years +	Dynamic equilibrium
Jagger [39](2015)	2000/02-2009/11	UK	SRH and LLI	Birth, 65 and 85 years	Expansion

LLI=limiting long-standing illness; SRH=self-rated health
Findings are general conclusions which not consider differences between age groups, genders and severity levels of health indicators.

terial presented in preceding chapters. At the end of each research paper, a paragraph synthesising the main findings and their relevance for the aims and objectives of the thesis is also presented. Each paper was prepared as a standalone article and therefore there is, inevitably, some repetition. The research papers already published or in press but published online (I, II and III) are presented in the format of the journal where they have been published. One research paper (IV) is prepared for publication but not yet submitted, and is presented in the format of a regular chapter, and its content and structure are the same as planned to be submitted to the journal, but the numbering of tables, figures and appendix follows the order of the thesis, and relative references are reported at the end of the thesis along with all the other references.

Figure 1.5 provides a visual representation of the conceptual development of the work and how all the papers come together within the thesis. Some of the findings of the systematic literature review (Research Paper I, blue boxes) have motivated the development of the thesis. Gaps and controversial findings that emerged from the review have been addressed in Research Paper II (green boxes), Research Paper III (orange boxes) and Research Paper IV (yellow boxes).

The thesis is organised in three main parts, as described below.

Part I contains the research background and rationale of the thesis. It includes this

introductory chapter; Chapter Two, which consists of Research Paper I, a systematic literature review, published in PLOS ONE [41], that synthesizes studies on demographic and socioeconomic inequalities in health expectancy in the older population. Chapter Three presents the conceptual framework for disability and Chapter Four outlines the research setting, explains in details the design and content of the data used in the thesis and discusses the advantages and disadvantages of the ELSA with respect to the aims of the project.

Part II consists of three chapters including Research Paper II, III and IV, which hold the quantitative analyses of the thesis. Research Paper II (Chapter Five) is published in the European Journal of Epidemiology [42] and investigates the association between disability and mortality focusing on gender differences in this relationship. Research Paper III (Chapter Six) is in press in the Disability and Health Journal [43] and aims at identifying a disability gradient and detecting a meaningful number of disability levels. Research Paper IV (Chapter Seven) is planned to be submitted to the European Journal of Epidemiology, and explores DFLE over the first decade of the 21^{st} century, for each grade of disability as quantified in Research Paper III.

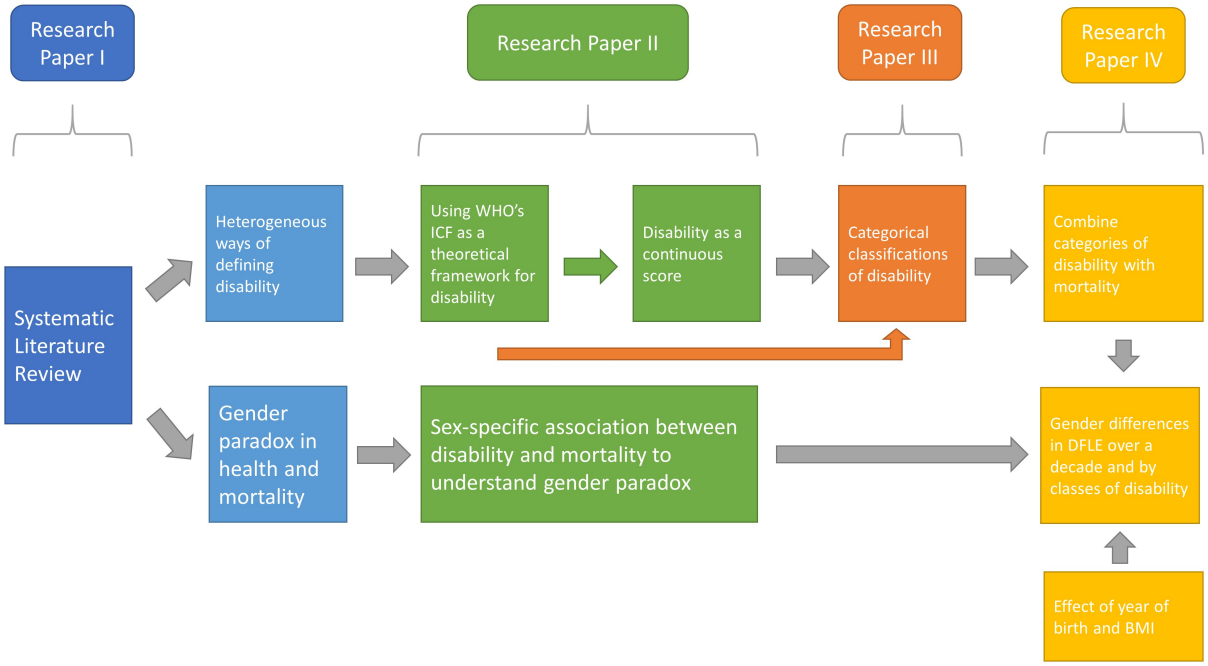
Part III includes the discussion of the work presented in this thesis (Chapter Eight). It brings together all the findings of the thesis and highlights its main contributions as well as its limitations, pointing at avenues for future research.

The last part of the thesis consists of the main appendix, which contains ancillary information, not necessary to understand the contents of papers and chapters, but providing further details on methodological and technical elements of the thesis. It also includes other disseminated work, such as posters and slides presented at international conferences.

1.4 Role of the candidate

The candidate contributed to conceptualizing all research questions and conducted all analyses in the systematic review and research papers. The candidate wrote the first draft of each paper, edited the draft based on co-authors' comments, and was primarily responsible for the final draft and submission of each article. She disseminated the PhD research work

Figure 1.5: Schematic of conceptual development of the thesis



at international conferences, in the form of oral presentations and posters.

1.5 Ethical clearance

Ethical approval for the research included within this thesis was granted by the London School of Hygiene and Tropical Medicine (LSHTM) ethics committee.

1.6 Funding

The candidate was awarded a three-year scholarship by the Economic and Social Research Council (ESRC) complemented with the Advanced Quantitative Methods (AQM) Enhanced Stipend, which covered research degree fees and an annual stipend. Most of the costs for the PhD research, including attendance of courses, participation at conferences to present PhD research and open access for the PhD publications were covered by Research Training Support Grant (RTSG) provided by the ESRC Bloomsbury DTC. Attendance at conferences was also funded by the LSHTM Faculty of Epidemiology and Population Health; as well as

the Society for Social Medicine provided financial support to participate at the 59th Annual Scientific Meeting, where Research Paper II was presented and a preliminary version of its abstract printed in the Journal of Epidemiology and Community Health[44].

Chapter 2

Systematic Literature Review

2.1 Preamble to literature review

Research Paper I consists in a systematic literature review where the first objective of this thesis is achieved:

- **Research Objective 1:** *To systematically review research that investigates socioeconomic and demographic inequalities in health and mortality using summary measures of population health.*

The motivation for conducting a systematic literature review was (i) to capture the current state of the art on inequalities in health expectancy among the older population; (ii) guide the development of the thesis in the light of findings of the selected studies, in order to investigate unexplored questions and provide new evidence to elucidate controversial results.

The content of the systematic literature review presented in this chapter allow to understand the dimensions of the phenomenon of demographic and socioeconomic inequalities in healthy ageing, identify which research questions have already been asked and answered, and acknowledge gaps and limitations in the existing literature.

2.2 Research Paper I

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Benedetta Pongiglione
Principal Supervisor	Bianca De Stavola
Thesis Title	Gender inequality in healthy ageing: a study of the English older population over a decade

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Plos One		
When was the work published?	June 26, 2015		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	

Student Signature: Bpongiglione **Date:** 1/11/2016

Supervisor Signature: B L De Stavola **Date:** 1/11/2016

PLOS ONE

RESEARCH ARTICLE

A Systematic Literature Review of Studies Analyzing Inequalities in Health Expectancy among the Older Population

Benedetta Pongiglione[1]*, Bianca L. De Stavola[1], George B. Ploubidis[2]

1 Medical Statistics Department, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, University of London, London, United Kingdom, **2** Centre for Longitudinal Studies, Institute of Education, London, United Kingdom

* benedetta.pongiglione@lshtm.ac.uk

OPEN ACCESS

Citation: Pongiglione B, De Stavola BL, Ploubidis GB (2015) A Systematic Literature Review of Studies Analyzing Inequalities in Health Expectancy among the Older Population. PLoS ONE 10(6): e0130747. doi:10.1371/journal.pone.0130747

Editor: Alessandra Marengoni, Karolinska Institutet, ITALY

Received: January 29, 2015

Accepted: May 23, 2015

Published: June 26, 2015

Data Availability Statement: All relevant data are within the paper, the reference list and its Supporting Information files.

Funding: All relevant data are within the paper, the reference list and its Supporting Information files. The authors have declared that no competing interests exist. This work has been supported by the Economic and Social Research Council (ESRC)+3 Studentship grant award (URL: http://www.esrc.ac.uk/). The grant number is ES/J500021/1. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Aim

To collect, organize and appraise evidence of socioeconomic and demographic inequalities in health and mortality among the older population using a summary measure of population health: Health Expectancy.

Methods

A systematic literature review was conducted. Literature published in English before November 2014 was searched via two possible sources: three electronic databases (Web of Science, Medline and Embase), and references in selected articles. The search was developed combining terms referring to outcome, exposure and participants, consisting in health expectancy, socioeconomic and demographic groups, and older population, respectively.

Results

Of 256 references identified, 90 met the inclusion criteria. Six references were added after searching reference lists of included articles. Thirty-three studies were focused only on gender-based inequalities; the remaining sixty-three considered gender along with other exposures. Findings were organized according to two leading perspectives: the type of inequalities considered and the health indicators chosen to measure health expectancy. Evidence of gender-based differentials and a socioeconomic gradient were found in all studies. A remarkable heterogeneity in the choice of health indicators used to compute health expectancy emerged as well as a non-uniform way of defining same health conditions.

Conclusions

Health expectancy is a useful and convenient measure to monitor and assess the quality of ageing and compare different groups and populations. This review showed a general

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

agreement of results obtained in different studies with regard to the existence of inequalities associated with several factors, such as gender, education, behaviors, and race. However, the lack of a standardized definition of health expectancy limits its comparability across studies. The need of conceiving health expectancy as a comparable and repeatable measure was highlighted as fundamental to make it an informative instrument for policy makers.

Introduction

Between 2000 and 2050 the proportion of the world's population over 60 years is expected to double from about 11 percent to 22 percent, according to WHO estimates [1]. Global ageing has a major influence on disability trends. Higher disability rates among older people reflect an accumulation of health risks across a lifespan of disease, injury, and chronic illness [2]. The rapid increase in population ageing and the longer survival to chronic conditions has raised the question of how healthily the gained years of life will be spent. Life expectancy has been used as an indicator of population health for long time, but with the completion of the "Epidemiological transition" and infectious diseases replacing chronic diseases [3], mortality has ceased to be as tied to health as it was before and life expectancy does no longer capture population's health. This has led towards the development of a new measure, and in the 1960s the concept of health expectancy was first introduced [4,5]. Health expectancy (HE) is a summary measure of a population's health (SMPH) that expresses the average number of years that a person can expect to live in "full health"[6]. This very basic definition of HE contains both the strength and limitation of this indicator.

The main advantage of health expectancy is that it captures both the "quantity" and "quality" of life dimensions of health. As it combines information on mortality and health, it is more informative than life expectancy to forecast expected healthcare costs and assistance needs, as well as to plan pension systems and set age of retirement in a sensitive way; and from a societal point of view, it helps to anticipate possible changes in social participation and social inclusion at older ages. Information on health expectancy can be analyzed together with information on life expectancy, and their trends related to each other and compared to assess whether the proportion of life spent unhealthily is expanding due to the longer survival of individuals suffering from chronic diseases [7], or compressing due to the postponement in the onset of morbidity [8], or is associated with a redistribution of diseases resulting in a dynamic equilibrium [9]. HE and the other SMPH have been also largely used to compare different populations and identify and quantify health inequalities within and between sub-groups; for example between genders, social classes, behavior-based groups, or across different countries. Comparisons in HE of different populations can be used to evaluate the performance of different health systems and to identify the determinants of inequalities between populations [10].

On the other hand, one of the drawbacks of using this measure descends from the fact that the whole concept of health expectancy depends on the interpretation given to "full health". This has generated a variety of ways of expressing health expectancy, differing depending on the health indicators used to capture "full health". Commonly used terms are disability-free life expectancy, active life expectancy, healthy life expectancy and years of healthy life. These measures are not directly comparable and one is not preferable to another. Authors often define health depending on the availability of the information they have access to or on the objective they want to pursue. The problems arising from this lack of standardization have already been pointed out [11] and efforts to find a global and homogeneous definition of health and

disability at international level are growing. In 2001 the World Health Organization (WHO) proposed a conceptual framework for describing functioning and disability: the International Classification of Functioning, Disability and Health (ICF), which conceives and organizes disability as a combination of three components: Impairment, Activity and Participation. Descending from ICF, a more parsimonious conceptualization and measurement of disability was proposed by the Washington Group on Disability Statistics, which uses a short set of six questions to assess disability [12]. Following the same direction, an increasing number of health surveys-such as the Health and Retirement Study (HRS), English Longitudinal Study of Ageing (ELSA), Survey on Health Ageing and Retirement in Europe (SHARE)- run in different countries and sharing similar intents have been trying to harmonize themselves and ask the same questions for the most important health items, in order to facilitate the comparability of studies at international level. Among these health indicators the most commonly used are activities of daily living (ADLs), which are basic activities that are necessary to independent living, such as eating, bathing and dressing; instrumental activities of daily living (IADLs) which are activities that involve aspects of cognitive and social functioning, including for example shopping and cooking; and self-rated health (SRH) which is a general self-evaluation of health status and is usually assessed on a scale ranging from excellent to poor. The efforts made to standardize the concept of health, however, have not been transferred yet to build an indicator of health expectancy that is as much homogeneous and universal. The heterogeneity of the measures falling under the wide umbrella term of "health expectancy" will be illustrated while presenting the results of this review and the implications of this lack of agreement will be discussed further after showing the current state of affairs.

The widespread use of SMPH to measure and compare population's health and the need of shed some light on the complexity and the uncertainty surrounding the concept of health expectancy have motived this systematic literature review. More specifically, the review seeks to collect and give systematic appraisal of studies exploring socioeconomic and demographic inequalities in health expectancy, focusing on the older population. Within this general aim, the first objective was to appraise the extent of such inequalities and observe which factors have been considered and which remain unexplored. The second objective was to assess the strengths and limitations of using HE to study inequalities in healthy survival, especially in terms of comparability of results across different studies and populations.

In the next section our search strategy is described and the processing and selection of references are explained. Then, findings are organized and presented by type of inequality considered, type of health indicator used and methods applied to estimate health expectancy. Finally, gaps in understanding HE are outlined and limitations acknowledged.

Methods

Search strategy

Two possible sources were considered for the literature review: 1) electronic databases, and 2) references in selected articles. Electronic databases were the main source: we used Web Of Science, Medline and Embase.

The search was developed to combine terms referring to outcome, exposure and participants of interest (Table 1). The general term *health expectancy* was declined using a number of expressions all referring either to physical or general health, such as disability-free life expectancy (DFLE), active life expectancy (ALE), healthy life expectancy (HLE). As the family of SMPH is quite large and heterogeneous, given the size of the literature that uses this measure, we decided to focus only on physical and self-reported health combined with mortality. We did not include studies focused on mental functions and cognitive impairments, even if these may

Table 1. Search strategy: keywords and MeSH terms for systematic literature review.

Concept	Keywords[a]	MeSH terms[b]
1. Outcome: Health Expectancy	"health* life expectanc*" OR "active life expectanc*" OR "disability-free life expectanc*" OR "health* life year*" OR "Health adjusted life expectancy" OR "disability adjusted life expectancy" OR "disability life expectancy" OR "dependent life expectancy" OR "life expectancy with disability*" OR "health* expectancy*	**Medline:** Life expectancy, Mortality, Health Status Indicators, Health Status, Health Status Disparities, Disabled Persons, "Activities of Daily Living", Life Tables, Morbidity. **Embase:** life expectancy, mortality, health, health status, health disparity, disabled person, daily life activity, life tables, morbidity, disability
2. Participant: Older population	"old* age*" OR "old* adult*" OR "old* population" OR "old* people" OR "aged 50" OR "aged 60" OR "aged 65" OR "elderly people" OR "elderly population" OR "elderly" OR "senior*"	**Medline:** Aged, Middle Aged, Adult, Aged, 80 and over, Age Factors, Aging. **Embase:** aged, middle aged, adult, very elderly, aging
3a. Exposure—Socioeconomic subgroups	"social class*" OR "education*" OR "socioeconomic factors" OR "socioeconomic status" OR "socioeconomic position" OR "Occupation*" OR "income" OR "employ*" OR "housing tenure" OR "deprivation area" OR "deprivation index" OR "poverty area*" OR "area* of deprivation"	**Medline:** Socioeconomic Factors, Occupations, Income, Educational Status, Employment, Social Class, Social Conditions. **Embase:** socioeconomics, occupation, income, education, adult education, employment status, social class, social status
3b. Exposure—Demographic subgroups	"sex" OR" gender" OR "marital status" OR "parental status" OR "fertility" OR "number of children" OR "race" OR "ethnicit*"	**Medline:** Sex Factors, Family Characteristics, Family Relations, Child, Ethnic Groups, Population Groups, Male, Female, Gender Identity. **Embase:** sex, parenthood, marriage, cohabitation, ethnic group, ethnicity, male, female, gender

[a] used in Web of Science, Medline and Embase.

[b] used in Medline and Embase.

* truncation symbol.

Searches combined with AND: 1 AND 2 AND 3a, 1 AND 2 AND 3b.

doi:10.1371/journal.pone.0130747.t001

cause disability at later stages. The target population-the study participants- of this search was older individuals. We did not use any specific age threshold to define old age, but general expressions commonly applied to identify older adults and the most frequently used age groups, such as 65+, were included among the keywords. The exposures of interest were socio-economic and demographic variables. We used general terms, such as 'socioeconomic factors', as well as specific components, e.g. education or income, and a wide range of demographic terms including gender, marital and parental status, race and ethnicity. Table 1 shows text-words and-where available- MeSH terms used in the systematic search.

Selection strategy

Titles and abstracts of all references identified in the search were screened applying exclusion and inclusion criteria. The references shortlisted from this first screening were read in their entirety and selection criteria re-applied to full texts. The lists of references of the resulting studies were checked to ensure that all relevant articles were included in the search. Inclusion and exclusion criteria were set according to the definitions of the outcomes, exposures and participants and depending on the type of study (Table 2). As for the outcome, studies which did not combine mortality and health in a single indicator were not included, as well as those focused on mental health or cognitive-impairment free life expectancy. Analyses considering only the impact of specific diseases on HE (e.g. diabetes) were not included either. Studies focused on general populations, not exclusively targeted on the elderly, were included only if estimates of HE were provided also for selected old ages. Finally, studies had to consider at least one risk factor (exposure). Most of the studies stratified their sample by gender, and this was considered satisfactory to meet the inclusion criterion. When data were available only at the macro level (e.g. municipality, region) and an ecological approach adopted, references were

Table 2. Selection strategy: inclusion and exclusion criteria.

Concept	Inclusion criteria	Exclusion criteria
Outcome: Health Expectancy	Studies estimating any of the outcomes listed in Table 1	Excluded: studies not centred on the outcome of interest (i.e. mortality, health status, cognitive-impairment free life expectancy)
Participants: Older population	Studies focused on older population; studies on general population where estimates of health expectancy were provided also for selected old age groups	Excluded: studies not focused on older adult population
Exposure: a. Socioeconomic subgroups. b. Demographic subgroups	Analysis considering any of the risk factors listed in Table 1	Macro level analysis, ecological studies
Type of study	English, quantitative studies, reviews	Conference abstracts, commentaries, reports and editorials

doi:10.1371/journal.pone.0130747.t002

discharged. All quantitative studies and reviews meeting the criteria were included, while conference abstracts, commentaries and editorials excluded.

Analysis

The findings of the studies included in this review were synthesized in a narrative format, and organized adopting different perspectives. Data were extracted using a customized template based on the PICOS statement and developed in Microsoft Excel including the following items: author, year, setting and country, participants, exposure, outcome, measure of outcome, methods, and source of data (S1 Table). Consequently, we organized findings according to the exposure of interest. Then, we distinguished studies according to the definition of HE they adopted and the measures used to estimate it. Finally, we organized findings based on the methods used to estimate HE. Results were reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [13] (S2 Table). The review protocol was not registered.

Methods of estimating Health Expectancy

Two approaches are usually adopted to estimate HE: these are cross-sectional or longitudinal methods. The first typically employs prevalence-based life tables, also known as Sullivan's Method [5]. The second is based on incidence rather than prevalence and often relies on increment-decrement life table methods, or multistate life tables (MSLT).

Sullivan (1971) provided the first calculation of DFLE. The Sullivan's method is the most commonly used (see for example Cambois et al.[14]; Crimmins and Saito[15]) because of its simplicity and because it can be applied also to cross-sectional data. Forty-five of the references collected in this systematic search used Sullivan's method. These are cross-sectional studies and studies comparing HE at different points in time.

The other approach to estimate HE is based on MSLT. MSLT takes into account the possibility of returning to a state of health (or to a state of disability) using incident rates instead of prevalence. In situations where longitudinal data are available, the incidence-based MSLT is preferred because it most accurately reflects the impact of current conditions (i.e., disability onset, recovery, and mortality) on the evolution of the target population. The MSLT model is an extension of the simple period life table that underlies standard life expectancy estimates, and it is the preferred method in analysis of health changes over time (Cai [16], p.4). Thirty-six references in our search used MSLT to compute HE.

When prevalence and incidence remain constant between periods Sullivan's method and multistate life table have been found to produce similar results [17]. When either prevalence or

incidence vary, Sullivan's method may underestimate (or overestimate) health expectancy because it produces estimates based on past (as opposed to current) probabilities of becoming unhealthy.

Results

From the combination of the three databases, 256 references were identified and screened after deduplication (Fig 1). After reviewing titles and abstracts and applying inclusion and exclusion criteria, 110 references were read in full text and 90 selected (a list of the 20 studies considered for inclusion, but finally excluded is provided in S3 Table). Checking the lists of references of these studies, we identified 6 additional articles that were also included. Of the 96 references composing this systematic review, 33 were focused only on gender-based inequalities; the remaining 63 considered gender along with other exposures.

Findings from the 96 selected references were organized according to two leading perspectives: the type of inequalities considered and the health indicators chosen to measure health expectancy.

Inequalities in Health Expectancy

Gender-based inequalities. All studies selected in this literature review, but one, stratified individuals by age and gender (only Crimmins and colleagues [18] did not include gender in their analysis). All results showed evidence of a 'gender paradox'. The gender paradox in health and mortality was first observed in the mid 1970s [19,20] and consists in the finding that women live longer than men, but tend to have worse health than males. Gender difference in health expectancy can be assessed as the difference between males and females in the number of healthy years or in the proportion of healthy years on total life expectancy. All studies found that females spend a larger proportion of their life in disability than men, but when it came to compare women and men in terms of the healthy-years gap, findings differed across studies. In most cases women were found to live more years in good health than men at every age, but to have a smaller proportion of health expectancy due to their longer survival. For example, in the United States, in 1997 TLE at age 65 was 18.6 years for women and 12.6 years for men, and ALE was 16 and 11.2 years respectively, meaning that women could expect to live 86 percent of their life active and men 88.9 percent [21]. On the contrary, few studies [22,23] found that women's HE was shorter than men's HE not only as a proportion of TLE, but also expressed as number of expected healthy years. For example, Sauvaget et al. [24] found that TLE at age 75 in the UK in 1993 was 10.6 years for women and 9.1 years for men, and ALE was 3.1 years for women (corresponding to 29.2 percent of TLE) and 4.6 years for men (corresponding to 50.5 percent).

Whether gender inequalities remain constant as age increases or shrinks at older ages is not clear. According to Minicuci and Noale [25] the 'gender paradox' applies to each age group and for any severity of disability; Konno and colleagues [26] found that TLE is longer among women than men until the age of 70 years, but gender-gap in life expectancy declines at older ages, while the gender-gap in the proportion of ALE on TLE increases at older ages, with men enjoying at every age over 90% of their life spent being active and women's proportion of active life falling from 90% at the age of 65 to 79% at the age of 85. On the contrary, other studies [27–29] found evidence of shrinking gender gap in TLE and HLE at older ages.

Many arguments have been proposed to explain the 'gender paradox' in mortality and disability. In general, gender differences are largely due to mortality differences favoring women, rather than differences in the onset of disability [30]. Higher disability prevalence among women may be a function of longer survivorship in disability rather than higher incidence of

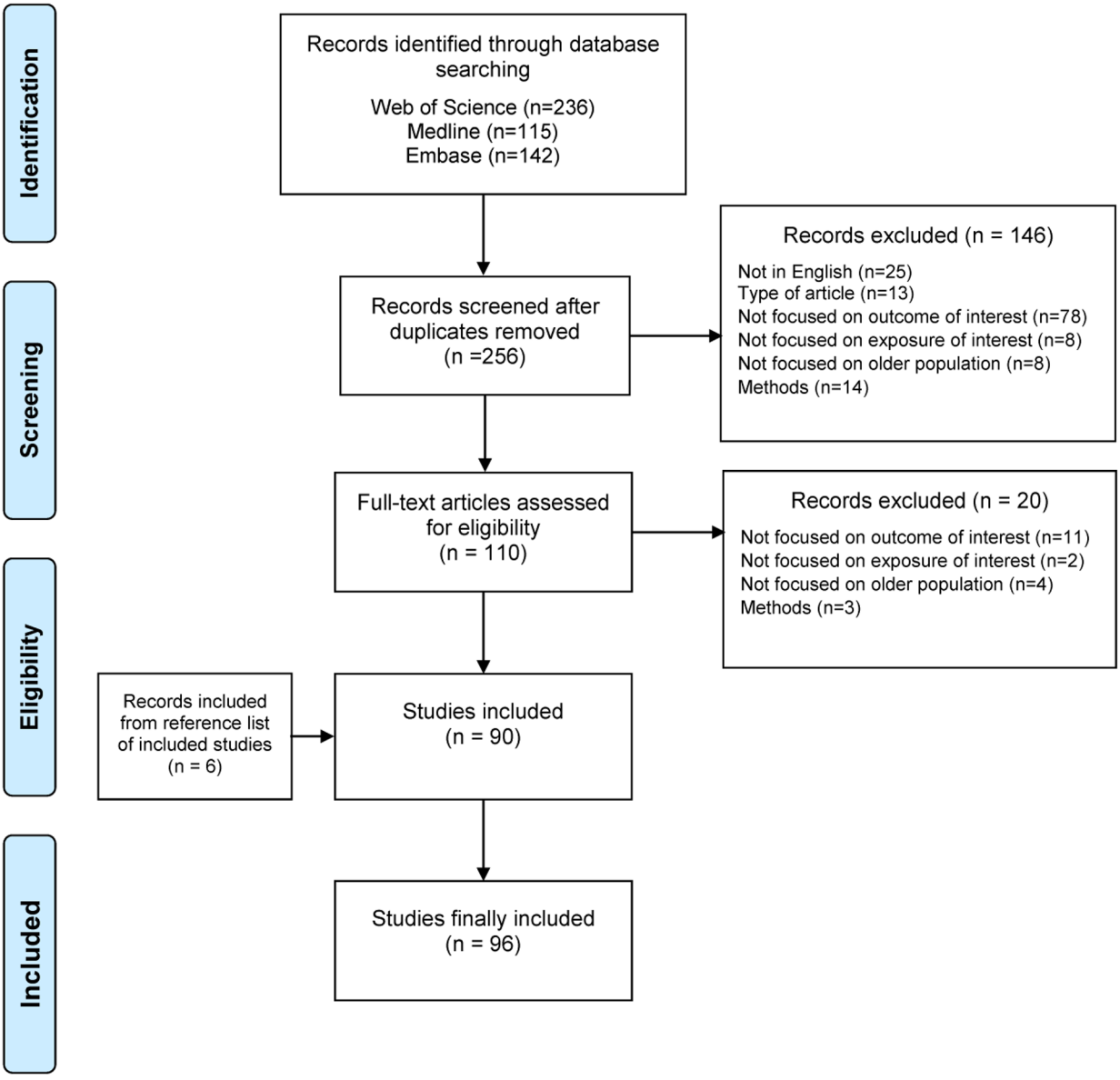


Fig 1. Flow diagram of systematic review.

doi:10.1371/journal.pone.0130747.g001

disability [31,32], in that women have higher prevalence of nonfatal but disabling diseases and men have higher prevalence of fatal diseases and chronic diseases strongly related to mortality [22]. Chang et al. [31] ascribed Japanese women disadvantage in disabilities to gender inequalities in socioeconomic status and disease profiles. The same explanation has been recently adopted for gender differences in health and functional status in Latin America [27]. Robine

and Cambois [33] attributed gender differences in ALE at age 65 to gender difference in ALE at earlier ages, and found that the percentage differences were much higher at age 65 than at earlier ages. Muangpaisan et al.[34] found evidence of gender paradox, but when HLE was assessed through SRH, women and men had similar proportions of life expectancy reporting "being in good health". This could suggest that women might accept some limitations of health status better than men. Results produced by Ishizaki et al.[28]showed women having longer TLE as well as longer years of dependency in activities of daily living (ADLs) as well as instrumental activities of daily living (IADLs) than men of the same age, but the proportions of physically active life expectancy (PALE) to TLE at any age did not greatly differ between men and women. Jitapunkul et al.[35] found evidence of gender difference in long-term DFLE, but very similar proportions of life expectancy spent unable to self-care between men and women. Tsuji et al.[36] found a marked gender difference in the process of developing and progressing disability, where men experienced disability at a younger ages and at a faster rate than women; the authors argued that this gender-based difference could be attributable in part to the differences between women and men in prevalence and incidence of diseases, where women were found for example to have significantly higher prevalence of arthritis and osteoporosis than men, while stroke incidence was significantly higher in men, as well as heart disease. All these findings would suggest that the strength and validity of gender differentials in HE may vary and depend on the health indicators used to estimate HE and may be explained by a variety of mechanisms, such as inequalities in socioeconomic position and social support and differences in health behaviors and disease profile, all of which would lead to different prevalence of fatal and nonfatal diseases in women and men.

Overall, the gender paradox is widely acknowledged and validated across different studies set in different countries, both high income and low and middle income countries (LMIC). For example, some of the results from our literature review showed evidence of the gender paradox in Brazil[27,29,37–39], Mexico[40], Hong Kong[41], Japan[31,36,42,43], Singapore[44], China [45–50], Bangladesh[23], Thailand[32,34,51–53], United States[22,30,54], Denmark[55–57], England[58,59], Italy[25,60], Bulgaria[61], Turkey [62], France[14,63,64] and across Europe [33,65–68].

Race-based inequalities. The studies of racial inequalities in HE included in our search were all set in the United States [15,30,54,69,70] and are mainly focused on Black and White differentials. The common finding is that Whites, as compared to any other ethnicity, enjoy more years in good health, but the gap reduces at older age. Crimmins et al.[30] distinguished Black and non-Black population and found generally non-significant race differences in TLE in the older populations; however Blacks have lower ALE than non-Blacks because of race differences in disability onset and recovery, deriving from socioeconomic inequalities. There is a debate about the existence of a black-white mortality crossover and at what age it would take place. Some argue that racial crossover is due to age misstatement by survey respondents[71]. Crimmins and colleagues [54] found evidence of mortality crossover by age 85 resulting in higher TLE as well as in expected life free of bed disability. Other studies [69,70] found that after age 75 black men and women have an advantage over Whites in both TLE and ALE.

When socioeconomic position (SEP) is controlled for, racial inequalities are larger among lower socioeconomic groups, with low SEP Blacks being the most disadvantaged group [15]. According to Land et al.[70], after stratifying by education, Whites' advantage in the corresponding unstratified comparisons tends to narrow and Blacks' disadvantage to decrease. According to Guralnik et al.[69], SEP has a greater association with TLE and ALE than race. Gender inequalities are also consistent across the different race groups: as their white counterpart, black women live longer than black men and spend a higher proportion of their life in disability [30,69].

Socioeconomic position-based inequalities. SEP can be measured via different indicators or their combination. The studies included in this literature review identify SEP using alternatively education, income, occupational or social class, area level deprivation, geographical area of residence (urban and rural), housing tenure, social support, or their combination [31,47,72,73]. Whatever the indicator was, all findings agreed that individuals belonging to lower SEP have shorter TLE and enjoy fewer years in good/active health. On the other hand, the strength of this association may depend on the indicator used to measure SEP. There is no agreement on which measure determines the largest SEP-based gap in HE and TLE. Some results suggest that housing tenure and education-based inequalities have the strongest impact, while the inequalities due to income were found not to be significant [72,73]. Other research found income having some strong influences and occupation little impact [31].

The strength of the association between SEP and HE varied across countries. It was substantial in the United States[15], while results pertaining to other countries, such as Indonesia [74], revealed a weaker association. In general, in LMIC high SEP is associated with an expansion of TLE that is accompanied by an expansion of years lived with disability.

It was commonly found that socioeconomic differentials in TLE and HE were larger for men than for women. One of the possible explanations is that women's SEP may be more a function of their household's SEP than their individual characteristics [72,73]. Accordingly, Matthews et al. [73] measured women's SEP through their husband's job. Chan et al.[31] also claimed socioeconomic inequalities between genders as one of the causes for gender-based inequalities in disability life expectancy (DLE).

SEP-based differences in HE has been generally found to exceed SEP-based differences in TLE, suggesting that inequalities are most pervasive with respect to quality rather than quantity of remaining life [15,64,72,73]. Another common finding is that of a widening of the SEP gap at increasing age [31,72].

Education-based inequalities. Since most studies measured SEP through education, special attention was dedicated to HE gap deriving from education-based inequalities. The main advantage of measuring SEP by education is that in most cases individuals have completed their educational path by their early adulthood and this indicator is therefore stable across ages. The vast majority of the studies included in this review confirmed the advantage of being highly educated in terms of both TLE and HLE [15,38,72,73,75]. The consistency of these associations seemed partly to be dependent on the country where the studies were set. Using data from São Paulo and urban areas in Mexico, Beltran-Sanchez and Andrade[76] found some indication of the role of education in influencing HE and TLE in Mexico and Brazil, but no significant educational differences in transition probability (incidence and recovery from disability, and mortality) (p. 827). The analysis of Hidajat et al.[74], showed that education among older Indonesians was associated with increases in life expectancy that was accompanied by longer life with functional problems. On the other hand, studies set in the US [69] found that education impacted both TLE and ALE-and to a greater extent than race- although the differences between higher and lower educated individuals diminished at older ages[77]. In other studies, using multiple indicators for SEP, education was found to be the most strongly predictive measure for HLE and TLE [31,39,72]. Montez and Hayward [78] studied whether educational attainment mediated and moderated the health consequences of early-life conditions. They found that early-life experiences were associated with TLE and ALE, even after adjusting for education. On the other hand, more years of education predicted more years of life and active life, regardless of childhood context (p. 431).

Behavior-based inequalities. It is generally recognized that lifestyle factors are associated with both morbidity and mortality. These factors, in particular smoking, alcohol consumption, overweight and obesity, and physical activity, have also been found to be associated with SEP

in several populations [79,80] and this represents one of the mechanisms through which SEP might influence health and mortality. However, overall, not many studies have explored the consequences of healthy behaviors on HE among the older population. Results have shown that smoking has a stronger association with mortality than morbidity, while overweight and obesity, and to a lesser extent alcohol consumption, mostly with disability [81]. This difference can be explained by the fact that a high body mass index (BMI), a measure of obesity, is more likely to be associated with non-lethal disabling diseases, whereas smoking is more strongly related to a number of fatal diseases, such as cancer. Weak association of overweight and obesity with mortality but significant association with morbidity was also found in several studies [82–85]; whilst another work [86] showed that mortality rates of overweight and obese participants were not only similar but sometimes better than those of normal weight ones. Flegal and colleagues [87] undertook a meta-analysis of 8 large studies to understand the reasons behind this BMI and mortality paradox. The choice of the reference category for computing the effect measures and the cut points used to define the categories were found to influence estimates and statistical power.

The negative effect of smoking on ALE was confirmed in all studies [77,88,89]. Also differences between heavy and light current smokers and recent and long-term quitters were found in terms of both TLE and ALE. [89]. Only one study in this review considered the association between physical activity and HE [88]. The authors combined this factor with smoking habits and found that the negative effect of inactivity on survival and length of disabled life was comparable or even higher than the effect of smoking.

Health indicators used to measure Health Expectancy

One of the difficulties in comparing studies on HE derives from the heterogeneity of its definition. HE is a fairly generic measure that can refer to physical as well as cognitive status. This systematic review was focused only on the former, expressed both in terms of general health and disability. Disability however can be defined using a variety of indicators, and in some studies more than one indicator was applied. In the following paragraph and tables we present studies according to the measures adopted to define HE.

Table 3 shows the different health indicators applied to measure HE and the corresponding definition of HE given by the studies included in this literature review. Most of selected references measured HE using ADLs. Forty-two studies adopted exclusively this indicator and in most cases HE was named either DFLE or ALE; other studies used the term 'Life Expectancy without ADL restrictions', or 'Years of life without functional disabilities', or 'Life Expectancy without functional problems'. Ishizaki et al.[28] used the expression 'Physically Active Life Expectancy' when HE was measured by ADLs, and 'Instrumentally Active Life Expectancy' when IADLs were applied. In other cases (seven references) ADLs and IADLs were combined together and HE was named either DFLE or ALE. Fourteen studies used SRH to measure HE and name this latter either 'Healthy Life expectancy', 'Years of Healthy Life' (YHL) or 'Life Expectancy in good health'; Lievre et al. (2007) [66] used the term 'Healthy Working Life Expectancy' (HWLE), because their work was targeted on working population aged between 50 and 70. Other researchers applied various indicators of functioning and mobility, such as the Barthel ADL index which describes ADLs and mobility using 10 variables[90], and the Chula ADL index [91], composed of five items and conceived to be used in low-middle income countries; other studies referred to the mobility indicators selected by Nagi (1967) [92]that express sensory-motor functioning of the organism, and are indicated by limitations in activities such as walking, climbing, bending, reaching, hearing, etc.; other authors measured HE by sensory function limitations, such as hearing, seeing, walking, etc. The Global Activity Limitation

Table 3. Type of Health Expectancy by the most commonly used health indicators.

Health indicator (n studies[a])	Type of Health Expectancy	Reference
Activity of daily living (ADL) **(42)**	Disability-free Life Expectancy (DFLE)	Al Snih et al. [82]; Andrade et al. [27]; Cambois et al. [63]; Campolina et al.[94]; Chan et al.[31]; Cheung and Yip [41]; Hayward et al. [95]; Karcharnubarn et al. [53]; Klijs et al. [81]; Matthews et al. [73]; Minicuni et al. [96]; Mutafova et al. [61]; Peres et al. [97]; Santos Camargos et al. [38]; Walter et al. [85]
	Active Life Expectancy (ALE)	Branch et al. [98]; Diehr et al. [86]; Ferrucci et al. [88]; Gu et al. [45]; Gurlanik et al. [69]; Izmirlian et al. [77]; Jiawiwatku et al. [32]; Kai [99]; Katz et al. [100]; Laditka and Laditka [101]; Land et al.[70]; Matthews et al. [93]; Reyes-Beaman et al. [40]; Reynolds et al. [102]; Reynolds and McIlvane [83]; Reynolds et al.[84]; Rogers et al. [103]; Sauvaget et al. [104]; Tian et al.[89]; Tsuji et al. [42]; Yi et al. [105]; Yi et al. [50]
	Physically Active Life Expectancy (PALE)	Ishizaki et al. [28]
	LE without ADL restrictions	Cambois et al. [64]; Jagger et al. [65]
	Years of life without functional disabilities	Yong et al. [44]
	LE without functional problems	Brayne et al. [59]
Instrumental activity of daily living (IADL) **(1)**	Instrumentally Active Life Expectancy (IALE)	Ishizaki et al. [28]
ADL + IADL **(7)**	Disability-free Life Expectancy (DFLE)	Minicuci and Noale[25]; Jagger et al.[106]; Crimmins et al. [18]
	Active Life Expectancy (ALE)	Yong and Sayto [107]; Konno et al.[26]; Hayward et al. [108]; Cai and Lubitz [22]
Various combination of items including: ADL, IADL, Barthel and Chula ADL index, mobility indicators (i.e. NAGI, sensory function limitations, etc. . .) **(24)**	Disability-free Life Expectancy (DFLE)	Crimmins et al. [54]; Jagger et al. [109]; Jitapunkul et al. [35]; Manton et al. [110]; Minicuci and Noale [60]; Muangpaisan et al. [34]; Sagardui-Villamor et al. [111]; Santos Camargos et al. [39]; Sauvaget et al. [24]; Tareque et al. [23]; Cambois et al. [63]; Beltran-Sanchez and Andrade [76]
	Active Life Expectancy (ALE)	Jitapunkul et al. [52]; Kaneda et al.[72]; Manton et al. [112]; Montez and Hayward [78]; Tsuji et al. [36]; Zimmer et al. [47]; Hidajat et al. [74]
	Healthy Life expectancy (HLE)	Szwarcwald et al. [113]
	LE without mobility limitations	Jeune and Bronnum-Hansen [57]; Karcharnubarn et al. [53]
	Functional Independence LE	Liu et al.[114]; Cambois et al. [64]
Self-rated health (SRH) **(14)**	Healthy Life expectancy (HLE)	Gu et al.[45]; Jitapunkul and Chayovan[51]; Karcharnubarn et al. [53]; Mutafova et al. [61]; Ofstedal et al. [115]; Santos Camargos et al.[29]; Yong and Saito [43]
	Years of Healthy Life (YHL)	Diehr et al.[116]; Diehr et al. [86]
	LE in good health	Bronnum-Hansen[55]; Cambois et al. [64]; Jagger et al.[65]; Jeune and Bronnum-Hansen [57]
	Healthy Working Life Expectancy (HWLE)	Lievre et al. [66]

(*Continued*)

Table 3. (*Continued*)

Health indicator (n studies[a])	Type of Health Expectancy	Reference
Global activity limitation index (GALI) **(5)**	Disability-free Life Expectancy (DFLE)	Majer et al.[67]; Cambois et al. [14]
	Healthy Life expectancy (HLE)	Crimmins and Saito [15]
	Active Life Expectancy (ALE)	Robine et al. [33]
	LE without GALI	Cambois et al. [64]
	Years of life w/ and w/o diseases	Yong et al.[44]; Cambois et al. [64]
Chronic conditions **(8)**	Disease-free Life Expectancy	Gu et al. [45]
	LE free of impairment	Jagger and Matthews [58]
	Chronic Morbidity-free Life Expectancy (CMFLE)	Jagger et al. [65]; van den Bos et al.[117]; Cheung et al. [41]; Deeg et al.[118]

[a] number of studies does not sum up to the total number of studies in the review, as some studies are counted more than once because several measure of HE were estimated.

doi:10.1371/journal.pone.0130747.t003

Index (GALI)—which measures long-standing severe disability through a single question such as "For at least the past six months, to what extent have you been limited because of a health problem in activities people usually do?"- was used in five studies. Eight studies measured HE in terms of chronic morbidity, consisting in various chronic or long-term diseases or condition. Finally some authors computed multiple types of HE using different health measures [41,45,53,57,61,63–65,93], this is why the total number of studies included in Table 3 (see column 2) is higher than the total number of studies included in the review. Table 4 presents estimates of HE computed in a sample of four of these studies, to show the variability of HE depending on the health measures adopted as indicator. In all these works TLE and HE were measured at age 65, almost in the same years (2000 and 2002). In all studies one of the measures used to estimate HE was SRH, the lowest estimate was 5.8 years for women in Thailand and the highest 10.3 years for women in Denmark. When health was measured using various questions for ALDs, the proportion of HE on TLE was the highest, for example for men estimates ranged from 84 percent in French men to 95 percent in Thailand (in this study HE was named LE without self-care activity limitations), and for women 77 percent in France and 94.5 in Thailand.

Discussion

This literature review aimed both at providing a systematic appraisal of studies that investigate socioeconomic and demographic inequalities in HE and giving a critical assessment of this measure. To our knowledge, this is the first review on aggregate/macro level measures of health expectancy in the older population to have been conducted systematically. A literature review on active life expectancy was published in 2002 [71] but it was targeted on women and was not systematic. Crimmins and Cambois [119] collected a number of studies focused on socioeconomic inequalities in HE, but their work was aimed at providing a theoretical explanation of HE and promoting the use of this indicator to design and implement policies.

With respect to the first objective of the review, two main results emerged. The first is the heterogeneity of the indicators and methods used to measure health expectancy and the

Table 4. Example of studies estimating multiple types of HE at age 65.

Study—Setting, Year	HE	Women		Men	
		years[a]	% of TLE[a]	years[a]	% of TLE[a]
Cambois et al.—France, 2002/03	TLE	20.5	100	16.9	100
	w/o ADL restrictions	15.8	77	14.2	84
	w/o GALI	12.2	59	11.2	66
	in good perceived health	8.2	40	8	47
	w/o functional limitations	6.6	32	6.7	40
	w/o chronic diseases	5.4	26	5.4	32
Gu et al. (2009)—China, 2002	TLE	16.85	100	14.05	100
	w/o ADL restrictions[1]	15.03	89.19	13.08	93.11
	in good perceived health	7.74	45.93	7.17	51.06
	w/o chronic diseases	6.27	37.2	5.69	40.54
Jeune et al. (2008)—Denmark, 2000	TLE	18.1	100	15	100
	w/o mobility limitations[2]	11.9	65.6	12.4	82.4
	w/o GALI	9.5	52.3	8.9	59.1
	in good perceived health	10.3	56.6	9.7	64.3
	w/o communication restriction[3]	15.3	84.3	12.9	85.7
Karcharnubarn et al. (2013)—Thailand, 2002	TLE	17.1	100	15.8	100
	w/o self-care activities limitations[4]	16.1	94.5	15.2	95.8
	w/o mobility limitations[5]	5.4	31.5	8.1	51.5
	in good perceived health	5.8	34.1	6.9	43.4

[a] Different decimal points according to rounding precision reported in the studies.

[1]Katz's ADL index.

[2]walk 400 m without resting, walk up or down a staircase, carry 5 kg.

[3]read ordinary newspaper print, hear normal conversation, speak with minor or major difficulty.

[4]including only feeding dressing and bathing.

[5]Squatting, carrying thing 5kg, walking 1km, climbing stair2–3 flights, taking a bus/ship alone.

doi:10.1371/journal.pone.0130747.t004

multiplicity of terms chosen to name it. It was generally observed that DFLE and ALE were used interchangeably, while HLE mainly referred to life expectancy in good SRH, although this was not always the case. Even when HE referred to the same dimension of health (e.g. functional limitations) this was measured using very different indicators. Therefore, the different levels of HE reported across studies may partly depend on the way healthy and unhealthy conditions were defined and not just on the risk factors considered for the analysis. Some studies computed multiple types of HE using different health measures [41,45,53,57,61,63–65,93]. This offered the opportunity to check to what extent the estimates of HE varied depending on the health measure adopted as indicator. In all these studies, the largest proportion of TLE lived healthily was observed when health was measured by limitations in ADLs, and the smallest when chronic morbidity were used [41,45,64,65], while intermediate estimates were obtained when using SRH. The other finding is the homogeneity of the results in terms of the types of inequalities found, such as gender, education, behavior-based disparities. It is interesting then to consider these two findings –the heterogeneity of the concept of health expectancy and the homogeneity of the results by type of inequality- together. Regardless of how health expectancy was measured, women were found to have longer life expectancy and spend a larger proportion of their life in poor health or with disability than men, in every study. Similarly, the existence of an educational gradient affecting positively both life expectancy and health

expectancy was confirmed by all studies that considered such exposure; as well as belonging to higher socioeconomic classes was found to lead to longer and healthier life. Findings were consistent across studies and results homogeneous in terms of the direction of inequalities; nevertheless, the actual numbers of years of health expectancy that each category may expect to live was not comparable across different studies. For example, it was possible to generally say that women live a shorter proportion of their life without disability than men, but it was not meaningful to report the precise year gap.

The other aim of the review was to give a critical appraisal of HE as summary measure of population health and discuss the convenience of using it to study inequalities in health and mortality. Two considerations are presented with this matter. First, one of the features in the computation of HE is that analyses are driven by and constrained to the availability of mortality rates or life tables for the specific groups. This has led to the study of inequalities among certain groups only. Hence we found that papers concentrated on investigating inequalities in HE by gender and SEP. However, how gender and SEP may influence HE is still controversial. Some hypotheses have been advanced to explain gender inequalities, for example explaining sex differences as a consequence of the differences in the prevalence of fatal and non-fatal diseases among men and women, and partly ascribing the existence of the 'gender paradox' to inequalities in socioeconomic characteristics between men and women. On the other hand, the causes of SEP-based disparities are still unclear. One of the possible mechanisms explaining this association are health behaviors; but not many researchers have been able to study their role among older populations and if they did so, have focused only on some risk factors, such as smoking and overweight, whilst other behaviors such as physical activity and diet have remained almost unexplored. The second consideration –also mentioned in the introduction, before the current state of affairs in literature was presented—pertains to the complexity of the concept of health and disability and the difficulty of measuring it. The findings of this review confirmed the limitations in comparing different studies because of the variety of health indicators used to measure HE, the different estimation methods and the multiplicity of definitions for HE. There is not agreement in the way of defining and conceiving HE and there is not a preferred interpretation either. Furthermore, currently there not seems to be efforts put in place to address this problem.

A final aspect that emerged from this systematic search pertains to the way in which institutionalized population has been considered across various studies. Institutionalized individuals represent an important component of older populations. Data used to estimate HE are mainly drawn from surveys that usually do not include these individuals. Some researchers combined different sources of data to cover the whole (total) population[54,93,106,120], others included hypotheses on the health condition of institutionalized population, assuming for example that they were either as healthy as non-institutionalized individuals[33,65] or all impaired[64]; other studies simply did not include this section of the population into their analysis. This heterogeneous approach to institutionalized population is another possible source of bias to take into consideration when comparing results from different studies.

Conclusion

The empirical evidence for the joint progress of morbidity and mortality are weak and scarce; and results are contradictory [7–9]. The evolution of health in respect to mortality concerns everyone as an individual perspective and involves the society as a whole –both in high and low middle income countries- because the way in which the older population ages determines the needs for health care and social protection, and impacts the availability of resources for younger generations. Health expectancy is a useful and convenient way of monitoring and assessing

the quality of ageing. It allows comparing different groups and populations and identifying disadvantaged categories in order to detect specific factors that can help to promote healthy ageing. This review showed a general agreement of results obtained in different studies with regard to the existence of inequalities associated with several factors, such as gender, education, behaviors, race. Some studies considered more than one factor at the time and found a sort of intersectionality of disparities, identifying particularly frail categories. Interventions should be addressed both to protect specific categories at risk and to generally increase the proportion of life that individuals spend in good health, without disability and being independent. This would mean work in the direction of postponing the onset of disability in the general population, as well as providing external help to carry out basic and complex activities (i.e. ADLs and IADLs) after disability has occurred. To make health expectancy an informative instrument to monitor the quality of ageing and learn class from the past and from the comparison with other populations, this indicator has to be comparable and repeatable. We advocate the need of standardizing health expectancy, both as a concept and in its measurement. This direction has already been taken in surveys by asking same questions for measuring certain health items and in research by assessing the validity of indicators across countries and validating different measures with each other [121,122]. What is still missing is to transfer and apply this to a summary indicator. General guidelines should be drawn to clarify the concept of health expectancy and more research carried out to investigate how best it could be computed. The advantage of using health expectancy to capture population health has produced a vast body of literature that has used this indicator to provide evidence of the existence of inequalities. To use these evidence efficiently and provide more in the future in a consistent way and to make health expectancy a global and informative instrument for policy makers, it needs to be conceived and estimated in a standardized and universal way.

Supporting Information

S1 Table. Descriptive characteristics of included studies.
(XLSX)

S2 Table. PRISMA guidelines checklist.
(DOC)

S3 Table. Full population of studies considered for inclusion.
(XLSX)

Author Contributions

Conceived and designed the experiments: BP BLD GBP. Analyzed the data: BP BLD GBP. Wrote the paper: BP BLD GBP.

References

1. United Nations, Department of Economic and Social Affairs, Population Division (2013) World Population Ageing 2013.
2. World Health Organization, the World Bank (2011) World report on disability. Geneva: WHO Press.
3. Omran AR (1971) The epidemiologic transition. A theory of the epidemiology of population change. Milbank Memorial Fund Quarterly 49: 509–538. PMID: 5155251
4. Sanders BS (1964) Measuring Community Health Levels. American Journal of Public Health & the Nation's Health 54: 1063–1070.
5. Sullivan D (1971) A Single Index of Mortality and Morbidity. HSMHA Health Reports 86: 347–354. PMID: 5554262

6. World Health Organization (WHO) (2013) Healthy life expectancy at birth. Available: http://apps.who.int/gho/indicatorregistry/App_Main/view_indicator.aspx?iid=66.
7. Kramer M (1980) The rising pandemic of mental disorders and associated chronic diseases and disabilities. Acta Psychiotrica Scandinaveca 62: 282–297. PMID: 7468150
8. Fries JF (1980) Aging, natural death, and the compression of morbidity. Gerontological Perspecta 1: 5–8.
9. Manton KG (1982) Changing concepts of morbidity and mortality in the elderly population. Milbank Mem Fund Q Health Soc 60: 183–244. PMID: 6919770
10. Murray CJ, Salomon JA, Mathers C (2000) A critical examination of summary measures of population health. Bulletin of the World Health Organization 78: 981–994. PMID: 10994282
11. Saito Y, Robine J-M, Crimmins EM (2014) The methods and materials of health expectancy. Statistical Journal of the IAOS 30: 209–223.
12. Washington Group on Disability Statistics (2010) Census Questions on Disability Endorsed by the Washington Group. Available: http://www.cdc.gov/nchs/data/washington_group/WG_Short_Measure_on_Disability.pdf.
13. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6: e1000097. doi: 10.1371/journal.pmed.1000097 PMID: 19621072
14. Cambois E, Robine JM, Hayward MD (2001) Social inequalities in disability-free life expectancy in the french male population, 1980–1991. Demography 38: 513–524. PMID: 11723948
15. Crimmins EM, Saito Y (2001) Trends in healthy life expectancy in the United States, 1970–1990: gender, racial, and educational differences. Social Science & Medicine 52: 1629–1641.
16. Cai L (2013) The cost of an additional disability-free life year for older Americans: 1992–2005. Health Serv Res 48: 218–235. doi: 10.1111/j.1475-6773.2012.01432.x PMID: 22670874
17. Robine JM, Mathers CD (1993) Measuring the compression or expansion of morbidity through changes in health expectancy. In: Robine JM, Mathers CD, Bone MR, Romieu I, editors. Calculation of Health Expectancies: Harmonization, Consensus Achieved and Future Perspectives. pp. 269–286.
18. Crimmins EM, Hayward MD, Hagedorn A, Saito Y, Brouard N (2009) CHANGE IN DISABILITY-FREE LIFE EXPECTANCY FOR AMERICANS 70 YEARS OLD AND OLDER. Demography 46: 627–646. PMID: 19771948
19. Nathanson CA (1975) Illness and the feminine role: a theoretical review. Social Science & Medicine 9: 57–62.
20. Verbrugge LM (1982) Sex differentials in health. Public Health Reports 97: 417–437. PMID: 6750677
21. Newman AB, Brach JS (2001) Gender gap in longevity and disability in older persons. Epidemiologic Reviews 23: 343–350. PMID: 12192741
22. Cai L, Lubitz J (2007) Was there compression of disability for older americans from 1992 to 2003? Demography 44: 479–495. PMID: 17913007
23. Tareque MI, Begum S, Saito Y (2013) Gender Differences in Disability-Free Life Expectancy at Old Ages in Bangladesh. Journal of Aging and Health 25: 1299–1312. doi: 10.1177/0898264313501388 PMID: 24013854
24. Sauvaget C, Tsuji I, Aonuma T, Hisamichi S (1999) Health-life expectancy according to various functional levels. Journal of the American Geriatrics Society 47: 1326–1331. PMID: 10573441
25. Minicuci N, Noale M, group I (2005) Influence of level of education on disability free life expectancy by sex: the ILSA study. Experimental Gerontology 40: 997–1003. PMID: 16199134
26. Konno K, Katsumata Y, Arai A, Tamashiro H (2004) Functional status and active life expectancy among senior citizens in a small town in Japan. Archives of Gerontology and Geriatrics 38: 153–166. PMID: 14698494
27. Andrade DFC, Guevara PE, Lebrao ML, de Oliveira Duarte YA, Santos JLF (2011) Gender differences in life expectancy and disability-free life expectancy among older adults in Sao Paulo, Brazil. Women's health issues: official publication of the Jacobs Institute of Women's Health 21: 64–70. doi: 10.1016/j.whi.2010.08.007 PMID: 21185991
28. Ishizaki T, Kai I, Kobayashi Y, Imanaka Y (2002) Functional transitions and active life expectancy for older Japanese living in a community. Archives of Gerontology and Geriatrics 35: 107–120. PMID: 14764349
29. Santos Camargos MC, Machado CJ, Rodrigues RN (2008) Sex differences in healthy life expectancy from self-perceived assessments of health in the City of Sao Paulo, Brazil. Ageing & Society 28: 35–48.

30. Crimmins EM, Hayward MD, Saito Y (1996) Differentials in active life expectancy in the older population of the United States. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 51: S111–S120.

31. Chan A, Zimmer Z, Saito Y (2011) Gender Differentials in Disability and Mortality Transitions: The Case of Older Adults in Japan. Journal of Aging and Health 23: 1285–1308. doi: 10.1177/0898264311408417 PMID: 21653794

32. Jiawiwatkul U, Aekplakorn W, Vapattanawong P, Prasartkul P, Porapakkham Y (2012) Changes in Active Life Expectancy Among Older Thais: Results From the 1997 and 2004 National Health Examination Surveys. Asia-Pacific Journal of Public Health 24: 915–922. doi: 10.1177/1010539511409923 PMID: 21622481

33. Robine JM, Jagger C, Cambois E (2002) European perspectives on healthy aging in women. Journal of Women & Aging 14: 119–133.

34. Muangpaisan W, Assantachai P, Intalapaporn S, Richardson K, Brayne C (2011) Health expectancies in the older Thai population. Archives of Gerontology and Geriatrics 53: 3–7. doi: 10.1016/j.archger.2010.05.012 PMID: 20570376

35. Jitapunkul S, Kunanusont C, Phoolcharoen W, Suriyawongpaisal P, Ebrahim S (2003) Disability-free life expectancy of elderly people in a population undergoing demographic and epidemiologic transition. Age and Ageing 32: 401–405. PMID: 12851183

36. Tsuji I, Sauvaget C, Hisamichi S (2002) Health expectancies in Japan gender differences and policy implications for women. Journal of Women & Aging 14: 135–148.

37. Belon AP, Lima MG, Barros MBA (2014) Gender differences in healthy life expectancy among Brazilian elderly. Health and Quality of Life Outcomes 12.

38. Santos Camargos MC, Machado CJ, Rodrigues RDN (2007) Disability life expectancy for the elderly, city of Sao Paulo, Brazil, 2000: Gender and educational differences. Journal of Biosocial Science 39: 455–463. PMID: 16707040

39. Santos Camargos MC, Machado CJ, Rodrigues RN (2008) Life expectancy among elderly Brazilians in 2003 according to different levels of functional disability. Cadernos De Saude Publica 24: 845–852. PMID: 18392362

40. Reyes-Beaman S, Jagger C, Garcia-Pena C, Munoz O, Beaman PE, et al. (2005) Active life expectancy of older people in Mexico. Disability and Rehabilitation 27: 213–219. PMID: 16025748

41. Cheung KSL, Yip PSF (2010) Trends in healthy life expectancy in Hong Kong SAR 1996–2008. European Journal of Ageing 7: 257–269. PMID: 21212818

42. Tsuji I, Minami Y, Fukao A, Hisamichi S, Asano H, et al. (1995) Active life expectancy among elderly Japanese. Journals of Gerontology—Series A Biological Sciences and Medical Sciences 50: M173–M176.

43. Yong V, Saito Y (2009) Trends in healthy life expectancy in Japan: 1986–2004. Demographic Research 20: 467–494.

44. Yong V, Saito Y, Chan A (2011) Gender differences in health and health expectancies of older adults in Singapore: an examination of diseases, impairments, and functional disabilities. Journal of cross-cultural gerontology 26: 189–203. doi: 10.1007/s10823-011-9143-0 PMID: 21416332

45. Gu D, Dupre ME, Warner DF, Zeng Y (2009) Changing health status and health expectancies among older adults in China: Gender differences from 1992 to 2002. Social Science & Medicine 68: 2170–2179.

46. Liu J, Chen G, Song X, Chi I, Zheng X (2009) Trends in Disability-Free Life Expectancy Among Chinese Older Adults. Journal of Aging and Health 21: 266–285. doi: 10.1177/0898264308328978 PMID: 19104033

47. Zimmer Z, Kaneda T, Tang Z, Fang X (2010) Explaining Late Life Urban vs. Rural Health Discrepancies in Beijing. Social Forces 88: 1885–1908. PMID: 21151893

48. Lai D (2009) A Comparative Study of Handicap-free Life Expectancy of China in 1987 and 2006. Social Indicators Research 90: 257–265.

49. Lai DJ, Lee LM, Lee ES (2000) Effects of handicap on life expectancy: the case of China. Public Health 114: 330–335. PMID: 11035451

50. Yi Z, Vaupel JW, Xiao Z, Zhang C, Liu Y (2001) The healthy longevity survey and the active life expectancy of the oldest old in China. Population (Paris) 56: 109–132.

51. Jitapunkul S, Chayovan N (2000) Healthy life expectancy of Thai elderly: Did it improve during the soap-bubble economic period? Journal of the Medical Association of Thailand 83: 861–864. PMID: 10998838

52. Jitapunkul S, Krungkraipetch N, Kamolratanakul P, Dhanamun B (2001) Dependence and active life expectancy of the elderly population living in the central region of Thailand. Journal of the Medical Association of Thailand 84: 349–356. PMID: 11460935

53. Karcharnubarn R, Rees P, Gould M (2013) Healthy life expectancy changes in Thailand, 2002–2007. Health & Place 24: 1–10.

54. Crimmins EM, Saito Y, Ingegneri D (1989) Changes in the life expectancy and disability-free life expectancy in the United States; Robine JM, Blanchet M, Dowd JE, editors. 199–207 p.

55. Bronnum-Hansen H (2005) Health expectancy in Denmark, 1987–2000. European Journal of Public Health 15: 20–25. PMID: 15788799

56. Bronnum-Hansen H, Juel K, Davidsen M (2006) The burden of selected diseases among older people in Denmark. Journal of Aging and Health 18: 491–506. PMID: 16835386

57. Jeune B, Bronnum-Hansen H (2008) Trends in health expectancy at age 65 for various health indicators, 1987–2005, Denmark. European Journal of Ageing 5: 279–285.

58. Jagger C, Matthews F (2002) Gender differences in life expectancy free of impairment at older ages. Journal of Women & Aging 14: 85–97.

59. Brayne C, Matthews FE, McGee MA, Jagger C, Mrc C (2001) Health and ill-health in the older population in England and Wales. Age and Ageing 30: 53–62. PMID: 11322674

60. Minicuci N, Noale M (2005) Disability free life expectancy in older Italians. Disability and Rehabilitation 27: 221–227. PMID: 16025749

61. Mutafova M, vandeWater HPA, Perenboom RJM, Boshuizen HC, Maleshkov C (1997) Health expectancy calculations: A novel approach to studying population health in Bulgaria. Bulletin of the World Health Organization 75: 147–153. PMID: 9185367

62. Yardim N, Akgun S, Mollahaliloglu S, Basara B (2009) Life expectancies and health adjusted life expectancy (HALE) at 60 years old, population in Turkey. Turk Geriatri Dergisi 12: 111–117.

63. Cambois E, Blachier A, Robine JM (2013) Aging and health in France: an unexpected expansion of disability in mid-adulthood over recent years. European Journal of Public Health 23: 575–581. doi: 10.1093/eurpub/cks136 PMID: 23042230

64. Cambois E, Laborde C, Romieu I, Robine JM (2011) Occupational inequalities in health expectancies in France in the early 2000s: Unequal chances of reaching and living retirement in good health. Demographic Research 25: 407–435.

65. Jagger C, Weston C, Cambois E, Van Oyen H, Nusselder W, et al. (2011) Inequalities in health expectancies at older ages in the European Union: findings from the Survey of Health and Retirement in Europe (SHARE). Journal of Epidemiology and Community Health 65: 1030–1035. doi: 10.1136/jech.2010.117705 PMID: 21471138

66. Lievre A, Jusot F, Barnay T, Sermet C, Brouard N, et al. (2007) Healthy working life expectancies at age 50 in Europe: A new indicator. Journal of Nutrition Health & Aging 11: 508–514.

67. Majer IM, Nusselder WJ, Mackenbach JP, Kunst AE (2011) Socioeconomic inequalities in life and health expectancies around official retirement age in 10 Western-European countries. Journal of Epidemiology and Community Health 65: 972–979. doi: 10.1136/jech.2010.111492 PMID: 21106546

68. Romero-Ortuno R, Fouweather T, Jagger C (2014) Cross-national disparities in sex differences in life expectancy with and without frailty. Age and Ageing 43: 222–228. doi: 10.1093/ageing/aft115 PMID: 23917483

69. Guralnik JM, Land KC, D. B, G.G. F, L.G. B (1993) Educational status and active life expectancy among older blacks and whites. N Engl J Med 329: 110–116. PMID: 8510687

70. Land KC, Guralnik JM, Blazer DG (1994) ESTIMATING INCREMENT-DECREMENT LIFE-TABLES WITH MULTIPLE COVARIATES FROM PANEL-DATA—THE CASE OF ACTIVE LIFE EXPECTANCY. Demography 31: 297–319. PMID: 7926190

71. Laditka SB, Laditka JN (2002) Recent perspectives on active life expectancy for older women. Journal of Women & Aging 14: 163–184.

72. Kaneda T, Zimmer Z, Tang Z (2005) Socioeconomic status differentials in life and active life expectancy among older adults in Beijing. Disability and Rehabilitation 27: 241–251. PMID: 16025751

73. Matthews RJ, Jagger C, Hancock RM (2006) Does socio-econornic advantage lead to a longer, healthier old age? Social Science & Medicine 62: 2489–2499.

74. Hidajat MM, Hayward MD, Saito Y (2007) Indonesia's social capacity for population health: the educational gap in active life expectancy. Population Research and Policy Review 26: 219–234.

75. Melzer D, McWilliams B, Brayne C, Johnson T, Bond J, et al. (2000) Socioeconomic status and the expectation of disability in old age: estimates for England. Journal of Epidemiology and Community Health 54: 286–292. PMID: 10827911

76. Beltran-Sanchez H, Andrade FC (2013) Educational and sex differentials in life expectancies and disability-free life expectancies in Sao Paulo, Brazil, and urban areas in Mexico. Journal of Aging & Health 25: 815–838.

77. Izmirlian G, Brock D, Ferrucci L, Phillips C (2000) Active life expectancy from annual follow-up data with missing responses. Biometrics 56: 244–248. PMID: 10783802

78. Montez JK, Hayward MD (2014) Cumulative Childhood Adversity, Educational Attainment, and Active Life Expectancy Among US Adults. Demography 51: 413–435. doi: 10.1007/s13524-013-0261-x PMID: 24281740

79. Schrijvers CT, Stronks K, van de Mheen HD, Mackenbach JP (1999) Explaining educational differences in mortality: the role of behavioral and material factors. American Journal of Public Health 89: 535–540. PMID: 10191797

80. Stowasser T, Heiss F, McFadden D, Winter J (2011) "Healthy, Wealthy and Wise?" Revisited: An Analysis of the Causal Pathways from Socioeconomic Status to Health. Investigations in the Economics of Aging: National Bureau of Economic Research, Inc. pp. 267–311.

81. Klijs B, Mackenbach JP, Kunst AE (2011) Obesity, smoking, alcohol consumption and years lived with disability: a Sullivan life table approach. Bmc Public Health 11.

82. Al Snih S, Ottenbacher KJ, Markides KS, Kuo YF, Eschbach K, et al. (2007) The effect of obesity on disability vs mortality in older Americans. Archives of Internal Medicine 167: 774–780. PMID: 17452539

83. Reynolds SL, McIlvane JM (2009) The impact of obesity and arthritis on active life expectancy in older Americans. Obesity 17: 363–369. doi: 10.1038/oby.2008.534 PMID: 19023280

84. Reynolds SL, Saito Y, Crimmins EM (2005) The impact of obesity on active life expectancy in older American men and women. Gerontologist 45: 438–444. PMID: 16051906

85. Walter S, Kunst A, Mackenbach J, Hofman A, Tiemeier H (2009) Mortality and disability: the effect of overweight and obesity. International Journal of Obesity 33: 1410–1418. doi: 10.1038/ijo.2009.176 PMID: 19786964

86. Diehr P, O'Meara ES, Fitzpatrick A, Newman AB, Kuller L, et al. (2008) Weight, mortality, years of healthy life, and active life expectancy in older adults. Journal of the American Geriatrics Society 56: 76–83. PMID: 18031486

87. Flegal KM, Kit BK, Graubard BI (2014) Body mass index categories in observational studies of weight and risk of death. Am J Epidemiol 180: 288–296. doi: 10.1093/aje/kwu111 PMID: 24893710

88. Ferrucci L, Izmirlian G, Leveille S, Phillips CL, Corti MC, et al. (1999) Smoking, physical activity, and active life expectancy. American Journal of Epidemiology 149: 645–653. PMID: 10192312

89. Tian X, Tang Z, Jiang J, Fang X, Wu X, et al. (2011) Effects of Smoking and Smoking Cessation on Life Expectancy in an Elderly Population in Beijing, China, 1992–2000: An 8-Year Follow-up Study. Journal of Epidemiology 21: 376–384. doi: 10.2188/jea.JE20110001 PMID: 21747208

90. Mahoney FI, Barthel DW (1965) Functional Evaluation: The Barthel Index. Maryland State Medical Journal 14: 61–65. PMID: 5832033

91. Jitapunkul S, Kamolratanakul P, Ebrahim S (1994) The meaning of activities of daily living in a Thai elderly population: development of a new index. Age & Ageing 23: 97–101.

92. Nagi SZ (1976) An epidemiology of disability among adults in the United States Milbank Memorial Fund Quarterly Health and Society 54: 439–467. PMID: 137366

93. Matthews FE, Miller LL, Brayne C, Jagger C, Mrc C (2006) Regional differences in multidimensional aspects of health: findings from the MRC cognitive function and ageing study. Bmc Public Health 6.

94. Campolina AG, Adami F, Santos JL, Lebrao ML (2013) Effect of eliminating chronic diseases among elderly individuals. Revista de Saude Publica 47: 514–522. PMID: 24346564

95. Hayward MD, Hummer RA, Chiu C-T, Gonzalez-Gonzalez C, Wong R (2014) Does the Hispanic Paradox in US Adult Mortality Extend to Disability? Population Research and Policy Review 33: 81–96. PMID: 25821283

96. Minicuci N, Noale M, Leon Diaz EM, Gomez Leon M, Andreotti A, et al. (2011) Disability-Free Life Expectancy: A Cross-National Comparison Among Bulgarian, Italian, and Latin American Older Population. Journal of Aging and Health 23: 629–681. doi: 10.1177/0898264310390940 PMID: 21220352

97. Peres K, Jagger C, Matthews FE, Mrc C (2008) Impact of late-life self-reported emotional problems on Disability-Free Life Expectancy: results from the MRC Cognitive Function and Ageing Study. International Journal of Geriatric Psychiatry 23: 643–649. PMID: 18044794

98. Branch LG, Guralnik JM, Foley DJ, Kohout FJ, Wettle TT, et al. (1991) Active life expectancy for 10,000 Caucasian men and women in three communities. Journals of Gerontology 46: M145–M150. PMID: 2071836

99. Kai I (1993) Active life expectancy and its determinants among the elderly in Japan. Bulletin of the Physical Fitness Research Institute 0: 19–27.

100. Katz S, Branch LG, Branson MH, Papsidero JA, Beck JC, et al. (1983) ACTIVE LIFE EXPECTANCY. New England Journal of Medicine 309: 1218–1224. PMID: 6633571

101. Laditka SB, Laditka JN (2014) More Education May Limit Disability and Extend Life For People With Cognitive Impairment. American Journal of Alzheimers Disease and Other Dementias 29: 436–447. PMID: 24413536

102. Reynolds SL, Haley WE, Kozlenko N (2008) The impact of depressive symptoms and chronic diseases on active life expectancy in older Americans. American Journal of Geriatric Psychiatry 16: 425–432. doi: 10.1097/JGP.0b013e31816ff32e PMID: 18448853

103. Rogers RG, Rogers A, Belanger A (1989) Active life among the elderly in the United States: multistate life-table estimates and population projections. Milbank Quarterly 67: 370–411. PMID: 2534561

104. Sauvaget C, Jagger C, Arthur AJ (2001) Active and cognitive impairment-free life expectancies: results from the Melton Mowbray 75+ health checks. Age and Ageing 30: 509–515. PMID: 11742781

105. Yi Z, Danan G, Land KC (2004) A new method for correcting underestimation of disabled life expectancy and an application to the chinese oldest-old. Demography 41: 335–361. PMID: 15209044

106. Jagger C, Matthews R, Matthews F, Robinson T, Robine JM, et al. (2007) The burden of diseases on disability-free life expectancy in later life. Journals of Gerontology Series a-Biological Sciences and Medical Sciences 62: 408–414. PMID: 17452735

107. Yong V, Saito Y (2012) Are There Education Differentials in Disability and Mortality Transitions and Active Life Expectancy Among Japanese Older Adults? Findings From a 10-Year Prospective Cohort Study. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 67: 343–353. doi: 10.1093/geronb/gbs029 PMID: 22421807

108. Hayward MD, Crimmins EM, Saito Y (1998) Cause of death and active life expectancy in the older population of the United States. Journal of Aging and Health 10: 192–213.

109. Jagger C, Matthews R, Melzer D, Matthews F, Brayne C, et al. (2007) Educational differences in the dynamics of disability incidence, recovery and mortality: Findings from the MRC Cognitive Function and Ageing Study (MRC CFAS). International Journal of Epidemiology 36: 358–365. PMID: 17255347

110. Manton KG, Stallard E, Corder L (1997) Education-specific estimates of life expectancy and age-specific disability in the U.S. elderly population: 1982 to 1991. J Aging Health 9: 419–450. PMID: 10182387

111. Sagardui-Villamor J, Guallar-Castillon P, Garcia-Ferruelo M, Banegas JR, Rodriguez-Artalejo F (2005) Trends in disability and disability-free life expectancy among elderly people in Spain: 1986–1999. Journals of Gerontology Series a-Biological Sciences and Medical Sciences 60: 1028–1034. PMID: 16127108

112. Manton KG, Gu X, Lowrimore GR (2008) Cohort changes in active life expectancy in the US elderly population: Experience from the 1982–2004 National Long-Term Care Survey. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 63: S269–S281.

113. Szwarcwald CL, da Mota JC, Damacena GN, Sardinha Pereira TG (2011) Health Inequalities in Rio de Janeiro, Brazil: Lower Healthy Life Expectancy in Socioeconomically Disadvantaged Areas. American Journal of Public Health 101: 517–523. doi: 10.2105/AJPH.2010.195453 PMID: 21233437

114. Liu X, Liang J, Muramatsu N, Sugisawa H (1995) Transitions in functional status and active life expectancy among older people in Japan. Journals of Gerontology Series B-Psychological Sciences & Social Sciences 50: S383–394.

115. Ofstedal MB, Zimmer Z, Cruz G, Chan A, Chuang YL (2004) Self-assessed health expectancy among older adults: A comparison of six Asian settings. Hallym International Journal of Aging 6: 95–117.

116. Diehr P, Newman AB, Jackson SA, Kuller L, Powe N (2002) Weight-modification trials in older adults: What should the outcome measure be? Current Controlled Trials in Cardiovascular Medicine 3. PMID: 12473163

117. van den Bos GAM, van der Maas PJ (1993) Social inequalities in the basic components of health expectancy: chronic morbidity, disability and mortality.; Robine J, Mathers C, Bone M, Romieu I, editors. Paris: Colloques INSERM/John Libbey Eurotext. 193–201 p.

118. Deeg DJH, Portrait F, Lindeboom M (2002) Health profiles and profile-specific health expectancies of older women and men: The Netherlands. Journal of Women & Aging 14: 27–46.

119. Crimmins E, Cambois E (2003) Social inequalities in health expectancy. Determing health expectancies. In: Robine J, Jagger C, Mathers C, Crimmins E, Suzman R, editors. Determing health expectancies. Chichester: John Wiley & Sons, Ltd: 111–126.

120. Jagger C, Cox B, Roy S.L, TEAM TE (2007) Health expectancy calculation by the Sullivan method: a practical guide, 3rd edition. EUROPEAN HEALTH EXPECTANCY MONITORING UNIT, France.

121. Berger N, Van Oyen H, Cambois E, Fouweather T, Jagger C, et al. (2015) Assessing the validity of the Global Activity Limitation Indicator in fourteen European countries. Bmc Medical Research Methodology 15.

122. Larson C, Mercer A (2004) Global health indicators: an overview. CMAJ Canadian Medical Association Journal 171: 1199–1200. PMID: 15534313

2.3 Discussion

2.3.1 Potential impact of ignoring grey literature

This systematic literature review was based on published articles and book chapters, collected via three electronic databases (Web of Science, Medline and Embase), and references in selected articles. This means that grey literature was not considered and this must be taken into account when interpreting the results.

The main source of grey literature on health expectancy comes from national websites and government reports, as many countries regularly produce and report estimates of health expectancy of their population. Excluding this evidence is likely to impact in particular the results on gender inequalities, because national statistics most commonly are available by sex and age groups. Estimates on health expectancies in some cases are available by region, and less frequently governments calculate health expectancy by socioeconomic or educational groups, or by ethnicity. The Office for National Statistics (ONS), for example, produces yearly estimates of HLE and DFLE, by gender and age, and by deprivation area, but data stratified by educational level or socioeconomic position are not available[1]. In New Zealand, the Ministry of Health and Statistics reported estimates of HE by ethnicity[2], and in the US, HE calculated by the National Center for Health Statistics are available by race group[3]. The European Health and Life Expectancy Information System (EHLEIS) has recently produced country reports on health expectancies and for each European country collected a list of publications and reports on health expectancies [4].

Therefore, the systematic literature review presented in this chapter must be read as an exhaustive summary of published works on health expectancies which focused on the older population, bearing in mind that additional evidence is available from grey literature sources, mainly consisting in government reports and national statistics. The exclusion of these sources may have had an impact particularly on the evidence reported with regards to gender inequalities. However, one of the main findings of this review was that all 95[5] studies,

[1] http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies

[2] http://www.health.govt.nz/system/files/documents/publications/independent-life-expectancy-new-zealand-2013-jul15-v2.pdf

[3] https://www.cdc.gov/nchs/data/misc/pophealth.pdf

[4] EHLEIS Country Reports Issue 10 (Technical_report_2017_4_1, EHLEIS team, April 2017)

[5] Only one study did not consider gender in the analysis [45].

which were set in different countries, included evidence of the existence of a gender paradox in health and mortality. I believe that this fact supports the validity and generalisability of this phenomenon, despite the potential loss of information arising from the exclusion of grey literature. At the same time, cross-country comparisons would benefit from including this source of information. However, it has to be pointed out that the problem of comparing results produced in different settings, i.e. using different measures of health and methods, would affect also the results derived from grey literature. Furthermore, grey literature is not peer reviewed and therefore the quality and reliability of the presented evidence is uncertain.

2.3.2 Conclusions

Some findings of this systematic literature review have been particularly relevant for the development of the thesis and are highlighted here.

1. **Heterogeneous definitions of disability:** as pointed out in Research Paper I, there is no consensus for defining health, and discrepancies occur also for measuring disability. A significant part of this thesis is devoted to conceptualise and measure disability. The conceptualization requires choosing and embracing a theoretical approach to define disability. To this aim, as described in Chapter Three, I have chosen a specific theoretical model of disability and relied on it consistently throughout the thesis. As a result, in all research papers, disability is interpreted and conceived in the same way. Its measurement derives from its conceptualization, and variables are selected accordingly. Different methods are adopted to model disability depending upon the research questions of each study. The theoretical conceptualisation of disability is presented in Chapter Three and challenges related to its empirical measurement are addressed in Research Papers II and III, and discussed in Chapter Eight.

2. **Gender paradox in health and mortality:** all studies included in the review agree on the finding that women live longer than men but spend larger proportions of their life in poor health -however health is defined- which is knwon in literature as gender paradox in health and mortality. At the same time, there is discord around the causes of the phenomenon and uncertainty over whether the gender paradox will continue to exist in the future. Both findings have significantly influenced the development of the

thesis. One of the explanations proposed to understand the gender paradox maintains that the gender gap is largely due to mortality differences favouring women, rather than differences in the onset of disability [46]. According to this hypothesis, females would have higher prevalence of nonfatal but disabling diseases, whilst males would be affected by higher prevalence of fatal diseases and chronic diseases strongly related to mortality [47]. Such a debate has inspired the research question of Research Paper II, which tries to explain the gender paradox in health and mortality asking whether the association of disability with mortality differs between women and men.

The choice of performing all the analyses of this thesis separately for men and women, although it is a common practice in most studies, has also been partly dependent on the special relevance of gender in our research. Only in Research Paper II, disability was measured for the whole sample, but in the sensitivity analysis its measurement was replicated separately by sex.

Chapter 3

Disability

3.1 Background

Conceptualizing and measuring disability is one of the main challenges of this thesis. In the systematic literature review presented in Research Paper I, it was pointed out how broad and heterogeneous its definition and measurement are across different studies. In Part II of this thesis, the work of three further papers is presented, and to put them into context, this chapter describes the theoretical approach that I have employed for measuring disability and gives the background for the development of the conceptualization of disability.

These components are included in Research Papers II-IV as follows. In Research Paper II, the theoretical framework of disability adopted and used throughout the thesis and the measurement model conceived accordingly are presented. In Research Paper III, I assess whether there exists a gradient in severity of disability and investigate how many levels are relevant for both descriptive and health policy planning purposes. In Research Paper IV, I use the results of the previous papers by examining how the selected categorisation of disability relates to mortality in order to provide empirical evidence for (or against) alternative theories of population health change over a decade specifically with regards to the English setting.

3.1.1 Models of Disability

A general definition of disability is included in the first chapter of the WHO's World report on disability [48], produced jointly with the World Bank. *"Disability is complex, dynamic, multidimensional, and contested"*. This definition, as general as inclusive, discloses the comprehensive approach and the WHO's intention of going beyond the models most commonly used for describing disability. The most frequently mentioned models are the medical and social models.

The **medical model** of disability recognises disability as a physical or mental impairment pertaining to the individual and directly caused by disease, trauma or other health conditions. Solutions to disability are found by focusing on the person and require medical care in the form of individual treatment. Figure 3.1 graphically describes the main features of the medical model and highlights the individual dimension of disability.

Figure 3.1: Medical model of disability

Source: `http://www.making-prsp-inclusive.org/en/6-disability/61-what-is-disability/611-the-four-models.html` adapted from Harris and Enfield [49, p. 172]

The **social model** of disability views disability as an incapability or impossibility of disabled subjects to interact with society, and therefore the causes of disability are identified in external factors and interpreted as societal barriers rather than individual limitations [50]. The social model distinguishes between impairment and disability. Impairment is meant as temporary or permanent conditions with which an individual lives, which may affect their appearance, or the functioning of their mind or body, and cause pain, fatigue, affect

communications, or interfere with mental capacity. Disability is not a condition a person has, but rather something he/she experiences due to external factors, and as such is often avoidable. For example, if a wheelchair user is unable to get into a building because of some steps, the social model of disability sees the steps as the disabling barrier, and a ramp would be the solution to get access into the building immediately. The medical model, in contrast, would suggest that this is because of the wheelchair, rather than the steps. Figure 3.2 illustrates the main features of the social model and shows the main social drivers that determine disability.

Figure 3.2: Social model of disability

Source: http://www.making-prsp-inclusive.org/en/6-disability/61-what-is-disability/611-the-four-models.html adapted from Harris and Enfield [49, p. 172]

There is a variety of other models used to describe and explain disability, although not as popular as the medical and social models. Models differ depending on how (and whether) medical, social, political, cultural, economic and other factors are taken into account and interpreted to understand disability. Among the existing models, the **charity model** sees disability as a deficit and people with disabilities victims of their impairment and dependent on other persons. The **rights-based** model is similar to the social model and is based on the concepts of empowerment, which means that people with disabilities have the right to participate in the society as active stakeholders, and of accountability, that refers to the duty of society to implement these rights.

The medical and social models have often been considered antithetical, although the

medical model does not discourage staging inclusive interventions towards disabled people, but, simply, it relates disability to the individual and does not consider society as part of or solution to the condition. The medical model is limited by not looking at social factors that influence the experience of the person. In so doing, it puts the responsibility of disability on the person alone. Similarly, the social model in no way rejects the idea of a person seeking medical intervention to minimise the impact of their impairment, but does not attribute the causes of disability to individual conditions. The problem, in this case, is that this view sometimes is unrealistic. Someone with dementia, for example, will not be able to participate fully no matter what social structures are put into place. The two models, in fact, can be considered complementary rather than mutually exclusive. The International Classification of Functioning Disability and Health (ICF) combines the different features of the medical and social models and adopts a comprehensive view on disability. For this reason, it is also described as a "bio-psycho-social model".

3.1.2 Conceptual schemes of disability preceding the ICF

In the past decades, three main conceptual schemes have offered a theoretical interpretation of disability preceding the development of the WHO's ICF framework. A conceptual scheme is a rudimentary scientific model that guides terminology, measurement and hypotheses. It is the basic architecture on which research, policy, and clinical care are built [51]. These schemes are, in chronological order, the "disablement model" proposed by Nagi [52] in 1965, the International Classification of Impairments Disabilities and Handicaps (ICIDH) [53] developed in the 1970s and issued in 1980, and the "disablement process" proposed by Verbrugge and Jette [51] in 1994.

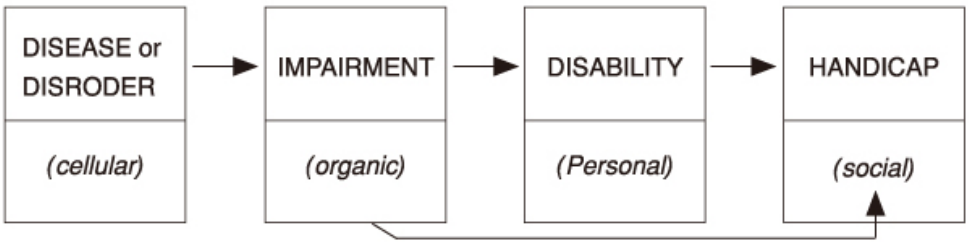
In the 1970s, sociologist Saad Nagi shifted the approach to disability from a medically based to a socially based interpretation, changing the focus from the physical consequences of pathologies to the dynamic process that leads to functional consequences. He renamed this process as "disablement" to highlight its dynamic nature. He defined disability as an "expression of a physical or a mental limitation in a social context, a gap between the individual's capabilities and the demands created by the physical and the social environment" and described disablement through four concepts: active pathology, impairment, functional

limitation, and disability.

The ICIDH described the stages in the clinical course of a disease, from the onset of disease to the consequent appearance of impairment, disability and handicap -not necessarily in this order (see figure 3.3). To illustrate figure 3.3 with an example, a child with coeliac disease, who is functionally limited, may be able to live a fairly normal life and not suffer activity restriction; he could nevertheless suffer disadvantage by virtue of his inability to partake of a normal diet [54, pag. 30]. This example shows the possibility of interrupting the sequence at any stage. Therefore, one can be impaired without being disabled, and disabled without being handicapped. The concepts of impairment and disability cover essentially the same scope as Nagi's disability, while the handicap component did not have a parallel concept in Nagi's disablement process. According to the ICIDH, impairment consisted of a deviation from normal organ functioning, and was defined as "any loss or abnormality of psychological, physiological, or anatomical structure or function"; disability was defined as "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being". An impairment does not necessarily lead to a disability, for the impairment may be corrected. Finally, handicap was described as "a disadvantage for a given individual, resulting from an impairment or a disability, which limits or prevents the fulfilment of a role that is normal (depending on age, sex, and social and cultural factors) for that individual." Handicap considers the person's participation in their social context.

The third model was proposed in 1994 by Verbrugge and Jette [51] and was presented as a

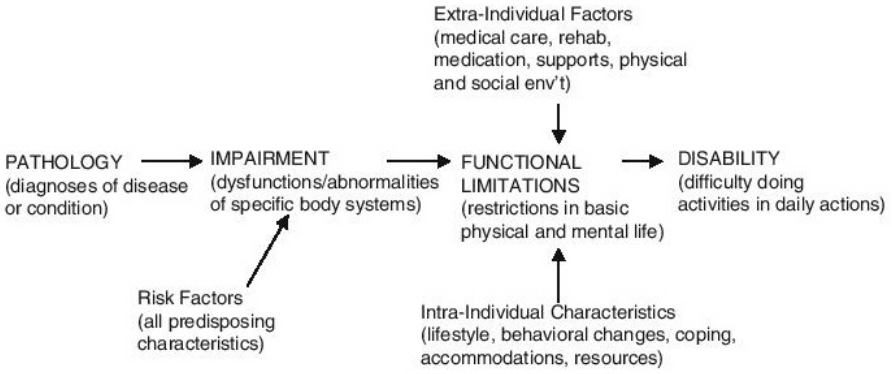
Figure 3.3: ICIDH's stages



socio-medical model, called the "disablement process". It described how chronic and acute conditions affect functioning in specific body systems, generic physical and mental actions,

and activities of daily life, as well as the personal and environmental factors that speed or slow disablement. It maintained the Nagi's original concepts but extended his model by specifying dimensions of disability and including in the model relationships between sociocultural factors and personal factors and the core disablement concepts. The term process is described as "the dynamics of disablement, that is, the trajectory of functional consequences over time and the factors that affect their direction, pace, and pattern of change" as described in figure 3.4

Figure 3.4: Model of the Disablement process



Source: A model of the "Disablement Process" proposed by Verbrugge and Jette [51]

3.1.3 International Classification of Functioning, Disability and Health

The ICF framework is concisely described is Research Paper II and mentioned in Research Paper III and IV. In this section it is explained and presented in more detail.

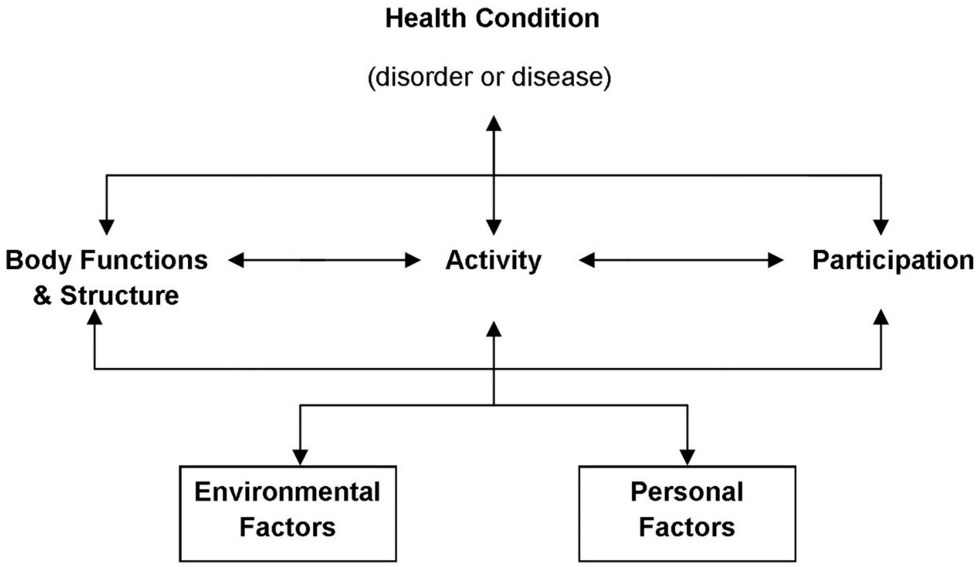
The ICF was officially endorsed by all 191 WHO Member States in the 54^{th} World Health Assembly on 22 May 2001, as the international standard to describe and measure health and disability. It was developed through a long process involving academics, clinicians and people with disability [48]. It represents an advancement of the previous ICIDH (1980). Some criticism was raised within the academic literature against the limitations of clinical and other uses of the ICIDH [55]. Impairment, disability and handicap categories were poorly defined and problems in classifying disabled individuals into a certain group arose because the categories did not cover the whole spectrum of disability or conversely

because they overlapped. Moreover, the ICIDH model was not adequate to describe all types of disability, and problems occurred especially with mental health conditions. Finally, external factors were completely absent in the ICIDH model. The ICF advances the ICIDH model especially addressing this last point: the inclusion of environmental factors in creating disability is the main difference between the new ICF and the previous ICIDH.

According to the ICF, disability is composed of three interconnected areas, consisting of impairment, activity limitations and participation restrictions. The impairment category is also called body functions and structure and consists of problems with or alteration of physical and mental domains, such as significant deviation or loss. Activity limitations are difficulties in executing basic tasks or activities that a person does on a daily basis, for example difficulties in moving or reading. Participation restrictions are problems in interacting in the society due to disability, for example restricted participation in employment and using public transport. As shown in figure 3.5, disability consists in any or all these three dimensions. Health conditions, on the other hand, are underlying conditions that may cause disability or appear during disability. The disability mechanism is determined by contextual factors which interact with health conditions. Contextual factors are environmental and personal factors. The former consist of the physical, social and attitudinal environments in which people live and conduct their lives. These are either barriers to or facilitators of the person's functioning. They are not within the person's control, such as family, work, government agencies, laws, and cultural beliefs. Personal factors may include gender, age, race, lifestyles, habits, education and profession. They represent influences on functioning particular to the individual which are not represented elsewhere in ICF. An example of this is when an individual cannot get a job due to lack of qualifications, rather than any difficulty in functioning or problem in the environment. To exemplify the dynamics illustrated in figure 3.5, let us consider the example of experiencing a cataract in one or both eyes as a health condition, which is not disability per se, but may lead to blindness (i.e. body function) and therefore to limited mobility (i.e. activity limitation) and may prevent the individual from working (i.e. participation restriction). The described mechanism, however, can be dramatically different depending, for example, upon the availability of information in format such as braille and reading tools, such as voice over text

(i.e. environmental factor) and/or retraining for another career (i.e. personal factor).

Figure 3.5: Representation of the ICF.



Source: World Health Organization, Geneva 2011 [48]

3.2 Measurement challenges

The ICF is an international standard to measure disability, and enhances comparability of studies and different settings. The drawback of the ICF's comprehensiveness and broad definition is the difficulty it imposes on measuring disability. First, it is necessary to identify adequate variables to capture the three areas of functioning (impairment, activity limitations and participation restrictions). Then, to combine them together and produce a single indicator of disability. Furthermore, there is no international standard for how this should be done. No gold standard questionnaire exists. The widely promoted questionnaires instead focus on components (e.g. the Washington Group focuses on activities, see next section).

3.2.1 Washington Group Questions

To facilitate the comparison of data on disability cross-nationally, the Washington Group (WG) on Disability Statistics was formed in June 2001 -less than a month after the ICF was

officially endorsed by the World Health Assembly- with the intent of providing statistical and methodological support at international level. The WG has developed, tested internationally, and adopted a short and an extended set of disability measures suitable for use in censuses, sample-based national surveys, or other statistical formats. The WHO's ICF was used as the basic framework for the development of these two sets.

The "short" set of questions was developed with the idea of balancing the importance of collecting information on all aspects of the disablement process, and the difficulty of doing so in censuses or surveys not dedicated to disability. As already pointed out, the survey used for this thesis offers a wide set of questions covering different dimensions of health, and therefore it has been possible to go beyond the WG short set of questions selected to measure disability. However, in Research Paper III, the WG questions are used to validate the disability indicator, and therefore they are briefly described here.

The short set of questions consists of six questions which identify the key variables to measure disability according to the ICF framework. Answers to each question are multiple-choice and include "no difficulty", "yes - some difficulty", "yes - a lot of difficulty" and "cannot do at all". The six WG questions are listed in Table 3.1 and the most similar questions available in ELSA are presented in the adjacent column. The selection of questions was done in order to (i) represent the majority, but not all persons with limitations in basic activity functioning in any one nation; (ii) represent the most commonly occurring limitations in basic activity functioning within any country; (iii) capture persons with similar problems across countries [56].

3.2.2 Measurement of ICF-based disability

In this thesis, disability has been measured in accordance with the ICF framework using data from the English Longitudinal Study of Ageing (ELSA) (see Chapter Four). First, I selected a broad number of self-reported questions covering the impairment, activity limitations and participation restrictions domains. Variables were collected from the Computer-Assisted Personal Interview (CAPI) questionnaire with the requirement of being available in other waves and asked in a similar way. I identified a list of variables of interest and classified them

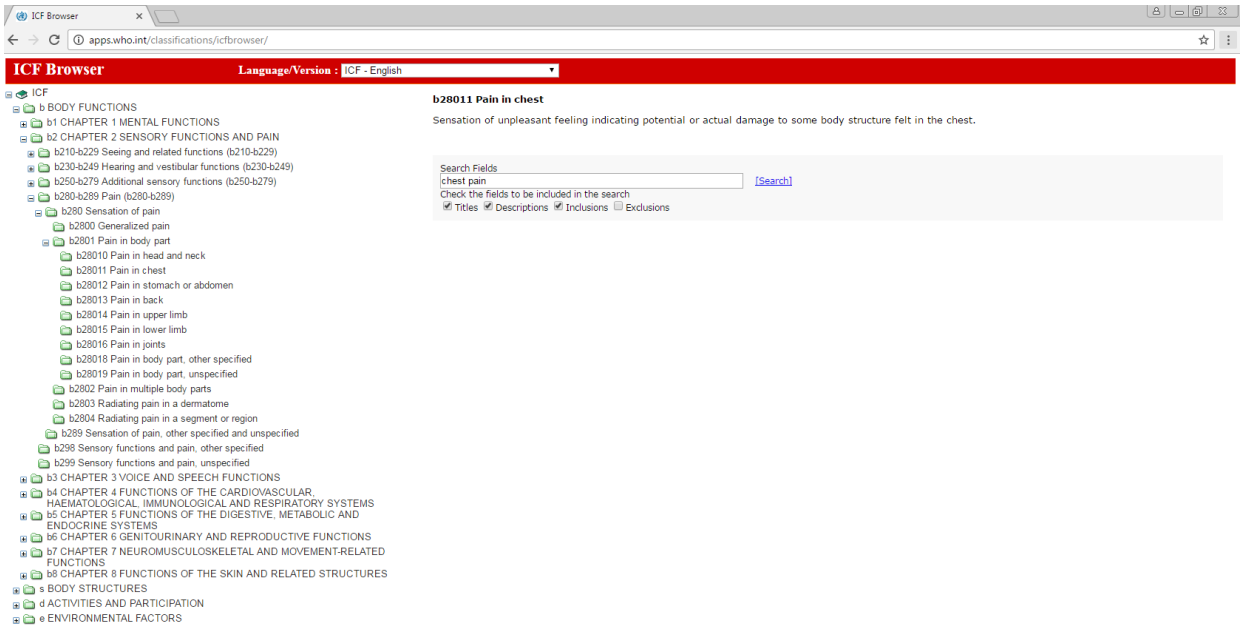
Table 3.1: Short set of WG questions

WG Questions	**Questions asked in ELSA**
1) Do you have difficulty seeing even if wearing glasses?	Is your eyesight (using glasses or corrective lens as usual) [excellent to poor]
2) Do you have difficulty hearing even if using a hearing aid?	Is your hearing (using a hearing aid as usual) [excellent to poor]
3) Do you have difficulty walking or climbing stairs?	By yourself and without using any equipment, how much difficulty do you have walking for a quarter of a mile?
4) Do you have difficulty remembering or concentrating?	How would you rate your memory at the present time?
5) Do you have difficulty with (self-care such as) washing all over or dressing?	Problems in at least one activity among: dressing, bathing or showering, eating, using the toilet
6) Using your usual language, do you have difficulty communicating (for example understanding or being understood by others)	Do you find it difficult to follow a conversation if there is background noise, such as TV, radio or children playing?

in one of the disability domains using the ICF browser. The ICF browser (see figure 3.6) is an online browser where any conditions included in the ICF framework are classified, coded and defined. Each domain is organised in a hierarchical structure of four levels, from more general to more specific classifications [57]. As described in Research Paper II, two authors (independently from each other) classified each variable into one of the ICF categories, i.e. impairment, activity limitations and participation restriction. In case of disagreement there was discussion, and if needed a third person was involved and the ultimate classification was based upon common agreement. The final list of selected variables was fully classified and coded within the ICF hierarchical scheme (see table B.1 in the appendix to the thesis at Chapter Ten). Second, I produced an indicator of disability. The measurement of disability, based on the selected items, was one of the aims of Research Papers II and III. Methods and results are fully described in the manuscripts presented at Chapters Five and Six. A criticism of this approach, however, is that it is dependent on the questions asked in the survey, and in this case in ELSA.

Unlike the approach adopted in this thesis, however, often only a much smaller set of variables is available to capture disability. In a number of studies run by the International Centre for Evidence in Disability (ICED), mainly in Low Middle Income Country (LMIC)

Figure 3.6: WHO ICF browser



Example of search of *"chest pain"* and corresponding classification (b28011)

[58, 59], a comprehensive population-based survey methodology compatible with the ICF was developed. The authors distinguished and compared three approaches for measuring disability, and explored their relationship within the context of the ICF, and assessed how the different approaches to measuring disability inter-related. (i) The first approach consisted in a single direct question on disability, such as "do you have a disability?". This is known to lead to under-reporting, due to stigma and cultural perceptions of disability, and is not considered adequate [59, 60] (ii) The second approach was based on multiple self-reported functional limitations in core domains of function. In this case the focus is on the "activities" component of the ICF. (iii) The third approach was to objectively measure clinical impairments or presence of specific, potentially disabling, health conditions. This approach focuses on the "body function and structure" (i.e. impairment) component of the ICF. Impairment data alone, however, do not capture how the individual functions in his or her environment and the overall disability experience. The authors found that the single question on disability leads to significant under-reporting. Using a self-reported activity limitation tool alongside clinical tools to measure specific impairments and health conditions showed a high proportion of participants screening positive to moderate/severe clinical

impairments and health conditions but not reporting significant activity limitations. The results of the study just described exemplify the complexities posed by measuring disability, even when a specific conceptual framework is selected. Similarly, as remarked in the Handbook of Aging and the Social Sciences, in Chapter Four [61], some research has shown a lack of correspondence between individual characteristics that might generally be interpreted as indicators of having a disability, and the perception that one actually has a disability. For example, a study by Iezzoni et al. [62] examining perceptions of disability among people with lower-extremity mobility difficulties showed that many people with serious mobility problems do not view themselves as disabled, so that among persons with major mobility problems, 70.8% perceived themselves as disabled, and among manual wheelchair users the percentage was 80.5%.

For explanatory purposes, the table presented as part of the ICED work [58] to summarize the three approaches for measuring disability is reproduced below (table 3.2), adapting its content to the setting of this thesis. Examples of possible variables available in ELSA for each of the approaches are listed, and the pros and cons (columns 3 and 4) are those identified by authors and are general considerations which determine the choice of one method over another depending on circumstances (e.g. census vs health survey).

Table 3.2: Approaches for measuring disability (adapted from ICED (2014))

Type	**Example in ELSA**	**PROs**	**CONs**
Direct Questioning	Do you have any long-standing illness, disability or infirmity? By longstanding I mean anything that has troubled you over a period of time, or that is likely to affect you over a period of time?	Rapid and limited space needed	Underreport (stigma and lack of self-identification)
Self-reported for specific components	**Impairment:** How good is your eyesight for seeing things at a distance? **Activity:** By yourself and without using any equipment, how much difficulty do you have walking for a quarter of a mile? **Participation:** Why don't you use public transport more often? *My health prevents me*	Simple to administer, information on experience and impact of the condition	Does not assist planning for assessment of services/needs.
Objectively measured indicators	**Impairment:** grip strength measured using the Smedley dynamometer. **Activity:** Observed and timed walking without/with help of another person or using support.	Information on the type of impairment, severity and causality for intervention	Resource intensive, only one component of impairment/activity is captured.

3.3 Short summary

In this chapter, the theoretical background for understanding disability was presented, including the main conceptual schemes that have preceded the development of the WHO's ICF framework, which currently is the predominant framework for disability [63]. Disability is a broad concept, and is influenced by impairment, environmental and personal factors. Focusing on any single component of disability will mean that the whole picture cannot be evaluated. Therefore, in the thesis I focus on the full ICF framework for disability, and will be interested in impairment, activities and participation.

Alongside the theoretical approach to disability adopted in this thesis, I introduced some challenges that arise when tackling its measurement. Findings from some of the research investigating the relationship between distinct measures of disability were presented, and showed that single direct questions on disability, multiple self-reported items of impairment, activity limitations and participation restriction, and objectively measured indicators often do not match, and persons having activity and participation restrictions and moderate or severe clinical impairments are not uniquely identified. For each measurement approach, some of the variables available in the ELSA were presented (tables 3.1 and 3.2). The wide range of information on health covered in the ELSA questionnaires allows adopting different approaches to measure disability, but there are also some limitations in translating certain theoretical schemes of disability using this survey. In particular, when attempting to reproduce the short set of six questions selected by the WG, there was some inconsistency for the question pertaining to communication. This domain was assumed to be measured by a variable capturing hearing conditions rather than communication problems. A sensitivity analysis was conducted in the Supplementary Information to Research Paper III (Chapter 6.3) excluding this domain from the short set of questions, and it was found that results were not affected by whether our proxy for communication was or was not included.

The measurement of disability and the set of self-reported variables selected to capture impairment, activity limitations and participation restrictions are described in details in Part II of the Thesis (Chapters Five and Six).

Chapter 4

Research Background

This thesis relies entirely on secondary data from the English Longitudinal Study of Ageing (ELSA). In each empirical research paper (i.e. Research Papers II, III and IV), an overview of the data and a description of the sample selected for the study are provided succinctly due to the word limits imposed by journals. In this chapter, ELSA is presented in more detail and descriptive analyses of ELSA's representativeness of its target population, in particular in terms of disability and mortality, are performed.

4.1 English Longitudinal Study of Ageing

4.1.1 Background and rationale

In the mid-1980s, anticipating the rapid ageing of the population known as the Baby Boom generation, scientists at the National Institute on Aging (NIA) and elsewhere recognized the need for developing surveys focused on older populations which would be able to adequately address contemporary ageing-related issues by adopting a multi-disciplinary approach. Consequently, the Health Retirement Study (HRS) was launched in 1992, with the purpose of collecting information on income, work, assets, pension plans, health insurance, disability, physical health and functioning, cognitive functioning, and health care expenditures among approximately 20,000 Americans over the age of 50 [64]. The success of the HRS and the growing relevance of population ageing worldwide led to the development of a number of

"sister" studies in several countries[1], including both middle and high-income countries.

One of the sister studies was ELSA. The ELSA has been established to collect longitudinal multidisciplinary data from a representative sample of the English population aged 50 and older (`http://www.elsa-project.ac.uk/about-ELSA`), to explore the dynamics of ageing, to inform policy debates and for enhancing comparative analysis with similar studies, such as the HRS. The ELSA has been developed through a collaboration between University College London, the Institute of Fiscal Studies and the National Centre for Social Research, with academics at the Universities of Cambridge, Nottingham and East Anglia and from the HRS. Funding for data collection for the early waves of the study was provided by the NIA and a consortium of British Government Departments.

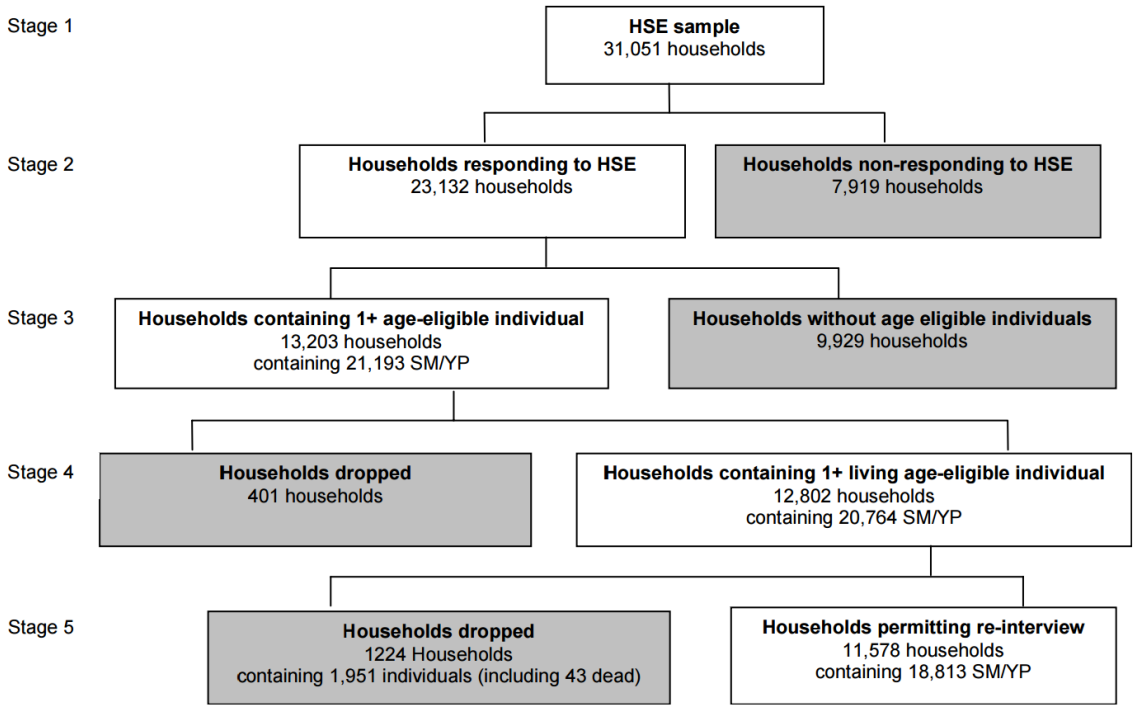
4.1.2 Sample design

The ELSA sample was designed to represent people aged 50 and over, living in private households in England and was selected from households that had previously responded to the Health Surveys for England (HSE) in 1998, 1999 or 2001 [65]. Figure 4.1 shows the selection process from the three-year HSE samples to the ELSA sample at wave 1. Background information about the HSE and description of the two-stage sampling design used to select the HSE sample are available in the ELSA Technical Report [65]. The process of selecting the ELSA sample comprises five stages as shown in figure 4.1. The first consists of the sample of households issued for the HSE. The final ELSA sample, descending from the HSE sample, consists of households that responded to the HSE (Stage 2), including at least one age-eligible individual (Stage 3), who, according to administrative records, was alive (Stage 4) and gave permission to be re-contacted in the future (Stage 5).

Therefore, households included in the ELSA contained at least one adult of 50 years or older who had agreed to be re-contacted at some time in the future. Within households there were three different types of individuals eligible to be invited to take part in the ELSA study:

[1]English Longitudinal Study of Ageing (ELSA), Mexican Health and Aging Study (MHAS), Survey of Health, Ageing and Retirement in Europe (SHARE), New Zealand Health, Work and Retirement Survey, Korean Longitudinal Study of Aging, WHO's Study on Global Ageing and Adult Health (SAGE), SHARE Israel, The Irish Longitudinal Study on Ageing (TILDA), Longitudinal Aging Study in India (LASI), the Chinese Health and Retirement Survey (CHARLS), the Japanese Study of Aging and Retirement (JSTAR), the Northern Ireland Cohort for the Longitudinal Study of Ageing (NICOLA)

Figure 4.1: The ELSA sample definition



Source: ELSA Technical Report [65, p. 9] `http://www.ifs.org.uk/elsa/report03/w1_tech.pdf`
SM=sample member; YP=younger partner

- Core sample members: individuals living within the household at the time of the HSE interview and born on or before 29 February 1952. The individuals were eligible if they were living in a private residential address in England at the time of the ELSA interview.

- Younger partners: the cohabiting spouses or partners of core sample members, who were living within the household at the time of the HSE interview and were born after 29 February 1952. They were invited to take part if they were still living with an eligible core sample member.

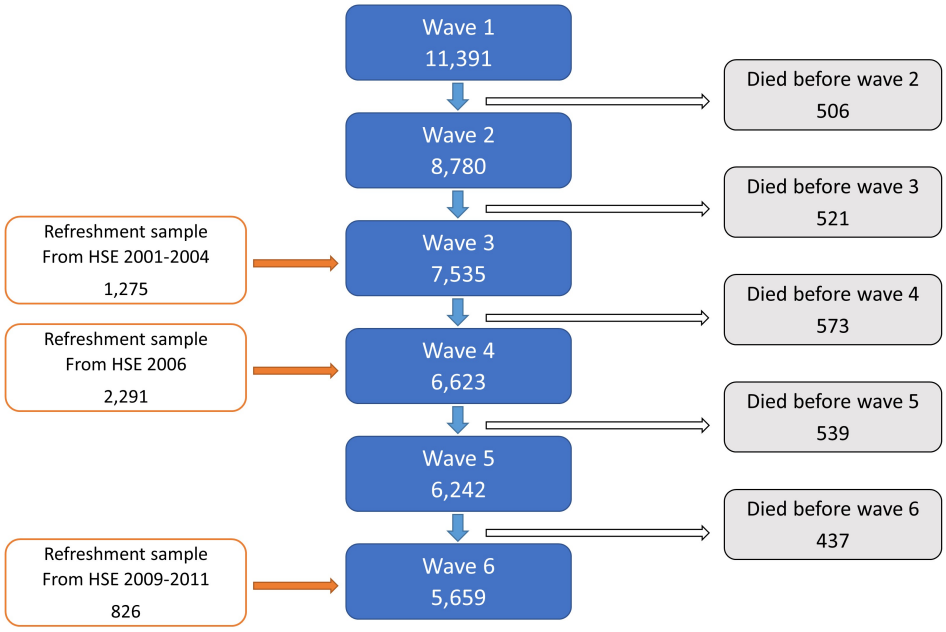
- New partners: the cohabiting spouses or partners of core sample members at the time of the first ELSA interview, who had joined the household since the HSE interview.

Only eligible core sample members who responded to the ELSA survey were part of the baseline sample for the analyses in this thesis. Younger and new partners were excluded

from analysis.

Respondents were first interviewed in 2002/2003. Since then, they have been approached for interviews every two years for a total of six waves through 2012/2013. When the study started, participants were aged 50+ years; with the progression of the study, it became necessary to refresh the sample to maintain an appropriate proportion of participants in their early 50s and facilitate cross-cohort comparison at later waves. Hence at waves 3, 4 and 6 the ELSA sample was refreshed selecting new participants respectively from HSE 2001-2004, 2006 and 2009-2011 who were previously too young to join ELSA in 2002. Figure 4.2 describes how the core members were expanded and followed up from wave 1 to wave 6.

Figure 4.2: Overview of data collection in ELSA Waves 1 to 6



Sample sizes are for the core members.

4.1.3 Survey content

At every wave, the study participants respond to a face-to-face interview consisting of a Computer-Assisted Personal Interview (CAPI) and a Self-Completion (SC) questionnaire. Respondents are asked about a broad range of topics related to family and work, economic

issues, physical and mental health, social and psychological factors, behaviour, cognition and biology. Every two waves (i.e. waves 2, 4 and 6) a nurse visit is carried out to collect biomarkers and more detailed measures of physical functioning. At wave 3 respondents were invited to participate in a life-history interview collecting information about events that had occurred previously in their lives; at wave 5 a risk and time preference module, designed to measure preferences for willingness to bear risk in pursuit of possible rewards versus risk aversion, was added and answered by respondents aged 50-75. Table C.1 in the appendix gives a brief overview of the content of the questionnaire (CAPI and SC) by module.

4.1.4 Advantages and disadvantages of using ELSA

Using ELSA data for the analyses presented in this thesis provides three central advantages. First, available data of ELSA cover a period of about ten years (from 2002/03 to 2012/13) allowing longitudinal analyses to assess population changes occurred over the course of the first decade of the 21^{st} century. The second advantage of using ELSA is that it provides a large number of self-reported health indicators and every two waves the core members are offered a nurse visit where objective biological information and measures of physical functioning are collected. Therefore, self-reported measures of health can be combined and validated with objectively measured variables. Third, ELSA data are linked to national mortality register and this allows combining disability questions asked in the survey with mortality data from register for all respondents who consented to link their data to external sources.

Alongside with strengths, some challenges emerge using ELSA. First, as in all longitudinal studies, attrition is an ongoing issue also in ELSA. Attrition poses a threat to the representativeness of longitudinal studies. If missingness is Missingness At Random (MAR), then observed variables account for selection. If principled methods are employed exploiting the drivers of missingness, estimated effects will be unbiased. On the contrary, if the data are not missing ar random (Non Missingness At Random (NMAR)), results will be biased. Maximising the plausibility of the MAR assumption increases the likelihood of a survey being representative of its target population despite the presence of missing data. Therefore, attrition is a threat, but can be minimized under reasonable assumptions in most cases.

Secondly, the use of the HSE as a sampling frame has some advantages, but also disadvantages. Some of the drawbacks are reported in the ELSA technical report; however, one of the main limitations is that ELSA is targeted on individuals living in private dwellings, and therefore individuals living in non-private accommodation, including institutionalised populations, are not included in the sample. This limits the representativeness of the sample and prevents extending results to the general population. However, individuals interviewed at baseline and living in private households who moved into a residential care home or similar establishment over the course of the study are followed up and an "institution interview" is sought. This partly extends the representativeness of ELSA longitudinally. Many national surveys do not include individuals living in institutions, and this was highlighted among the results of the systematic literature review in Chapter Two. I acknowledge and comment on this limitation in the discussion and provide evidence of the implication of excluding/including institutionalised population in terms of representativeness and generalizability to total population in section 4.2 of this chapter.

4.2 Analysis of ELSA representativeness

As outlined in the previous sections, ELSA is representative of individuals living in private dwellings and any English institutionalised individuals of the same age are not included at baseline, but only considered at later stages if the original participants move into a residential care home or similar establishment over the course of the study. Therefore, ELSA's representativeness partly changes during its evolution.

In this context, the main focus of this section is to understand how the exclusion of the institutionalised population affects the representativeness of the ELSA with regards to the general English population aged 50+. In particular, I assess whether studies based on the ELSA are valid only for a specific target population (i.e. non-institutionalised English individuals aged 50+) or can be partly extended to the general population (i.e. all English individuals aged 50+, regardless of residence type).

The issue is particularly relevant when dealing with mortality and morbidity/disability data. This is because information on each of these dimensions often comes from different sources. Information on mortality is usually available from censuses and national registries

for the entire population, while individual health records are less widely available. Therefore, commonly, studies combine mortality data obtained from national registers with morbidity data derived from representative surveys -that often do not include institutionalised populations. In recent years, a number of longitudinal studies have targeted the older population and adopted a multi-disciplinary approach to collecting information on socioeconomic position, demographic characteristics and health conditions. They are designed to follow up health conditions of participants over time and often to also gather prospective information on deaths. This represents a powerful source of information, because for each respondent data on mortality and health are jointly available. Problems however arise when follow up time is limited, or when using the last waves of the study as a baseline for which follow up information is not yet available.

Given this context and the research content of this thesis, I compared mortality and disability among the ELSA respondents and the general English population aged 50+ using data of the Office for National Statistics (ONS), in order to appraise the differences between the survey sample and census-based population. Census data were derived from national register for the years 2002 and 2012, and ELSA data came from wave 1 (2002/2003) and wave 6 (2012/2013). The ultimate goal was to understand to what extent the ELSA followed-up sample differs from the general English population. In particular, the relevance of this analysis for the outcomes of this thesis derives from the fact that: (i) mortality and disability were the variables of interest in each of the three empirical papers. (ii) In Research Papers II and III, mortality data came from the survey and was the dependent variable to be related to disability; in Research Paper IV the outcome of interest was an aggregate measure of mortality and disability, and moratlity rates were derived from population estimates and deaths collected by the ONS in the national registers.

In the next sections, I describe the approaches used to compare ELSA with national statistics, show the results and discuss the implications.

4.2.1 Methods

Sample

For the purposes of this analysis, I used the sample composed of core ELSA members at wave 1. This consisted of 11,391 individuals. Among them I selected only respondents who consented to link their data to mortality records of the National Health Service Central Register (NHSCR) to compare ELSA mortality rates with the national mortality rates. This led to 10,771 observations, with 544 respondents excluded because either they did not give their permission or other death-related information (i.e. year of death) were not available, and 76 respondents because they died the same year they entered the study (note that information on time of death was available only by year). To compare disability prevalence in ELSA with national statistics, I used ELSA core members interviewed at wave 1 who answered the selected disability question, corresponding to 11,382 observations while national disability data were obtained from the 2001 Census communal establishment data in England.

Measures

Mortality

Two different sources of information were used to obtain age-specific mortality rates.

1. National register data (from the ONS): mortality rates were calculated from data on the mid-year population of England and the number of deaths occurred from 2003 to 2011, for each year, stratifying by sex and age.

2. ELSA data: mortality rates were estimated on the ELSA participants, using data covering the period from wave 1 (2002-03) to wave 6 (2012-13) and including deaths registered after wave 5 and before wave 6 was issued, excluding deaths occurred the same year the survey participant was interviewed. This corresponded to 2,288 deaths linked to NHSCR.

Disability

Disability was defined as having a limiting long-lasting illness. This definition was chosen because the same question was asked both in ELSA and in the 2001 Census. In ELSA,

respondents answered two questions: "Do you have any long-standing illness, disability or infirmity? By longstanding I mean anything that has troubled you over a period of time, or that is likely to affect you over a period of time?". If they replied yes, they were asked "(Does this/Do these) illness(es) or disability(ies) limit your activities in any way?". In the 2001 Census, each person living in a household or resident in a communal establishment was asked "Do you have any long-term illness, heath problem or disability which limits your daily activities or the work you can do? (include problems which are due to age)".

Analysis

To compare the ELSA mortality rates with the national mortality rates, period- and age-specific Standardised Mortality Ratio (SMR) were calculated by sex. SMRs are the ratio of observed deaths in the study group to the expected deaths in the study group if they had the same mortality rates as the general population. To compute SMRs the age groups in the range 50-64 years were pooled together due to few deaths observed, and the 10-year follow-up of ELSA was split into two periods, 2003-2007 and 2008-2011. Approximate 95% confidence intervals are given by the range $\left[\frac{SMR}{EF}, SMR \times EF\right]$. The error factor, EF, is equal to $exp\left(\frac{1.96}{\sqrt{D}}\right)$, where D is the number of observed deaths in a particular sex, age and period combination. A chi-square test was performed to test whether each sex-, period- and age-specific SMR differed significantly from one. For disability, age- and sex-specific prevalence in the ELSA sample and in the general population when excluding or including people living in institutional residence were compared descriptively.

4.2.2 Results

Of the 10,771 respondents selected for the first set of comparisons, 54.6 percent were women and 45.4 percent men. All age groups were similarly represented across genders, with more females than men aged 80 plus. The deaths registered during follow-up among the ELSA participants concerned men more than women: 11 percent of men aged under 70 at wave 1 (374) and 52 percent of men aged 70+ at wave 1 (790) died, while women's percentages were 6.5 (254) and 43.5 (870), respectively.

Table 4.1 shows the estimated SMRs by age groups, gender and period and their re-

spective 95% confidence intervals. The chi-squared test, compared with the χ^2 distribution with one-degree of freedom, was produced to test whether the SMRs of each age group and overall differ significantly from one (not shown in the table). A SMR significantly different from one and lower than one would indicate mortality rate is lower among ELSA participants, and a value greater than one would indicate mortality rate is higher among ELSA participants. Results revealed some differences between the observed and expected number of deaths, mostly occurring in the first time period. Looking at the SMRs for all ages, in 2003/2007 the SMR was lower than one for females (SMR=0.76, 95% CI 0.73-0.86, p-value < 0.001) as well as for males (SMR=0.9, 95% CI 0.83-0.97, p-value = 0.0068). In 2008/2011, no SMR was significantly different from one (apart from men aged 70-74), across all age groups and both for men and women, indicating that during these periods (corresponding to wave 6) the ELSA mortality rates reflected accurately the general population mortality rate.

Table 4.1: SMRs by age, gender and period.

Women

Age	**2003/2007**				**2008/2011**			
	Obs. deaths	**Exp. deaths**	**SMR**	**95% CI**	**Obs. deaths**	**Exp. deaths**	**SMR**	**95% CI**
50-64	55	61	0.9	(0.69,1.17)	26	29.7	0.91	(0.62,1.33)
65-69	35	48.5	0.72	(0.52,1)	37	34.7	1.04	(0.75,1.44)
70-74	54	74.7	0.72	(0.55,0.94)	45	53.1	0.85	(0.63,1.13)
75-79	99	111	0.89	(0.73,1.09)	83	78.8	1.05	(0.85,1.31)
80-84	113	151.4	0.75	(0.62,0.9)	110	110.3	1.01	(0.84,1.21)
85+	191	274.6	0.7	(0.6,0.8)	276	277.8	0.99	(0.88,1.11)
Total [a]	547	721.2	0.76	(0.73,0.86)	577	584.3	0.99	(0.91,1.07)

Men

Age	**2003/2007**				**2008/2011**			
	Obs. deaths	**Exp. deaths**	**SMR**	**95% CI**	**Obs. deaths**	**Exp. deaths**	**SMR**	**95% CI**
50-64	83	80.4	1.03	(0.83,1.28)	44	37.2	1.18	(0.88,1.59)
65-69	59	71.6	0.82	(0.64,1.06)	46	45.8	1.03	(0.77,1.37)
70-74	99	102.7	0.96	(0.79,1.17)	54	71.7	0.75	(0.58,0.98)
75-79	120	130.2	0.92	(0.77,1.1)	102	93.5	1.11	(0.92,1.35)
80-84	133	147.8	0.9	(0.76,1.07)	110	110.5	0.97	(0.8,1.17)
85+	143	176.4	0.81	(0.69,0.95)	171	171.6	1	(0.86,1.16)
Total [a]	637	709.1	0.9	(0.83,0.97)	527	530.2	0.99	(0.91,1.08)

[a] Calculated without weighting the age-specific SMRs
CI=confidence interval

Table 4.2 shows the disability prevalence and relative 95% confidence interval estimated in the ELSA and in the national population, by gender and age. Looking at the total sample for men and women, as expected, the proportions of individuals with disability were similar across ELSA respondents and the census participants living in private households, while the ELSA percentages were smaller when compared with the total population, including institutionalised people. With this regard, it is worth to remark that the gap was larger for women. Looking at age-specific prevalences, they were similar only in the age group 60-64. The fact that ELSA participants were more disabled than the English population living in private households at younger ages and less disabled at older ages may partly depend on the fact that in the Census respondents were openly advised to consider also problems due to age as disability. Attrition in the ELSA can be another explanation too, which could be assessed, for example, estimating prevalences after multiple imputation (MI) with baseline variables used as auxiliary variables to maximise the plausibility of the MAR assumption [66].

Table 4.2: Prevalence of limiting long-lasting illness in ELSA and English population

Age	**ELSA**				**2001 census people in HH**		**2001 census tot. pop.**	
	Men %	**(95 %CI)**	**Women %**	**(95 %CI)**	**Men**	**Women**	**Men**	**Women**
50-59	27	(25.1; 29)	29	(27.2; 30.9)	22.1	23.3	22.4	23.5
60-74	37.1	(33.8; 40.4)	31	(28; 34.1)	36.2	31.4	36.6	31.6
75-84	39.3	(37.3: 41.3)	41.2	(39.4; 43.1)	46.3	46.8	47.1	48.1
85+	58.2	(50.8; 65.2)	53.8	(47.9; 59.5)	66.7	73.4	69.7	77.7
Total [a]	35	(33.8; 36.4)	35.9	(34.7; 37.1)	35.4	37.8	36.2	39.4

[a] Calculated without weighting the age-specific prevalence
CI=confidence interval; HH=households

4.3 Discussion

The first part of this chapter offers a detailed description of the data used in this thesis and presents the sample design, survey content and advantages and disadvantages of using the ELSA for the purposes of this thesis. In the second part, the representativeness of ELSA is assessed, mainly in a descriptive way, in particular comparing disability prevalence

and mortality rates in the English general population and in the ELSA sample. This investigation is fundamental for the progression of this thesis and the interpretation of its results.

Specifically, the first finding concerns the comparability of ELSA mortality rates with the mortality rates of the English population. The relevance of this finding represents at the same time an important contribution to assess ELSA's representativeness -in fact to my knowledge, no other study so far has analysed this aspect- and a crucial point for understanding the extent of the generalizability of my results. Some differences between ELSA mortality rates and mortality rates of the general English population emerged, both for men and women, but only in the first period of observation (2003 - 2007). Over time, the estimated ELSA mortality rates converged to those from national statistics (that included institutionalised population). This could partly be due to the fact that throughout follow-up, ELSA respondents who moved into communal establishment remained part of the study, and therefore the ELSA sample partly changed and became closer to the total English population.

Another aspect that emerged was gender-specific mortality differences in the comparison between ELSA and the national data. In the first period, both men and women experienced mortality rates lower than the general English population, but women's SMR was smaller than males' SMR (overall male SMR for 2003/2007 equal to 0.9, p-value= 0.0068; female SMR for 2003/2007 equal to 0.76, p-value < 0.001). For disability, the computed national prevalence, which included institutionalised populations, was higher than in the ELSA sample, and the gap was larger for women compared to men. This finding may help to interpret the larger differences in mortality with respect to the general population that were observed among women. The proportion of institutionalised women was considerably high at older ages (around 21 percent in 2001) and almost all of older institutionalised women had disability problems. This could explain the lower mortality rates of ELSA females compared to general female population (because at baseline there were not institutionalised respondents in the study), and why the gap in mortality was larger for women than for men (the proportion of institutionalised men in 2001 was 11.5 percent).

This piece of research was intended to assess ELSA's representativeness ando under-

stand the extent of the generalizability of the findings of this thesis. It emerged that findings can be extended to the general population of English adults living in private households, which is the ELSA target population at baseline. Some convergence over time to the general population including institutionalised individuals emerged as well. Results suffer from some limitations, however. With regards to the inclusion of the institutionalised population, data were available only for 2001, while ELSA data pertained to 2002 and 2003. Moreover, although questions about limiting long-lasting illness were asked in the ELSA and the Census in a substantially similar manner, there were a few differences: in the ELSA respondents were asked two questions while in the Census it was only one; in the Census there was a specific reference to age-related problems while in the ELSA there was not. This may have had an effect in the way people self-reported their condition. Finally, the evidence presented in this chapter is mainly descriptive and therefore conclusions are mostly considerations to take into account when interpreting and discussing the findings presented in the next chapters.

Part II

Research Papers

Chapter 5

Association between disability and mortality

5.1 Preamble

In research Paper II, I investigate the association of disability measured at baseline with mortality observed over a decade.
Research Paper II addresses objectives 2 and 3 of this thesis:

- **Research Objective 2:** *To theoretically conceptualise and empirically measure disability:* Disability is conceived according to the theoretical framework presented in Chapter Three, the WHO's International Classification of Functioning Disability and Health (ICF). The variables selected to its measurement are the basis upon which disability is measured in the other research papers of this work.

- **Research Objective 3:** *To explore gender inequalities in the effect of disability on mortality:* Through this objective, I enter the debate on the gender paradox in health and mortality [67, 68]. Assessing whether disability affects mortality differently for men and women, I seek to shed light on why women live longer than men but spend larger proportion of their life in disability.

5.2 Research Paper II

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Benedetta Pongiglione
Principal Supervisor	Bianca De Stavola
Thesis Title	Gender inequality in healthy ageing: a study of the English older population over a decade

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	European Journal of Epidemiology		
When was the work published?	May 13, 2016		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	

Student Signature: BPongiglione **Date:** 1/11/2016

Supervisor Signature: B De Stavola **Date:** 1/11/16

Eur J Epidemiol
DOI 10.1007/s10654-016-0160-8

MORTALITY

Disability and all-cause mortality in the older population: evidence from the English Longitudinal Study of Ageing

Benedetta Pongiglione[1] · Bianca L. De Stavola[1] · Hannah Kuper[2] · George B. Ploubidis[3]

Received: 15 December 2015 / Accepted: 5 May 2016

Abstract Despite the vast body of literature studying disability and mortality, evidence to support their association is scarce. This work investigates the role of disability in explaining all-cause mortality among individuals aged 50+ who participated in the English Longitudinal Study of Aging. The aim is to explain the gender paradox in health and mortality by analysing whether the association of disability with mortality differs between women and men. Disability was conceived following the International Classification of Functioning, Disability and Health (ICF), proposed by the WHO, that conceptualizes disability as a combination of three components: impairment, activity limitation and participation restriction. Latent variable models were used to identify domain-specific factors and general disability. The association of the latter with mortality up to 10 years after enrolment was estimated using discrete-time survival analysis. Our work confirms the validity of the ICF framework and finds that disability is strongly associated with mortality, with a time-varying effect among men, and a smaller constant effect for women. Adjusting for demographic, socioeconomic and behavioural factors attenuated the association for both sexes, but overall the effects remained high and significant. These findings confirm the existence of gender paradox by showing that, when affected by disability, women survive longer than men, although if men survive the first years they appear to become more resilient to disability. Sensitivity analyses suggested that the gender paradox cannot be solely explained by gender-specific health conditions: there must be other mechanisms acting within the pathway between disability and mortality that need to be explored.

Keywords Disability · Mortality · Ageing · Gender

Abbreviations

ADLs	Activities of daily living
DTSA	Discrete-time survival analysis
IADLs	Instrumental activities of daily living
ICF	International Classification of Functioning, Disability and Health
LRT	Likelihood ratio test
SEP	Socioeconomic position

Introduction

In 2001 the World Health Organization (WHO) developed a conceptual framework for describing functioning and disability: the International Classification of Functioning, Disability and Health (ICF). One of the aim of the ICF was to provide a common set of instruments to measure disability to standardize this concept and its use in international studies. The ICF conceives difficulties with human functioning as three interconnected areas (see Fig. 1). This

Electronic supplementary material The online version of this article (doi:10.1007/s10654-016-0160-8) contains supplementary material, which is available to authorized users.

✉ Benedetta Pongiglione
benedetta.pongiglione@lshtm.ac.uk

[1] Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, UK

[2] International Centre for Evidence in Disability, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, UK

[3] Centre for Longitudinal Studies, Department of Social Science, UCL Institute of Education, University College London, 55–59 Gordon Square, London WC1H 0NU, UK

Published online: 13 May 2016

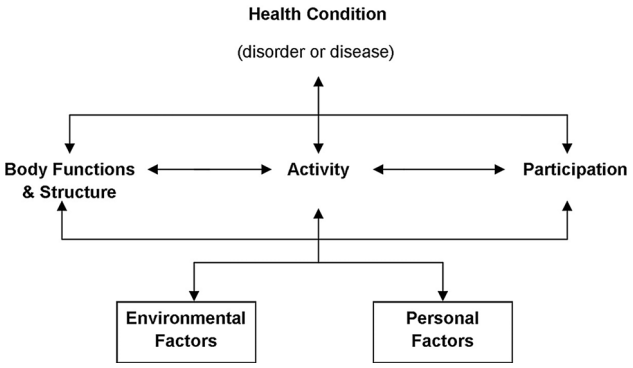


Fig. 1 Representation of the International Classification of Functioning, Disability and Health (ICF). *Source*: World Health Organization Geneva 2002, 'Towards a Common Language for Functioning, Disability and Health: ICF'

is impairments that are problems in body function or alterations in body structure; activity limitations that are difficulties in executing daily activities such as walking or eating; and participation restrictions that are problems with involvement in any area of life—for example, facing discrimination in employment due to disability [1, p. 5]. Disability refers to difficulties encountered in any or all three areas of functioning.

The ICF is considered the dominant conceptual framework for describing functioning and disability [2]. Nevertheless, it is not yet widely used in research relating or combining disability and mortality. Dale and colleagues [2] examined the relationship between disability and mortality conceiving disability according to the ICF's framework, focusing on women aged 60–79 years. A key aspect in studying disability and mortality, however, is related to gender differences. The gender paradox in health and mortality is well known in the literature. It was first observed in the mid-1970s [3, 4] and reflects the finding that women live longer than men, but tend to have more disability than males. Many theories have been proposed to explain the 'gender paradox' in mortality and disability, among which the most prevalent is that women may have higher prevalence of nonfatal but disabling diseases and men have higher prevalence of fatal and chronic diseases strongly related to mortality. Some researchers [5, 6] hypothesize that higher disability prevalence among women may be a function of longer survival in disability rather than higher incidence of disability.

With our work we seek to contribute to the debate of the gender paradox in health and mortality by (1) showing whether the association between disability and mortality differs between men and women (2) proposing possible explanations of why it may occur. More specifically, we measure disability among the older population using data from the English Longitudinal Study of Ageing (ELSA), and empirically test with a measurement model the construct validity of the WHO's ICF. Based on this comprehensive interpretation of disability, we then apply discrete-time survival analysis (DTSA) to study the impact of disability measured at baseline on mortality observed over the course of a decade, and assess whether and how this association changes over time, stratifying the analysis by gender.

Materials and methods

Data source and sample

This study used data drawn from the first wave of the English Longitudinal Study of Ageing (ELSA), which took place in 2002/2003. Briefly, ELSA core members are a representative sample of the noninstitutionalized population, living in England, who were aged 50 years or older at the time of interview. 11,391 core-member respondents were recruited at wave 1. For our analysis, we included all participants who had complete records on all disability items, leaving us with a sample of 9715. At the time of interview, respondents were asked to give their permission to link their data to the National Health Service Central Register (NHSCR) mortality records. For those who gave their consent, information on mortality was available by year from 2002 to 2011. Interviews were done using computer-assisted interviewing and self-completion questionnaires.

Measures

Death

The primarily outcome of this analysis was deaths occurred from 2002 to 2011. As time of death was available only by year, binary time-specific event indicators were created for each period of observation (ten intervals). For some respondents (n = 358) status of death was available but time of death was unknown; in this case information were partially retrieved looking whether respondents took part in the following surveys; if they were interviewed in later waves, they were assumed to be alive at least until the year of the last survey they responded; otherwise they were considered lost to follow-up and their event indicators treated as missing. This way three patterns of observations were possible: (1) survivors or censored: individuals who did not experience the event and were followed-up for all time-periods of observation; (2) dead: individuals who experienced the event at some point during the period of observation; (3) lost to follow-up: individuals who dropped out the study before it ended.

Disability

Variables describing disability were selected according to the WHO's ICF framework, in order to construct the impairment, activity limitation and participation restriction components. Consulting the WHO's ICF browser, one author selected all possible disability items from the questionnaire to be included in the measurement model; the list was screened in agreement with another author and selected items were classified in a double-blind fashion in one of the three components; in case of disagreement a third opinion was sought for the final classification. Inter-rater agreement for classification of selected items was measured using the kappa statistic [7]. A total of fifty items were selected from the questionnaire to construct the ICF model: 19 for impairment, 20 for activity limitation and 11 for participation restriction (Supplementary Table 1). Impairment was described by variables such as self-rated eyesight and hearing, chronic conditions such as high blood pressure and arthritis, and questions about pain. Activity limitation was assessed by questions on ADLs and mobility functions, for example climbing flights of stairs or walking 100 yards. Finally, participation included questions on instrumental activities of daily living (IADLs), and various limitations due to health problems, such as using public transports or working. Variables were all either dichotomous (i.e. yes/no answer) or ordered categorical, for example ranging from 'excellent' to 'poor', from 'never' to 'always' and from 'no difficulty' to 'unable'. A list of the questions asked for each item and possible answers is available in the appendix (Supplementary Table 1).

Confounders

A number of potential confounders known to be related to disability and mortality from the literature (see for example [8–13]) were accounted for in the survival models. These included basic demographic characteristics, such as age at wave 1, marital status and household size; socioeconomic position (SEP) measured through education, income, wealth and occupation; socioeconomic background represented by father's occupation when respondent was 14; health-related behaviours including smoking, drinking and physical activity; and presence of limiting long-lasting illness. In sensitivity analyses, objective measures of health were also introduced as additional confounders in the analyses that used the information collected at wave 2 (2004/2005) where health measures were assessed during the nurse visit with survivors up to that wave included in the analysis. Four observer-measured indicators were selected. These were blood assays for inflammation, blood clotting and cholesterol—all known to be associated with risk of heart disease- and a measure of respiratory functioning. The inflammatory activity in the body was measured by the level of C-reactive protein (CRP); blood clotting by a protein called fibrinogen; cholesterol is a type of fat present in the blood and was assessed as total cholesterol. Respiratory functioning was measured by Forced Vital Capacity (FVC), which is the volume of air that can forcibly be blown out after full inspiration; three measurements were taken of FVC, and we used the highest technically satisfactory reading.

Analysis

The analysis was carried out in two steps. First we estimated factor scores for disability using a latent variable model, then we used the stored factor scores in survival analysis.[1]

Measurement model

For the first step, a three factor first-order model was first fit to assess the ICF structure using the items selected for each ICF component, i.e. impairment, activity limitation and participation restriction.

Since all observed items were either categorical or binary, the fitted model can be formulated as follows. Categorical/binary observed indicators (y_{ij}) are related to continuous latent variable (η_j) via a normal ogive response model, such that:

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* > \tau_i \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $y_{ij}^* = \beta_i + \lambda_i \eta_j + \varepsilon_{ij}$ for i = 1, …, I_j (I_j being the number of observed indicators for latent variable j) and j = 1, …, J (J being the number of individuals). We also assume that

$$\eta_j \sim N(0, \sigma^2),\ \varepsilon_{ij} \sim N(0,1),\ \text{covariance}\left(\eta_j, \varepsilon_{ij}\right) = 0$$

where σ^2 is the variance of the latent measure. For simplicity, here we refer to unidimensional model; for more general notation see Rabe-Hesketh and Shrondal [14].

Model (1) can be equivalently expressed as:

$$\Pr\left(y_{ij} = 1 | \eta_j\right) = \Pr\left(y_{ij}^* > \tau_i | \eta_j\right) = \Phi\left(\beta_i + \lambda_i \eta_j\right)$$

$$\Phi^{-1} \Pr\left(y_{ij} = 1 | \eta_j\right) = \beta_i + \lambda_i \eta_j$$

[1] One-step analysis was performed as a robustness check. It consists of estimating the measurement model using the disability items at baseline and jointly performing a discrete time survival analysis for the 10-year period, without storing factor scores (first step) and then introducing them in the survival model (second step). Both analyses returned very similar results, therefore, for practical reasons only the results from the two-step analysis are reported here (results from the one step analysis available from corresponding author).

where $\Phi(\cdot)$ is the cumulative standard normal distribution and Φ^{-1} is the probit link.

Modification indices (MIs) were examined to improve model fit. MIs quantify the decrease of the χ^2 goodness of fit measure when the corresponding parameter is freed; they indicate whether any of the observed items should be correlated above and beyond their assumed relationships with latent factors. As this test's recommendations are directly motivated by the data and not by theoretical considerations [15, p. 491], we used them to suggest improvements but did not tie model specification on their values.

The best fitting first order model that reflects the ICF structure described impairment, activity limitation and participation restriction and was improved by adding an extra factor for eyesight within the impairment component. Based on this construct and reflecting the WHO conceptualization, we fitted a second order model, where disability was the second order factor and impairment, eyesight, activity limitation and participation restriction were the first order factors. However the model presented some inconsistencies.[2] To deal with that, we decided to conceptualize disability in a general-specific model where the observed items are explained by one general factor disability- and domain-specific factors (see Fig. 2). Both the general and the specific factors were linked to the observed items as described above, and all factors were assumed to be uncorrelated with each other.

For identification purposes, both models (first order and general-specific) were defined constraining all factor variances to be equal to one, and allowing the error terms of the manifest items 'pain in chest' and 'pain' to correlate. Model estimation was performed using only complete records via weighted least squares means and variance adjusted (WLSMV) [16].[3]

Model fit was assessed using the Root Mean Square Error of Approximation (RMSEA) which assesses absolute fit, and two comparative indices, Comparative Fit Index (CFI) and Tucker–Lewis Index (TLI), which compare the model with the unrealistic null model of uncorrelated items. Fit is typically considered 'good' if the RMSEA is below 0.05 and the CFI and TLI are above 0.90 [14, p. 86].

Discrete-time survival analysis (DTSA)

Data were set in a way to carry out DTSA in a general latent variable framework [17]. A binary time-specific event indicator was created for each of the ten time periods, with the probability of an event occurring during an interval denoted by h(j), j = 1, …, 10, and referred to as the hazard probability for that time period [17].[4] The first step was to fit a crude mortality risk model that included the 10 binary time-specific event indicators of death,[5] with no predictors (including no intercept) or in other words to estimate the interval-specific risks (i.e. the probabilities for each time interval, analogous to separate intercepts in a regular regression model).

These probabilities were then related to covariates through a logit link function—that is, logistic regression—so that the effect of a covariate on the timing of death is parameterized by its effect on the log odds of an event during a given time interval [18]. For a single covariate x, its effect on the probability of event occurrence in period j is expressed in terms of the log odds ratio (log OR) β_j[6]:

$$\text{logit } h(j) = \log\left(\frac{h(j)}{1-h(j)}\right) = -\tau_j + \beta_j x$$

$$h(j) = \frac{1}{1+\exp(\tau_j - \beta_j x)}$$

Then, we evaluated whether the corresponding logORs were constant over the 10 intervals (i.e. $\beta_{j=}$ β for all j, equivalent to the proportionality assumption), separately for each of the covariates, by introducing each covariate in the model (i.e. assuming a time invariant effect) and then including an interaction between the covariate and time (i.e. allowing for time varying effects) tested, using the log-likelihood ratio test (LRT). For disability, we double-checked whether its effect was time-varying controlling first only for age and then for the complete set of selected confounders.

Finally, we fitted models that includes the confounders sequentially, by group. In the baseline model we considered the effect of disability on mortality without controlling for

[2] Second-order model had a good fit (CFI = 0.945, TLI = 0.942, RMSEA = 0.042), but presented some problems: activity measured disability very poorly and its factor loading had an extreme value and was not significant (28.2 and 95 % CI [−120.3, 176.8]; p value = 0.71). At wave 2, the value was even more extreme and the model did not converge.

[3] Maximum likelihood estimator would have been too cumbersome given the large number of dimensions to be integrated.

[4] *Hazard probability* is the term used in Muthen's and Masyn's paper. The authors defined *the sample-estimated hazard probability for time period j as the number of events that are observed to occur in time period j divided by the total number of subjects at risk in time period j* (p. 33). In the context of our analysis, we will also be using the term *mortality risk* instead of hazard probability.

[5] In a general latent variable framework, the likelihood for a latent class model with binary indicators gives the probability of the event indicator being equal to one; in Mplus it is a (negative) "threshold" which defines the cut-point in the latent variable distribution for the switch from 'category' 0–1, and it is estimated for each time interval (i.e. her we estimate ten thresholds).

[6] The notation is the one used in Muthen and Masyn [17].

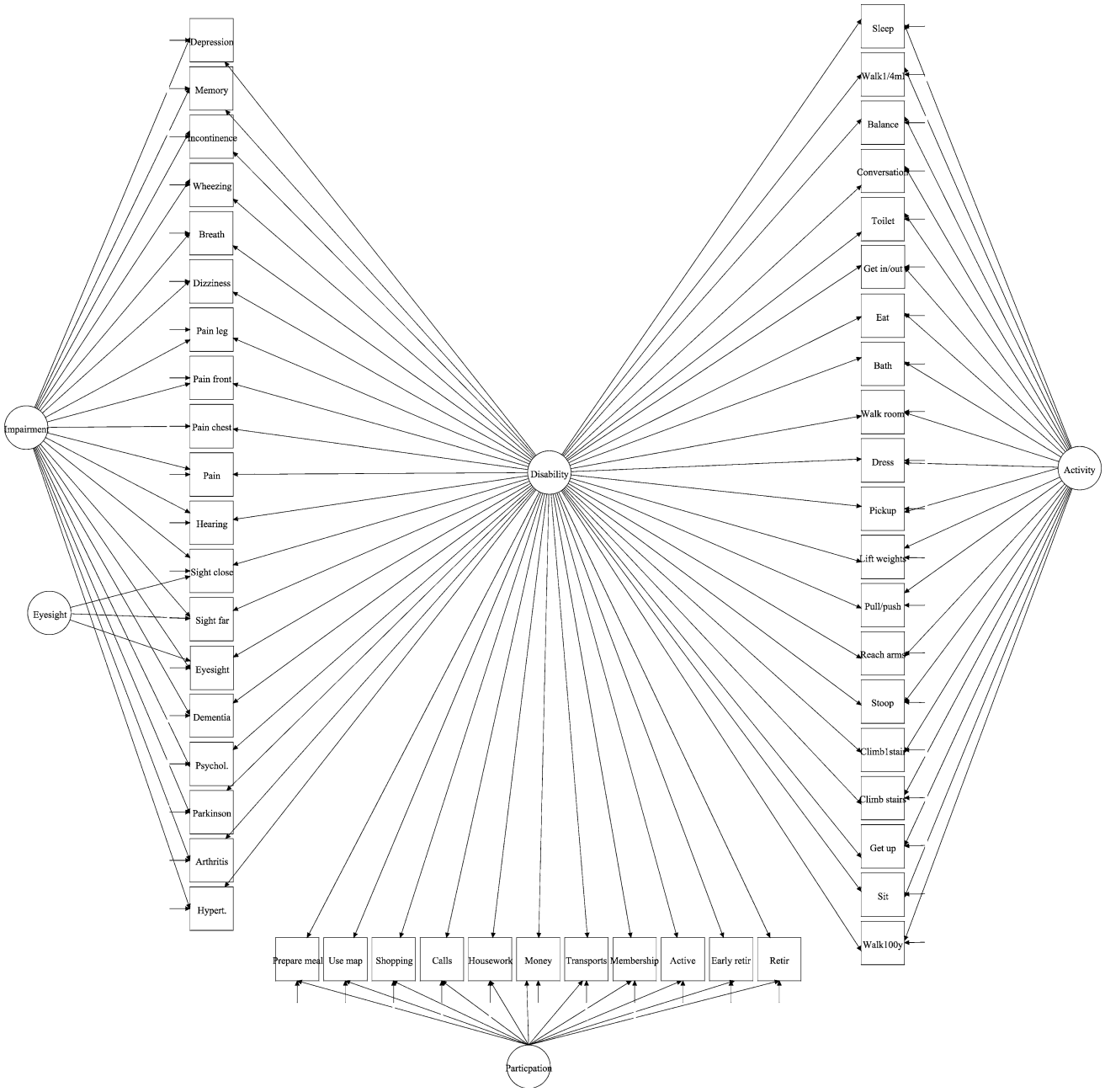


Fig. 2 General-specific measurement model. Names of observed variables (*rectangles*) are those listed in Table 1

any confounders but age; and in the full model all potential confounders were added, including long-lasting illness and health-related behaviours (all measured at wave 1).

Events indicators were treated as missing in correspondence of time intervals that followed the time when the event occurred or when the individual was lost to follow-up. Missingness was assumed to be at random (MAR) which for this model corresponds to uninformative loss to follow-up; FIML estimation with robust standard errors (MLR) was used [17]. When we added confounders, we incurred in missing values for these x variables; however only 4 % of data were missing, corresponding to three main missing patterns. When confounders were added into the model, complete case analysis (CCA) was carried out. However, this way adjusted for age analyses and adjusted for all confounders analyses were based on different numbers of observations; to deal with this problem, we first repeated the age adjusted models on the same numbers as those for the fully adjusted analyses, and secondly we the fully adjusted model was re-run using FIML in order to have the same sample size as in age adjusted models. Details on missing data patterns and results for CCA and

for regressions using FIML are provided in the appendix (Supplementary Tables 2 and 3).

Sensitivity analysis

A number of robustness checks were implemented in order to assess whether gender differences in the association between disability and mortality were driven by gender differences in prevalence of specific disabling diseases. In the first instance, we accounted for the fact that self-reported measures of health may not capture specific diseases and there may be a gender effect in the probability of reporting health limitations. To account for this potential bias, observer-measured health indicators were additionally considered as potential confounders. To this aim, we replicated the analysis including only respondents interviewed at wave 1 who took part in the following survey and using information on physical conditions measured during the nurse visit at wave 2. Four observer-measured indicators were selected and added as confounders in DTSA based on data from wave 2.

With the same rationale, but using a different approach, we also re-estimated the measurement model for disability dropping the items describing health/body functions (i.e. hypertension, arthritis, dementia, Parkinson, psychological problems and depression) originally included within the impairment component, to make sure that differences in mortality were not led by body functions and structures whose prevalence is more likely to differ between men and women.

To test whether the measurement model differed for males and females, we also re-estimated the factor scores for disability running separate analyses for men and women, and then testing whether there was heterogeneity by sex (we used a multiple group analysis for the total sample assuming strong invariance). The survival analysis model was also refitted using these new disability scores.

Finally, to account for possible differences across age groups, the original measurement model—as described in the previous paragraph—was re-estimated via multiple group analysis, without stratifying by gender. Then, we run DTSA using the resulting disability factor score and stratifying the sample by age group (i.e. 50–64, 65–74, 75+). Additionally and separately, we also re-run DTSA including an interaction term for age and disability (as measured in the baseline model).

Results

Sample

Of the 9715 respondents 46 % were men (4455) and 54 % women (5260). Over the course of the study, 21 % of male and 16 % of female respondents died (Supplementary Table 4). Demographic and socioeconomic characteristics are shown in the appendix (Supplementary Table 4). In general, demographic characteristics were quite similar between females and males; the average age of men and women was 64.4 and 64.8 years respectively with more women than men being aged 75+ (19.5 % of females compared to 17.4 % of males); higher proportions of women were widowed as expected due to their longer life expectancy. Men reported higher SEP in all indicators, e.g. higher education, income, occupational class. On the other hand, women had healthier behaviours, reporting higher proportions in those that never smoked as well as lower percentage of heavy drinkers. Finally, among respondents survived at wave 2, men had a more healthy profile than women with regards to all biomarkers and almost same level of inflammation.

Measurement model

The final agreed list of disability variables (kappa statistic for inter-rater agreement equal to 0.85) consisted of 50 items (19 impairments, 20 activities and 11 participations—Supplementary Table 1). The prevalence of these variables was higher for women than men (Table 1), with the exception of difficulty in communicating (conversation) and being engaged in social activity (active), and to a lesser extent in visual functioning. Descriptive statistics show that more men than women died, but women overall had more disability problems than men at baseline.

Following this classification, a latent variable model appropriate for the nature of the indicators was implemented. A first-order multidimensional model was first estimated, and its fit was rather poor (see Table 2). Some items presented high modification indices both for factor loadings and covariances among measurement errors. In particular, eyesight items (which are self-rated eyesight, being able to seeing at distance and close) presented high modification indices for both factor and covariances among their measurement errors. Rather than allowing the errors of the eyesight items to correlate, we introduced within the impairment factor an extra eye-specific latent factor to explain eye-items variability, producing a sort of general-specific model within the multidimensional first order model. The resulting model fit was highly satisfactory (Table 2). Standardized factor loadings λ_{ij}, which express the strength of the association between the indicators and latent variables, by rule of thumb are considered satisfactory when $|\lambda_{ij}| > 0.4$ [19]. Standardized factor loadings obtained from the first-order model showed that 13 out of 19 indicators of impairment were strongly associated with this factor; 19 out of 20 indicators of activity were strongly associated with this factor and 8 out of 10 indicators with

Table 1 Prevalence of health indicator by gender

Disability items	Men %	Women %
Hypertension	36.6	38.6
Arthritis	25.5	38.1
Parkinson	0.4	0.4
Psycho problems	5.8	8.8
Dementia	0.4	0.2
Self-rated eyesight (less than good)	12.8	15.5
Eyesight at distance (less than good)	6.4	9.8
Eyesight close (less than good)	9.9	10.8
Hearing	23.8	28.4
Troubled with pain	34.1	40.5
Pain in chest	33.3	27.4
Pain across the front of chest	11.5	7.5
Pain in leg	28.3	30.6
Dizziness	11.8	16.5
Shortness of breath	32.1	42.5
Shortness of breath with wheezing	14.5	14.8
Incontinence	8.3	20.8
Self-rated memory (less than good)	32.3	30
Depression	14.1	18.3
Walking 100 yards	11.4	11.1
Sitting for 2 h	13	14.8
Getting up	22.2	28.1
Climbing stairs	28.8	41.2
Climbing 1 flights of stairs	11.6	15.3
Stooping	31.1	38.4
Reaching arms	9.2	12
Pulling/pushing	12.3	20.8
Lifting weights over 10 lb	15.9	31.9
Picking up 5p coin	4.5	5.2
Dressing	14.1	11.4
Walking across room	2.6	2.8
Bathing	9.9	12.5
Eating	1.3	1.8
Getting in/out bed	6	6.3
Toileting	3	3.1
Following conversation	40.8	28.1
Keeping balance	18.5	24.9
Walking quarter mile	25.1	28.9
Restless sleep	34.7	45.3
Preparing hot meal	3.4	3.7
Using map	2.2	6.2
Grocery shopping	6.1	9.9
Making calls	1.8	0.9
Housework	13.2	16.4
Managing money	1.9	1.6
Using transports	5.4	7.9
Being member of any org.	28.8	32.1
Doing activity	34.5	22.9
Early retirement (due to health)	7	3.7
Retirement (due to health)	3.1	4.8

participation factor. Particularly high were the factor loadings for activity, in most cases larger than 0.75 (Supplementary Table 5).

Based on the first-order measurement model described above, a general-specific model was fitted to identify the latent disability structure (Fig. 2). Goodness of fit (GoF) indicators are presented in Table 2. The distribution of the disability factor score, derived from the general-specific model, is shown in Fig. 3, by gender. The distributions are approximately Gaussian (Fig. 3); with that for males slightly more right-skewed than that for females, meaning that, compared to women, fewer men had high disability score. The average score of disability was higher for women; on a range going from −1.72 to 3.36, the female average score was 0.165, whilst on a range from −1.72 to 2.88 the male average score was equal to −0.025, i.e. 0.19 units lower (p value <0.001). When controlling for various chronic conditions not included in disability measure and for self-reported long-lasting illness, the mean difference in disability between women and men remained the same (0.19, p value < 0.001).

Discrete-time survival analysis

1775 respondents died over the course of the observation period, 53 % were men and 47 % women. Overall, mortality rate was 0.56 % for men and 0.34 % for women in the first interval (first year of follow up since 2002) and almost 3 % in the last interval (3.1 and 2.9 % for men and women respectively), with a relatively steadily increasing trend during the observation period. The Kaplan–Meier survival curve in Fig. 4 illustrates the survival curves by quartile of disability, separately by gender. The estimated survival curves are lower as the severity of disability increases, both for women and men. Male disadvantage in mortality is observed across each disability quartile and widens over time; the gap in mortality between men and women is more pronounced for the two most disabled groups. In particular, 56.5 % of men having the highest disability level survive to the end of the 10-year period, while the equivalent survivors percentage for women is 67.4 %.

To evaluate whether the effect of the pre-defined confounders on mortality were time-varying we introduced in the model each variable separately with/without its

Table 2 Goodness of fit

Model	CFI[a]	TLI[b]	RMSEA[c]
(1) First order model (3 factors)	0.873	0.867	0.067
(2) Fist order model (3 factors + eyesight component)	0.945	0.942	0.042
(3) General-specific model	0.956	0.952	0.039

[a] Comparative Fit Index

[b] Tucker–Lewis Index

[c] Root mean square error of approximation

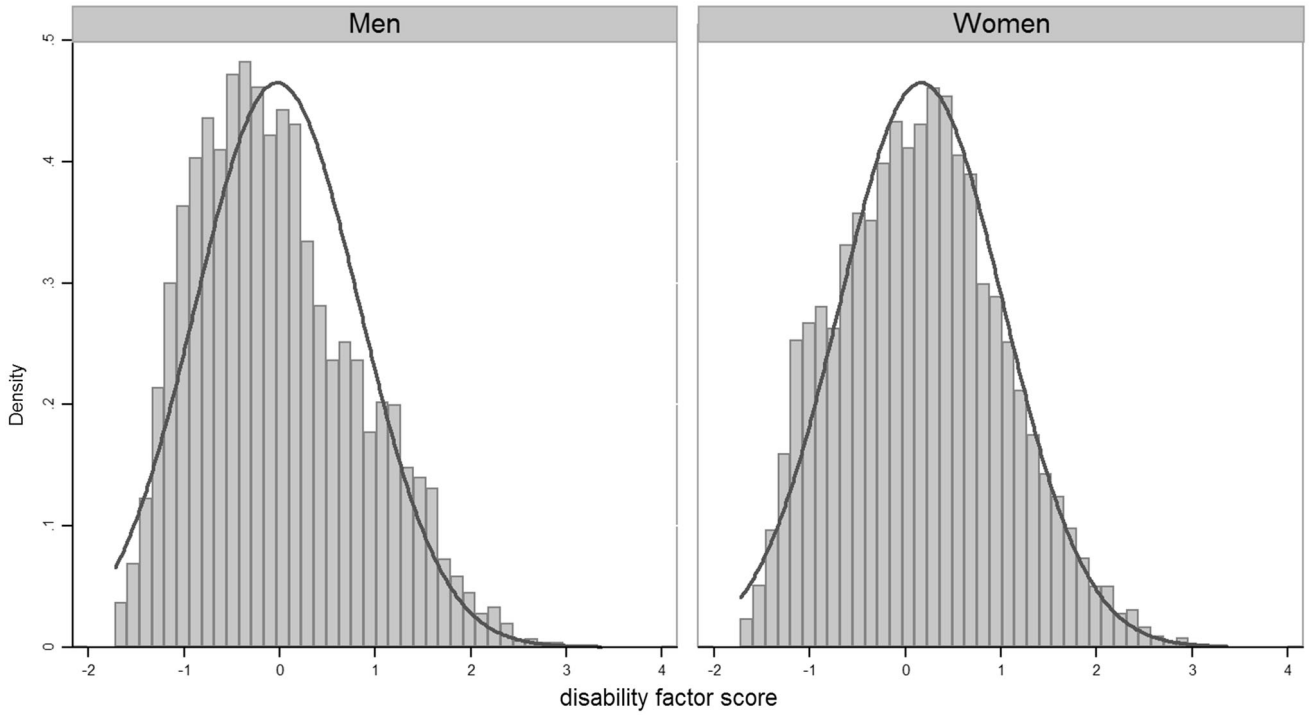


Fig. 3 Disability factor score by gender

interaction with time, while controlling for age. The constant proportional hazard assumption (i.e. time-invariant effect) was rejected for age and physical activity, but the latter only for men (Supplementary Table 6 for LRT test results). To assess the proportionality assumption for the predicted disability score we performed separate LRTs for its interaction with time, first controlling only for age, and then adjusting for all confounders. In both cases, disability was found to have time-varying effects for men and a time-invariant effect for women.

The parameter estimates for disability (expressed on the odds ratio scale) are shown in Table 3. For men, the time-specific disability odds ratios estimated controlling only for age (Model 1) were all significantly greater than 1, albeit decreasing over time. Although we did not observe a continuously declining trend, the test for trend showed evidence of a linear trend ($X^2_8 = 17.54$, p value $= 0.025$). The estimated disability OR corresponding to the first time period (2002) was 3.4 (95 % CI 2.12, 5.38), which means that for one-unit (1 SD in the latent score) increase in disability score the expected increase in the odds of mortality was by a factor of 3.4. Over subsequent time intervals the estimated ORs declined, but remained significantly higher than 1. Interestingly the estimated ORs dropped substantially immediately after the first period, from 3.4 to 2 in the following period; then the decline became more gradual. With regards to women, as we did not reject the proportionality assumption, the disability effect on mortality was estimated assuming a time-invariant effect, leading to a single estimated OR of 1.65 (95 % CI 1.51, 1.81; Model 1).

Table 3 also reports the estimated disability odds ratios by gender, obtained from fitting the model fully adjusted for demographic, socioeconomic and behavioural factors,

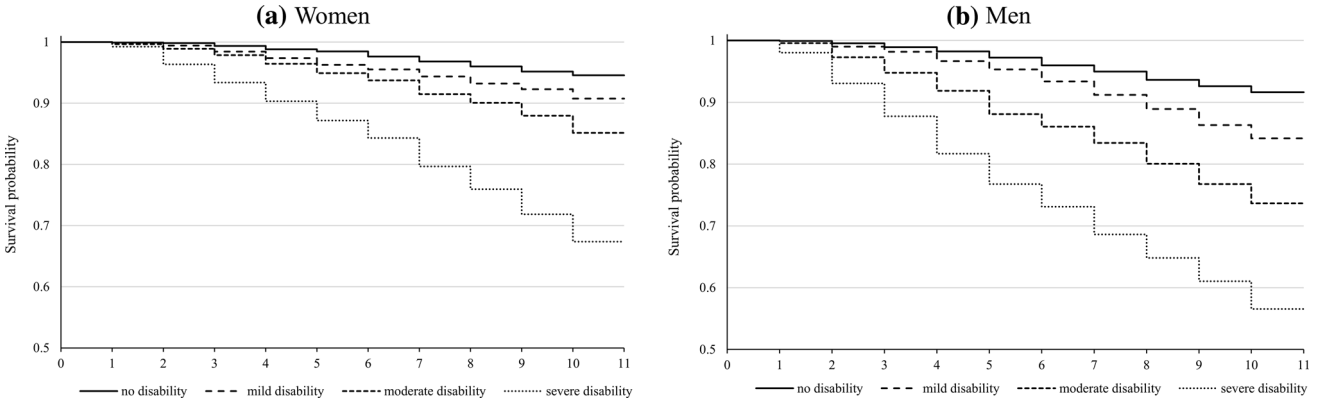


Fig. 4 Kaplan–Meier survival estimate, by disability component and gender. Results are presented by gender but disability factor score is estimated for the pooled sample

father's occupation and limiting long-lasting illness. For men, the estimated disability OR for time interval 1 decreased from 3.4 in the age-adjusted model to 2.2 in the fully adjusted model. The effect of confounders seemed particularly strong in this first interval, and although the estimated ORs in the following intervals were all smaller compared to those of model 1, they were all significant (at 5 % significance) with the exception of those for interval 6, 7 and 8. Among women, the estimated time-invariant effect of disability on mortality moderately declined after controlling for confounders, dropping from 1.65 to 1.36 (95 % CI 1.21–1.54). As a sensitivity analysis we also checked for a moderating effect of age and found a significant interaction of age and disability for men, such that the impact of disability measured at baseline becomes smaller as men age, while for women the interaction was not significant. When stratifying the analysis by age group, after age 75 the results for men disappear and disability OR decreases across age groups for women only (Supplementary Table 7).

When observer-measured health indicators were considered as potential confounders, DTSA was performed using the respondents interviewed at wave 1 who took part in the following survey, which was nurse-led and included collection of biomarkers. The results are shown in Table 4. The fully adjusted model was replicated first (columns 1 and 2), and then inflammation, blood clotting, cholesterol and respiratory functioning were added among the confounding variables (column 3 and 4). Among women the time-invariant effect of disability on mortality slightly decreased when controlling for observer-measured health indicators, whilst for men the estimated time-varying effect of disability was no longer significant both when adjusting or not adjusting for the biomarkers. (The results of other sensitivity analyses are not presented here, but available in the appendix and commented in the discussion section).

Discussion

Our study provides evidence on the association between mortality and disability in the older population and how this differs between men and women. Consistent with previous research, survival was found to be higher for women than men, whereas women had higher prevalence of disability. When looking at the relationship of disability at baseline with mortality observed over a decade later, the present study revealed: (1) increasing odds of dying as the baseline disability score increased, both for women and men with the association being stronger among the latter; and (2) decreasing association over time for men, as the impact of baseline disability on their mortality decreased with longer survival; (3) no variation over time for women, as the effect of disability remained constant over the 10-year period of observation.

With regard to men, the most striking result was the dramatic drop in the effect of disability on mortality from baseline period to the following year (2.2–1.8 per 1 standard unit change in disability score): disability in men, compared to women, seemed to have a stronger association with mortality in the very short rather than in the long term, when their estimated ORs converged to those in women. This could mean that men become more resilient to disability the longer they survive, and therefore that the effect of disability on their mortality in the long-run becomes less pronounced. Alternatively it could mean that disability is measured differently in men and women. However, as discussed in the next paragraphs, when we investigated this by extending the disability measurement model we found no evidence to support this explanation. For women, the impact of disability was found to be constant over time and overall the effect was smaller than that experienced by men. This is in accordance with the gender paradox in morbidity and mortality, and shows that in fact women spend a higher proportion of their life in disability because

Table 3 Disability odds ratios for mortality

	Time interval since disability measurement (years)	Model 1[a]		Model 2[b]	
		OR[c]	95 % CI[d]	OR[e]	95 % CI[d]
Males	1	3.381***	(2.12; 5.38)	2.237***	(1.27; 3.95)
	2	2.038***	(1.55; 2.69)	1.789***	(1.27; 2.52)
	3	2.157***	(1.68; 2.76)	1.875***	(1.39; 2.54)
	4	2.114***	(1.69; 2.65)	1.424***	(1.09; 1.86)
	5	1.826***	(1.45; 2.3)	1.445***	(1.09; 1.91)
	6	1.296*	(1; 1.69)	1.04	(0.77; 1.4)
	7	1.557***	(1.22; 1.98)	1.304*	(0.98; 1.74)
	8	1.499***	(1.18; 1.9)	1.305*	(0.99; 1.72)
	9	1.571***	(1.24; 1.99)	1.375**	(1.04; 1.82)
	10	2.083***	(1.62; 2.67)	1.955***	(1.46; 2.62)
Females	Time-invariant effect	1.654***	(1.51; 1.81)	1.365***	(1.21; 1.54)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[a] Model 1: model adjusted for age only

[b] Model 2: fully adjusted model: adjusted for age, demographic and socioeconomic confounders, father's occupation and long-lasting illness

[c] Test for linear trend $\chi^2(8) = 17.54$, p value = 0.025

[d] SE estimated from pooled logistic regression

[e] Test for linear trend $\chi^2(8) = 15.96$, p value = 0.043

Table 4 Disability odds ratios for mortality, wave 2

	Time since disability measurement (years)	Model 1[a]		Model 2[b]	
		OR	95 % CI[c]	OR	95 % CI[c]
Males	1	2.403*	(0.97; 5.94)	2.316*	(0.93; 5.77)
	2	1.649*	(0.95; 2.87)	1.598	(0.91; 2.8)
	3	0.985	(0.58; 1.66)	0.952	(0.56; 1.61)
	4	1.559*	(0.92; 2.64)	1.519	(0.89; 2.58)
	5	1.343	(0.8; 2.25)	1.297	(0.77; 2.18)
	6	1.151	(0.69; 1.92)	1.114	(0.66; 1.87)
	7	0.821	(0.52; 1.29)	0.796	(0.51; 1.25)
	8	1.48	(0.8; 2.74)	1.468	(0.79; 2.72)
Females	Time-invariant effect	1.435**	(1.12; 1.83)	1.331**	(1.04; 1.71)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sample size males = 1897; females = 2162

[a] Fully adjusted model

[b] Fully adjusted model + observer-measured indicators

[c] SE estimated from pooled logistic regression

they survive longer with disability, suggesting that higher disability prevalence among women may be a function of longer survivorship with disability rather than higher incidence of disability.

Along with evidence confirming the existence of the gender paradox among the English population aged 50+, we sought possible explanations of why it may occur. To address this question, we adopted three different strategies, whose results are discussed below. (1) In this study, we interpreted disability as a general phenomenon that may affect men and women to a different extent, rather than intend gender differences in disability depending on the definition of disability itself. Accordingly, disability was measured on the pooled sample. To investigate whether gender may instead affect the measurement itself of disability, we replicated the latent variable measurement model considering men and women separately and also running a multiple group analysis in the pooled sample (results are presented in Supplementary Table 8). The resulting latent measure of disability was in both cases

substantially similar to the results obtained from the pooled sample and results of DTSA were the same as those obtained in the original model. This suggests that the different impact of disability on mortality for men and women does not depend on gender-specific features of disability. (2) Additionally, since men are known to suffer more than women from fatal conditions, such as heart disease, and these conditions may not be captured by self-reported indicators, we also considered the confounding effect of observer-measured health indicators (measured at wave 2). We expected that after controlling for these indicators the effect of disability on mortality would decrease and the drop to be larger for males than females. Among women disability continued to exert a similar effect, while for men we found no evidence of an association between disability and mortality at wave 2. This discrepancy of results between sexes might be explained by the fact that the sub-sample of survivors to wave 2 was likely to be different for men and women, with the male sub-sample consisting of a more highly selected—less disabled—group than the equivalent females. Differences in terms of survival between men and women were not unexpected. What is surprising is that the consequences of male disadvantage in mortality and advantage in disability were visible already after 2 years from the beginning of the observation. (3) Finally, we also re-estimated the general-specific model for disability dropping some impairment items that described health functions, to make sure gender differences in mortality were not led by body functions and structures that may affect men and women differently. Again, the latent measure of disability obtained dropping these variables was very similar to the one obtained in the original measurement model, and the results of DTSA (Supplementary Table 9) essentially depicted the same patterns found using the original measure of disability. All the sensitivity analyses suggest that the observed differences in the association between disability and mortality in men and women are not driven only by gender-specific health conditions and body structures.

A complementary objective of the study was to provide a comprehensive definition of disability in order to test empirically the construct validity of the WHO's ICF framework when applied to the older population. After explorative investigations, disability was conceived as a general independent factor, and impairment, activity and participation as separate specific factors. The results of our study suggest that the three ICF components can be detected using the questions asked in ELSA, and indeed the first order factor model had a good fit. When it came to relate these parts with the concept of disability, disability was conceived as a single construct common to all individual indicators, explaining some proportion of their covariation; while the specific domains, i.e. impairment, eyesight, activity limitation and participation restriction, explain additional covariation among observable indicators. Detailed explanation of why we chose a general-specific model, may be found in the appendix (Supplementary Material B).

Finally, we highlight the strengths and weaknesses of this work. Strengths of the study include the availability of representative of the older population of England longitudinal dataset and the availability of various disability indicators that allowed us to reliably capture the ICF conceptualisation of disability. On the other hand some potential limitations should be considered while interpreting our results. There were no questions on the onset of disability, therefore it was not possible to estimate how long respondents survived from the actual disability onset. However, adjusting for pre-existing long-lasting limiting illness accounted, at least in part, for pre-existing disability; and this enabled us to consider the effect of disability at baseline (wave 1) on mortality as independent from any pre-existing disability/illness. A key point of this study, which represents both a strength and limitation, was that disability (and all confounders) was only measured at the study onset. This way, we did not know how disability had already impacted on health and mortality nor how it evolved over the observation period. This limited our understanding of its relationship with mortality. Nevertheless, the baseline effect can still be interpreted net of any effect that disability change over time on mortality might have had. Moreover, one of the advantages of measuring disability and all confounders at baseline is that, while keeping the model simple, we do not incur reverse-causality problems. Another limitation—as in most observational studies—is bias due to unmeasured confounders and/or residual confounding that might still bias the association under study. We acknowledge this as a potential source of bias, although we believe the most relevant confounders were taken into account.

Conclusion

The present work contributes to the debate on the gender paradox in health and mortality by showing that women spend a larger proportion of their life in disability because they survive longer with disability. We also enrich the discussion on possible explanations of why this occurs and show that gender differences in the association between disability and mortality are not driven only by gender-specific health conditions and body structures. There must be some other mechanisms acting within the pathway between disability and mortality that make women survive with disability better than men. Future studies should focus

on exploring these mechanisms to fully understand the gender paradox in health and mortality.

References

1. World Health Organization, World Bank. World report on disability. Geneva, Switzerland: World Health Organization; 2011.
2. Dale C, Prieto-Merino D, Kuper H, et al. Modelling the association of disability according to the WHO International Classification of Functioning, Disability and Health (ICF) with mortality in the British Women's Heart and Health Study. J Epidemiol Community Health. 2011;66(2):170–5.
3. Nathanson CA. Illness and the feminine role: a theoretical review. Soc Sci Med. 1975;9(2):57–62.
4. Verbrugge LM. Sex differentials in health. Public Health Rep. 1982;97(5):417–37.
5. Chan A, Zimmer Z, Saito Y. Gender differentials in disability and mortality transitions: the case of older adults in Japan. J Aging Health. 2011;23(8):1285–308. doi:10.1177/0898264311408417.
6. Jiawiwatkul U, Aekplakorn W, Vapattanawong P, Prasartkul P, Porapakkham Y. Changes in active life expectancy among older thais: results from the 1997 and 2004 National Health Examination Surveys. Asia Pac J Public Health. 2012;24(6):915–22. doi:10.1177/1010539511409923.
7. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46. doi:10.1177/001316446002000104.
8. Gjonca E, Tabassum F, Breeze E. Socioeconomic differences in physical disability at older age. J Epidemiol Community Health. 2009;63(11):928–35.
9. Lang I, Guralnik J, Wallace RB, Melzer D. What level of alcohol consumption is hazardous for older people? Functioning and mortality in U.S. and English national cohorts. J Am Geriatr Soc. 2007;55(1):49–57.
10. Lima-Costa MF, Steptoe A, Cesar CC, De Oliveira C, Proietti FA, Marmot M. The influence of socioeconomic status on the predictive power of self-rated health for 6-year mortality in English and Brazilian older adults: the ELSA and Bambui cohort studies. Ann Epidemiol. 2012;22(9):644–8.
11. Liu IY, LaCroix AZ, White LR, Kittner SJ, Wolf PA. Cognitive impairment and mortality: a study of possible confounders. Am J Epidemiol. 1990;132(1):136–43.
12. Silva VDL, Cesse EÂP, Albuquerque MDFPM. Social determinants of death among the elderly: a systematic literature review. Rev Bras Epidemiol. 2014;17:178–93.
13. Wahrendorf M, Reinhardt JD, Siegrist J. Relationships of disability with age among adults aged 50 to 85: evidence from the United States, England and Continental Europe. PLoS ONE. 2013;8(8):e71893. doi:10.1371/journal.pone.0071893.
14. Rabe-Hesketh S, Skrondal A. Classical latent variable models for medical research. Stat Methods Med Res. 2008;17:5–32.
15. Mueller RO, Hancock GR. Best practices in structural equation modeling. In: Osborne JW, editor. Best practices in quantitative methods. Thousand Oaks: Sage; 2008. p. 488–508.
16. Brown T. Confirmatory factor analysis for applied research. New York: Guilford Press; 2006.
17. Muthén B, Masyn K. Discrete-time survival mixture analysis. J Educ Behav Stat. 2005;30(1):27–58.
18. Petras H, Masyn KE, Buckley JA, Ialongo NS, Kellam S. Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual-and context-level influences. J Educ Psychol. 2011;103(1):223.
19. Stevens J. Applied multivariate statistics for the social sciences. Hillsdale: Erlbaum; 1992.

5.3 Supporting information to Research Paper II

This section provides additional details of the methods and analyses reported in Research Paper II. Some of these details were already included in the Supplementary material of the paper, available online on Springer website (http://link.springer.com/article/10.1007%2Fs10654-016-0160-8). They are reproduced here, together with further material, in particular concerning the sensitivity analysis mentioned in the paper.

Rationale for choosing a general-specific model

Disability was conceived as a general independent factor, and impairment, activity and participation as separate specific factors. This means that disability was assumed to explain the part of variation of the observable items that was not captured by specific factors, and it was assumed not to be correlated with them. The general-specific model differs from higher-order models in that first-order factors are not subsumed by the higher order factor but are, instead, uncorrelated and distinct. A second-order model would have appeared to be more coherent with the WHO's conceptualization of disability, but in fact the choice of a general-specific model had both theoretical and empirical advantages. General-specific models are potentially applicable when (i) there is a general factor that is hypothesised to account for the commonality of the items; (ii) there are multiple domain specific factors, each of which is hypothesized to account for the unique influence of the specific domain over and above the general factor; (iii) there may be an interest in the domain specific factors as well as the common factor that is of focal interest [69, p. 190]. The last two points in particular, justify our choice of applying a general-specific model. As a matter of fact, the innovative and fundamental feature of the ICF is that it allows us to disentangle the various disability components in order to estimate separately their impact on different spheres of life. The empirical foundation for preferring a general-specific model was that in the multidimensional first order model that was firstly estimated, it was observed that impairment (I), activity (A) and participation (P) were very highly correlated (A with I=0.989; A with P=0.988; I with P=0.963), whilst the eyesight factor was very specific (correlation of eyesight factor with I=0.075; with A=0.051; with P=0.049). This means that the eyesight factor captured some variance which was unique to its specific items, i.e.

the factor loadings for the eye items onto the eyesight factor remained high, implying that there was a very strong specific variance that was not related to disability. For all these reasons a general-specific model was preferred to a second-order model. For completeness, Table 5.1 shows the model fit indicators for three alternative models: the multiple correlated factor model, the second-order factor model and the general-specific factor model. Fit was best for the latter model, but overall fit was very good in all models.

Table 5.1: Goodness of fit statistics specific to the set of alternative models considered

	Chi-square	**d.f.**	**CFI**	**TLI**	**RMSEA**
1) Multiple correlated factor model [a]	21265.823	1167	0.945	0.942	0.042
2) Second order factor model	21441.391	1168	0.945	0.942	0.042
3) General-specific model	15595.549	1121	0.961	0.957	0.036

[a] i.e. First order model with 3 factors and eyesight component
d.f.=degree of freedom; CFI=comparative fit index; TLI=TuckerLewis index; RMSEA=root mean square error of approximatione

ICF disability items

Table 5.2: Disability items agreed by three authors

(1) Impairment	
Has a doctor ever told you that you have (or have had) any of the conditions on this card?	High blood pressure or hypertension
	Arthritis (including osteoarthritis , or rheumatism)
	Parkinson's disease
	Any emotional, nervous or psychiatric problems
	Dementia, organic brain syndrome, senility or any other serious memory impairment
Is your eyesight (using glasses or corrective lens as usual) [excellent, very good, good, fair, poor, blind]	
How good is your eyesight for seeing things at a distance, like recognising a friend across the street (using glasses or corrective lens as usual)? [excellent, very good, good, fair, poor, blind]	
How good is your eyesight for seeing things up close, like reading ordinary newspaper print (using glasses or corrective lens as usual)? [excellent, very good, good, fair, poor, blind]	
Is your hearing (using a hearing aid as usual) [excellent, very good, good, fair, poor]	
Are you often troubled with pain?	
How often do you have problems with dizziness when you are walking on a level surface? [never, sometimes, often, very often, always, no walk]	
Have you ever had a severe pain across the front of your chest lasting for half an hour or more?	
Have you ever had any pain or discomfort in your chest?	
Are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill?	
Have you ever had attacks of shortness of breath with wheezing?	
This might not be easy to talk about, but we would like to ask you about incontinence. During the last 12 months, have you lost any amount of urine beyond your control?	
How would you rate your memory at the present time? [excellent, very good, good, fair, poor]	
Do you get pain or discomfort in either of your legs which comes on when you walk? [no, yes, can't walk]	
Much of the time during the past week, you felt depressed?	

(2) Activity Limitations

Please tell me whether you have any difficulty doing each of the everyday activities on this card. Exclude any difficulties that you expect to last less than three months. Because of a health problem, do you have difficulty doing any of the activities on this card?	Walking 100 yards
	Sitting for about two hours
	Getting up from a chair after sitting for long periods
	Climbing several flights of stairs without resting
	Climbing one flight of stairs without resting
	Stooping, kneeling, or crouching
	Reaching or extending your arms above shoulder level
	Pulling or pushing large objects like a living room chair
	Lifting or carrying weights over 10 pounds
	Picking up a 5p coin from a table
Please tell me if you have any difficulty with these because of a physical, mental, emotional or memory problem. Again exclude any difficulties you expect to last less than three months. Because of a health or memory problem, do you have difficulty doing any of the activities on this card?	Dressing, including putting on shoes and socks
	Walking across a room
	Bathing or showering
	Eating, such as cutting up your food
	Getting in or out of bed
	Using the toilet, including getting up or down
Do you find it difficult to follow a conversation if there is background noise	
How often do you have problems with keeping your balance when you are walking on a level surface? [never, sometimes, often, very often, always, no walk]	
By yourself and without using any special equipment, how much difficulty do you have walking for a quarter of a mile? [no difficulty, some difficulty, much difficulty, unable]	
(Much of the time during the past week), your sleep was restless?	

(3) Participation Restrictions

Please tell me if you have any difficulty with these because of a physical, mental, emotional or memory problem. Again exclude any difficulties you expect to last less than three months. Because of a health or memory problem, do you have difficulty doing any of the activities on this card?	Preparing a hot meal
	Using a map to figure out how to get around in a strange place
	Shopping for groceries
	Making telephone calls
	Doing work around the house or garden
	Managing money, such as paying bills and keeping track of expenses
Do you use public transport? If no Why don't you use public transport more often?[My health prevents me]	
What were your reasons for retiring? [Own ill health]	
What were your reasons for taking early retirement? [Own ill health]	

Study sample

The sample used in the Research Paper II is based on core-member respondents recruited at wave 1. Of the 11,391 interviewed, only participants who had complete records on all disability items selected for measuring disability were included. This corresponded to 9,715 observations. The choice of selecting participants with complete baseline information on the disability items arose from the need for keeping the first stage of the analysis simple. The choice was supported, first, by the comparison of the distribution of the main variables in complete versus incomplete records. Table 5.3 presents descriptive statistics of the main socio-demographic variables used in this work for all core-members at wave 1 and for the selected sample. In the Supplementary material provided with Research Paper II (`http://link.springer.com/article/10.1007%2Fs10654-016-0160-8#SupplementaryMaterial`), the data for respondents having complete records for disability items (right columns) are available stratified by gender (Supplementary Table 4). Data did not show any systematic differences between respondents included into the analysis, and those excluded. In the second place, including all participants in the measurement model, i.e. those without complete records of the disability variables, would have required either to: (i) use Full Information Maximum Likelihood (FIML) with MLR[1] , which implies an assumption of Missingness At Random (MAR). (ii) Use Pairwise deletion with Weighted least squares means and variance (WLSMV). For the first option, given the high number of parameters of the model, using MLR would have been very time consuming[2] and it would have still required to make assumptions on the missingness mechanisms; the second option, although less time consuming, would have required a stricter form of missing mechanism (Missingness Completely At Random (MCAR)), depending on whether the outcome is associated with missingness (in this case the disability items are the outcomes though, the latent variable is the "predictor"). The difference between MAR and MCAR is that missing data patterns are independent of the observed data under MCAR, while they are not independent of the observed data under MAR.

[1] Mplus option for maximum likelihood estimation with robust standard errors.

[2] Montecarlo integration would have been needed and a very high starting value (i.e. 5000) necessary to make the model converge.

Table 5.3: Descriptive statistics for all core-members and selected sample

Variables	Core-member (n=11,391)	respondents	Respondents with complete records for disability items (n=9,715)	
	N	%	N	%
Sex				
Male	5,186	45.53	4,455	45.86
Female	6,205	54.47	5,260	54.14
Total	11,391	100	9,715	100
Age				
50-64	5,854	51.39	5,193	53.45
65-74	3,181	27.93	2,724	28.04
75+	2,356	20.68	1,798	18.51
Marital status				
single	630	5.53	515	5.3
married	7,570	66.47	6,651	68.47
divorced or separated	1,195	10.49	992	10.21
widowed	1,994	17.51	1,556	16.02
Total	9,714	9,714	9,714	100
Parental status				
Yes	9,300	81.65	7,955	81.88
No	2,090	18.35	1,760	18.12
Total	11,390	100	9,715	100
Household size	11,391	2.03	9,715	2.04
Ethnicity				
White	9,504	97.89	9,504	97.89
non-White	205	2.11	205	2.11
Total	9,709	100	9,709	100
Education				
high educate	2,500	22.01	2,279	23.47
middle educated	3,997	35.18	3,584	36.9
low educated	4,864	42.81	3,849	39.63
Total	11,361	100	9,712	100
Income				
1st quintile	2,178	20.01	1,861	20
2nd quintile	2,177	20	1,861	20
3rd quintile	2,178	20.01	1,861	20
4th quintile	2,176	19.99	1,862	20.01
5th quintile	2,177	20	1,859	19.98
Total	10,886	100	9,304	100
Wealth				
1st quintile	2,241	20.03	1,915	20.01
2nd quintile	2,236	19.98	1,915	20.01
3rd quintile	2,242	20.03	1,914	20
4th quintile	2,235	19.97	1,915	20.01
5th quintile	2,237	19.99	1,913	19.99
Total	11,191	100	9,572	100
Occupation				
Managerial or professional	2,854	25.6	2,605	27.25
intermediate	3,945	35.39	3,445	36.04
Routine or technical	4,349	39.01	3,509	36.71
Total	11,148	100	9,559	100
Father's job at 14				
professional or managerial	3,216	28.67	2,834	29.51
skilled	3,167	28.23	2,711	28.22
unskilled	1,494	13.32	1,296	13.49
other	3,341	29.78	2,764	28.78
Total	11,218	100	9,605	100
Smoking				
never	3,990	35.56	3,481	35.83
ex smoker	5,232	46.64	4,568	47.02
current	1,997	17.8	1,666	17.15
Total	11,219	100	9,715	100
Drinking				
heavy drinker	489	4.36	427	4.4
normal drinker	7,195	64.15	6,432	66.22
non-drinker	3,532	31.49	2,854	29.38
Total	11,216	100	9,713	100
Physical activity				
very active	6,967	62.13	6,276	65
normally active	3,274	29.2	2,763	28.45
inactive	972	8.67	674	6.94
Total	11,213	100	9,713	100
Chronic Illness				
Yes	6,471	56.85	5,397	55.56
No	4,911	43.15	4,316	44.44
Total	11,382	100	9,713	100

When Discrete-Time Survival Analysis (DTSA) for the fully adjusted model was run, as pointed out in the published paper, observations with missing values for the x variables (i.e. the independent variables, that in my model include disability and all confounders) were discharged. Only 4% of these data were missing, corresponding to three main missing patterns. This is illustrated in Table 5.4. A total of 9,295 observations have complete records, with the most common missing data pattern arising from missing only own occupation ($N = 155$), both income and wealth ($N = 140$) and only father's occupation ($N = 106$). For these variables, simple logistic regression models were run where the dependent variable was the indicator of each variable being missing and the independent variables were some key demographic variables and education for which information was complete. All demographic variables and disability dummy were found predictive of missingness of each of these variables (Table 5.5).

Table 5.4: Missing Data Pattern

N	**Marital status**	**Alcohol**	**Chronic illness**	**Physical activity**	**Education**	**Ethnicity**	**Father's job**	**Income**	**Wealth**	**Occupation**
9295	1	1	1	1	1	1	1	1	1	1
155	1	1	1	1	1	1	1	1	1	0
140	1	1	1	1	1	1	1	0	0	1
106	1	1	1	1	1	1	0	1	1	1
6	1	1	1	1	1	0	1	1	1	1
2	1	0	1	1	1	1	1	1	1	1
2	1	1	0	1	1	1	1	1	1	1
2	1	1	1	1	0	1	1	1	1	1
2	1	1	1	1	1	1	0	0	0	1
1	0	1	1	1	1	1	1	1	1	1
1	1	1	1	0	1	1	0	1	1	1
1	1	1	1	0	1	1	1	1	1	1
1	1	1	1	1	0	1	0	1	1	1
1	1	1	1	1	1	1	1	0	0	0

Table 5.5: Mutually adjusted estimated odds ratios of missingness in wave 1 by variable affected by missingness; estimated by logistic regression

Logistic regression coefficients	**Income**	**Wealth**	**Occupation**	**Father's job**
Age (years)	0.98**	0.98**	1.04***	1.01
Sex (M vs F)	1.62***	1.62***	8.79***	1.03
Disability (yes vs no)	0.74***	0.74***	1.2*	0.96
Education -middle vs high	1.03	1.03	1.3	1.97**
Education -low vs high	0.68	0.68	3.04***	2.44***

The fully adjusted model of DTSA presented in the published paper (table 3) was based on complete Complete Case Analysis (CCA). Below (Table 5.6), the results from the same model, but estimated using FIML are reported, so that the sample size was the same as that of the model where age was the only predictor. The results obtained using FIML were very close to those of Table 3 of Research Paper II, both for men and women.

Table 5.6: Odds ratios estimated by DTSA using FIML

	Time since disability measurement (years)	**OR**	**95% CI**
MALES	1	2.271***	(1.15,3.39)
	2	1.679***	(1.09,2.27)
	3	1.85***	(1.27,2.43)
	4	1.389**	(1.02,1.76)
	5	1.485***	(1.08,1.89)
	6	1.01	(0.73,1.29)
	7	1.344*	(0.95,1.74)
	8	1.315*	(0.95,1.68)
	9	1.412***	(1.07,1.76)
	10	1.916***	(1.34,2.49)
FEMALES	time-invariant effect	1.365***	(1.2,1.53)

CI=confidence interval; OR=odds ratio
$***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Factor loadings

Table 5.7 reports the factor loadings obtained from the first order measurement model with three factors, corresponding to the three domains of disability identified by the WHO's ICF, and the eyesight component. By rule of thumb, if the absolute value of the standardised loading is greater than 0.4, the variable is considered relevant for the particular factor.

Table 5.7: First order measurement model's factor loadings [a]

Impairment	**Estimate (S.E.)**
Hypertension	0.24 (0.015)
Arthritis	***0.615 (0.011)***
Parkinson	0.375 (0.048)
Emotional or psychiatric problems	0.219 (0.021)
Dementia	0.377 (0.058)
Hearing	0.315 (0.011)
Sight	0.366 (0.011)
Sight at distance	0.357 (0.011)
Sight close	0.356 (0.011)
Pain	***0.727 (0.009)***
Chest-pain	***0.454 (0.013)***
Pain across the front of chest	***0.467 (0.017)***
Pain in legs	***0.752 (0.008)***
Dizziness	***0.72 (0.011)***
Shortness of breath	***0.781 (0.007)***
Breath with wheezing	***0.436 (0.016)***
Incontinence	***0.439 (0.016)***
Memory problems	0.251 (0.011)
Depression	***0.439 (0.015)***

Eyesight	**Estimate (S.E.)**
Sight	***0.808 (0.005)***
Sight at distance	***0.774 (0.006)***
Sight close	***0.793 (0.005)***

Activity	Estimate (S.E.)
Walking 100 yards	***0.937 (0.005)***
Sitting for about 2 hours	***0.686 (0.011)***
Getting up from a chair	***0.774 (0.008)***
Climbing several flights of stairs	***0.837 (0.006)***
Climbing one flight of stairs	***0.889 (0.006)***
Stooping- kneeling or crouching	***0.796 (0.007)***
Reaching or extending arms above shoulder	***0.693 (0.012)***
Pulling or pushing large objects	***0.877 (0.006)***
Lifting or carrying weights over 10 pounds	***0.862 (0.006)***
Picking up a 5p coin from a table	***0.644 (0.017)***
Getting dressed	***0.801 (0.009)***
Walking across room	***0.875 (0.011)***
Bathing	***0.851 (0.008)***
Eating	***0.697 (0.021)***
Getting in and out of bed	***0.826 (0.01)***
Toileting	***0.766 (0.015)***
Following conversation	0.31 (0.014)
Balance	***0.79 (0.007)***
Walking quarter of mile	***0.922 (0.004)***
Restless sleep	***0.416 (0.013)***

Participation	Estimate (S.E.)
Preparing a hot meal	***0.874 (0.011)***
Using a map to get oriented	***0.538 (0.021)***
Shopping for groceries	***0.926 (0.007)***
Making telephone calls	***0.527 (0.029)***
Doing work around house or garden	***0.935 (0.006)***
Managing money	***0.615 (0.025)***
Using public transport	***0.885 (0.009)***
Being member of any organization	0.262 (0.015)
Engaging in any activities	***0.474 (0.014)***
Early retirement due to health problems	0.324 (0.022)
Retirement due to health problems	0.385 (0.023)

[a] Factor loadings larger than 0.4 are reported in bold and italics

Likelihood Ratio Test (LRT)

To evaluate whether the effect of each predictor on mortality was constant over the 10 intervals or varied with time, likelihood ratio tests were performed, separately for each of the covariates, by comparing the model including the only covariate (i.e. assuming a time invariant effect) and the model including an interaction between the covariate and time (i.e. allowing for time-varying effect). Results are reported in table 5.8.

Table 5.8: Likelihood Ratio Test

Variable	**Males**			**Females**		
	LR Chi square	**df**	**p-value**	**LR Chi square**	**df**	**p-value**
Disability	24.46	9	0.0036	11.21	9	0.2619
	17.02[a]	9	0.0484	7.84[a]	9	0.5507
Age	23.86	9	0.0045	26.58	9	0.0016
Marital status	33.52	27	0.1804	19.85	27	0.837
Household size	11.26	9	0.2584	9	9	0.4372
Parental status	9.55	9	0.388	9.17	9	0.4217
Ethnicity	3.63	6	0.7267	0.78	3	0.8536
Education	9.27	18	0.9531	18.97	18	0.3936
Occupation	7.96	18	0.9792	22.18	18	0.224
Income	36.52	36	0.4445	26.57	35	0.8464
Wealth	19.52	36	0.9885	25.8	36	0.896
Smoking	19.28	18	0.3749	12.76	18	0.8057
Drinking	14.3	18	0.7097	16.1	17	0.5169
Physical activity	47.65	18	0.0002	24.19	18	0.149
Father's job	20.52	27	0.8081	25.02	27	0.5733
Chronic illness	12.38	9	0.1928	6.74	9	0.664

[a] from full model

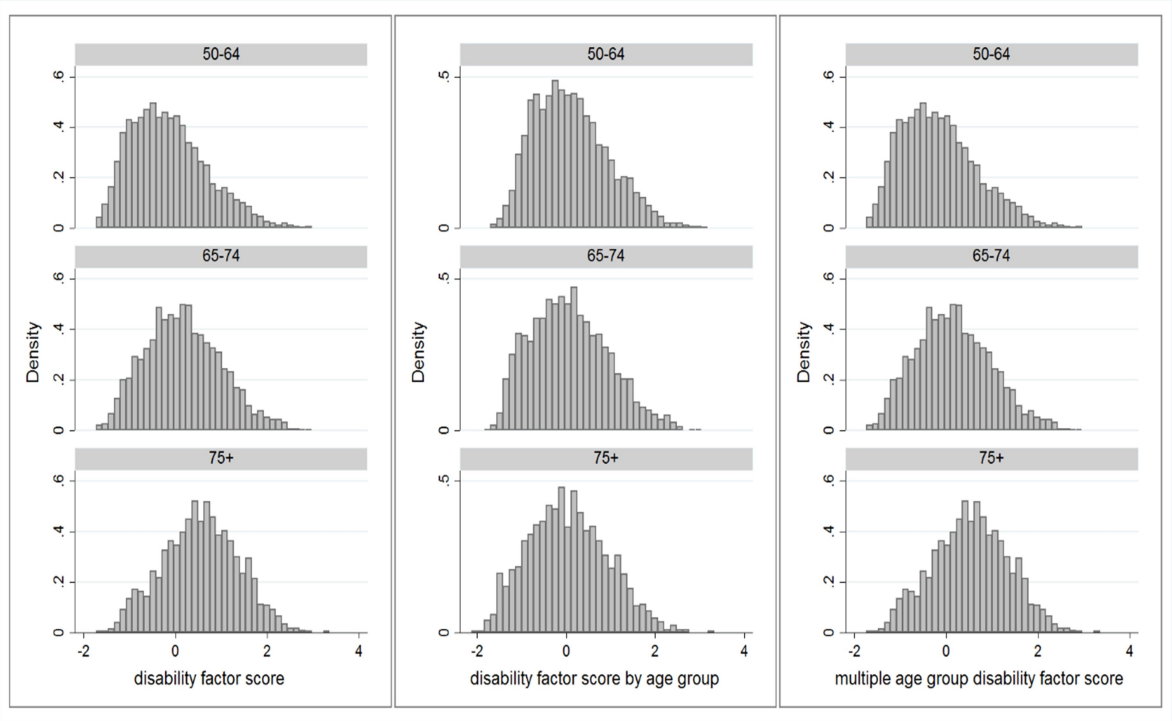
Disability factor scores by age

As a sensitivity analysis, the disability factor score was re-estimated in the general-specific model via multiple group analysis in the whole sample, where the grouping variable is categorical age (50-64 years, 65-74 years and 75+ years). As in the original model, I did not stratify by sex. The factor score derived from this model was then used as the disability explanatory variable in the survival analysis, which was run separately for the three age groups. The results are reported in table 5.9. In this section, I explain in more details why this sensitivity analysis was performed, and present graphically the disability factor scores obtained from the multiple age group model, and also the factor scores additionally estimated stratifying the sample by age group. The graphical representation was not mention in the paper and is not available in the online appendix.

The focus of Research Paper II was on the older population, defined as adults aged 50 years and older. Within this group, age could potentially have an effect in the estimates of the disability factor scores. Therefore, its effect was reassessed by first estimating the factor scores separately by age bands (i.e. 50-64, 65-74 and 75+) and then for the total sample using multiple group analysis (in both cases without stratifying by gender). This is shown in figure 5.1. On the left hand side column, the distribution of the disability factor scores, as estimated in the paper, is shown by age group; in the middle column, the distributions of the factor scores computed stratifying by age group are reported; and on the right hand side the distributions of the disability factor scores estimated by the multiple age group model are also shown by age category. In each model and for each age group, the factor score distributions were very similar to that produced by the original measurement model. In the model stratified by age (middle column), the oldest age group (75+) presented more variation and its disability factor scores had a larger range of values (range of disability factor score for group 75+ was -2.1 to 3.3, s.d. 0.88; while for group 50-64 years from -1.6 to 3.2, s.d. 0.83; and for group 65-74 years -1.7 to 3, s.d. 0.86).

Finally, to check whether the survival analysis results were the same in men and women of younger and older age groups, the disability factor scores used in the paper were replaced with the factor scores from the measurement model fitted using multi-group analysis stratifying the sample by age group. Results are shown in Table 5.9. For women, the effect of

Figure 5.1: Disability factor scores by age group, for various measurement models



Left column reports distribution of disability factor score, as estimated in the paper for each age group. Middle column reports distribution of disability factor score estimated separately for each age group. Right column reports distribution of disability factor score estimated using multiple group analysis.

disability on mortality was stronger in the youngest age group compared to groups aged 65-74 years and 75+ years, whose results were similar to each other. For men, results were similar between the age groups 50-64 years and 65-74 years, while the impact of disability on mortality was found to be weaker among the oldest age group. If men aged 75+ are considered as survivors -as those who made it at least until age 75- this result can be in line with the interpretation of the finding that disability has a time-varying effect in men, which is stronger in the short term and then declines and converges to the level observed among women. From table 5.9 we observe that for men aged 75+ years at wave 1, the impact of disability was roughly as mild as for women. The survival function for this group is implicitly conditional to be alive until age 75; among men, this may mean that survivors made it until the age of 75 because they were a subsample more resilient to disability and therefore the effect of disability on mortality among this group is similar to that experienced by women.

Table 5.9: Odds ratios estimated by DTSA by gender and age group using disability estimated via multiple group analysis by age group

	Time since disability measurement (years)	50-64		65-74		75+	
		OR	**95% CI[a]**	**OR**	**95% CI[a]**	**OR**	**95% CI[a]**
MALES	1	3.622**	(1.11;11.83)	4.076**	(1.1;15.1)	1.283	(0.63;2.61)
	2	1.714	(0.71;4.16)	1.696*	(0.91;3.14)	1.741***	(1.15;2.65)
	3	2.232**	(1.2;4.16)	1.928***	(1.2;3.1)	1.616**	(1.02;2.57)
	4	1.604**	(1;2.57)	1.413	(0.87;2.28)	1.132	(0.75;1.71)
	5	1.405	(0.8;2.47)	1.507*	(0.95;2.4)	1.34	(0.89;2.01)
	6	0.971	(0.48;1.95)	0.866	(0.55;1.36)	1.252	(0.8;1.97)
	7	1.894**	(1.04;3.43)	1.19	(0.73;1.94)	1.129	(0.75;1.7)
	8	2.156***	(1.23;3.77)	1.147	(0.74;1.78)	1.058	(0.69;1.62)
	9	1.267	(0.73;2.21)	1.488*	(0.95;2.33)	1.354	(0.87;2.1)
	10	2.081***	(1.26;3.44)	1.632*	(1;2.67)	2.07***	(1.24;3.45)
FEMALES	time-invariant	1.580***	(1.2;2.08)	1.286**	(1.04;1.59)	1.267***	(1.08;1.48)

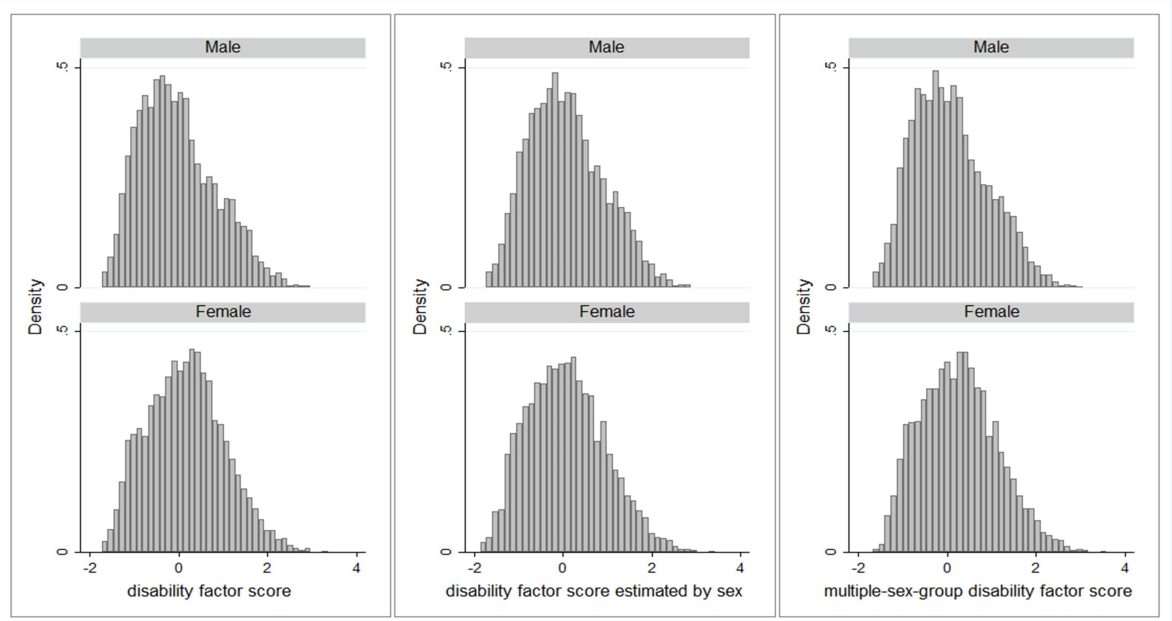
[a] SE estimated from pooled logistic regression
CI=confidence interval; OR=odds ratio
$***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Disability factor scores by gender

Another sensitivity analysis pertained to the potential gender effect in the measurement of disability. For this reason, the general-specific measurement model was replicated (i) stratifying by gender; (ii) via multiple sex group analysis. Similarly to how results were presented for age in the previous section, figure 5.2 illustrates the distribution of factor scores by sex derived from the measurement model used in the published paper (i.e. the model run for the total sample) (on the left), derived from the measurement model run separately for men and women (in the middle), and derived from the measurement model via multiple sex group analysis (on the right). Distributions were very similar across models, with the women having greater frequency of large values, confirming that they suffer from disability more than men. The range of values of factor scores was smaller for men; it ranged from -1.7 to 2.9, -1.7 to 2.8 and from -1.7 to 3 respectively for each model, while female ranges were -1.7 to 3.4, -1.9 to 3.4 and -1.6 to 3.6. When disability was measured stratifying by gender (central graphs), male and female results were closer than those obtained using sex-group analysis (graphs on the right).

Table 5.10 reports the results of the DTSA obtained using the factor scores described

Figure 5.2: Disability factor scores by gender



Left column reports distribution of disability factor score, as estimated in the paper, by sex.
Middle column reports distribution of disability factor score estimated separately for men and women.
Right column reports distribution of disability factor score estimated using multiple group analysis, by sex.

above. Column 1 is for the fully-adjusted model with factor scores estimated separately by gender, and column 2 those obtained from multi-group analysis. Both models produced results very close to each other in terms of direction and magnitude of the association of disability with mortality. Results were also very close to those presented in the published paper, reassuring, as stated, that the disability measurement model does not suffer from a gender effect in its measurement, and therefore the different effect of disability on mortality observed between men and women did not depend on how disability was measured.

Disability factor score and health conditions

The last part of the sensitivity analysis was probably the most important and the most critical. Disability scores were re-estimated excluding some of the variables originally selected. Their selection was motivated by the WHO's ICF guidelines, and following agreement between the three authors in charge of classifying disability items. However, some of these variables can also be considered as health conditions. This was the case for hypertension,

Table 5.10: Odds ratios estimated by DTSA using disability measurement by gender and via multiple group analysis by gender

	Time since disability measurement (years)	(1)		(2)	
		OR	**95% CI[a]**	**OR**	**95% CI[a]**
MALES	1	2.351***	(1.29; 4.28)	2.341***	(1.3; 4.22)
	2	1.824***	(1.28; 2.61)	1.836***	(1.29; 2.62)
	3	1.966***	(1.43; 2.7)	1.945***	(1.42; 2.66)
	4	1.472***	(1.11; 1.94)	1.457***	(1.11; 1.92)
	5	1.479***	(1.11; 1.97)	1.477***	(1.11; 1.97)
	6	1.04	(0.76; 1.42)	1.039	(0.76; 1.41)
	7	1.325*	(0.99; 1.78)	1.319*	(0.98; 1.77)
	8	1.317*	(0.99; 1.75)	1.302*	(0.98; 1.73)
	9	1.406**	(1.06; 1.87)	1.395**	(1.05; 1.86)
	10	2.013***	(1.48; 2.73)	1.989***	(1.47; 2.69)
FEMALES	time-invariant	1.357***	(1.21 ; 1.53)	1.364***	(1.21; 1.54)

[a] SE estimated from pooled logistic regression
CI=confidence interval; OR=odds ratio
$***p < 0.01$, $**p < 0.05$, $*p < 0.1$
(1) fully adjusted model
(2) fully adjusted model controlling also for chronic conditions

arthritis, dementia, Parkinson, psychological problems and depression, all classified within the impairment domain. The rationale for implementing this sensitivity analysis, within the context of the aims of Research Paper II, was to make sure that gender-differences in the association between disability and mortality were not driven by some body functions and structures whose prevalence is more likely to differ between men and women. In a more general perspective, the comparison of measurement models including (i.e. original model) and excluding these variables would reassure that the validity of disability measurement was not jeopardised by the inclusion of items classifiable either as impairment or health conditions.

Figure 5.3 illustrates the distribution of disability factor scores for the total sample obtained from the original model (on the left), and from the model where items about health conditions were dropped (on the right). Observations with complete records on all disability items excluding health condition items were 9,720 (i.e. five observations more than in original model). The mean of the original disability factor scores was 0.0781 (s.d. 0.862, min -1.72, max 3.36), the mean of the disability factor scores calculated without HC items was 0.085 (s.d. 0.851, min -1.6, max 3.33). The analysis was also repeated for the

same set of observations used throughout the paper (i.e. 9,715) and the results were the same. The Pearson's correlation between disability factor scores including and excluding health conditions was equal to 0.9955.

Figure 5.3: Disability factor scores including and excluding health conditions

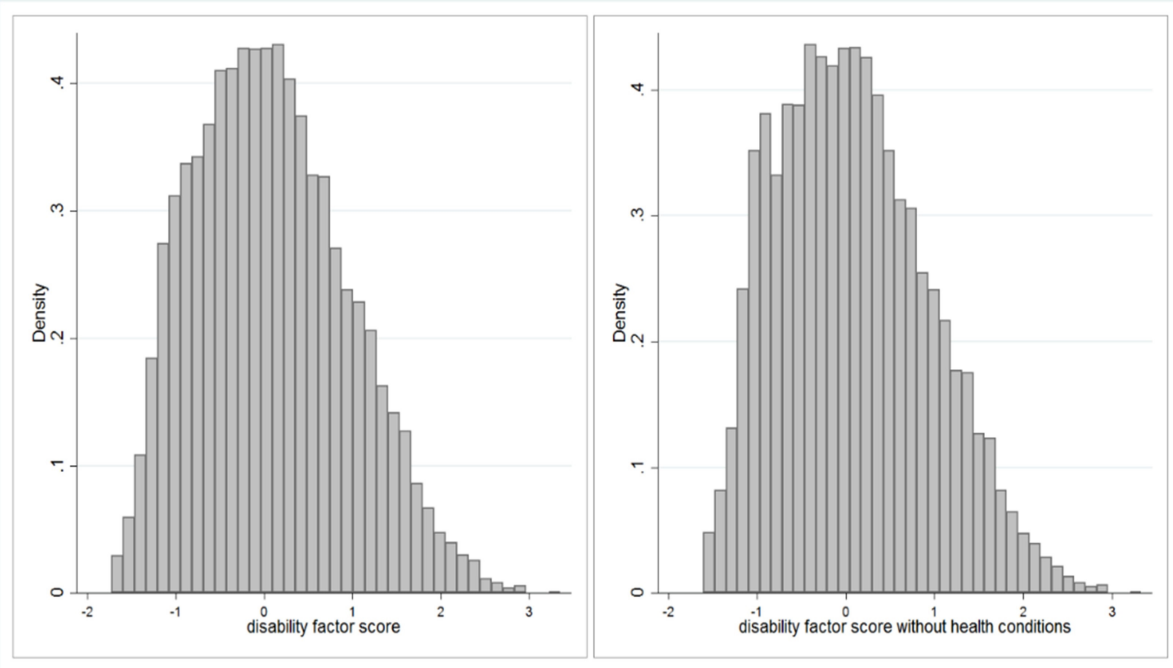


Table 5.11 reports the results of the DTSA obtained using the disability factor score derived from the measurement model not including health conditions. In column 1 the ORs and 95% CIs pertain to the fully-adjusted model, and in column 2 the health conditions excluded from the measurement model were also added as confounders. Again, results were very similar to those presented in the manuscript. Controlling for the health conditions reduced the effect of disability on mortality only mildly.

Conclusions from these findings are twofold. On the one hand, the results reassure us that the inclusion of some variables that can be considered either as health conditions or impairments does not affect the validity of the disability measurement, and also that their role in the association with disability is really minor compared to disability itself (as identified by the core variables). This was not unexpected. According to the WHO's ICF framework, health conditions underlay disability and often their onset precedes and causes the onset of disability, which therefore is a development in the condition and as such its

impact on mortality is stronger. The second conclusion is that, although prevalence of these conditions differs between men and women (e.g. 25.5% of men and 38% of women suffered from arthritis; 14% of men and 18% of women had depression; 6% of men and 9% of women had psychological problems. See table 1 of the manuscript), this did not drive the differences observed across gender in the association of mortality and disability.

Table 5.11: Odds ratio estimated by DTSA using disability measurement not including chronic conditions

	Time since disability measurement (years)	**(1)**		**(2)**	
		OR	**95% CI[a]**	**OR**	**95% CI[a]**
MALES	1	2.19***	(1.24,3.85)	2.14***	(1.21,3.78)
	2	1.8***	(1.28,2.54)	1.76***	(1.24,2.49)
	3	1.87***	(1.38,2.53)	1.82***	(1.34,2.47)
	4	1.44***	(1.1,1.88)	1.4**	(1.07,1.84)
	5	1.46***	(1.1,1.93)	1.42**	(1.07,1.88)
	6	1.04	(0.77,1.41)	1.01	(0.74,1.37)
	7	1.33*	(0.99,1.77)	1.28*	(0.96,1.72)
	8	1.32**	(1,1.75)	1.28*	(0.97,1.69)
	9	1.38**	(1.04,1.82)	1.33**	(1,1.77)
	10	1.96***	(1.45,2.64)	1.9***	(1.41,2.56)
FEMALES	time-invariant	1.39***	(1.23,1.56)	1.38***	(1.23,1.56)

[a] SE estimated from pooled logistic regression
CI=confidence interval; OR=odds ratio
$***p < 0.01$, $**p < 0.05$, $*p < 0.1$
(1) fully adjusted model
(2) fully adjusted model controlling also for chronic conditions

Addedum to Research Paper II

As additional robustness check, I considered a model for the pooled sample that interacts disability with gender and formally examines whether there are differences in the relationship between disability and mortality in men and women. A significant interaction between gender and disability would suggest a different effect of disability on mortality between men and women, and therefore reinforce the finding of Research Paper II that disability is more fatal for men than women in the short term, but the association tends to converge to the level observed over time in women. The sensitivity analysis was done considering both the time-invariant as well as the time-varying effect of disability on mortality and

the time-invariant and time-varying effect of its interaction with gender. The most general model (i.e. without constraining disability and its interaction with gender to have a time invariant effect) confirmed the results presented in Research Paper II. The time-varying interaction coefficient was only marginally significant at time one and had a negative coefficient (equal to -0.189, p-value=0.063), meaning that at time one women with disability had lower probability of dying compared to men. For deaths that occurred 2+ years after disability was measured, the interaction coefficients were not significant, suggesting that the effect of disability on mortality was not different for men and women. This reinforces the findings obtained stratifying the sample by sex, where it was found that disability has a time-varying effect for men which is stronger in the short time and then tends to converge to womens level. Time-varying and time-invariant coefficients of interaction terms are presented in table 5.12.

Table 5.12: Odds Ratios of interaction between disability and gender for mortality

Time since disability measurement (years)	**OR**	**95% CI**
1	0.828*	(0.68; 1.01)
2	0.94	(0.77; 1.15)
3	0.768	(0.54; 1.1)
4	0.793	(0.56; 1.12)
5	0.9	(0.63; 1.28)
6	1.009	(0.68; 1.49)
7	1.031	(0.73; 1.45)
8	1.118	(0.76; 1.63)
9	1.155	(0.81; 1.64)
10	0.818	(0.56; 1.19)
time invariant effect	0.907*	(0.81 ;1.01)

CI=confidence interval; OR=odds ratio
$***p < 0.01$, $**p < 0.05$, $*p < 0.1$

5.4 Conclusions

The key findings of Research Paper II concern gender differences in the association between disability at baseline and mortality observed over a decade.

Of particular interest is the finding that among men disability appeared to be very strongly associated with mortality in the short term and then the effect tended to converge to that observed among women, for whom disability had a time-invariant effect. This may be key to interpret and forecast future trends in health and mortality, and I make explicit reference to this point in Research Paper IV, when I discuss gender differences in trends in Disability-Free Life Expectancy (DFLE) in England over the past decade.

In Research Paper II disability was measured as a continuous variable and the disability odds of dying increased as the baseline disability score increased, both for women and men. To answer the research question asked in the paper, considering disability as a quantitative variable was more advantageous than using it as a binary or categorical variable. However, for the progression of this thesis, it is needed to classify disability as a discrete variable, in order to combine it with mortality and estimate DFLE. The classification of disability as a categorical variable is tackled in Research Paper III, and the results are then used for answering the research questions asked in Research Paper IV.

Research Paper II has been published in the European Journal of Epidemiology in 2015 [42]. A preliminary version of its abstract was published in the Journal of Epidemiology and Community Health [44]. The work was presented as a poster during the Research Degree Students' Poster Day at London School of Hygiene and Tropical Medicine (2016) and awarded the prize of best poster (available in the appendix, figure 9.2). It was also presented in its preliminary stage at the 59^{th} Annual Scientific Meeting of the Society for Social Medicine in September 2015 and in its final stage at the international conference of the Population Association of America in April 2016. The slides of the most updated presentation are also included in the appendix (figure 9.1).

Chapter 6

Levels of disability

6.1 Preamble

Research Paper III addresses objective 4 and contributes to the debate on the usefulness and relevance of adopting a finer categorization of disability, and discusses why it is important to go beyond a binary definition of disability, as it is currently common in public health research and evaluation.

- **Research Objective 4:** *To identify an optimal number of severity levels of disability based on empirical evidence*

To achieve this objective, I start from the continuous score of disability estimated in Research Paper II and convert it into a binary measure. Then, adopting a different approach, disability is classified as a multi-categorical variable. Lastly, the alternative specifications of disability levels are compared to select the most useful classification for health evaluation. This process corresponds to address two research questions:

- How severe does continuous disability need to be for a person to be classified as disabled?

- Would a finer classification of disability better capture clustering of disability items and of heterogeneities in later disease and mortality rates?

Disability is measured according to the theoretical framework used in Research Paper II and illustrated in Chapter Three. Different approaches and methods are used to

operationalise disability as a binary and multi-categorical summary variable, but they are all based on the same set of variables identified in Research Paper II according to the International Classification of Functioning Disability and Health (ICF)'s framework.

6.2 Research Paper III

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Benedetta Pongiglione
Principal Supervisor	Bianca De Stavola
Thesis Title	Gender inequality in healthy ageing: a study of the English older population over a decade

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Disability and Health Journal
Please list the paper's authors in the intended authorship order:	Pongiglione B., Ploubidis G.B., De Stavola B.L.
Stage of publication	**In press**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	

Student Signature: Benedetta Pongiglione **Date:** 25 January 2017

Supervisor Signature: Bianca L De Stavola **Date:** 25 January 2017

Disability and Health Journal xxx (2017) 1–9

Contents lists available at ScienceDirect

Disability and Health Journal

journal homepage: www.disabilityandhealthjnl.com

Levels of disability in the older population of England: Comparing binary and ordinal classifications

Benedetta Pongiglione [a, *], George B. Ploubidis [b], Bianca L. De Stavola [a]

[a] Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, UK
[b] Centre for Longitudinal Studies, Department of Social Science, UCL Institute of Education, University College London, 55-59 Gordon Square, London, WC1H 0NU, UK

ARTICLE INFO

Article history:
Received 13 October 2016
Received in revised form
19 January 2017
Accepted 21 January 2017

Keywords:
Disability
International classification of functioning disability and health
England
Older population
Latent classes
Latent growth model

ABSTRACT

Background: Recent studies suggest the importance of distinguishing severity levels of disability. Nevertheless, there is not yet a consensus with regards to an optimal classification.
Objective: Our study seeks to advance the existing binary definitions towards categorical/ordinal manifestations of disability.
Methods: We define disability according to the WHO's International Classification of Functioning, Disability and Health (ICF) using data collected at the baseline wave of the English Longitudinal Study of Aging, a longitudinal study of the non-institutionalized population, living in England. First, we identify cut-off points in the continuous disability score derived from ICF to distinguish disabled from non-disabled participants. Then, we fit latent class models to the same data to find the optimal number of disability classes according to: (i) model fit indicators; (ii) estimated probabilities of each disability item; (iii) association of the predicted disability classes with observed health and mortality.
Results: According to the binary classification criteria, about 32% of both men and women are classified disabled. No optimal number of classes emerged from the latent class models according to model fit indicators. However, the other two criteria suggest that the best-fitting model of disability severity has four classes.
Conclusions: Our findings contribute to the debate on the usefulness and relevance of adopting a finer categorization of disability, by showing that binary indicators of disability averaged the burden of disability and masked the very strong effect experienced by individuals having severe disability, and were not informative for low levels of disability.

1. Introduction

A large body of research on the conceptualization and measurement of disability has been published, accompanied by reviews of alternative disability models (with corresponding measurement methods[1]) and studies of the methodology for the measurement of disability. The challenges faced in this field derive from the fact that disability is often measured by self-reported responses to survey questions,[2,3] with problems including the wording of the questions, the time periods for which disability is reported, and the difficulties in administering surveys (see for example Freedman,[4] Jette[5]). On the other hand, the challenge of coding survey data in studies of disability, and of choosing between binary or more refined classifications of disability has received remarkably less attention. At the same time, the relevance of identifying meaningful classification of disability is becoming clearer, while there is no general consensus in the literature with regards to the optimal number of disability grades/levels. For example, a recent study set in the UK[6] and studying trends in disability-free life expectancy has shown that the increase in number of years with any disability was higher for periods with mild disability than for those with moderate or severe disability. This suggests how important it is to distinguish different levels of severity of disability when assessing disability and mortality trends. Another example is the recent study of Wolf and colleagues (2015)[7] that identified three distinctive trajectories of disability, which differed with respect to their pace of decline. The authors acknowledged as a limitation of their work that they relied

* Corresponding author. Tel.: +44 7949131608.
E-mail addresses: benedetta.pongiglione@lshtm.ac.uk (B. Pongiglione), g.ploubidis@ucl.ac.uk (G.B. Ploubidis), Bianca.DeStavola@lsthm.ac.uk (B.L. De Stavola).

http://dx.doi.org/10.1016/j.dhjo.2017.01.005

on a binary definition of disability and that a more finely graded measure might have led to different trajectories.

An attempt to categorize disability by severity level was made by Manton and Gu (2005).[8] They identified, among the older US population, six groups with distinctive aspects of disability ranging from active to frail and a seventh group comprising of nursing home residents. By looking at disability trends over seventeen years, the authors claimed that simply considering the 'per annum decline in chronic disability in the US elderly population (..) masks variation in per annum changes in the prevalence of disability and institutionalized residents' (Manton and Gu (2005), p.32[8]). Changes were found to be different across disability categories and by identifying where they occurred, the authors calculated changes in Medicare costs and related savings in population shifts from a severely disabled category to the non-disabled group.

The lack of equivalent evidence for the UK and, more generally, the scarcity of studies on more refined disability classifications, along with the potential policy relevance of addressing this gap, have motivated our work. Although the correct balance between the need for identifying a more refined disability grouping and the risk of over-classification is not easily achievable, this paper attempts to explore the advantages and disadvantages of binary versus multi-categorical classifications, by examining their discriminant power in terms of different health outcomes.

This study seeks to (1) produce binary and multi-categorical classifications of disability; (2) identify an optimal number of categories of disability using alternative criteria; (3) examine whether a multi-categorical classification may be advantageous compared to a binary.

To do so, we first measure disability as a continuous score and ask "How severe does continuous disability need to be for a person to be classified as disabled?". The rationale for finding a threshold for a binary classification, comes from the need for summarising the information, for example to provide population estimates of disability prevalence.

Splitting the population into two categories, however, may be too simplistic. Therefore, the second question we ask is "Would a finer classification of disability better capture clustering of disability items and of heterogeneity in later disease and mortality rates?". To answer this question, we identify boundaries among categories of disability severity in terms of their association with health function and mortality, adopting a similar approach as the one adopted by Serlin and colleagues[9] who delineated different levels of cancer pain severity. The rationale behind assessing the association between disability and selected health outcomes was derived from three hypotheses about why people with disabilities may have poorer mental and physical health than their non-disabled peers. First, the experience of living with a disability could lead to mental health problems and worse physical conditions; second, people with mental and physical health problems could be more likely to subsequently become disabled; and third, other factors, such as socioeconomic circumstances, might independently increase the risk of disability and mental ill health. Given that we measure disability at baseline and health outcomes over the course of the following ten years, we mainly explore the first hypothesis, supported by evidence on the association between disability and mental health,[10] as well as studies investigating the association between physical disability (or physical activity)[11,12] and muscle strength and quality and functional limitation.

To address these research questions, we rely upon the WHO's conceptualization of disability as defined in the International Classification of Functioning, Disability and Health (ICF),[13] and use a common set of disability items selected according to the ICF conceptual framework to derive first a continuous score of disability and relative cut-off points, and then to derive a multi-categorical disability summary performing latent class analysis (LCA).

2. Material and methods

2.1. Data source and sample

This study used data drawn from the English Longitudinal Study of Ageing (ELSA). Briefly, ELSA is a longitudinal study of the non-institutionalized population, living in England, who were aged 50 years or older at the time of interview. 11,391 core-member respondents were recruited at wave one in 2002/2003. For our analysis, we included all participants who had complete records on all disability items at wave 1, leaving us with a sample of 9715 (in the Appendix, section A, we justify this choice). We take advantage of the longitudinal nature of the study and also use outcome measures from the following five waves of the survey, i.e. up to 2010/2011. A detailed description of the ELSA cohort profile has been released recently.[14]

2.2. Measures

2.2.1. Disability

Variables describing disability were chosen according to the WHO's ICF framework. In this study we rely on the screening and selection carried out in a previous work.[15] There, a total of fifty items was selected to model disability. Briefly, body function and structure were measured using variables such as self-rated eyesight and hearing, chronic conditions such as high blood pressure and arthritis, and questions about pain; activity limitations were measured through activity of daily living (ADLs) and mobility functions; instrumental activities of daily living (IADLs) and various limitations due to health problems, such as in working, were selected for assessing participation restrictions. Variables were all either dichotomous (i.e. yes/no answer) or ordered categorical. A detailed description of the items and selection process is available elsewhere.[15]

2.2.2. Health measures and death

Information on deaths that occurred from 2002 to 2011 was freely available, and for respondents who gave their consent to link their data to the National Health Service Central Register (NHSCR) time of death by year was also disclosed.

A number of health indicators were considered and selected to cover different spheres of health, including measures of mental health and anthropometric measures for physical domains. Here we present one outcome for each group, although analyses were replicated for more indicators and are available upon request. Mental health function was measured at every wave using eight items of the CES-D scale and treated as a continuous variable ranging from 0 to 8. Anthropometric measurements for physical functioning were assessed during the nurse visits, which took place every two waves, i.e. the second, fourth and sixth waves. Physical functioning was assessed through grip strength in the dominant hand and was estimated by the average of three measurements done using the Smedley dynamometer.

2.2.3. Confounders

When assessing the association of disability with mortality and selected health outcomes, we controlled for a number of confounders, all of them measured at wave 1. We only considered these early measures to avoid the issues arising from later values of these confounders lying on the causal pathway between disability and mortality: controlling for them would remove some of the association between exposure (disability) and outcome.[16] The confounders in the model included demographic characteristics such

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

as ethnicity, marital and parental status and household size; socioeconomic position measured through income, wealth, occupation and education; health-related behaviours including smoking, drinking and physical activity; and the presence of limiting long-lasting illness and socioeconomic background represented by father's occupation when respondents were fourteen.

2.3. Analysis

The methods used are presented in separate paragraphs, according to the research question they address. All analyses were carried out separately for men and women.

2.3.1. Binary disability

In order to capture disability as a binary variable, we used the continuous score estimated in our previous study from the battery of 50 items described above using a latent variable measurement model. A full description of the model is available elsewhere.[15]

Here we intended to find a threshold in this continuous score that optimally discriminates disabled and non-disabled individuals. To do so, we adopted two approaches:

1. We follow the WHO's strategy[13] and replicated the WHO's approach looking at the average disability score observed among those reporting at least one limitation in any of the six disability questions selected by the Washington Group (WG).[17] The six disability domains identified as crucial by the WG include problems in seeing, hearing, walking or climbing steps, remembering or concentrating, washing all over or dressing and communicating, for example understanding or being understood.
2. We used the Receiver Operating Characteristic (ROC) methods[18] to assess the agreement of our score with an external gold standard. The external gold standard we chose consists of receiving health or disability benefits. The cut-off in the continuous disability score was then chosen using two alternative criteria[19] known as the point on the curve closest to the $(0, 1)$ (where specificity $= 1$ and sensitivity $= 1$), and the Youden index.

The disability prevalence resulting from these approaches were then compared with national statistics on the proportion of disabled people in the UK in 2002. National data were collected in the General Lifestyle Survey.[20]

2.3.2. Multi-categorical disability

To produce categorical measures of disability, LCA was performed, using all 50 binary and categorical variables previously identified as indicators of disability. An individual is assigned to be member of a class according to his/her highest probability of being in that class, even though an individual may have several classes to which he/she is a partial membership.[21] We explore models with two through six latent classes of a latent variable (the algebraic notation is available in the Appendix, section B). The choice of the number of classes, i.e. of best-fitting model to represent categorical disability, is based on three sets of criteria. The first set consists of statistical indicators, including entropy, the Bayesian Information Criterion (BIC) and the bootstrap likelihood ratio test (BLRT).[22] The second set of criteria consisted of comparisons of each model's estimated probabilities of endorsing a disability item across classes and the assessment of whether they highlight informative differences. The third criterion used external validation of each model's predicted disability classes in terms of their association with mortality and health, as explained in the next paragraph.

2.3.3. Association of disability with mortality and health

We considered the association between each version of categorical disability, i.e. with classes from two to six classes, and each of the following outcomes: mortality by 10 years since entry, and the longitudinal trajectories of grip strength and mental health.

Association between disability and mortality was assessed using discrete-time survival analysis (DTSA) through pooled logistic regression models,[23] with measures of effect expressed as odds ratios.

The association with health outcomes was parameterised using latent growth models (LGM).[24] Fig. 1 illustrates the conceptual model of the association of latent categorical disability and mental health observed at each wave (i.e. six-time points) selected here as an illustrative outcome, controlling for all confounders (age and all the demographic and socioeconomic variables described in the previous section). The algebraic notation to express the top part of Fig. 1 (i.e. latent growth model) can be found in the Appendix, section C. Equivalent models were fitted for all the other health outcomes.

Missingness in outcome observations was assumed to be missing at random (MAR) and maximum likelihood estimation was used. This means that we assume that missingness in outcome data was explained by observed outcomes at other waves and the variables included in the model. These variables however also suffer from missingness; for this reason, we only include respondents that have complete confounder and exposure data.

3. Results

3.1. Sample characteristics

Of 9715 total participants included in the analyses, 54% were women, and the overall average age was 64 years (64.4 and 64.8 for men and women respectively). Over the course of the study (from 2002 to 2011), 21% of males and 16% of female respondents died corresponding to 1775 respondents in total. Of the 50 disability items, women had higher prevalence than men in almost all variables, with the exception for example of difficulty in communicating and being engaged in social activity. Prevalence of all disability indicators by gender are available in Appendix Table A1 and details on summary statistics of health outcomes data across waves are available in Table 1.

3.2. Binary disability

According to national statistics for the 45 + UK population in 2002, 32.8% of men and 32.9% of women have disability.[20]

The continuous disability score developed in a previous work[15] ranged from −1.71 to 2.85 among men and from −1.85 to 3.42 among women (see Fig. 2). The average score among women was 0.08 and 0.09 among men.

To reproduce the WHO's strategy for finding the disability cut-off, we estimated first the proportions of respondents reporting at least one limitation in the six WG activities. 62.4% of men and 60.7% of women were found to have at least one limitation in the six WG activities. The mean disability score among the respondents belonging to this group was 0.44 and 0.47 for men and women respectively. Setting the cut-off point at these values led to 31.5% of male respondents and 31.7% of female respondents being classified as having disability.

When we used ROC analysis to set the threshold for distinguishing disabled and non-disabled individuals, we found that the cut-off for men was 0.51, according to both criteria, i.e. the Youden index and minimization of r. For women, the cut-off was 0.50 when using the minimum value of r and 0.58 based on the

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

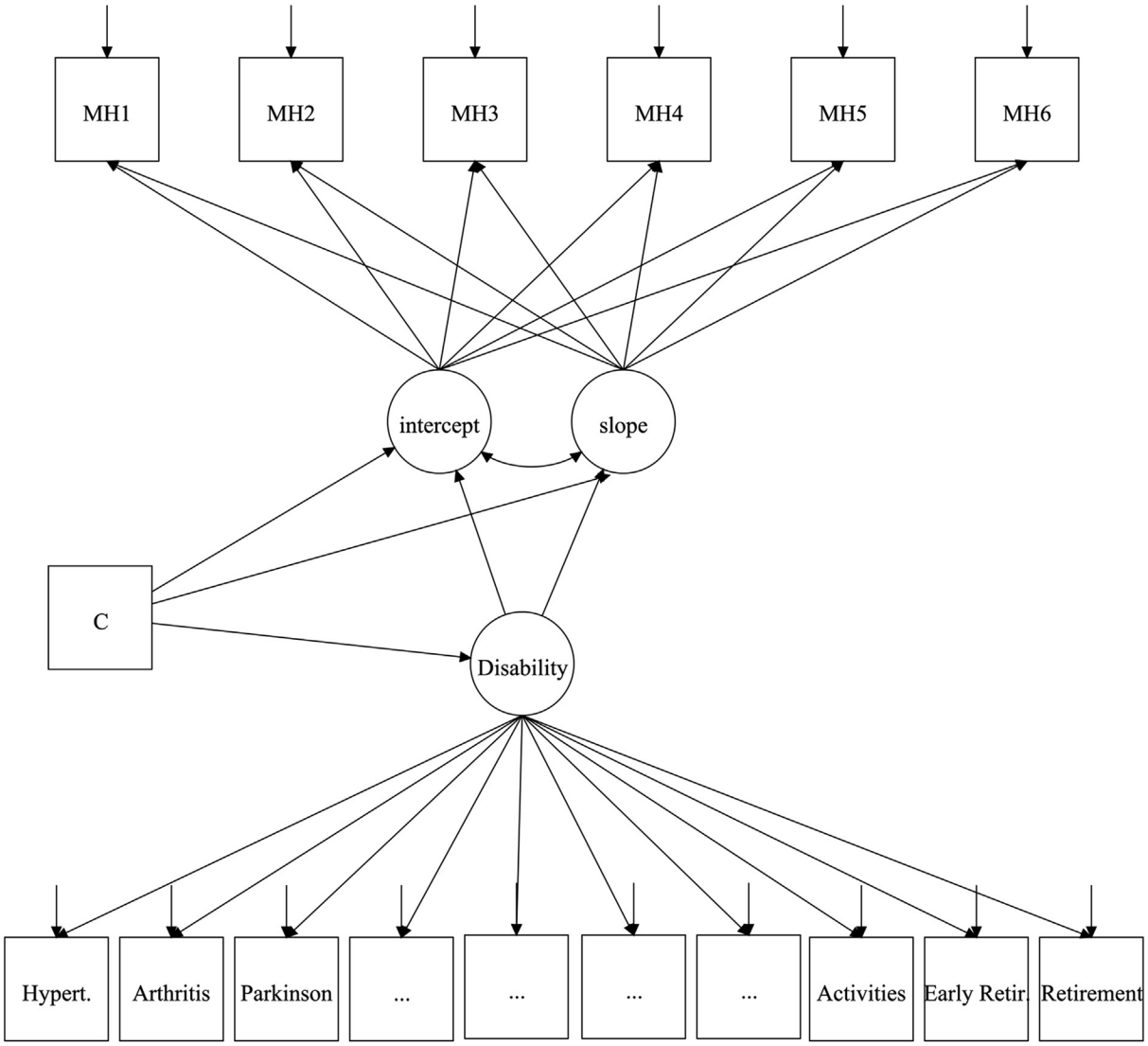


Fig. 1. Conceptual model including measurement model of disability and latent growth model for mental health measured on six occasions. MH = Mental Health; C = confounders (it includes all confounders listed in the text); " …" only some of the 50 disability items were reported to ease the reading of the model.

Youden index. According to these values, 29.2% of male respondents and either 30.3% -according to Youden index- or 27.9% -according to minimum value of *r*-of women were considered disabled.

A graphical summary of these results is provided in Fig. 2. It shows the distribution of the continuous disability score by gender and the cut-off points obtained using the two approaches (with two results presented when using the ROC method for women).

3.3. Categorical disability

Latent class models were fitted separately by gender; their fit

Table 1
Sample size and mean scores of health outcomes by wave at measurement.

	Mental health score[a]		Grip strength[b]		Coagulation[c]	
	n	Mean score (s.d.)	n	Mean score (s.d.)	n	Mean score (s.d.)
Wave 1	9715	1.51 (1.92)				
Wave 2	7743	1.52 (1.91)	6839	29.4 (11.43)	5382	3.2 (0.72)
Wave 3	6630	1.46 (1.92)				
Wave 4	5792	1.36 (1.86)	5021	28.1 (11.29)	3513	3.4 (0.56)
Wave 5	5389	1.48 (1.91)				
Wave 6	4889	1.28 (1.8)	4238	26.9 (10.55)	2966	3 (0.54)

[a] Scale from 0 (no symptoms in any CES-D item).
[b] Measured in kg and averaged over three measurements.
[c] Measured in g/l.

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

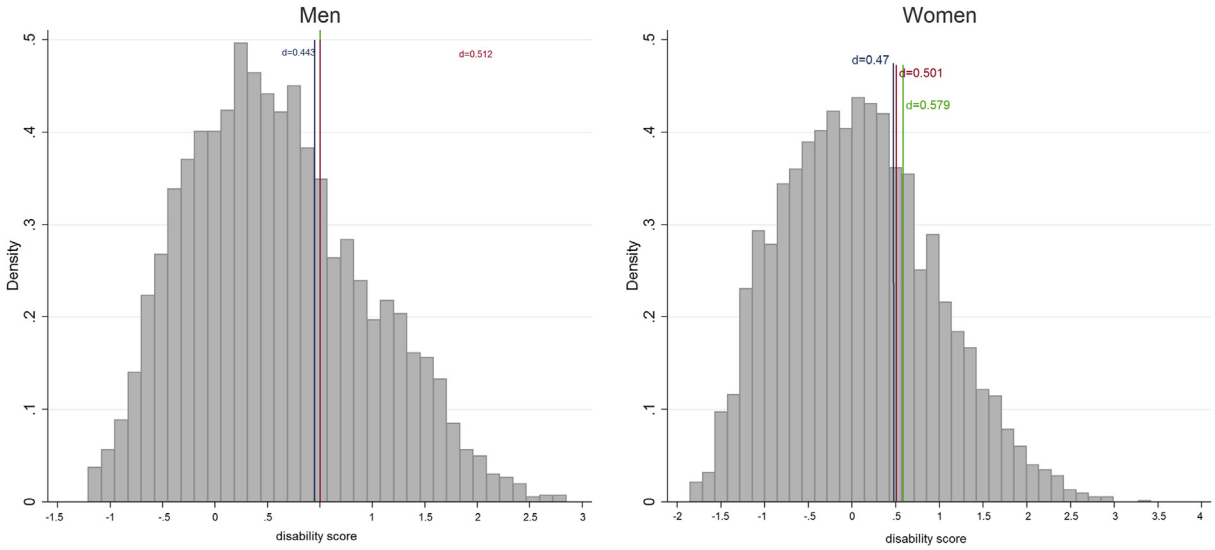


Fig. 2. Observed distributions of continuous disability score and cut-off points identified by the two approaches, by gender. Blue line: cut-off according to WHO's approach. Red line: cut-off according to ROC curve, minimazing *r*. Green line: cut-off according to ROC curve, maximizing *j*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

indicators are presented in Appendix (Table A2). Overall, all models had entropy higher than 0.88, indication of good allocation quality for both for men and women. For both genders, the highest entropy was observed in the model with three classes, whilst the lowest BIC was in the 6-classes model. When assessing the BLRT, problems with local maxima occurred, therefore BLRT was not used as a criterion for model selection. The proportions of respondents assigned to different categories of disability by each model are reported in Table A3 in the appendix.

The estimated probabilities of a positive item response for each disability class provide a first description of each disability class. They are illustrated below in Fig. 3a and b for 3-class and 4-class models and available in Appendix for 5-class and 6-class models (Fig. A2). These probabilities were quite distinct across disability groups in the 3-class model both for men and women. We labelled the category with lowest probabilities "no disability", the intermediate "mild disability" and the group with highest probability "severe disability". The largest differences were found for the items concerning pains and mobility, such as walking a quarter of mile, or climbing stairs. In the 4-class model, categories were labelled as "no disability", "mild disability", "moderate disability" and "severe disability". "No disability" and "mild disability" overall presented similar estimated probabilities, but some differences were still noticeable among impairment items. "Moderate disability" had higher estimated probability than "no disability" and "mild disability" for all items, and the largest differences were observed, again, for some of the impairment variables, such as having pain, and in the mobility items. Severe disability had the highest estimated probabilities for all items and the gap with moderate disability was particularly large with regards to the activity domain as well as for items describing participation. In models with five and six classes (see Appendix, Fig. A2), the groups having the lowest levels of disability (i.e. groups 1, 2 and 3) tended to have very similar estimated probabilities, suggesting that there was not much difference in the endorsement of disability items among these groups.

3.4. Association of alternative specifications of disability levels with health and mortality

Before showing whether alternative specifications of disability appear to explain some outcome variation over time and in which direction, we briefly present the results of the association of each alternative specification of disability measured at baseline with health outcomes measured at wave 6; corresponding tables by gender are available in the appendix (Table A4). We use this analysis as a preliminary step to see whether disability affects health measured after a ten-year lag, regardless of how this relationship changed over this interval, to facilitate the interpretation of the results from the LGM presented below. In synthesis, there seemed to be a severity gradient, where those belonging to the most disabled group performed the worst compared to non-disabled, but the number of significant disability levels strongly varied depending on the outcome considered.

Table 2 shows the association of binary and multi-categorical specifications (i.e. 3-class and 4-class models) of disability with the estimated growth parameters of selected outcomes, controlling for the complete set of confounders, while model estimated means of latent growth factors and the results for 5-class and 6-class models are available in the appendix (Tables A5 and A6 respectively). Binary disability (however specified) (first three rows) was associated with a significant worse intercept (i.e. mean outcome at baseline) for disabled versus not disabled men and women. The trajectories of grip strength in women as well as the trajectories of mental health in men appeared to converge with those of non-disabled ones, as shown by the opposite sign of the disability effect on the intercepts and slopes estimated in the growth model.

To assess the existence of a severity gradient of disability, we looked at multi-categorical disability and observe whether there was evidence of heterogeneity or linearity of associations. A disability gradient was found both in the 3-class and 4-class models. In the 3-class model, at baseline (as measured by the intercept coefficient estimated in the LGM) those having mild and severe disability presented lower grip strength, higher mental health problems and higher probability of dying compared to non-disabled, with the disadvantage being larger for those suffering from severe disability. In the 4-class model, all disability categories presented worse health conditions and higher odds of dying at each time point compared to no disability. The size of the intercept coefficients of moderate disability in the 4-class model was close to that of mild disability in the 3-class model. Results for 5 and 6-class models are available in the appendix and overall indicate that the

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

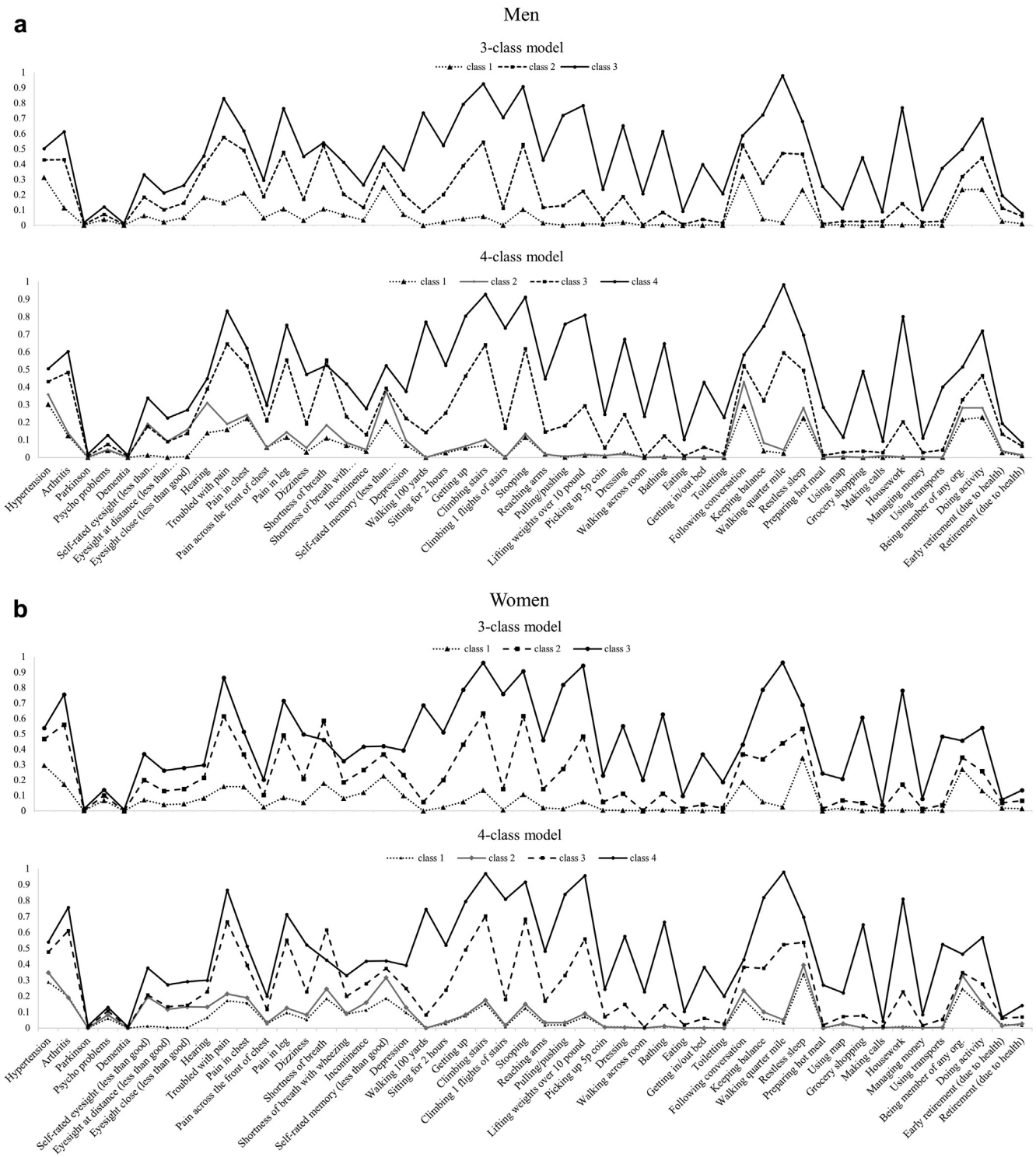


Fig. 3. a Probability of each disability item estimated by the 3 and 4-class models, men. b Probability of each disability item estimated by the 3 and 4-class models, women.

intercepts and slopes of all outcomes among individuals assigned to disability groups 2 and 3 were not significantly different from those assigned to group 1 (i.e. no disability), and for higher severity of disability groups, the differences in grip strength, mental health at the first wave and odds of dying relative to non-disabled group was progressively larger as disability increased (intercept coefficients).

Therefore, comparing the results of the growth models for each disability specification, we observe that estimates for binary disability in terms of size were in between the coefficients observed for mild and severe disability in multi-categorical disability. Intercept coefficients of the association between mild disability in the 3-class model and each outcome were very close to those of moderate disability in the 4-class model; and intercept coefficients of severe disability in 3-class and 4-class models were close to each other. The lowest disability group in the 4-class model (i.e. mild disability) presented significant worse health and mortality conditions

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

Table 2
Association of disability groups with mortality and a number of mental and physical health indicators for a selection of models[a].

Women						
Specification of disability level		Mean grip strength		Mental Health		Death
		Intercept	Slope	Intercept	Slope	β coef
2 classes (WHO)	disabled	−3.387***	0.259***	1.099***	−0.006	1.456***
		(−3.9; −2.88)	(0.13; 0.39)	(0.98; 1.22)	(−0.04; 0.03)	(1.219; 1.740)
2 classes (ROC curve min r)	disabled	−3.439***	0.236***	1.108***	−0.008	1.486***
		(−3.96; −2.92)	(0.11; 0.37)	(0.98; 1.23)	(−0.05; 0.03)	(1.244; 1.774)
2 classes (ROC curve max j)	disabled	−3.702***	0.285***	1.115***	−0.016	1.452***
		(−4.23; −3.17)	(0.15; 0.42)	(0.99; 1.24)	(−0.05; 0.02)	(1.217; 1.732)
2 classes - LCA	disabled	−3.68***	0.281***	1.223***	−0.025	1.407***
		(−4.21; −3.15)	(0.15; 0.42)	(1.1; 1.35)	(−0.06; 0.01)	(1.18; 1.678)
3 classes - LCA	mild	−2.092***	0.053	0.827***	0.014	1.333***
		(−2.56; −1.62)	(−0.06; 0.17)	(0.71; 0.94)	(−0.02; 0.05)	(1.094; 1.624)
	severe	−5.398***	0.363***	1.971***	−0.072***	1.779***
		(−6.16; −4.63)	(0.17; 0.56)	(1.8; 2.15)	(−0.12;−0.02)	(1.390; 2.276)
4 classes - LCA	mild	−0.594**	0.003	0.191***	−0.002	1.212
		(−1.12; −0.07)	(−0.12; 0.13)	(0.07; 0.31)	(−0.04; 0.03)	(0.941; 1.563)
	moderate	−2.604***	0.118*	0.94***	0.003	1.477***
		(−3.14; −2.07)	(−0.01; 0.25)	(0.81; 1.07)	(−0.03; 0.04)	(1.176; 1.855)
	severe	−5.829***	0.32***	1.994***	−0.082***	1.850***
		(−6.69; −4.97)	(0.1; 0.54)	(1.8; 2.18)	(−0.14;−0.02)	(1.406; 2.433)
Men						
Specification of disability level		Mean grip strength		Mental Health		Death
		Intercept	Slope	Intercept	Slope	Beta
2 classes - WHO	disabled	−2.396***	0.038	1.164***	−0.071***	1.548***
		(−3.2; −1.59)	(−0.17; 0.24)	(1.05; 1.28)	(−0.11; −0.04)	(1.306; 1.836)
2 classes - ROC curve	disabled	−2.452***	0.022	1.175***	−0.073***	1.529***
		(−3.28; −1.62)	(−0.19; 0.24)	(1.06; 1.29)	(−0.11; −0.04)	(1.291; 1.810)
2 classes - LCA	disabled	−3.115***	0.029	1.311***	−0.068***	1.439***
		(−4.02; −2.21)	(−0.21; 0.27)	(1.19; 1.44)	(−0.11; −0.03)	(1.213; 1.708)
3 classes - LCA	mild	−1.867***	0.09	0.854***	−0.055***	1.349***
		(−2.67; −1.06)	(−0.11; 0.29)	(0.74; 0.97)	(−0.09; −0.02)	(1.131; 1.608)
	severe	−4.691***	0.288*	1.978***	−0.147***	1.726***
		(−5.94; −3.44)	(−0.05; 0.62)	(1.81; 2.14)	(−0.2; −0.09)	(1.375; 2.165)
4 classes - LCA	mild	−0.669*	−0.046	0.234***	0.002	1.087
		(−1.45; 0.12)	(−0.24; 0.14)	(0.13; 0.34)	(−0.03; 0.03)	(0.885; 1.335)
	moderate	−2.22***	0.041	1***	−0.054***	1.529***
		(−3.13; −1.31)	(−0.19; 0.27)	(0.87; 1.13)	(−0.09; −0.02)	(1.252; 1.868)
	severe	−5.474***	0.426**	2.179***	−0.19***	1.978***
		(−6.84; −4.11)	(0.06; 0.79)	(2; 2.36)	(−0.25; −0.13)	(1.544; 2.533)

β coef, β coefficient; 95% confidence intervals in brackets (); ***p < 0.01, **p < 0.05, *p < 0.1.
[a] All models are adjusted for all confounders.

compared to no disability group, but the disadvantage was smaller than in the other disability groups. Such a small but significant disadvantage was not captured when identifying only 3 categories of disability.

Finally, regarding the rate of change over time, as captured by the slopes, in most cases multi-categorical definitions of disability did not seem to identify significant differences in changes in health conditions over time. In the few cases where there was evidence of association, we observed a 'protective effect' of disability, where the worsening of the condition was slower among disabled -whatever the level of disability was-compared to their non-disabled counterpart, indicating a convergence of the trajectories over time.

4. Discussion

In this paper we sought to examine whether a finer classification of disability in the 50 + English population may be useful for both descriptive and health policy planning purposes. To the best of our knowledge, this is the first study attempting to identify categories of disability based on empirical evidence, rather than upon a priori theoretical classification, for the English population. Along with the identification of the most adequate number of disability categories, the results produced in this study also help to understand the relevance of classifying disability correctly and describe the main characteristics of each category. Our conclusion suggests that the best classification of disability consists of four classes. While acknowledging that some arbitrariness and subjectivity affect this conclusion, there are also multiple sources of evidence lending support to our decision. Below we elaborate why we think that the optimal number of disability levels is four and discuss why it is important to go beyond a binary definition of disability.

The optimal number of disability classes was chosen based on: (i) fit indices; (ii) estimated probabilities of disability items conditional on class membership; and (iii) external validation based on the association of disability classes with health and mortality. The first criterion did not point to a preferred model. When we looked at probabilities of endorsing disability items for each class, some evidence for making a decision emerged. We wanted to identify as many grades of disability as each group presented a different probability of having problems with each disability item. The support to the 4-class model came considering and comparing all the class models, and therefore observing that in the 5 and 6-class models four patterns appeared, and with 4 classes the distinction between the two lowest levels was small, but still appreciable. The third criterion confirmed this conclusion by showing the existence of a disability gradient associated with mortality and health

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

consisting of four levels of disability. The identification of three categories of disability and a no-disability group (i.e. 4-class model) enabled us to capture the strong effect of severe disability as well as the intermediate impact of moderate disability, and also the small but significant disadvantage in health and mortality experienced by those affected by mild disability compared to the non-disabled. This small but significant effect experienced by the mildly disabled had a correspondence in the finding that the item profiles of non-disabled and mildly disabled in the 4-class model were very similar to each other, but with some appreciable differences, especially in the sensory functions in the impairment domain. Therefore, our results suggest that the best classification of disability has four classes, consisting of "no disability", "mild disability" which presents characteristics more similar to no disability than the other disability levels, and "moderate" and "severe disability".

In the introduction, we pointed out the relevance of going beyond a binary definition of disability, and mentioned recent studies[6,7] that have stressed the importance of identifying more finely graded measures of disability. Our findings contribute to this debate by showing the loss of information due to too broad categorizations of disability, compared to a more refined scale of severity. Using the 4-class model as a reference, we observed that the magnitude of the association between binary disability and all outcomes lied in between the results observed for the groups suffering from the most severe disability (moderate and severe). In the growth model for the relationship between disability and health outcomes, the intercept coefficients of binary disability were just slightly higher than the intercept coefficient of moderate disability. The binary indicator averaged the burden of disability and masked the very strong effect experienced by individuals having severe disability, and was not informative for low levels of disability. As shown by Jagger et al. (2015),[6] life expectancy with mild-disability is expanding, meaning that the number of years expected to be lived in mild-disability is increasing over time and the proportion of life expected to be lived without mild disability has decreased. If adequate grades of severity are not identified, it is not possible to monitor low disability, and equally it is impossible to assess the strong impact of severe disability.

Finally, we report some limitations that affect this work. First, we used estimated probabilities of class membership as a covariate in regression analysis and this ignores misclassification error. Various 3-step methods have been proposed to account for this.[25,26] We acknowledge this here, but it was beyond the scope of this work to explore in details this technical aspect. Moreover, the fact that entropy was high and adequate separation between the identified latent classes was found suggests that results would be very similar. Second, while latent variable modelling was in many respects a natural approach to measure disability, one issue is that it is data dependent which asks for further research to compare findings across different settings. A possible step to validate our results would be to replicate the analysis using the ELSA's sister studies, such as the Health and Retirement Study (HRS), Survey of Health, Ageing and Retirement (SHARE), which offer a substantially similar set of variables and are targeted on the same type of population of ELSA (adults aged 50+) but in other countries.

Lastly, we mention a feature of the study which limits the interpretation of results and calls for future research. While we investigated the association between disability and health outcomes to assess the grades of disability, we came across interesting results, especially concerning the effect of disability over time. Particularly unexpected was the finding that in some cases the health status of disabled individuals seemed to deteriorate at a slower pace compared to their non-disabled counterparts (where slope coefficients were significant and with opposite sign compared to intercept coefficients). We hypothesize that it may be due to the fact that non-disabled people are more likely to become disabled over time, and this would explain why slope coefficients were significant and suggesting a protective effect of disability especially among the most disabled group, which cannot become any more disabled, and therefore the impact of their level of disability at baseline on health and mortality over time is more likely to remain the same. The nature of the data limits a more detailed investigation of this finding, and most research on the association between disability and anthropometric measures for physical domains and mental health is cross-sectional limiting the conclusions that can be drawn about causality. This opens the doors to further research to robustly explore our hypothesis, for example treating disability as a time varying variable and observing its trajectory and its impact on health.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

Charges for publication would be supported by the Economic and Social Research Council (grant number ES/J500021/1).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.dhjo.2017.01.005.

References

1. Palmer M, Harley D. Models and measurement in disability: an international review. *Health Policy Plan*. 2012;27(5):357–364.
2. Wolf D. *Chapter 4-late-life Disability Trends and Trajectories. Handbook of Aging and the Social Sciences*. eighth ed. San Diego: Academic Press; 2016:77–99.
3. Wolf DA, Hunt K, Knickman J. Perspectives on the recent decline in disability at older ages. *Milbank Q*. 2005;83(3):365–395.
4. Freedman VA. Implications of asking "ambiguous" difficulty questions: an analysis of the second wave of the asset and health dynamics of the oldest old study. *J Gerontol Ser B Psychol Sci Soc Sci*. 2000;55(5):S288–S297.
5. Jette AM. How measurement techniques influence estimates of disability in older populations. *Soc Sci Med*. 1994;38(7):937–942.
6. Jagger C, Matthews FE, Wohland P, et al. A comparison of health expectancies over two decades in England: results of the cognitive function and ageing study I and II. *Lancet*. 2015;387(10020):779–786.
7. Wolf DA, Freedman VA, Ondrich JI, Seplaki CL, Spillman BC. Disability trajectories at the end of life: a "countdown" model. *J Gerontol B Psychol Sci Soc Sci*. 2015;70(5):745–752.
8. Manton KG, Gu X. Disability declines and trends in medicare expenditures. *Ageing Horiz*. 2005;2(1):25–34.
9. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain*. 1995;61(2):277–284.
10. Honey A, Emerson E, Llewellyn G. The mental health of young people with disabilities: impact of social conditions. *Soc Psychiatry Psychiatr Epidemiol*. 2011;46(1):1–10.
11. Feyzi V, Hasheminejad N, Jafariroodbandi AS, Zolala F, Asadi A. Relationship between hand grip strength and functional disability of the elderly living in nursing homes in Kerman, Southeastern Iran. *Int J Occup Hyg*. 2016;8(1):39–44.
12. Hairi NN, Cumming RG, Naganathan V, et al. Loss of muscle strength, mass (sarcopenia), and quality (specific force) and its relationship with functional limitation and physical disability: the concord health and ageing in men project. *J Am Geriatr Soc*. 2010;58(11):2055–2062.
13. World Health Organization. *World Bank. World Report on Disability*. Geneva, Switzerland: World Health Organization; 2011.
14. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort profile: the english longitudinal study of ageing. *Int J Epidemiol*. 2013;42(6):1640–1648.
15. Pongiglione B, De Stavola BL, Kuper H, Ploubidis GB. Disability and all-cause mortality in the older population: evidence from the english longitudinal study of ageing. *Eur J Epidemiol*. 2016;31(8):735–746.

16. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med*. 2013;32(9):1584–1618.
17. Centers for Disease Control and Prevention, Washington Group on Disability Statistics 2015 [cited 2016 25 Feb], Available from: http://www.cdc.gov/nchs/washington_group/wg_questions.htm.
18. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29–36.
19. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatr*. 2011;48(4):277–287.
20. Office for National Statistics, General Lifestyle Survey Overview, a report on the 2011 General Lifestyle Survey 2013 [cited 2016 April 22, 2016], Available from: http://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/compendium/generallifestylesurvey/2013-03-07/chapter7generalhealthgenerallifestylesurveyoverviewareportonthe2011generallifestylesurvey.
21. Vermunt JK. Multilevel latent class models. *Sociol Methodol*. 2003;33(1):213–239.
22. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Model A Multidiscip J*. 2007;14(4):535–569.
23. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham heart study. *Stat Med*. 1990;9(12):1501–1515.
24. Muthén LK, Muthén BO. *Growth Modeling with Latent Variables Using Mplus: Introductory and Intermediate Growth Models*; 2010. Available from: http://www.statmodel.com/download/Topic3-v.pdf.
25. Bakk Z, Tekle FB, Vermunt JK. Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociol Methodol*. 2013;43(1):272–311.
26. Bakk Z, Vermunt JK. Robustness of stepwise latent class modeling with continuous distal outcomes. *Struct Equ Model A Multidiscip J*. 2016;23(1):20–31.

Please cite this article in press as: Pongiglione B, et al., Levels of disability in the older population of England: Comparing binary and ordinal classifications, Disability and Health Journal (2017), http://dx.doi.org/10.1016/j.dhjo.2017.01.005

6.3 Supporting Information to Research Paper III

This section contains the content of the supplemental material of Research Paper III, which is not included in the paper, but available online (`http://www.disabilityandhealthjnl.com/article/S1936-6574(17)30007-9/addons`). Sections, tables and figures are numbered according to the format of regular chapters, but I report also the corresponding numbered headings of the online material to facilitate the comparison with references in the text, as shown in tablerp3app.

Table 6.1: Numbered headings in the online supplemental material and Supporting Information to Research Paper III

Online supplemental material	**Supporting Information**
Not included	ROC curve
Not included	Sensitivity analysis of WG questions (Table 6.2
Section A	Supporting Information to Research Paper II (pag. 87)
Section B	Latent Class Analysis
Section C	Latent Growth Model
Fig. A1	Figure 6.3
Fig. A2	Figure 6.4
Table A1	Table 1 of Research Paper II
Table A2	Table 6.3
Table A3	Table 6.4
Table A4	Table 6.5
Table A5	Table 6.6
Table A6	Table 6.7

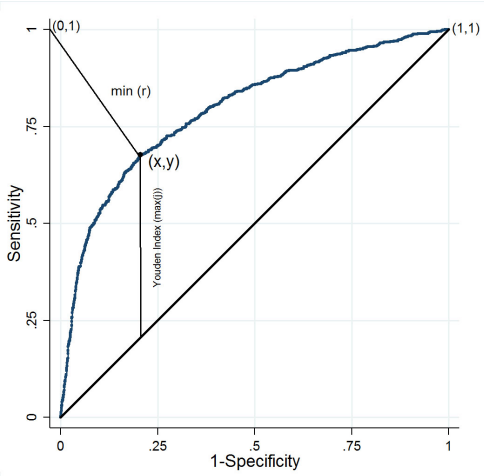
ROC curve

The ROC curve was used to check the agreement of the continuous score of disability with an external gold standard, consisting of receiving health or disability benefits. The cut-off in the continuous disability score was set based on two different criteria: (i) the Youden index that maximises the vertical distance (j) from line of equality to the point $[x, y]$ on the curve, and (ii) the minimal value of r:

$$r = \sqrt{(1 - specificity)^2 + (sensitivity - 1)^2} \tag{6.1}$$

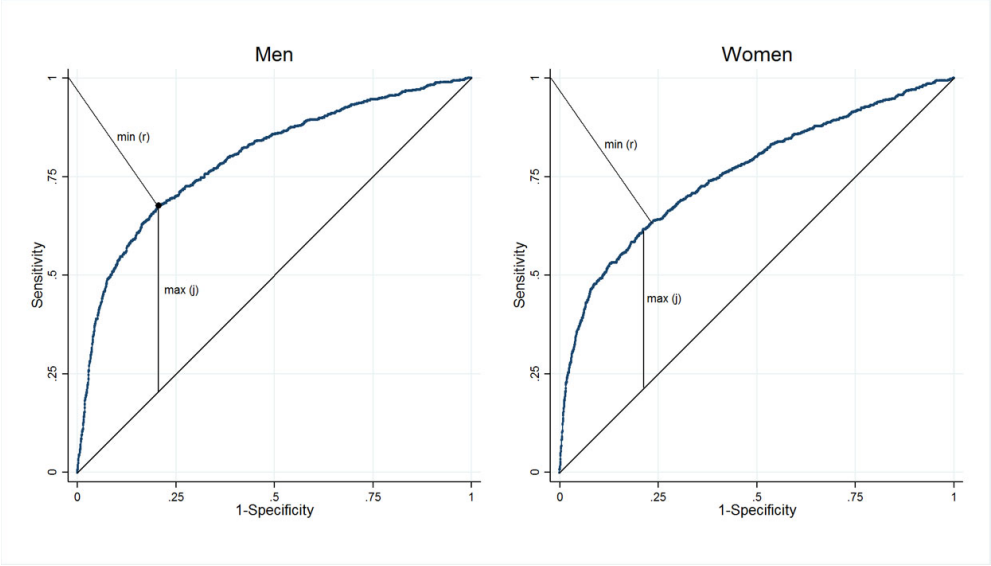
This is illustrated in figure 6.1 in a generic example.

Figure 6.1: ROC curve and its components



The cut-off points obtained for our sample, by gender, are illustrated in figure 6.2. For men, the Youden index and minimal r criteria identified the same cut-off point; while for women, the cut-off was different depending on the criterion.

Figure 6.2: ROC curve of receiving health or disability for the detection disability, by gender.



Sensitivity analysis of Washington Group questions

As mentioned in Research Paper III (p. 3), the six disability domains identified by the WG include problems in seeing, hearing, walking or climbing steps, remembering or concentrating, washing all over or dressing and communicating. The ELSA questionnaire includes questions related to these domains, but communicating was captured by a variable related more to hearing impairment rather than communication problems. Therefore, as a sensitivity analysis, I excluded this question and looked at the average disability score observed among those reporting at least one limitation in any of the five disability questions. The proportion of respondents reporting at least one limitation in the five WG activities was 61.5% of men and 59.1% of women. The mean disability score among the respondents of this group was 0.46 and 0.50 for men and women respectively. Setting the cut-off point at these values led to 31.0% of male respondents and 30.6% of female respondents being classied as having disability. Table 6.2 summarizes all different proportions of disabled respondents found according to each approach (i.e. WHOs approach and ROC curve) and national statistics, plus proportions obtained from sensitivity analysis for the WHOs approach.

Table 6.2: Binary classifications of disability

Approach	**Proportion**	
	Men	Women
WHO	31.5	31.7
WHO sensitivity analysis	*31*	*30.6*
ROC curve - min (r)	29.2	30.3
ROC curve - max (j)	29.2	27.9
ONS	32.8	32.9

Latent Class Analysis

For binary items and a categorical latent variable C_{k} with K classes ($C = k; K = 2, \ldots 6$), the marginal probability of observed item u_{j} ($withj = 1, 2 \ldots, 50$) being equal to 1 is

$$Pr(u_j = 1) = \sum_{k=1}^{K} Pr(C_k = k) Pr(u_j = 1 | C_k = k) \tag{6.2}$$

Where the second part of Eq. 6.2, $Pr(u_j = 1|C_k = k)$, denotes the conditional prob-

ability of the item being equal to 1 the class is equal to k and Pr(C_k=k) is the probability of class k.

Latent Growth Model

In the algebraic notation used in Structural Equation Model (SEM)s, LGM can be formulated as follows, with y_{it} representing the health outcome being observed in individual i at occasion t:

$$y_{it} = \lambda_{0t} intercept_i + \lambda_{1t} slope_i + e_{it} \tag{6.3}$$

where

$$intercept_i = \beta_{00} + \mathbf{Z}_i^T \boldsymbol{\beta}_{01} + u_{0i} \tag{6.4}$$

$$slope_i = \beta_{10} + \mathbf{Z}_i^T \boldsymbol{\beta}_{11} + u_{1i} \tag{6.5}$$

Eq. 6.3 is the first level equation, Eq.6.4 and 6.5 are the second level equations. $\mathbf{Z}$ is a vector of time-fixed variables that includes the disability levels defined by a given latent class model; and u_{0i} and u_{1i} are individual-level random effects which are assumed to be jointly normally distributed, with mean (0,0) and variance-covariance matrix Σ. The residual errors e_{ti} are also assumed to be normally distributed and, to aid identification, to have the same variance for all t. The parameters λ_{0t} are pre-defined factor loadings for the subject-specific random intercepts (and are fixed to be equal to 1), while the pre-defined factor loadings for the subject-specific random slopes are set to values that reflect the observation times of the outcomes. The assumption of linear growth for health outcomes implies fixing λ_{1t} =0,1,2,3,4,5 in the case of 6 time-points of observation. We assume linear growth for mental health and grip strength, while for coagulation non-linear growth models are fitted by freeing the factor loading for the last time point [1].

The model [2] estimated growth factor means are obtained from:

[1] There are only three observation time-points for grip strength and coagulation, while a minimum of four time-points is recommended for growth models (`http://www.statmodel.com/discussion/messages/14/20.html?1459099011`), so that the model has enough degrees of freedom to allow for modifications of the model. However, with three timepoints and a continuous outcome, the $H1$ model has 9 free parameters and the $H0$ model has 8 free parameters. We did not incur problems of model fit with grip strength outcome. When we estimated the factor loading for the slope of coagulation at time score 3 we fixed the covariance between the random slopes and intercepts to be 0 to achieve identifiability.

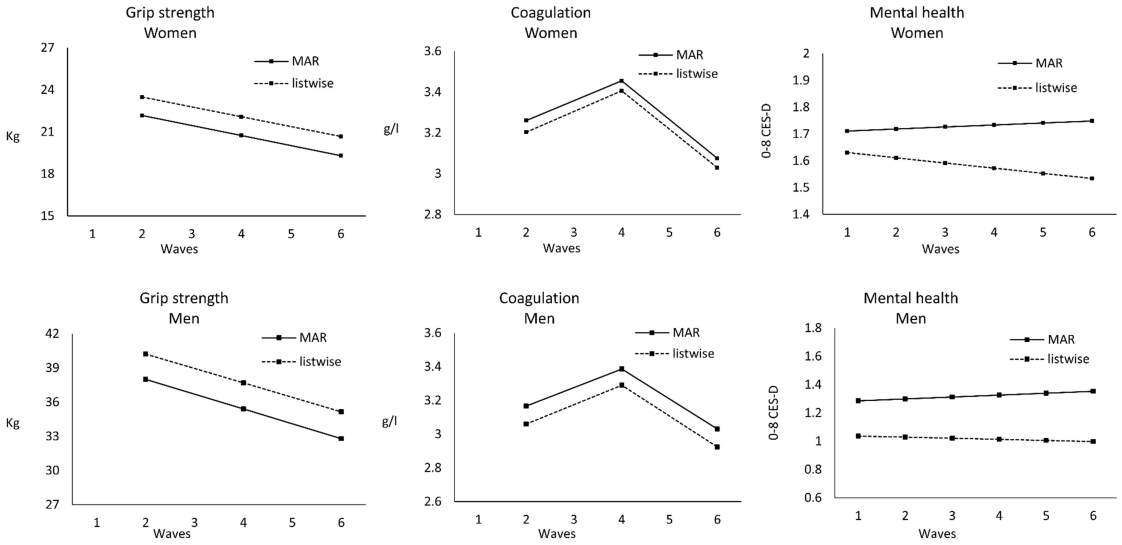
[2] Fully adjusted controlling for all confounders mentioned in the methods section.

$$\hat{E}(intercept_{\mathrm{i}}) = \hat{\beta}_{00} + \bar{\mathbf{Z}}_i^T \hat{\boldsymbol{\beta}}_{\mathbf{01}} \tag{6.6}$$

$$\hat{E}(slope_{\mathrm{i}}) = \hat{\beta}_{10} + \bar{\mathbf{Z}}_i^T \hat{\boldsymbol{\beta}}_{\mathbf{11}} \tag{6.7}$$

Supporting tables and figures

Figure 6.3: Mean grip strength, mental health score and coagulation estimated via latent growth models, by gender



Sample size for grip strength: women listwise=2,054 women MAR=3,817 men listwise=1,664 men MAR=3,301
Sample size for coagulation: women listwise=1,250 women MAR=3,083 men listwise=952 men MAR=2,665
Sample size for mental health: women listwise=2,526 women MAR=4,962 men listwise=1,952 men MAR=4,333

Figure 6.4: Probability of each disability item estimated by 5-class and 6-class models, by gender

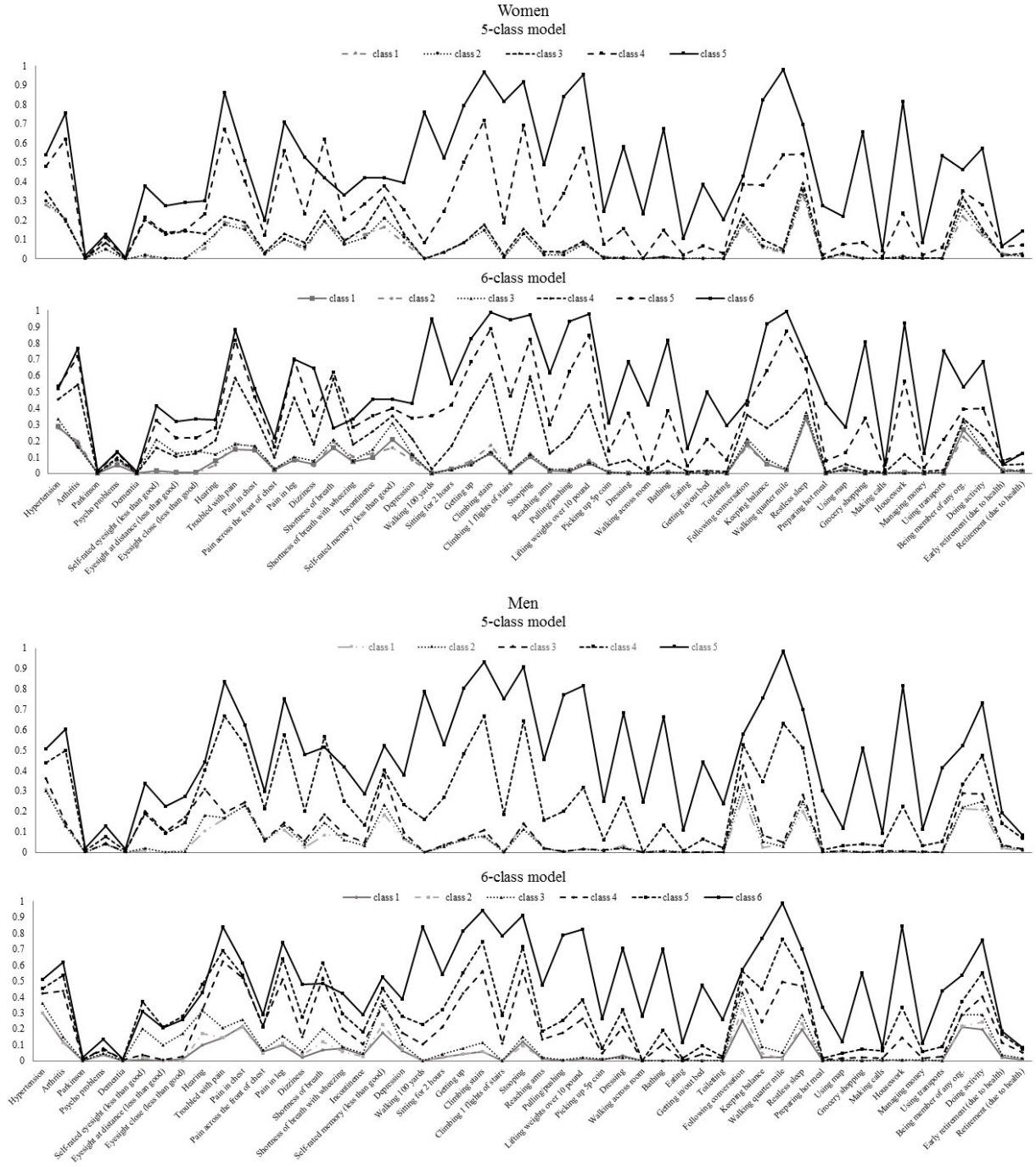


Table 6.3: Latent class models goodness of fit indicators, by gender

Men

# classes	BIC	entropy	free parameters
3	184507.25	0.914	242
4	181752.86	0.904	323
5	180330	0.901	404
6	179795.2	0.898	485

Women

# classes	BIC	entropy	free parameters
3	235,433.02	0.906	242
4	232,979.96	0.889	323
5	231,568.25	0.892	404
6	230,449.94	0.885	485

Table 6.4: Estimated latent class proportions according to their most likely latent class membership, by model and gender

# classes	3-class model		4-class model		5-class model		6-class model	
	men	women	men	women	men	women	men	women
1	62.9	53.6	43.8	35.9	17.6	13.7	16.8	13.2
2	24.8	33	24.9	23.8	28.1	24	26.6	21.7
3	12.3	13.4	20.6	28.8	24.5	22.9	24.9	19.8
4			10.8	11.6	19.5	28.1	12.6	25.7
5					10.4	11.3	10	13.9
6							9.2	5.6

Table 6.5: Association of disability groups measured at baseline with mental and physical health indicators measured at wave 6

Men

Specification of disability level	**Mean grip strength**		**Mental Health score**	
	β coef	**95% CI**	**β coef**	**95% CI**
2 classes (WHO)	-1.710***	(-2.715 - -0.705)	0.794***	(0.612 - 0.975)
2 classes (ROC curve)	-1.844***	(-2.896 - -0.792)	0.836***	(0.647 - 1.025)
2 classes	-0.979**	(-1.943 - -0.0147)	0.563***	(0.391 - 0.734)
3 classes	-2.448***	(-4.086 - -0.811)	1.489***	(1.203 - 1.775)
	-2.444***	(-3.610 - -1.277)	1.057***	(0.851 - 1.263)
4 classes	-0.486	(-1.347 - 0.374)	0.200**	(0.0476 - 0.353)
	-1.383**	(-2.488 - -0.277)	0.683***	(0.485 - 0.882)
	-2.558***	(-4.327 - -0.789)	1.506***	(1.194 - 1.818)
5 classes	-0.501	(-1.472 - 0.470)	0.0646	(-0.110 - 0.239)
	-0.668	(-1.710 - 0.375)	0.234**	(0.0490 - 0.420)
	-1.851***	(-3.123 - -0.578)	0.784***	(0.556 - 1.012)
	-2.903***	(-4.826 - -0.981)	1.549***	(1.209 - 1.888)
6 classes	-0.418	(-1.406 - 0.570)	0.0754	(-0.103 - 0.253)
	-0.69	(-1.737 - 0.357)	0.271***	(0.0843 - 0.458)
	-1.001	(-2.362 - 0.360)	0.648***	(0.404 - 0.892)
	-2.959***	(-4.632 - -1.286)	1.039***	(0.740 - 1.339)
	-2.942***	(-5.002 - -0.882)	1.525***	(1.159 - 1.890)

Women

Specification of disability level	**Mean grip strength**		**Mental Health score**	
	β coef	**95% CI**	**β coef**	**95% CI**
2 classes (WHO)	-1.780***	(-2.418 - -1.142)	0.970***	(0.776 - 1.163)
2 classes (ROC min r)	-1.942***	(-2.594 - -1.290)	0.999***	(0.802 - 1.196)
2 classes (ROC max j)	-1.973***	(-2.641 - -1.305)	1.024***	(0.822 - 1.226)
2 classes	-1.899***	(-2.571 - -1.226)	1.020***	(0.816 - 1.223)
3 classes	-1.378***	(-1.944 - -0.812)	0.790***	(0.617 - 0.962)
	-3.383***	(-4.360 - -2.405)	1.641***	(1.346 - 1.935)
4 classes	-0.672**	(-1.237 - -0.107)	0.149*	(-0.0252 - 0.324)
	-1.653***	(-2.294 - -1.012)	0.834***	(0.638 - 1.030)
	-4.430***	(-5.528 - -3.332)	1.643***	(1.312 - 1.975)
5 classes	-0.39	(-1.104 - 0.324)	0.0546	(-0.165 - 0.274)
	-0.855**	(-1.579 - -0.132)	0.235**	(0.0123 - 0.458)
	-1.811***	(-2.598 - -1.024)	0.940***	(0.700 - 1.180)
	-4.604***	(-5.796 - -3.412)	1.719***	(1.357 - 2.081)
6 classes	-0.141	(-0.878 - 0.596)	0.031	(-0.195 - 0.257)
	-0.703*	(-1.450 - 0.0451)	0.186	(-0.0437 - 0.417)
	-1.732***	(-2.518 - -0.946)	0.713***	(0.473 - 0.952)
	-3.023***	(-4.029 - -2.017)	1.464***	(1.159 - 1.769)
	-4.723***	(-6.431 - -3.016)	1.938***	(1.439 - 2.438)

CI=confidence interval; β coef=β coefficient ; 95% confidence intervals in brackets (); $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.
Parameter estimates (β coefficient) and p-values calculated by the linear model adjusted for all confounders.

Table 6.6: Model estimated means (and standard errors) of latent growth factors

	Grip strength[a]		**Mental Health score[a]**	
	intercept (s.e.)	**slope (s.e.)**	**intercept (s.e.)**	**slope (s.e.)**
Women	22.896 (0.119)	-0.719 (0.026)	1.711 (0.026)	0.007 (0.007)
s.d.	5.767	0.316	1.498	0.217
Men	39.308 (0.188)	-1.304 (0.041)	1.286 (0.025)	0.014 (0.007)
s.d.	8.614	0.609	1.364	0.176

[a] Estimated means are the same across all-class models. Values are reported once. s.e.=standard error; s.d.=standard deviation

Table 6.7: Association of disability groups with mortality and a number of mental and physical health indicators, 5 and 6-class model

Specification of disability level		Mean grip strength				Mental Health				Death	
		Women		Men		Women		Men		Women	Men
		Intercept (95% CI)	Slope (95% CI)	Intercept (95% CI)	Slope (95% CI)	Intercept (95% CI)	Slope (95% CI)	Intercept (95% CI)	Slope (95% CI)	β coef	β coef
5 classes	group 2	-0.006 (-0.66;0.65)	-0.078 (-0.23;0.08)	-0.583 (-1.52;0.36)	0.019 (-0.2;0.24)	0.085 (-0.07;0.24)	-0.014 (-0.06;0.03)	0.024 (-0.1;0.15)	0.023 (-0.01;0.06)	1.084 (0.757; 1.554)	0.998 (0.763; 1.306)
	group 3	-0.316 (-0.98;0.35)	-0.105 (-0.26;0.05)	-0.921* (-1.9;0.05)	-0.032 (-0.26;0.2)	0.2** (0.05;0.35)	0 (-0.04;0.04)	0.24*** (0.11;0.37)	0.016 (-0.02;0.05)	1.306 (0.919; 1.858)	1.026 (0.781; 1.347)
	group 4	-2.376*** (-3.05;-1.7)	0.027 (-0.14;0.19)	-2.442*** (-3.54;-1.34)	-0.003 (-0.27;0.26)	0.99*** (0.83;1.15)	0 (-0.05;0.05)	1.062*** (0.91;1.21)	-0.036 (-0.08;0.01)	1.513** (1.082; 2.115)	1.421** (1.083; 1.863)
	group 5	-5.658*** (-6.61;-4.7)	0.255** (0.01;0.5)	-5.807*** (-7.33;-4.28)	0.465** (0.07;0.86)	2.007*** (1.79;2.22)	-0.08** (-0.14;-0.02)	2.176*** (1.98;2.37)	-0.17*** (-0.23;-0.11)	1.954*** (1.354; 2.819)	1.944*** (1.427; 2.648)
6 classes	group 2	0.091 (-0.59;0.77)	-0.058 (-0.22;0.1)	-0.54 (-1.5;0.42)	0.02 (-0.2;0.24)	0.055 (-0.1;0.21)	-0.016 (-0.06;0.03)	0.012 (-0.12;0.14)	0.029 (-0.01;0.06)	1.06 (0.729; 1.539)	0.987 (0.744; 1.310)
	group 3	-0.177 (-0.87;0.51)	-0.092 (-0.25;0.07)	-1.139** (-2.12;-0.16)	-0.011 (-0.24;0.22)	0.13 (-0.03;0.29)	0.005 (-0.04;0.05)	0.293*** (0.16;0.43)	0.016 (-0.02;0.05)	1.249 (0.863; 1.807)	1.072 (0.812; 1.416)
	group 4	-1.751*** (-2.43;-1.07)	-0.051 (-0.22;0.11)	-1.81*** (-3.02;-0.6)	0.02 (-0.27;0.31)	0.792*** (0.63;0.95)	-0.007 (-0.05;0.04)	0.789*** (0.62;0.95)	-0.038 (-0.09;0.01)	1.431** (1.015; 2.017)	1.353** (1.005; 1.820)
	group 5	-4.02*** (-4.85;-3.19)	0.169 (-0.04;0.37)	-3.53*** (-4.91;-2.15)	0.005 (-0.35;0.36)	1.577*** (1.38;1.77)	-0.006 (-0.06;0.05)	1.327*** (1.14;1.51)	-0.016 (-0.07;0.04)	1.826*** (1.272; 2.623)	1.635*** (1.208; 2.213)
	group 6	-7.366*** (-8.62;-6.11)	0.644*** (0.3;0.99)	-5.966*** (-7.57;-4.36)	0.524** (0.1;0.95)	2.421*** (2.15;2.69)	-0.173*** (-0.26;-0.09)	2.241*** (2.03;2.45)	-0.179*** (-0.25;-0.11)	1.868*** (1.240; 2.814)	2.009*** (1.454; 2.775)

CI=confidence interval; β coef=β coefficient ; 95% confidence intervals in brackets (); $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

6.4 Conclusions

The main result of Research Paper III answers the second research question presented in the introduction of the paper (section 6.2.2) and in the preamble of this chapter (section 6.1) *"Would a finer classification of disability better capture clustering of disability dimensions and of heterogeneities in later disease and mortality rates?"*. In the context of our research, the best classification of disability consists of four classes, labeled "non-disability", "mild disability", "moderate disability" and "severe disability".

Some advantages of using multicategorical classifications of disability compared to binary classifications are also presented, and it is shown how dichotomous categorizations of disability -compared to a more refined scale of severity- mask the very strong effect of severe disability by averaging the burden of disability, without being informative on the impact of lower levels.

The results of Research Paper III offer strong evidence for separating low grades of disability from moderate and more serious grades of disability and are at the basis of development of Research Paper IV, where trends in Disability-Free Life Expectancy (DFLE) over a decade are explored distinguishing three levels of disability, i.e. mild, moderate and severe, as opposed to non-disability.

Evidence produced in Research Paper III indicate that identifying adequate categories of severity scale is needed to fully understand the paths along which populations are ageing.

Research Paper III is in press in the Disability and Health Journal [43]. A preliminary version of this work was presented in a seminar session of the Department of Medical Statistics at London School of Hygiene and Tropical Medicine (LSHTM) and a more advanced version in a seminar session of the MRC Unit for Lifelong Health and Ageing, University College of London. The slides of the most updated presentation are available in the appendix (figure 9.3).

Chapter 7

Trends in disability-free life expectancy between 2002 and 2012 in England

7.1 Preamble

In Research Paper IV, I address objectives 5 and 6 of this thesis. This means (i) assessing how Disability-Free Life Expectancy (DFLE) has evolved over a decade in England when distinguishing four levels of disability (including non-disability), and (ii) proposing possible explanations for observed changes in DFLE and differences across gender and severity levels.

- **Research Objective 5:** *To combine each level of disability with mortality in a single indicator of population health, DFLE, and use it to estimate how expectancies of healthy life have changed over a decade across the two genders and across different levels of disability.* Disability is measured according to the WHO's International Classification of Functioning Disability and Health (ICF) framework, as in Research Papers II and III, and the choice of distinguishing four levels of disability (i.e. non-disability and mild, moderate and severe disability) comes from findings of Research Paper III.
- **Research Objective 6:** *To explore possible demographic and behavioural factors explaining changes in DFLE observed over a decade.* Do achieve objective 6, the aggregate outcome of DFLE is replaced with the individual-level Years Lost to Disability (YLD)

to assess whether changes in YLD between waves 2 and 6 are associated with Body Mass Index (BMI). BMI is specifically chosen as a risk factor of interest because: (1) overweight and obesity are known to be associated with higher risk of becoming and remaining disabled, but have little or no effect on life expectancy; (2) the prevalence of overweight or obesity varies across cohorts, and cohort differences are of particular interest when comparing life expectancy and health expectancy. This is an exploratory analysis which open doors to further investigation. In the discussion of Research Paper IV as well as in the discussion of the thesis (Chapter Eight), I present my findings and propose and discuss avenues for future research.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Benedetta Pongiglione
Principal Supervisor	Bianca De Stavola
Thesis Title	Gender inequality in healthy ageing: a study of the English older population over a decade

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	European Journal of Epidemiology
Please list the paper's authors in the intended authorship order:	Pongiglione B., Ploubidis G.B., De Stavola B.L.
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	

Student Signature: Benedetta Pongiglione **Date:** 16 January 2017

Supervisor Signature: Bianca L De Stavola **Date:** 16 January 2017

7.2 Research Paper IV

Disability-free life expectancy between 2002 and 2012 in England: trends differ across genders and levels of disability

7.2.1 Abstract

Background: The aim of this work is to assess how disability-free life expectancy (DFLE) has evolved over the past decade in England distinguishing four levels of disability, and to propose possible explanations for observed changes over time and differences between genders and disability severity levels.

Methods: We used data from the English Longitudinal Study of Ageing and considered both cross-sectional and longitudinal samples, interviewed from 2002 to 2012 (at 6 waves). Disability was defined according to the WHO's International Classification of Functioning, Disability and Health, from which 4 classes were estimated (no disability, mild, moderate and severe), in correspondence to each wave, using latent class analysis. DFLE was estimated at the first and last wave by applying Sullivan's method, and years lost to disability (YLD) were estimated in a second stage to perform individual-level analyses of the relationship between changes in YLD between 2002 and 2012 and Body Mass Index (BMI) measured in 2002 and year of birth.

Results: Changes in DFLE observed between 2002 and 2012 differed across gender and disability classes. Severe and moderate disability declined for women, while their mild disability increased, indicating a dynamic equilibrium overall. Men experienced worse changes, with stable levels of severe disability and increasing moderate disability. There was evidence of modification of the effect of BMI by year of birth on changes in YLD, such that high BMI resulted particularly detrimental to younger cohorts.

Conclusion: Two conclusions emerge from these results: (i) It is important to distinguish between milder and more severe levels of disability because their trends seem to be divergent. (ii) The evidence of interaction between BMI and year of birth points towards the need for closely monitoring BMI in younger generations as this appears to be detrimental in terms of their disability experience in later life.

Key words: *Disability free life expectancy; expansion; compression; older population; England*

7.2.2 Introduction

Life expectancy has been used as an indicator of population health for a long time. More recently, with the completion of the "epidemiological transition" in high and low-middle income countries [70], mortality has ceased to be as tied to health as it was before, and life expectancy does no longer fully capture the health status of a population. From the 1960s, with the study of Sanders [71] and Sullivan [72], the assessment and monitoring of population health changes have shifted towards indicators that combine both mortality and morbidity (or disability); these are known as Summary Measures of Population Health (SMPH). With the development and adoption of new population health indicators, evaluations of their trends over time have emerged. Three distinct theories of population health changes have been proposed, namely: compression [18], expansion [16, 17] and dynamic equilibrium of morbidity and mortality [20]. It has often been remarked that empirical evidence supporting any of these theories is scarce. However, there has been increasing interest in the use of health expectancy indicators for public policy and planning and for the evaluation of public health programs over the last decades. Hence, the lack of support for any of the abovementioned theories is not only due to the scarcity of studies, but also to the heterogeneity and discordance of results. In 2003, the results of a decade's work on health expectancy of the Réseau Espérance de Vie en Santé (REVES) project was collected in a book and evidence on theories of population health changes were evaluated [73]. At chapter 18, combining the chronological series available for several European countries, from the 1980s and 1990s, Perenboom et al. [74] showed that Total Life Expectancy (TLE) has been increasing in European countries, but this was not always accompanied by a rise in health expectancy. Health expectancy has increased but not as much as TLE. A closer look indicated an increase in the number of years in mild ill health and a decreasing or stable situation for the number of years in moderate or severe ill health [75]. However, the evidence was not very strong and it is unclear whether the conclusions also hold for more recent years.

In the UK and England -where the present study is set- evidence is mixed. The UK is one of the few countries for which time-series of life expectancy and health expectancy have been available since the 1980s [34], and thus it has been possible to study trends over about

three decades. Nevertheless, no clear pattern has been found. Between 1981 and 1999, dynamic equilibrium of morbidity was found [34]. Perenboom et al. [74], collecting evidence from UK-based studies, reported that in the UK, between 1980 and 1994 there seemed to be an increase in DFLE for females aged 65 years [37], and handicap-free life expectancy increased between 1976 and 1991, but the trend reversed downward between 1991 and 1994 [38, 75]. A more recent study [40] investigated how various health expectancies have changed in England between 1991 and 2011, and showed that cognitive impairment compressed in absolute terms (i.e. supporting evidence of reduction), self-perceived health compressed in relative terms (i.e. increase in the proportion of life spent healthy), and disability evolved in dynamic equilibrium, with less severe disability increasing and more severe disability declining.

Looking at other recent studies focused on the older population and set in different countries (US [76, 77], and Sweden [78]), evidence varied across settings, but some common findings emerged as well. The studies that distinguished mild and severe forms of disability [76, 78], generally agreed in finding a decline in severe forms of disability and a rise in milder levels. The study by Freedman et al., set in the US between 1982 and 2011 [76], found that among women aged 65+ mild disability has increased, while severe disability has decreased in proportional terms, but increased in absolute value. The study by Sundberg et al. [78] of Swedish trends between 1992 and 2011, supported absolute and proportional compression of severe disability among women and expansion among men; it also showed absolute expansion of mild disability for men and women, and proportional stability of mild disability, with women experiencing an expansion in the last period of observation. No compression of disability across the life cycle, but some compression at age 65 years, was found in the US study that did not distinguish levels of disability [77].

As just described, health expectancies can be assessed in absolute terms or relative to trends in life expectancy [73]. Attention has been focused on these two alternative measures of health expectancy to understand the advancement in the process of healthy ageing. What has often been neglected, however, is the importance of the actual number of expected years with and without disability and the need for monitoring them over time, regardless of their comparison -in terms of proportions or differences- with life expectancy. This is because

estimates of expected years with and without disability are informative of the overall burden of disability.

Given these premises, this study intends to contribute to the debate on compression, expansion and dynamic equilibrium of mortality and morbidity by assessing how DFLE has evolved in England over a decade -hence providing additional evidence to support or challenge the prevailing theories of population health change- and produce evidence for absolute and proportional shifts in population health. We use data from the English Longitudinal Study of Ageing (ELSA), and therefore focus on adults aged 50 years and older, to estimate DFLE applying the Sullivan method. Doing so, our research provides new evidence by (i) updating results for the last decade in England among the non-institutionalised population aged 50 years and older; (ii) interpreting disability according to the ICF framework, which is a comprehensive approach to disability that includes impairments, activity limitations and participation restrictions, rather than focusing on specific domains separately; (iii) distinguishing severity levels of disability to better understand changes in DFLE; (iv) considering both longitudinal and cross-sectional samples to provide robust estimates; (v) exploring possible explanations for the observed dynamics of DFLE, modelling changes in a corresponding outcome measurable at the individual level, YLD.

This corresponds to the following specific objectives:

1. to estimate DFLE for four different levels of disability at two points in time a decade apart (2002 and 2012), separately by gender;

2. to compare changes in DFLE over time between men and women and across severity levels of disability;

3. to propose possible explanations for the estimated changes in DFLE over a decade, shifting the analysis to the individual level via estimates of YLD, and looking at individual's overweight and obesity status at baseline as an explanatory factor, as well as a modifier of the association between year of birth and YLD.

7.2.3 Data and Methods

Sample

We used data from the ELSA. The ELSA is a longitudinal study designed to collect longitudinal multidisciplinary data on health, social outcomes, wellbeing and economic circumstances from a representative sample of the English population aged 50 years and older living in private households. Originally, the sample was drawn from households that had previously responded to the Health Surveys for England (HSE) in 1998, 1999 or 2001. So far, six waves have been issued and every two waves -starting from the second- a nurse visit has taken place. The nurse interview involves measurements of physical function, anthropometric measurements and collection of blood samples. As the study progresses, the youngest groups are -as expected- depleted. Therefore, refreshment samples of participants aged 50+ have been included at wave 3, wave 4 and wave 6 of data collection. DFLE estimates for 2002 are based on core-member respondents at wave 1 who have complete records on all variables used to measure disability. This corresponded to 9,731 observations, 45.9% (4,462) men and 54.1% (5,269) women. Respondents were all aged 50 years and older, female mean age was equal to 64.8 years and male mean age was 64.4 years. For estimating DFLE in 2012 we used data from wave 6 and considered two alternative sample definitions. The first, which we refer to as the cross-sectional sample, consisted of core members[1] of wave 6 who have complete records on all disability variables measured at this wave; this included also the refreshment sample of wave 6 as well as refreshment samples from previous waves who participated in the last wave. This corresponded to 7,507 observations of which 4,173 women (55.6%) and 3,334 men (44.4%). Female mean age was equal to 67.4 years and male mean age was 67.3 years. The second definition consisted of respondents selected at wave 1 and interviewed again at wave 6, whether they did or did not take part in the surveys between the first and the sixth wave. We refer to this as the longitudinal sample. It corresponded to 4,602 observations, of which 44.3% (2,037) men and 55.7% (2,565) women. In this case, since respondents had been followed up for about ten years, the youngest group was aged 60 years, with women's mean age being 71.6 years and men's mean age 71.4 years.

[1] members are both individuals interviewed at wave 1 and followed up throughout each wave and participants included in refreshment samples at wave 3, 4 and 6.

Measure

Disability

Since the 1960s, disability has been increasingly interpreted through a disablement process, along which functional limitations expose to activity restrictions, with a hierarchy in the occurrence of restrictions [52]. As a result, disability has often been measured by activity limitations, most commonly using Activity of Daily Living (ADL) (e.g. Jagger et al. [79], Lazaridis et al. [80], Dunlop et al. [81]) or combining in hierarchical scales ADL and Instrumental Activity of Daily Living (IADL) [82] and mobility functions [83]. In this work, we adopt a more recent and comprehensive approach to conceptualize disability, elaborated in 2001 by the WHO, the ICF [48], which is currently the predominant theoretical model of disability [63, 84]. The ICF model derives from previous conceptual schemes of disability, the first of which was the "disablement model" proposed by Nagi in 1965 [52], in the 1980 the International Classification of Impairments Disabilities and Handicaps (ICIDH) [53] was issued, and the more recent "disablement process" was proposed in 1994 by Verbrugge and Jette [51]. These models introduced a new approach to conceive disability that was interpreted not only as a medical condition, but also in terms of its social implications. The ICF, while including concepts of disability very similar to those used by Nagi, has the advantage of being developed by an international organization after a long consultative process, and has been intended to become the predominant language to define disability [63]. The ICF views functioning along a continuum and attempts to replace previous terminology that implies distinctions between healthy and disabled individuals. According to this framework, disability consists of three main domains: "body-function and structure" (or impairments), "activity limitations" and "participation restrictions". Importantly, within the ICF, the terms function and disability are not used to label specific elements in the model but instead are used as umbrella terms in the same fashion that the term disablement is used within the Nagi framework [63]. The validity and applicability of the ICF to capture disability among the older population was tested in a previous work [42], which was based on the same data used in this paper (ELSA), and where a continuous score of disability covering impairments, activity limitations and participation restrictions was measured. We rely on this previous work for the selection and classification of the variables capturing each

of the domains [43]. In this setting, we add an extra criterion for the inclusion of items. Only items collected at each wave, from the first to the sixth, were included to measure disability. This corresponded to a battery of 42 items, sub-classified across the three domains, as follow. Body function and structure were measured by 12 variables including hypertension, arthritis, Parkinson, psychosocial problems, dementia, self-rated eyesight (including eyesight at distance and close) and hearing, being troubled with pain, incontinence and depression. Some of these items are most commonly considered health conditions -and as such not part of disability- and studied as forms preceding disability rather than its component [85, 86, 87, 88]. Previous sensitivity analyses compared disability estimates obtained including and not including hypertension, arthritis, Parkinson, psychosocial problems and dementia among the impairment domain and found no difference on the disability summary measures nor its effect on mortality[2] [42]. As a result, we decided to include these variables in the model of disability. We also performed sensitivity analysis excluding these items from disability measures. Results are presented in the Supporting Information to Research Paper IV and discussed in the discussion section. Activity was measured by 19 variables consisting in ten mobility functions such as walking 100 yards, sitting for two hours, climbing stairs; six ADLs, i.e. dressing, walking across room, bathing, eating, getting in/out bed, toileting; being able to follow a conversation and quality of sleep. For participation, 11 variables were selected: six IADLs, i.e. preparing hot meal, using map, grocery shopping, making calls, doing housework, managing money; being member of any organization and doing any social activity; and limitations due to health in using transports and working. Disability classes were estimated for both cross-sectional and longitudinal samples at each wave to assess the validity of stationarity assumption of the Sullivan method (see next paragraph). Results for waves 1 and 6 are presented in the text, while results for intermediate waves are available in the Supporting Information to Research Paper IV (Table 7.9).

Mortality

Mortality rates were estimated using estimates of the relevant English population and reported deaths in 2002 and in 2012, by sex and single year of age, obtained from the Office

[2]Pearson's correlation between disability scores was equal to 0.9955

for National Statistics (ONS)[3]. Mortality rates were produced by 5-year age groups.

ONS mortality rates pertain to the total population, while disability prevalence refers to ELSA's sample, which does not include institutionalised individuals. The problem of combining national data on the general population, with survey-based information often targeted on non-institutionalised populations is particularly crucial for older populations, as known in literature [41, 89]. A commonly applied option to provide population estimates of disability prevalence, adjusting for the population excluded from surveys, was proposed by Sullivan [72] and consists in adjusting estimates by assuming that the entire population of health-related institutions have disability. In this study, no assumptions were made, thus implicitly assuming same prevalence among institutionalised and non-institutionalised populations. A study aimed at testing a number of hypotheses for including institutionalised individuals in disability prevalence estimates, showed that for advanced age groups, the overestimation resulting from the Sullivan's hypothesis can be greater than the underestimation descending from the assumption of same prevalence in institutionalised and household populations [89]. However, there is no guarantee that in the context of our analysis the same result holds, and neither hypothesis can avoide bias completely. A further consideration is that, while at baseline (wave 1), ELSA's target population consisted in people aged 50 years and older living in private household, over the course of the study if respondents moved into a residential care home or similar, they were followed up and still included into the sample. Therefore, the ELSA samples at wave 1 and 6 are likely to be representative of slightly different populations. As a sensitivity analysis, presented in the Supporting Information to Research Paper IV (tables 7.11 and 7.12), we excluded from our sample at wave 6 "institution interviews", so that both samples at wave one and wave six were representative of the same population, consisting in non-institutionalised English adults aged 50+. Additionally, an analysis of the representativeness of ELSA's samples in terms of mortality, provided in the Supporting Information to Research Paper IV (table 7.8), found that the ELSA mortality converges to that of the general population over time. Implications and limitations of this aspect are further commented in the discussion.

[3] https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/adhocs/005676englishpopulationestimatesanddeathsbysexandsingleyearofage1993to2013

Body Mass Index

Overweight and obesity are known to be associated with higher risk of becoming and remaining disabled, but have little or no effect on life expectancy [90, 91, 92, 93]. At the same time, the prevalence of overweight or obesity varies across cohorts, and a UK study found that the probability of overweight or obesity in childhood was two to three times greater among younger cohorts (i.e. born after 1980s), and older generations were exposed to increases in the probability of overweight or obesity across adulthood [94]. Therefore, variations in BMI (which is the most commonly used measure for monitoring the prevalence of overweight and obesity) may be associated with changes in disability and mortality to different extent between younger and older cohorts. For this reason, in the last stage of analysis we considered BMI at baseline and its interaction with year of birth in the association with YLD.

BMI was calculated by dividing a person's weight in kilograms by the square of their height in metres. Measurements of weight and height were available from the nurse visits (available from wave 2). If height or weight could not be measured, then an estimate was obtained from a self-assessment of the respondent instead. In the following we use BMI measured at wave 2 and treated as a continuous variable.

Analysis

Measurement of disability

Disability is a complex and challenging process to study, especially when it develops over years or decades [95]. Researchers who study disability have adopted different approaches to examine steps along the pathway to disability. These include: (a) disentangling the different disability dimensions such as functional limitations and activity and participation restrictions following a hierarchical order. This approach has been efficient to explaine contrasted trends: for instance the increase in the years lived with functional limitations did not systematically translate into an increase in the years of activity restrictions [23]. (b) Identifying categories of disability based on its severity rather than its dimension [28, 43]. This implies considering disability as a continuum -rather than a hierarchical process- that, ideally, taps full ranges of ability [51], and low and high levels can be distinguished [28].

Each approach addresses a different question. For the purposes of our study, we chose the second, and arguably more parsimonious approach, and adopted the same process as a previous work [43] set in England, which measured disability according to the ICF using ELSA to identify the optimal number of disability classes among the older population. The study found that the best classification consisted of four classes ("non-disability", "mild disability", "moderate disability" and "severe disability"), such that each grade of disability was significantly different from the "non-disability" group in the association with health and mortality observed over a 10-year period; and each level presented a specific profile in impairment, activity limitations and participation restrictions (see figure 7.3 in the Supporting Information to research Paper IV). To replicate this result in our setting, we used Latent Class Analysis (LCA) to estimate four classes of disability at each of the 6 follow-up waves. For binary items and a categorical latent variable C with four classes ($k = 1, \ldots 4$), the marginal probability of observing item u_j (with $j = 1, 2 \ldots, 42$) being equal to one is

$$Pr(u_j = 1) = \sum_{k=1}^{4} Pr(C = k) Pr(u_j = 1 | C = k) \tag{7.1}$$

Where the second part of Eq. 7.1, $Pr(u_j = 1|C = k)$, denotes the conditional probability of the item being equal to one given that the class is equal to k and Pr(C=k) is the marginal probability of the class being k. The model was fitted at each wave, for both the cross-sectional and longitudinal samples, separately for men and women.

Disability-free life expectancy

DFLE was estimated using the Sullivan method. The Sullivan method employs a relatively simple modification of the conventional life table model to compute the expected duration of certain defined conditions of interest among the living population [72]. The health expectancy component reflects the current health of a real population adjusted for mortality levels independent of age structure. Sullivan's method relies on the assumption that a specific cohort observed at a certain age in a given year will be experiencing the same disability prevalence rates observed among the other age groups (i.e. other cohorts) in the same year. This is an extra stationarity assumption of the population, in addition to the three stationarity assumptions inherited from the period life table (i.e. the age-specific hazard rate is

constant over time, the birth rate is constant over time and the net migration rates are 0 at all ages) which are also assumed in the Sullivan methods [96]. Problems of under- or overestimation may occur when disability prevalence changes over time, although several studies demonstrated that Sullivan's method can be extended to estimate health expectancy without stationarity assumptions [96, 97]. To assess the plausibility of the stationarity assumption on disability, we estimated disability at each wave in order to check whether the age-specific disability prevalence were stable across waves, although this implicitly requires strong as well as untestable assumptions about the occurrence or non-occurrence of health or disability transitions between assessment times. When disability data are collected at long intervals, it is likely that aspects of the underlying disablement process are undetected [98, 99].

DFLE was estimated for each disability class, i.e. mild DFLE, moderate DFLE and severe DFLE. The prevalence of mild, moderate and severe disability was estimated in order to be mutually exclusive rather than cumulative, consequently disability-free years and years with disability sum up to TLE separately for each level of disability. For each class of disability, we also report the ratio of DFLE over TLE and express them as proportions. Finally, given that the Sullivan health expectancy is subject to random variation, 95% confidence intervals were calculated from the standard errors of the probability of each disability class [100].

Years lost due to disability

DFLE is an aggregate measure and as such it is not possible to model it in terms of individual level variables. To overcome this problem, we complement this SMPH with a similar indicator that can be measured both at individual or population level: YLD. YLD is one of the components used to compute Disability-Adjusted Life Years (DALY). DALYs are the sum of two time-specific dimensions: the present value of future years of lifetime lost through premature mortality, called Years of Life Lost (YLL), and the present value of years of future lifetime adjusted for the average severity (frequency and intensity) of any mental or physical disability caused by a disease or injury, which corresponds to YLD [101].

Dealing with a non-extinct longitudinal cohort study, estimating DALY (and YLD)

requires assumptions relating to the censoring pattern, which was assumed to be non-informative because the reason for censoring was the end of time of observation, which is unrelated with the outcome of interest.

Generally, on an individual basis and for each gender, the basic formula for calculating YLD is:

$$YLD(a, c) = DW_c * L_{c,a}$$

Such that YLD is a function of age (a) and type of condition (c) that in this case corresponds to disability class. DW_c is the disability weight for disability class c, it is a weight factor that reflects the severity of the disease on a scale from zero (perfect health) to one (equivalent to death); $L_{c,a}$ is the duration of disability c from age a until remission or death [102]. Further details on the computation of YLD are available in the Supporting Onformation to Research Paper IV.

We measured YLD at waves 2 and 6 and then used these individual estimates of YLD to investigate whether their variation from 2004 to 2012 could be explained in terms of changes in BMI in 2004, controlling for year of birth. We used wave 2 instead of wave 1 as starting point because BMI was measured only during nurse visits. To examine this, at waves 2 and 6, we first predicted YLD regressing it against the age of the participant at the time of each interview, and retained the residuals. Then, we modelled the difference between the age-adjusted residuals of YLD estimated in 2012 and 2004, separately in men and women, and regress them linearly on year of birth, BMI and their interaction. Because of the age-adjustment, the significance of the regressor year of birth is sufficient to test the compression (or expansion) of disability [103]. It does not allow though to establish whether the effect is a cohort or period effect [104].

7.2.4 Results

Disability distribution across waves

Tables 7.1 and 7.2 illustrate, respectively for men and women, the distribution of disability classes at wave 1 and wave 6 for cross-sectional and longitudinal samples, without standardising for age. At wave 1, around 37% of women and 46% of men were classified as

non-disabled and the most severe form of disability affected 12% of women and 9.5% of men. At wave 6, in the cross-sectional sample the percentages of respondents belonging to non-disabled group were larger than at wave 1 for women and smaller for men. When compared to proportions obtained from the longitudinal sample, cross-sectional percentages of non-disabled at wave 6 were larger both for males and females. This was most likely because of confounding by age, with members of the longitudinal sample older than members of the cross-sectional sample. Disability classes were also estimated at intermediate waves, both for longitudinal and cross-sectional samples (table 7.9 in the Supporting Information to Research Paper IV). Cross-sectional proportions did not differ remarkably across waves, but overall the percentages of respondents belonging to non-disabled group increased with time and the proportions of respondents belonging to the most severely disabled group were slightly higher in the first wave, both for men and women. For longitudinal samples, we did not observe an increase in proportions of non-disabled at later waves, instead they were constant for women and slightly declining for men.

Table 7.1: Disability classes in cross-sectional and longitudinal samples, waves 1 and 6, men

Disability level	**Wave 1**		**Wave 6**			
	Cross-sect.=Long.		**Cross-sect.**		**Long.**	
	n	**%**	**n**	**%**	**n**	**%**
Non-disabled	2070	46.4	1529	45.9	875	43
Low disabled	1138	25.5	815	24.5	500	24.6
Mildly disabled	831	18.6	670	20.1	455	22.3
Severely disabled	423	9.5	320	9.6	207	10.2
Total	4,462	100	3,334	100	2037	100

Disability-free life expectancy

In this section, we compare DFLE in 2002 and 2012, and estimates based on cross-sectional and longitudinal samples. Since the longitudinal sample consisted on a subsample of survivors interviewed at wave 1 and followed up to the sixth wave, we expected longitudinal estimates of DFLE to be higher (i.e. more years without disability) compared to cross-sectional ones. For each class of disability, we report the number of expected years without and with disability (DFLE and Disability Life Expectancy (DLE) respectively), and pro-

Table 7.2: Disability classes in cross-sectional and longitudinal samples waves 1 and 6, women

Disability level	Wave 1		Wave 6			
	Cross-sect=Long.		Cross-sect.		Long.	
	n	%	n	%	n	%
Non-disabled	1931	36.7	1653	39.6	924	36
Low disabled	1266	24	1030	24.7	639	24.9
Mildly disabled	1439	27.3	1040	24.9	721	28.1
Severely disabled	633	12	450	10.8	281	11
Total	5,269	100	4,173	100	2,565	100

portional results for DFLE over TLE. DFLE is presented by gender in tables 7.3 and 7.4 for both longitudinal and cross-sectional samples. DLE, which represents the absolute gap between DFLE and TLE, and proportional DFLE are shown in tables 7.5 and 7.6 only for the cross-sectional sample, and available in the Supporting Information to Research Paper IV for the longitudinal sample (tables 7.10).

Tables 7.3 and 7.4 show, respectively for men and women, TLE in 2002 and 2012 and expected years of life spent free of each level of disability (i.e. mild DFLE, moderate DFLE, severe DFLE) in 2002 and 2012. Cross-sectional and longitudinal samples coincided at wave 1 and therefore estimates of DFLE were the same in 2002. In 2012 estimates of DFLE for the longitudinal sample were available only from age 60, being the youngest respondents at wave 1 (i.e. ten years before) aged 50. In both samples, from 2002 to 2012 the expected number of years spent without any level of disability has increased, for both men and women. For example, a woman aged 50-54 in 2002 could expect to live 32.5 years, of which 28 years without severe disability, 23.1 years without moderate disability and 25 years without mild disability. In 2012 life expectancy of women aged 50-54 raised to 34.6 years and 30.3 years of those were expected without severe disability, and 25.5 years and 26.4 years without moderate and mild disability, respectively.

The longitudinal estimates of DFLE in 2012 presented slightly more years of life expectancy without severe and moderate disability compared to the corresponding cross-sectional estimates. Therefore in the longitudinal sample the increase in DFLE from 2002 to 2012 was larger, but not noticeably. This was in line with expectations, because the

longitudinal sample was composed of selected healthier individuals that did not die during the follow up.

Estimates of DFLE at wave one and wave six for the cross-sectional sample were also replicated using attrition weights for wave 6. Results were substantially similar to unweighted estimates, and are available in the Supporting Information to Research Paper IV (tables 7.15 and 7.16).

Table 7.3: Years of TLE and life expectancy without disability by age and gender in cross sectional and longitudinal samples. Men

Age	TLE		Severe DFLE			Moderate DFLE			Mild DFLE		
	2002	2012	2002	2012 cross-sect.	2012 long.	2002	2012 cross-sect.	2012 long.	2002	2012 cross-sect.	2012 long.
50	28.9	31.7	25.9 (25.6; 26.2)	28.4 (28.1; 28.8)	- -	23.2 (22.9; 23.6)	25.3 (24.9; 25.8)	- -	21.6 (21.2; 22)	24 (23.4; 24.5)	- -
55	24.5	27.2	21.7 (21.5; 22)	24.2 (23.8; 24.5)	- -	19.2 (18.9; 19.6)	21.2 (20.8; 21.7)	- -	18.5 (18.1; 18.8)	20.9 (20.5; 21.3)	- -
60	20.3	23	17.8 (17.5; 18.1)	20.1 (19.8; 20.5)	20.4 (20.1; 20.7)	15.5 (15.2; 15.9)	17.4 (17; 17.9)	17.6 (17.1; 18)	15.4 (15.1; 15.8)	17.7 (17.3; 18.1)	17.5 (17.1; 17.9)
65	16.5	19	14.4 (14.1; 14.6)	16.3 (16; 16.6)	16.5 (16.2; 16.9)	12.2 (11.9; 12.6)	14 (13.6; 14.4)	14.1 (13.6; 14.5)	12.5 (12.2; 12.8)	14.7 (14.3; 15.1)	14.6 (14.2; 14.9)
70	13	15.2	11.1 (10.9; 11.4)	12.7 (12.4; 13)	12.9 (12.6; 13.3)	9.4 (9; 9.7)	10.9 (10.5; 11.3)	10.9 (10.5; 11.3)	9.9 (9.6; 10.2)	12 (11.7; 12.4)	11.9 (11.5; 12.3)
75	10	11.9	8.2 (7.9; 8.4)	9.6 (9.3; 10)	9.8 (9.5; 10.2)	7.1 (6.8; 7.5)	8.3 (7.8; 8.7)	8.2 (7.8; 8.7)	7.8 (7.5; 8.1)	9.3 (9; 9.7)	9.2 (8.9; 9.6)
80	7.7	9	5.9 (5.5; 6.2)	7 (6.6; 7.4)	7.2 (6.8; 7.6)	5.4 (5; 5.7)	6 (5.5; 6.4)	5.9 (5.5; 6.4)	6 (5.7; 6.3)	7.4 (7.1; 7.8)	7.4 (7; 7.7)

95% confidence intervals in brackets ()

Table 7.4: Years of TLE and life expectancy without disability by age and gender in cross-sectional and longitudinal samples. Women

Age	TLE		Severe DFLE			Moderate DFLE			Mild DFLE		
	2002	2012	2002	2012 cross-sect.	2012 long.	2002	2012 cross-sect.	2012 long.	2002	2012 cross-sect.	2012 long.
50	32.5	34.6	28	30.3	-	23.1	25.5	-	25	26.4	-
			(27.7; 28.3)	(29.9: 30.7)	-	(22.7; 23.5)	(25: 26)	-	(24.6; 25.4)	(25.9: 26.8)	-
55	28	30	23.7	25.9	-	19.3	21.5	-	21.9	23	-
			(23.3; 24)	(25.5: 26.3)	-	(18.9; 19.7)	(21: 21.9)	-	(21.5; 22.2)	(22.6: 23.4)	-
60	23.6	25.5	19.6	21.8	22.2	15.8	17.6	17.7	18.7	19.8	19.4
			(19.3; 19.9)	(21.4: 22.2)	(21.9; 22.6)	(15.4; 16.2)	(17.2: 18.1)	(17.3; 18.2)	(18.3; 19)	(19.4: 20.2)	(19; 19.9)
65	19.4	21.2	15.7	17.8	18.2	12.6	14.1	14.3	15.6	16.6	16.3
			(15.4; 16)	(17.4: 18.1)	(17.8; 18.5)	(12.2; 12.9)	(13.7: 14.5)	(13.8; 14.7)	(15.3; 15.9)	(16.2: 17)	(15.9; 16.7)
70	15.5	17.1	12.1	14	14.3	9.6	10.9	11	12.7	13.6	13.3
			(11.8; 12.4)	(13.6: 14.3)	(14; 14.7)	(9.2; 9.9)	(10.4: 11.3)	(10.6; 11.5)	(12.4; 13)	(13.3: 14)	(13; 13.7)
75	12	13.4	9	10.5	10.8	7.3	8	8.1	9.8	10.9	10.7
			(8.7; 9.3)	(10.1: 10.8)	(10.4; 11.1)	(7; 7.7)	(7.6: 8.4)	(7.6; 8.5)	(9.5; 10.1)	(10.6: 11.2)	(10.4; 11.1)
80	9	10	6.3	7.3	7.7	5.5	5.7	5.6	7.5	8.5	8.3
			(6; 6.6)	(6.9: 7.8)	(7.3; 8)	(5.2; 5.9)	(5.2: 6.2)	(5.2; 6.1)	(7.2; 7.8)	(8.2: 8.8)	(8; 8.7)

95% confidence intervals in brackets ()

Tables 7.5 and 7.6 present complementary data to tables 7.3 and 7.4 for the cross-sectional sample, by level of disability, for men and women respectively. Years lived with each level of disability are reported by gender in 2002 and 2012, and the difference between the two periods ($\Delta DLE = DLE_{2012} - DLE_{2002}$) corresponds to absolute compression if negative (i.e. fewer expected years with disability in 2012 compared to 2002), and absolute expansion if positive. Proportions of DFLE on TLE in 2002 and 2012 are also reported, along with their difference (i.e. $\Delta\%DFLE = (DFLE_{2012}/TLE_{2012}) - (DFLE_{2002}/TLE_{2002})$). Opposite to the differences in absolute values, the differences in proportional DFLE correspond to proportional compression in case of positive values (i.e. proportion of life without disability larger in 2012 compared to 2002) and to expansion for negative values.

Among men, years of life with severe disability (top of table 7.5) have slightly increased, and symmetrically proportions of severe DFLE have declined only marginally, especially at younger ages. Life expectancy with moderate disability (middle table) has increased in absolute terms and declined as proportion of TLE; life expectancy with mild disability (bottom table) has increased in absolute terms (i.e. positive difference in DLE between 2012 and 2002), but its proportion on TLE has declined (i.e. positive difference in DFLE over TLE between 2012 and 2002). In general, however, absolute and proportional differences between 2002 and 2012 were small and confidence intervals overlapped. Among women the number of expected years with severe disability (top of table 7.6) has slightly declined in 2012, but only by about 0.3 years, and the proportion of severe DFLE has increased. The opposite was observed for mild DLE (bottom table), which has increased in absolute terms and as proportion of TLE (i.e. smaller proportion of life expectancy without mild disability). For moderate disability (middle table), changes varied across ages, with reduced years in disability and larger proportion of life free of disability at younger ages, and the opposite observed at older ages. As for men, overall, variations were quite small and in most cases, the confidence intervals overlapped.

All combined, these results pointed at identifying a dynamic equilibrium for women, while men experienced a worse pattern than females, because their years with any level of disability increased, although only slightly, and proportions of life without disability increased only for mild disability.

A final remark pertains to the fact that patterns in proportions of DFLE appeared to have a break al older ages in some cases: among women the proportion of life expectancy with moderate disability in 2012 compared to 2002 has decreased until age 70-74, but it has increased after the age of 75 compared to 2002; among men, the proportion of severe DFLE on TLE has increased, while the proportion of moderate DFLE has declined noticeably at ages older than 70 years.

Years lost due to disability

Lastly, we tried to interpret our findings and understand whether there is an effect of year of birth and BMI. Table 7.7 reports the results of the gender-specific linear regressions where the outcome consists in the difference between age-adjusted residuals of YLD at wave 6 and wave 2, and the exposures are sequentially, year of birth, BMI measured at wave 2, and their interaction. Positive values of the outcome correspond to year increases in age-adjusted YLD, i.e. YLD at wave 6 larger than YLD at wave 2; conversely, negative values correspond to a reduction in age-adjusted YLD at wave 6 compared with wave 2. There was no statistical evidence of a quadratic effect of BMI (p=0.89), and so it was treated linearly. When only year of birth was included in the model (column 1), it appeared to have a positive effect for men, such that for 1-year increase in year of birth the difference in YLD residuals increased by 0.04, meaning that younger cohorts experienced larger increase in YLD (i.e. more years lost to disability) than older cohorts. No year of birth effect was found for women. BMI was not significantly associated with the outcome either when controlling (column 3) or not controlling (column 2) for year of birth. This was seen for both men and women. The most interesting results are those in column 4. In this model an interaction term between continuous BMI and continuous year of birth was added and it was found to be significant (p=0.023 and 0.016, respectively for men and women). Results are explained graphically in figures 7.1 and 7.2. The graphs report the association between the outcome and one of the exposures (BMI in figure 7.1 and year of birth in figure 7.2) holding constant the other variable in the interaction term (year of birth in figure 7.1 and BMI in figure 7.2) at pre-selected values. Figure 7.1 highlights how the direction of the association between BMI and the age-adjusted difference in YLD changed across cohorts

Table 7.5: TLE and disability-related life expectancy measures for absolute and proportional changes, men

Age	TLE			Severe-disability LE					
				DLE		ΔDLE[a]	DFLE/TLE		Δ%DFLE[b]
	2002	2012	ΔTLE	2002	2012		2002	2012	
50	28.9	31.7	2.8	3.0 (2.7; 3.2)	3.3 (2.9; 3.6)	0.3 (-0.3;0.9)	89.7 (87.6; 91.8)	89.6 (85.6; 93.6)	-0.1 (-2.2;2)
55	24.5	27.2	2.7	2.8 (2.5; 3)	3.1 (2.7; 3.4)	0.3 (-0.3;0.9)	88.7 (86.6; 90.8)	88.7 (85.9; 91.5)	0 (-2.3;2.3)
60	20.3	23.0	2.7	2.5 (2.3; 2.8)	2.8 (2.5; 3.2)	0.3 (-0.3;0.9)	87.6 (85.2; 90)	87.7 (85.2; 90.2)	0.1 (-2.6;2.8)
65	16.5	19.0	2.5	2.1 (1.9; 2.4)	2.7 (2.3; 3)	0.6 (0;1.2)	87.1 (84.6; 89.6)	85.9 (83.3; 88.5)	-1.2 (-4.5;2.1)
70	13.0	15.2	2.2	1.9 (1.6; 2.1)	2.5 (2.2; 2.8)	0.6 (0;1.2)	85.5 (82.6; 88.4)	83.6 (80.4; 86.8)	-1.9 (-6.1;2.3)
75	10.0	11.9	1.9	1.8 (1.6; 2.1)	2.2 (1.9; 2.6)	0.4 (-0.2;1)	81.5 (77.7; 85.3)	81.2 (77.5; 84.9)	-0.3 (-6.1;5.5)
80	7.7	9.0	1.3	1.8 (1.5; 2.1)	2.1 (1.7; 2.5)	0.3 (-0.4;1)	76.4 (72.1; 80.7)	77.3 (72.9; 81.7)	0.9 (-7.9;9.7)

Age	TLE			Moderate-disability LE					
				DLE		ΔDLE[a]	DFLE/TLE		Δ%DFLE[a]
	2002	2012	ΔTLE	2002	2012		2002	2012	
50	28.9	31.7	2.8	5.6 (5.3; 6)	6.4 (5.9; 6.8)	0.8 (0;1.6)	80.5 (77.8; 83.2)	79.9 (74.6; 85.2)	-0.6 (-3.2;2)
55	24.5	27.2	2.7	5.2 (4.9; 5.6)	6.0 (5.6; 6.4)	0.8 (0;1.6)	78.6 (75.9; 81.3)	78 (74.4; 81.6)	-0.6 (-3.5;2.3)
60	20.3	23.0	2.7	4.8 (4.4; 5.1)	5.5 (5.1; 5.9)	0.7 (0;1.4)	76.5 (73.4; 79.6)	75.9 (72.6; 79.2)	-0.6 (-4;2.8)
65	16.5	19.0	2.5	4.2 (3.9; 4.6)	5.0 (4.6; 5.4)	0.8 (0.1;1.5)	74.3 (71.1; 77.5)	73.7 (70.4; 77)	-0.6 (-4.7;3.5)
70	13.0	15.2	2.2	3.6 (3.3; 4)	4.3 (3.9; 4.7)	0.7 (0;1.4)	72 (68.3; 75.7)	71.5 (67.6; 75.4)	-0.5 (-5.6;4.6)
75	10.0	11.9	1.9	2.9 (2.5; 3.2)	3.6 (3.2; 4)	0.7 (0;1.4)	71.4 (67; 75.8)	69.5 (65.2; 73.8)	-1.9 (-8.6;4.8)
80	7.7	9.0	1.3	2.3 (2; 2.7)	3.0 (2.6; 3.5)	0.7 (-0.1;1.5)	69.8 (65.1; 74.5)	66.5 (61.5; 71.5)	-3.3 (-13;6.4)

Age	TLE			Mild-disability LE					
				DLE		ΔDLE[a]	DFLE/TLE		Δ%DFLE[b]
	2002	2012	ΔTLE	2002	2012		2002	2012	
50	28.9	31.7	2.8	7.3 (6.9; 7.6)	7.8 (7.2; 8.3)	0.5 (-0.4;1.4)	74.8 (71.8; 77.8)	75.5 (69.8; 81.2)	0.7 (-2.2;3.6)
55	24.5	27.2	2.7	6.0 (5.7; 6.4)	6.3 (5.9; 6.7)	0.3 (-0.5;1.1)	75.4 (72.6; 78.2)	76.8 (73.1; 80.5)	1.4 (-1.5;4.3)
60	20.3	23.0	2.7	4.9 (4.5; 5.2)	5.3 (4.9; 5.7)	0.4 (-0.3;1.1)	76 (72.8; 79.2)	76.9 (73.6; 80.2)	0.9 (-2.4;4.2)
65	16.5	19.0	2.5	4.0 (3.7; 4.3)	4.3 (3.9; 4.6)	0.3 (-0.4;1)	75.9 (72.7; 79.1)	77.5 (74.4; 80.6)	1.6 (-2.2;5.4)
70	13.0	15.2	2.2	3.1 (2.8; 3.3)	3.2 (2.8; 3.5)	0.1 (-0.5;0.7)	76.5 (73; 80)	79.2 (75.7; 82.7)	2.7 (-1.9;7.3)
75	10.0	11.9	1.9	2.2 (1.9; 2.5)	2.5 (2.2; 2.9)	0.3 (-0.3;0.9)	78.1 (74.1; 82.1)	78.7 (74.8; 82.6)	0.6 (-5.3;6.5)
80	7.7	9.0	1.3	1.6 (1.3; 2)	1.6 (1.2; 2)	0 (-0.7;0.7)	78.5 (74.3; 82.7)	82.2 (78.2; 86.2)	3.7 (-4.5;11.9)

[a] $\Delta DLE = DLE_{2012} - DLE_{2002}$
[b] $\Delta\%DFLE = (DFLE_{2012}/TLE_{2012}) - (DFLE_{2002}/TLE_{2002})$
95% confidence intervals in brackets ()

Table 7.6: TLE and disability-related life expectancy measures for absolute and proportional changes, women

Age	TLE			Severe-disability LE					
	2002	2012	ΔTLE	DLE		ΔDLE[a]	DFLE/TLE		Δ%DFLE[b]
				2002	2012		2002	2012	
50	32.5	34.6	2.1	4.5 (4.2; 4.8)	4.2 (3.8; 4.6)	-0.3 (-1;0.4)	86.1 (83.9; 88.3)	87.7 (83.8; 91.6)	1.6 (-0.5;3.7)
55	28	30	2	4.3 (4; 4.7)	4.1 (3.7; 4.4)	-0.2 (-0.9;0.5)	84.5 (82.3; 86.7)	86.4 (83.8; 89)	1.9 (-0.5;4.3)
60	23.6	25.5	1.9	4 (3.7; 4.3)	3.7 (3.4; 4.1)	-0.3 (-1;0.4)	83.1 (80.5; 85.7)	85.4 (83; 87.8)	2.3 (-0.5;5.1)
65	19.4	21.2	1.8	3.7 (3.4; 4)	3.5 (3.1; 3.8)	-0.2 (-0.9;0.5)	80.9 (78.2; 83.6)	83.7 (81.2; 86.2)	2.8 (-0.5;6.1)
70	15.5	17.1	1.6	3.4 (3.1; 3.7)	3.2 (2.8; 3.6)	-0.2 (-0.9;0.5)	78 (74.8; 81.2)	81.4 (78.3; 84.5)	3.4 (-0.8;7.6)
75	12	13.4	1.4	3 (2.7; 3.3)	2.9 (2.6; 3.3)	-0.1 (-0.8;0.6)	74.9 (71; 78.8)	78.1 (74.6; 81.6)	3.2 (-2.3;8.7)
80	9	10	1	2.7 (2.4; 3.1)	2.7 (2.3; 3.1)	0 (-0.8;0.8)	69.6 (65.8; 73.4)	73.1 (69; 77.2)	3.5 (-4.4;11.4)

Age	TLE			Moderate-disability LE					
	2002	2012	ΔTLE	DLE		ΔDLE[a]	DFLE/TLE		Δ%DFLE[b]
				2002	2012		2002	2012	
50	32.5	34.6	2.1	9.4 (9; 9.9)	9.1 (8.6; 9.6)	-0.3 (-1.2;0.6)	71 (68.2; 73.8)	73.7 (68.5; 78.9)	2.7 (0;5.4)
55	28	30	2	8.7 (8.3; 9.1)	8.5 (8; 8.9)	-0.2 (-1.1;0.7)	68.9 (66.1; 71.7)	71.7 (68.3; 75.1)	2.8 (-0.2;5.8)
60	23.6	25.5	1.9	7.8 (7.4; 8.2)	7.9 (7.4; 8.3)	0.1 (-0.7;0.9)	66.9 (63.6; 70.2)	69.1 (65.9; 72.3)	2.2 (-1.2;5.6)
65	19.4	21.2	1.8	6.9 (6.5; 7.2)	7.1 (6.7; 7.6)	0.2 (-0.6;1)	64.6 (61.3; 67.9)	66.4 (63.2; 69.6)	1.8 (-2.2;5.8)
70	15.5	17.1	1.6	5.9 (5.6; 6.3)	6.3 (5.8; 6.7)	0.4 (-0.4;1.2)	61.7 (58; 65.4)	63.4 (59.6; 67.2)	1.7 (-3.2;6.6)
75	12	13.4	1.4	4.7 (4.3; 5)	5.4 (5; 5.8)	0.7 (-0.1;1.5)	61 (56.6; 65.4)	59.7 (55.5; 63.9)	-1.3 (-7.6;5)
80	9	10	1	3.5 (3.1; 3.9)	4.3 (3.9; 4.8)	0.8 (0;1.6)	61.3 (57.2; 65.4)	56.7 (52.1; 61.3)	-4.6 (-13.2;4)

Age	TLE			Mild-disability LE					
	2002	2012	ΔTLE	DLE		ΔDLE[a]	DFLE/TLE		Δ%DFLE[b]
				2002	2012		2002	2012	
50	32.5	34.6	2.1	7.5 (7.1; 7.9)	8.2 (7.7; 8.7)	0.7 (-0.2;1.6)	76.9 (74.3; 79.5)	76.3 (71.2; 81.4)	-0.6 (-3.1;1.9)
55	28	30	2	6.1 (5.8; 6.5)	7 (6.6; 7.4)	0.9 (0.1;1.7)	78.1 (75.6; 80.6)	76.7 (73.5; 79.9)	-1.4 (-4;1.2)
60	23.6	25.5	1.9	4.9 (4.6; 5.2)	5.7 (5.3; 6.1)	0.8 (0.1;1.5)	79.2 (76.3; 82.1)	77.6 (74.8; 80.4)	-1.6 (-4.5;1.3)
65	19.4	21.2	1.8	3.8 (3.5; 4.1)	4.6 (4.3; 5)	0.8 (0.1;1.5)	80.4 (77.6; 83.2)	78.2 (75.4; 81)	-2.2 (-5.5;1.1)
70	15.5	17.1	1.6	2.8 (2.5; 3.1)	3.5 (3.2; 3.9)	0.7 (0.1;1.3)	81.9 (79; 84.8)	79.4 (76.2; 82.6)	-2.5 (-6.4;1.4)
75	12	13.4	1.4	2.2 (1.9; 2.5)	2.5 (2.2; 2.8)	0.3 (-0.3;0.9)	81.8 (78.3; 85.3)	81.4 (78.1; 84.7)	-0.4 (-5.3;4.5)
80	9	10	1	1.6 (1.3; 1.8)	1.6 (1.2; 1.9)	0 (-0.6;0.6)	82.7 (79.5; 85.9)	84.6 (81.3; 87.9)	1.9 (-4.6;8.4)

[a] $\Delta DLE = DLE_{2012} - DLE_{2002}$
[b] $\Delta\%DFLE = (DFLE_{2012}/TLE_{2012}) - (DFLE_{2002}/TLE_{2002})$
95% confidence intervals in brackets ()

for both men and women, with the older cohort groups (born in 1915, 1925 and 1935) experiencing fewer years lost with increasing BMI, i.e. a protective effect of BMI (negative slopes), whilst increasing BMI appears to be detrimental for those born in 1955. Looking at the effect of year of birth moderated by BMI (figure 7.2), we find that year of birth was positively associated with the outcome for respondents classified as obese at wave 2, not correlated for those overweight and negatively associated for those who were normal-weight. Again, results were similar for women and men.

In synthesis, we observed an effect of year of birth only for men (column 1). This is in line with the increase in disability observed in their absolute and proportional DFLE. Then, when we included an interaction between year of birth and BMI, the fact that the coefficient was significant and positive means that (i) a year increase in year of birth makes the effect of BMI on the outcome larger: YLD increases with higher level of BMI, more strongly the younger the respondent is; (ii) a unit increase in BMI makes the effect of year of birth on the outcome larger: YLD increases with year of birth, more strongly the higher the BMI is.

The analysis was also replicated using attrition weights for wave 6 and results are available in the Supporting Information to Research Paper IV (table 7.17).

Table 7.7: Estimated coefficients from linear regression models of the difference between age-adjusted YLD residuals at wave 6 and wave 2 ($\hat{\epsilon_6} - \hat{\epsilon_2}$) (in years)

	(1)		**(2)**		**(3)**		**(4)**	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Year of birth[a]	0.0308**	0.0133			0.0297**	0.0143	0.0307**	0.0163
	(0.0127)	(0.0128)			(0.0131)	(0.0133)	(0.0131)	(0.0134)
BMI[a]			0.0183	-0.0131	0.0153	-0.0136	0.0113	-0.0152
			(0.0256)	(0.0206)	(0.0256)	(0.0206)	(0.0256)	(0.0206)
BMI*yob							0.0077**	0.0065**
							(0.0034)	(0.0027)
Constant	1,708	2,172	1,556	1,997	1,556	1,997	1,556	1,997

[a] Both variables were centred.
Standard errors in brackets; $*p < 0.1; **p < 0.05; ***p < 0.01$; yob=year of birth

Figure 7.1: Observed and predicted association between BMI and age-adjusted YLD difference, moderated by year of birth, by gender

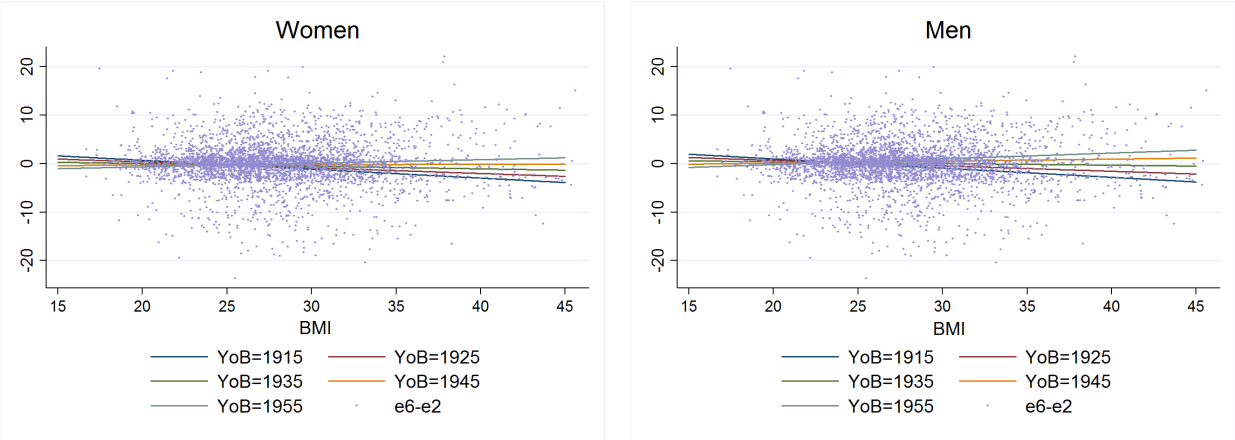
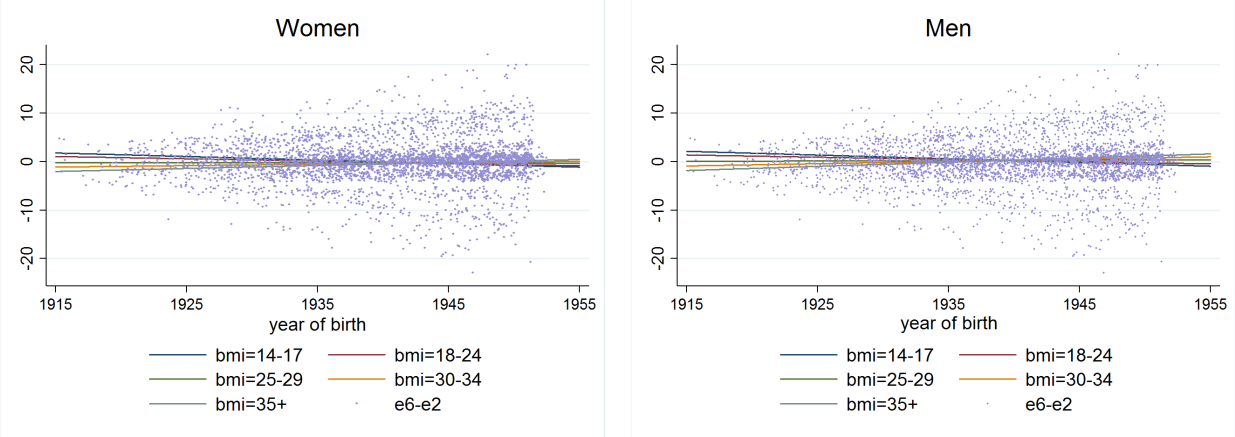


Figure 7.2: Observed and predicted association between year of birth and age-adjusted YLD difference, moderated by BMI, by gender



7.2.5 Discussion

Synthesis of findings

Our study adopted a comprehensive interpretation of disability which was derived from the WHO's ICF framework, and used a multi-categorical classification of disability that distinguished non-disabled from those with mild, moderate and severe forms of disability. We used it to study trends in category-specific DFLE over the past decade. We also proposed possible explanations for the observed changes in DFLE by changing the focus from the aggregate level of health expectancy to an equivalent individual-level outcome, corresponding to YLD. BMI and year of birth were explored as explanatory factors for changes in this quantity from 2004 to 2012. In the following we discuss our results, first interpreting the main findings in the light of the theories of population health change. Then we interpret the exploratory analyses performed to assess possible causes and mechanisms behind changes in DFLE. Finally, the implications and relevance of these findings for society and the health system are highlighted.

Interpreting evidence within theories of population health change

Results were different for men and for women. While men experienced larger increases in life expectancy than women -and the phenomenon of men catching up with female life expectancy has been previously reported and recently confirmed in England and Wales [4]- the increase in DFLE was very similar across genders. For men, severe DLE and severe DFLE on TLE roughly stayed constant, while moderate DLE increased both in absolute and proportional terms. Women experienced proportional and absolute decline of severe DLE, while their mild DLE increased both in absolute and proportional terms. Results were similar across most age groups, but not after around age 75 years, where the proportion of moderate DFLE on TLE declined among both men and women and severe DLE in women remained stable, while at younger ages years with severe disability declined, but only slightly.

When we compared results of cross-sectional and longitudinal samples at wave 6 (tables 7.3 and 7.4), we found the same direction of changes and similar estimates. The longitudinal sample performed slightly better than the cross-sectional respondents, with larger compression where the cross-sectional DFLE compressed and smaller expansion where it expanded.

This would suggest that the subset of survivors, both males and females, that make up the longitudinal study were healthier than the general sample and therefore, while surviving over the entire observation period, they experienced less disability.

Putting our results into context, we contributed to support the evidence that women are experiencing a compression of severe disability and expansion in milder levels, which corresponds to a general dynamic equilibrium. This was in line with what observed over the past two decades in the US [76], and in England [40]. The results of our work complement the study by Jagger et al. [40], and advance the understanding of current dynamics of healthy ageing in England. In fact, Jagger and colleagues provided evidence on trajectories in health expectancy considering separately various health indicators, including disability that was measured by ADLs and IADLs, and showed different paths depending on the dimension of health considered. This is extremely useful to address specific policies and intervene on the spheres of health that appear particularly at risk of deterioration. Given the complexity of the concept of disability, however it is often difficult to measure it independently from other dimensions of health. Moreover, relying on self-reported measures, one of the threats is that self-reporting bias may affect different spheres of health in different ways, and therefore the comparison of different domains may be biased as well. Therefore, assessing and combining results based on different measures of disability (i.e. ours and Jagger and colleagues' measures) can bring further understanding on the process of healthy ageing. Specifically, compared to Jagger and colleagues' measures of disability, our study used a broader interpretation and distinguished a more refined scale of severity, including also moderate levels. Results suggest that the main burden of disability in future years is likely to come from mild disability in women, and moderate disability in men. The identification of an intermediate grade of disability appears informative, especially in the comparison between men and women.

Another result which agrees with the existing literature, is that despite the improvement in DFLE experienced by women and the worsening conditions of men, the so-called "gender paradox in health and mortality" -according to which women live longer than men but spend larger proportions of their life with disability [67, 68]- continued to apply: both in 2002 and 2012 women had higher TLE than men and their proportion of DFLE over TLE

was smaller. The gap however has shrunken, and if the direction of changes will remain the same, men will catch up women by not only living longer, but also spending larger proportions of their lives with disability.

With regard to gender differences in changes in health expectancy, a critical aspect to bear in mind is that our estimates of DFLE are based on a disability measure that also includes some health conditions. Therefore, differences in diagnosis of specific health conditions between 2002 and 2012 might result in gender differences in disability measure. For this reason, we undertook a sensitivity analysis removing health condition from disability measure and found similar results to those produced including these items (tables 7.13 and 7.14 of the Supporting Information to Research Paper IV). This reassures that the influence of these variables was only modest and the gender gap that we observed was not (only) due to gender differences in prevalence and incidence of health condition over the past decade.

Explanations

Why did men do worse than women with regards to life expectancy with disability? Why did trends differ across levels of disability? Why did the direction of change differ across age groups? Before we describe findings relative to the last part of the analysis on interacting roles of year of birth and BMI in explaining variations in age-adjusted YLD, we propose and discuss possible answers to these questions. A possible interpretation to the worse trends in DFLE experienced by men compared to women is that one of the consequences of living longer lives is the possibility of living longer proportions of life in poor conditions (i.e. with disability). If men are still in the process of catching up with women in terms of survival pattern, it may be that their life expectancy is currently increasing because they are in the process of no longer dying due to disability, but largely surviving disability and consequently living longer with disability. This would explain the expansion of life expectancy with all levels of disability in men. At the same time, women, who were already more resilient to disability, may be experiencing a shift from severe forms of disability to milder conditions, possibly because of the success of preventive and curative medicine. In a previous study [42], disability at baseline was found to be positively predictive of mortality observed over a decade with the association being stronger for men, especially in the very short terms (i.e. within two years) while the effect of disability on mortality experienced by men was

found to converge to women's levels in the long term. This could mean that men become more resilient to disability the longer they survive, and therefore their life expectancy with disability is increasing relatively more than it does among women.

Moving away from pure speculations, we now focus on the last part of this work and discuss the exploratory analyses we undertook. We replaced the aggregate outcome of DFLE with the individual-level YLD and tried to assess whether changes in YLD between waves 2 and 6 were associated with year of birth and BMI at wave 2, and whether the two factors interacted with each other. The rationale behind the analysis came from the finding of expansion of mild disability opposed by compression of severe disability among women and therefore the consequent question of whether this was observed because severely disabled women had moved to milder forms of disability or because there had been some factors/mechanisms affecting different levels of disability in different ways. This consideration was combined with recent evidence that younger cohorts tend to be heavier than older cohorts [94], and therefore the two factors (i.e. year of birth and BMI) may interact. While acknowledging that our findings on this are quite exploratory, they suggest an interactive role of year of birth and BMI in changes in YLD, such that high BMI is particularly detrimental for younger generations. However, results were similar for men and women and therefore were not of use to explain gender differences in the trends of DFLE.

Limitations and strengths

Some limitations affect this work and must be borne in mind when interpreting the results. First of all, the cross-sectional data for 2002 does not include those in institutions whereas the 2012 data to an extent does. To overcome this limitation, we performed a sensitivity analysis excluding institutionalised respondents from ELSA sample at wave 6. Results based on this sample, available in the Supporting Information to Research Paper IV (tables 7.11 and 7.12), were the same as when participants in institutions were included. Therefore, the ELSA samples selected for our analysis both at wave one and six are representative of non-institutionalised population. Previous analyses have shown convergence in mortality patterns between the ELSA sample and the general English population. The national ONS mortality data used in this study reflects the total population and therefore includes individuals in institutions. The mortality rates for the very old age groups are therefore likely

to be too high for the oldest ELSA respondents, especially for the cross-sectional sample at wave 2. The second limitation, concerns BMI that was measured only at wave 2 and no information on onset or duration of overweight and obesity was considered. Younger cohorts have been found to become overweight much earlier in adulthood [94, 105] and this might explain why being overweight or obese was associated with increase in YLD for younger individuals. Another limitation comes from the fact that we dealt with non-extinct cohorts, and therefore incurred problems of censoring, which was assumed non-informative, and YLD was estimated based on very strong assumptions.

Our work has also some unique strengths. The identification of four levels of disability (including non-disability) allowed to capture finer differences in the diverging paths of DFLE between men and women. The generally agreed finding that severe forms of disability are not increasing was confirmed in our study, in accordance with previous evidence for England [40]. In this case, by identifying intermediate levels of disability we were able to describe the expansion of milder grades a step further, showing that men have experienced increasing level of moderate disability while women of milder forms. Another strength is that this study replicated the cross-sectional analyses on the longitudinal sample, allowing the comparison of results across two types of respondents, the former representative of the general English population aged 50+, the latter of survivors and as such presenting different probabilities of incurring disability. Therefore, it was not unexpected that the 2012 results on the longitudinal sample were slightly better than those from the cross-sectional data (expansion of disability was smaller and compression larger where observed). Nevertheless, the estimates of disability prevalence and DFLE were overall similar across the two samples. This is quite reassuring as it seems to indicate that attrition bias does not affect our measure of disability much, and thus the bias possibly introduced by the missing values affecting the YLD analyses may be only modest. Finally, all sensitivity analyses presented in the Supporting Information to Research Paper IV confirmed the results shown in the paper, strengthening the robustness of findings.

Implications for public health provision

This study offers robust evidence on the features and directions of ageing in contemporary England. Distinguishing mild, moderate and severe forms of disability allowed us to capture

specific patterns otherwise masked by averaging different trends. Distinguishing severe and moderate DLE levels, we were able to appreciate, for example, the decline of severe disability in women, although it was modest. Results were interpreted considering proportional changes along with absolute variations. All levels of disability life years have expanded, with the exception of severe disability for women, which stayed the same. This means that people with disability will need assistance for longer time and therefore the overall burden of disability on health system and families will increase. This is a very important finding, which would be ignored if focusing only on changes in DFLE in relative terms with changes in TLE.

To conclude, at least two central messages must be taken from this work, which have important implications for government as well as individuals, specifically for health service providers and family carers, i.e. the subjects in charge of supporting people with disability and incurring assistance costs. (i) It is helpful to distinguish between milder and more severe levels of disability because their trends seem to be divergent. Intermediate disability, on the other hand, appeared to behave fairly similarly to severe disability and therefore a mild-moderate-severe classification is not as key as the mild-severe categorization, although more informative. (ii) Although, this work did not show a causal effect, the evidence of a modifying effect of BMI and year of birth can be taken as a warning for closely monitoring BMI in younger generations and paying particular attention to avoiding an early onset of overweight and obesity.

7.3 Supporting Information to Research Paper IV

Years of life lost due to disability

Dealing with a non-extinct longitudinal cohort study, estimating DALY requires several assumptions to deal with censoring, which was assumed to be non-informative. Our focus was on the second dimension of DALY, YLD. Generally, for an individual who is in age category a with disability class c the basic formula for calculating his/her YLD is:

$$YLD(a, c) = DW_c * L_{a,c} \tag{7.2}$$

where DW_c is the disability weight corresponding to class c and $L_{a,c}$ is the number of years expected with disability c when in age group a. Since disability classes were mutually exclusive YLD was not estimated as a sum of different disability conditions, but independently for each disability level. The disability weights DW_c range from zero, representing perfect health, to one, representing death. They were determined at a meeting of experts in international health; we adopted those proposed by Murray [106]. The L component captures the duration of disability; its use required several assumptions. The first of these assumptions is that disability status does not change for the remaining of an individual's life expectancy, from the time when YLD is calculated, i.e. respectively from 2004 and from 2012. This could empirically be verified for the 2004 calculations, by monitoring whether respondents remained in the same disability condition as measured in 2004/2005 throughout the following waves, and therefore for a time span of eight years. An additional assumption regarded the duration of disability before it was assessed during the survey. We assumed that those answering "yes" to the question asked at wave 2 about having any long-standing illness, disability or infirmity, had been in the disability state recorded at wave 2 for at least two years before that interview, i.e. the time of the previous interview (wave 1). For YLD measured at wave 6, we proceeded in a similar way, assessing the duration of disability before wave 6 by checking if respondents were in the same disability status of wave 6 in all previous interviews, i.e. for 10 years.

Supporting tables and figures

Table 7.8: SMRs by age, gender and period.

Women

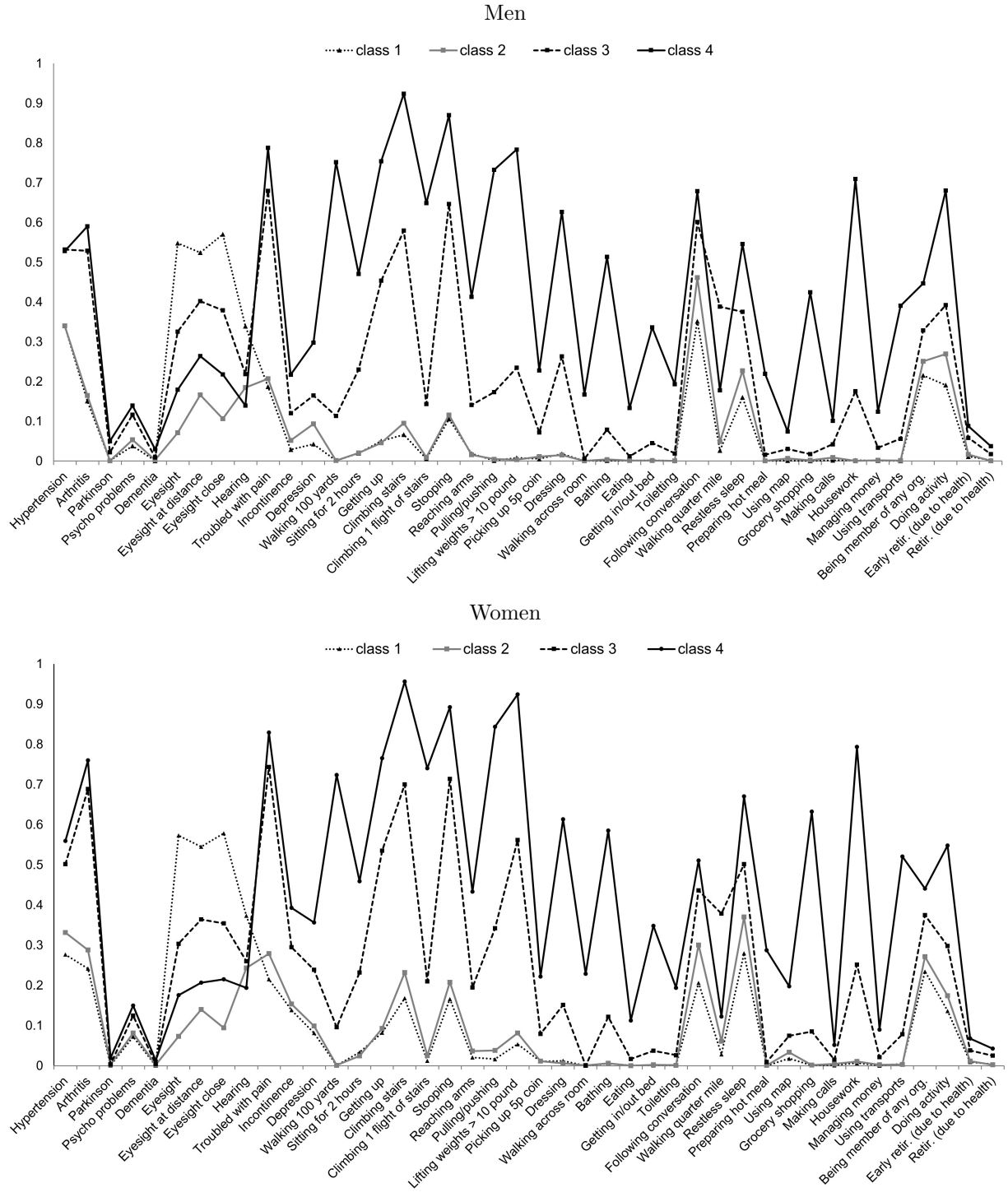
Age	2003/2007				2008/2011			
	Obs. deaths	Exp. deaths	SMR	95% CI	Obs. deaths	Exp. deaths	SMR	95% CI
50-64	55	61	0.9	(0.69,1.17)	26	29.7	0.91	(0.62,1.33)
65-69	35	48.5	0.72	(0.52,1)	37	34.7	1.04	(0.75,1.44)
70-74	54	74.7	0.72	(0.55,0.94)	45	53.1	0.85	(0.63,1.13)
75-79	99	111	0.89	(0.73,1.09)	83	78.8	1.05	(0.85,1.31)
80-84	113	151.4	0.75	(0.62,0.9)	110	110.3	1.01	(0.84,1.21)
85+	191	274.6	0.7	(0.6,0.8)	276	277.8	0.99	(0.88,1.11)
Total [a]	547	721.2	0.76	(0.73,0.86)	577	584.3	0.99	(0.91,1.07)

Men

Age	2003/2007				2008/2011			
	Obs. deaths	Exp. deaths	SMR	95% CI	Obs. deaths	Exp. deaths	SMR	95% CI
50-64	83	80.4	1.03	(0.83,1.28)	44	37.2	1.18	(0.88,1.59)
65-69	59	71.6	0.82	(0.64,1.06)	46	45.8	1.03	(0.77,1.37)
70-74	99	102.7	0.96	(0.79,1.17)	54	71.7	0.75	(0.58,0.98)
75-79	120	130.2	0.92	(0.77,1.1)	102	93.5	1.11	(0.92,1.35)
80-84	133	147.8	0.9	(0.76,1.07)	110	110.5	0.97	(0.8,1.17)
85+	143	176.4	0.81	(0.69,0.95)	171	171.6	1	(0.86,1.16)
Total [a]	637	709.1	0.9	(0.83,0.97)	527	530.2	0.99	(0.91,1.08)

[a] Calculated without weighting the age-specific SMRs
CI=confidence interval
Results are those presented at Chapter Four (table 4.1), and are re-proposed here to make Research Paper IV a standalone article.

Figure 7.3: Probability of each disability item estimated in the 4-class model, by gender



Figures are those presented in Research Paper III. They are included here to facilitate the reading of the Research Paper IV, conceived as a standalone article.

Table 7.9: Disability classes in cross-sectional and longitudinal samples at wave 2 to 5

Men

Disability level	Wave 2		Wave 3				Wave 4				Wave 5			
	Cross-sect=Long		Cross-sectional		Longitudinal		Cross-sectional		Longitudinal		Cross-sectional		Longitudinal	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
no disabled	1437	45	1383	43.5	1161	42.2	1664	46.4	1066	45	1,712	44.6	1191	43.32
low disabled	857	26.8	874	27.5	753	27.4	892	24.9	572	24.2	944	24.6	665	25.8
mildly disabled	623	19.5	600	18.9	555	20.2	707	19.7	496	21	779	20.3	509	19.7
severely disabled	279	8.7	325	10.2	282	10.3	320	8.9	234	9.9	407	10.6	290	11.2
Total	3,196	100	3,182	100	2,751	100	3,583	100	2,368	100	3,842	100	2,583	100

Women

Disability level	Wave 2		Wave 3				Wave 4				Wave 5			
	Cross-sect=Long		Cross-sectional		Longitudinal		Cross-sectional		Longitudinal		Cross-sectional		Longitudinal	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
no disabled	1467	37.7	1369	35.7	1152	34.8	1765	40	1113	37.7	1,686	40.1	1050	36.6
low disabled	976	25.1	1042	27.2	901	27.2	1062	24.1	748	25.3	926	22	658	22.9
mildly disabled	1012	26	977	25.5	849	25.6	1113	25.2	757	25.6	1059	25.2	771	26.9
severely disabled	440	11.3	450	11.7	411	12.4	472	10.7	335	11.3	532	12.7	389	13.6
Total	3,895	100	3,838	100	3,313	100	4,412	100	2,953	100	4,203	100	2,868	100

Table 7.10: Disability-related life expectancy measures for absolute and proportional changes, longitudinal sample

Men

Age	Severe-disability LE						Moderate-disability LE						Mild-disability LE					
	DLE		DLE	DFLE/TLE		%DFLE	DLE		DLE	DFLE/TLE		%DFLE	DLE		DLE	DFLE/TLE		%DFLE
	2002	2012		2002	2012		2002	2012		2002	2012		2002	2012		2002	2012	
60-64	2.5	2.6	0.1	12.4	11.2	-1.2	4.8	5.4	0.6	23.5	23.6	0.1	4.9	5.5	0.6	24	23.8	-0.2
65-69	2.1	2.4	0.3	12.9	12.9	0	4.2	4.9	0.7	25.7	25.8	0.1	4	4.4	0.4	24.1	23.3	-0.8
70-74	1.9	2.3	0.4	14.5	15.1	0.6	3.6	4.3	0.7	28	28.3	0.3	3.1	3.3	0.2	23.5	21.8	-1.7
75-79	1.8	2	0.2	18.5	17.1	-1.4	2.9	3.6	0.7	28.6	30.7	2.1	2.2	2.7	0.5	21.9	22.4	0.5
80+	1.8	1.8	0	23.6	20.4	-3.2	2.3	3.1	0.8	30.2	34.4	4.2	1.6	1.7	0.1	21.5	18.4	-3.1

Women

Age	Severe-disability LE						Moderate-disability LE						Mild-disability LE					
	DLE		DLE	DFLE/TLE		%DFLE	DLE		DLE	DFLE/TLE		%DFLE	DLE		DLE	DFLE/TLE		%DFLE
	2002	2012		2002	2012		2002	2012		2002	2012		2002	2012		2002	2012	
60-64	4	3.3	-0.7	16.9	12.8	-4.1	7.8	7.8	0	33.1	30.5	-2.6	4.9	6.1	1.2	20.8	23.8	3
65-69	3.7	3.1	-0.6	19.1	14.4	-4.7	6.9	7	0.1	35.4	32.9	-2.5	3.8	5	1.2	19.6	23.3	3.7
70-74	3.4	2.8	-0.6	22	16.4	-5.6	5.9	6.1	0.2	38.3	35.7	-2.6	2.8	3.8	1	18.1	22.3	4.2
75-79	3	2.6	-0.4	25.1	19.7	-5.4	4.7	5.3	0.6	39	39.8	0.8	2.2	2.7	0.5	18.2	20	1.8
80+	2.7	2.4	-0.3	30.4	23.8	-6.6	3.5	4.4	0.9	38.7	44.2	5.5	1.6	1.7	0.1	17.3	17	-0.3

Sensitivity analysis: removing respondents in institution at wave 6

At wave six, 9,169 core members were interviewed. Of these, 72 (0.79%) were living in a residential care home or similar establishment and had an institutional interview. The sample considered for the longitudinal analysis included 7,507 respondents and only five of them were in institution at the time of the interview, two men and three women. Given such a small proportion of respondents having an institutional interview at wave 6, the ELSA sample including institutionalised individuals is almost coincident with the sample in which these respondents are excluded. Tables 7.11 and 7.12 reports DFLE in 2012 for each severity level based on the disability prevalence in the samples including and excluding institutionalised individuals. As predicted, the expected years lived without any level of disability were almost the same whether institutionalised were, or were not, included. When observed, differences were negligible and life expectancies without severe forms of disability were higher when disability prevalence was based on the sample excluding institutionalised population.

Table 7.11: Life expectancy without disability in 2012 by age in cross-sectional samples, including and excluding institutionalised respondents at wave 6. Men

Age	Severe DFLE (95% CI)		Moderate DFLE (95% CI)		Mild DFLE (95% CI)	
	w institutionalised	w/o institutionalised	w institutionalised	w/o institutionalised	w institutionalised	w/o institutionalised
50	28.4 (28.1; 28.8)	28.6 (28.2;28.9)	25.3 (24.9; 25.8)	25.3 (24.8;25.7)	24.0 (23.4; 24.5)	24.0 (23.5;24.5)
55	24.2 (23.8; 24.5)	24.3 (24;24.6)	21.2 (20.8; 21.7)	21.2 (20.7;21.6)	20.9 (20.5; 21.3)	21.0 (20.6;21.4)
60	20.1 (19.8; 20.5)	20.3 (20;20.6)	17.4 (17; 17.9)	17.4 (17;17.8)	17.7 (17.3; 18.1)	17.7 (17.3;18.1)
65	16.3 (16; 16.6)	16.5 (16.1;16.8)	14.0 (13.6; 14.4)	13.9 (13.5;14.3)	14.7 (14.3; 15.1)	14.7 (14.4;15.1)
70	12.7 (12.4; 13)	12.9 (12.6;13.2)	10.9 (10.5; 11.3)	10.8 (10.4;11.2)	12 (11.7; 12.4)	12.1 (11.7;12.4)
75	9.6 (9.3; 10)	9.9 (9.5;10.2)	8.3 (7.8; 8.7)	8.1 (7.7;8.6)	9.3 (9; 9.7)	9.4 (9;9.7)
80	7.0 (6.6; 7.4)	7.2 (6.9;7.6)	6.0 (5.5; 6.4)	5.9 (5.4;6.3)	7.4 (7.1; 7.8)	7.5 (7.1;7.8)

w=with; w/o=without

Table 7.12: Life expectancy without disability in 2012 by age in cross-sectional samples, including and excluding institutionalised respondents at wave 6. Women

Age	Severe DFLE (95% CI)		Moderate DFLE (95% CI)		Mild DFLE (95% CI)	
	w institutionalised	w/o institutionalised	w institutionalised	w/o institutionalised	w institutionalised	w/o institutionalised
50	30.3 (29.9: 30.7)	30.5 (30.1;30.9)	25.5 (25: 26)	25.4 (24.9;25.9)	26.4 (25.9: 26.8)	26.3 (25.8;26.8)
55	25.9 (25.5: 26.3)	26.1 (25.7;26.5)	21.5 (21: 21.9)	21.4 (20.9;21.9)	23.0 (22.6: 23.4)	22.9 (22.5;23.4)
60	21.8 (21.4: 22.2)	22.0 (21.7;22.4)	17.6 (17.2: 18.1)	17.6 (17.1;18)	19.8 (19.4: 20.2)	19.8 (19.4;20.1)
65	17.8 (17.4: 18.1)	18.0 (17.6;18.3)	14.1 (13.7: 14.5)	14.0 (13.6;14.5)	16.6 (16.2: 17)	16.6 (16.2;16.9)
70	14.0 (13.6: 14.3)	14.2 (13.8;14.6)	10.9 (10.4: 11.3)	10.8 (10.4;11.2)	13.6 (13.3: 14)	13.6 (13.2;13.9)
75	10.5 (10.1: 10.8)	10.7 (10.3;11.1)	8.0 (7.6: 8.4)	7.9 (7.5;8.4)	10.9 (10.6: 11.2)	10.9 (10.5;11.2)
80	7.3 (6.9: 7.8)	7.6 (7.2;8)	5.7 (5.2: 6.2)	5.6 (5.2;6.1)	8.5 (8.2: 8.8)	8.4 (8.1;8.8)

w=with; w/o=without

Sensitivity analysis: removing health conditions from disability measures

As discussed in Research Paper IV, the ICF measure of disability used in this work also included some health conditions, which are known to affect men and women differently. Previous sensitivity analysis [42] has shown that gender differences in disability were not led by these conditions. However, when assessing changes over time, discrepancies between men and women in the diagnosis of specific health conditions may have taken place and this might result in gender differences in the disability measure driven only, or largely, by this specific component of disability. To address this potential problem, disability was re-estimated both at wave one and six removing health conditions (i.e. hypertension, arthritis, dementia, Parkinson, psychological problems and depression) and changes in DFLE were assessed in terms of the prevalence of this new measure of disability. Results are shown in tables 7.13 and 7.14 for men and women respectively. Estimates of DFLE in 2002 and 2012, as well as their absolute and proportional changes, were similar to those presented in tables 7.3 and 7.4 for the cross-sectional sample, where health conditions were included to measure disability prevalence. This suggests that the influence of these variables in the measurement of disability was only modest. Therefore, the changes in DFLE as presented in

the paper were not due to differences in diagnosis of specific health conditions between 2002 and 2012 that affect men and women differently, but actual diverging patterns of disability and mortality across the two sexes.

Table 7.13: Years of TLE and life expectancy without disability removing health conditions from disability measures. Men

Age	TLE		Severe DFLE (95% CI)		Moderate DFLE (95% CI)		Mild DFLE (95% CI)	
	2002	2012	2002	2012	2002	2012	2002	2012
50	28.9	31.7	25.9 (25.6;26.2)	28.6 (28.2;28.9)	23.3 (23;23.7)	25.3 (24.9;25.8)	21.6 (21.2;21.9)	23.9 (23.4;24.4)
55	24.5	27.2	21.7 (21.5;22)	24.3 (24;24.6)	19.3 (19;19.6)	21.2 (20.8;21.7)	18.4 (18.1;18.8)	20.9 (20.5;21.3)
60	20.3	23.0	17.8 (17.6;18.1)	20.3 (19.9;20.6)	15.6 (15.2;15.9)	17.4 (17;17.8)	15.4 (15.1;15.8)	17.6 (17.2;18)
65	16.5	19.0	14.3 (14.1;14.6)	16.4 (16.1;16.7)	12.3 (12;12.6)	13.9 (13.5;14.3)	12.5 (12.2;12.8)	14.7 (14.3;15)
70	13.0	15.2	11.1 (10.8;11.3)	12.8 (12.5;13.1)	9.4 (9.1;9.7)	10.8 (10.4;11.2)	10.0 (9.7;10.3)	12.0 (11.7;12.4)
75	10.0	11.9	8.1 (7.9;8.4)	9.8 (9.4;10.1)	7.1 (6.8;7.5)	8.2 (7.7;8.6)	7.8 (7.6;8.1)	9.4 (9;9.7)
80	7.7	9.0	5.8 (5.5;6.2)	7.1 (6.7;7.5)	5.4 (5;5.7)	5.9 (5.4;6.3)	6.0 (5.7;6.3)	7.4 (7.1;7.8)

Table 7.14: Years of TLE and life expectancy without disability removing health conditions from disability measures. Women

Age	TLE		Severe DFLE (95% CI)		Moderate DFLE (95% CI)		Mild DFLE (95% CI)	
	2002	2012	2002	2012	2002	2012	2002	2012
50	32.5	34.6	28.3 (27.9;28.6)	30.3 (30;30.7)	23.6 (23.2;24)	25.5 (25;26)	24.6 (24.2;25)	26.3 (25.8;26.8)
55	28.0	30.0	23.9 (23.6;24.2)	25.9 (25.5;26.3)	19.8 (19.4;20.2)	21.5 (21.1;22)	21.5 (21.2;21.9)	22.9 (22.5;23.3)
60	23.6	25.5	19.8 (19.5;20.1)	21.8 (21.5;22.2)	16.2 (15.8;16.6)	17.7 (17.3;18.2)	18.4 (18;18.7)	19.7 (19.3;20.1)
65	19.4	21.2	15.9 (15.5;16.2)	17.8 (17.4;18.2)	12.9 (12.5;13.3)	14.2 (13.7;14.6)	15.4 (15.1;15.7)	16.5 (16.2;16.9)
70	15.5	17.1	12.2 (11.9;12.5)	14.0 (13.6;14.4)	9.9 (9.5;10.2)	10.9 (10.5;11.4)	12.5 (12.2;12.8)	13.6 (13.2;13.9)
75	12.0	13.4	9.1 (8.7;9.4)	10.5 (10.1;10.9)	7.5 (7.1;7.9)	8.0 (7.6;8.5)	9.7 (9.4;10)	10.9 (10.5;11.2)
80	9.0	10.0	6.4 (6;6.7)	7.4 (7;7.8)	5.6 (5.3;6)	5.7 (5.3;6.2)	7.4 (7.1;7.7)	8.4 (8.1;8.8)

Sensitivity analysis: incorporation of attrition weights for wave 6

As a robustness check, the analyses of Research Paper IV were replicated incorporating attrition weights for wave 6. In particular, I estimated weighted prevalence of disability at waves one and six and recalculated DFLE accordingly; and then the linear model in which the dierence between age-adjusted YLD residuals at wave 6 and wave 2 was regressed on year of birth, BMI and their interaction was replicated using attrition weights. Tables 7.15 and 7.16 show new estimates of DFLE based on weighted prevalence of disability, corresponding to the unweighted estimates presented in tables 7.3 and 7.4 of Research Paper IV for the cross-sectional sample. Weighted regression coefficients corresponding to those presented in table 7.7 of Research Paper IV are shown below in table 7.17.

Table 7.15: Life expectancy without disability by age and gender for the cross-sectional sample, using attrition weights for wave 6 to estimate disability prevalence. Men

Age	Severe-disability LE (95% CI)		Moderate-disability LE (95% CI)		Mild-disability LE (95% CI)	
	2002	2012 cross-sectional	2002	2012 cross-sectional	2002	2012 cross-sectional
50	25.8 (25.5; 26.1)	28.2 (27.8; 28.5)	23.1 (22.7; 23.4)	25.2 (24.8; 25.7)	21.6 (21.2; 22)	24 (23.4; 24.5)
55	21.6 (21.4; 21.9)	23.9 (23.6; 24.2)	19.1 (18.8; 19.4)	21.2 (20.7; 21.5)	18.5 (18.1; 18.8)	21 (20.6; 21.4)
60	17.7 (17.5; 18)	19.9 (19.5; 20.2)	15.5 (15.1; 15.8)	17.3 (16.9; 17.7)	15.5 (15.1; 15.8)	17.8 (17.4; 18.1)
65	14.3 (14.1; 14.6)	16 (15.7; 16.4)	12.2 (11.9; 12.5)	13.9 (13.5; 14.3)	12.5 (12.2; 12.8)	14.8 (14.4; 15.2)
70	11.1 (10.8; 11.4)	12.5 (12.1; 12.8)	9.3 (9; 9.7)	10.8 (10.4; 11.2)	9.9 (9.6; 10.2)	12.1 (11.8; 12.5)
75	8.2 (7.9; 8.4)	9.5 (9.1; 9.8)	7.2 (6.8; 7.5)	8.2 (7.8; 8.6)	7.8 (7.5; 8.1)	9.4 (9.1; 9.8)
80	5.8 (5.5; 6.2)	6.8 (6.4; 7.2)	5.4 (5; 5.7)	6 (5.5; 6.4)	6 (5.7; 6.3)	7.5 (7.1; 7.8)

Table 7.16: Life expectancy without disability by age and gender for the cross-sectional sample, using attrition weights for wave 6 to estimate disability prevalence. Women

Age	Severe-disability LE (95% CI)		Moderate-disability LE (95% CI)		Mild-disability LE (95% CI)	
	2002	2012 cross-sectional	2002	2012 cross-sectional	2002	2012 cross-sectional
50	27.9	30.2	23	25.3	25	26.4
	(27.6; 28.3)	(29.8; 30.5)	(22.6; 23.4)	(24.8; 25.8)	(24.6; 25.4)	(25.9; 26.9)
55	23.6	25.7	19.3	21.3	21.8	23
	(23.3; 23.9)	(25.3; 26.1)	(18.9; 19.7)	(20.9; 21.8)	(21.5; 22.2)	(22.6: 23.4)
60	19.5	21.6	15.8	17.5	18.7	19.8
	(19.2; 19.9)	(21.3; 22)	(15.4; 16.2)	(17.1; 18)	(18.3; 19)	(19.5; 20.2)
65	15.7	17.6	12.6	14	15.6	16.7
	(15.4; 16)	(17.2; 18)	(12.2; 12.9)	(13.6; 14.4)	(15.3; 15.9)	(16.3; 17)
70	12.1	13.8	9.6	10.8	12.7	13.7
	(11.8; 12.4)	(13.4; 14.2)	(9.2; 10)	(10.4; 11.2)	(12.4; 13)	(13.3: 14)
75	9	10.3	7.3	7.9	9.8	11
	(8.7; 9.3)	(9.9; 10.7)	(7; 7.7)	(7.5; 8.4)	(9.5; 10.1)	(10.6; 11.3)
80	6.3	7.2	5.5	5.6	7.5	8.6
	(6; 6.7)	(6.8; 7.6)	(5.2; 5.9)	(5.2; 6.1)	(7.2; 7.7)	(8.3; 8.9)

Table 7.17: Estimated coecients from linear regression models of the dierence between age-adjusted YLD residuals at wave 6 and wave 2 ($\hat{e}_6 - \hat{e}_2$) (in years), using attrition weights for wave 6.

	1		2		3		4	
	Men	Women	Men	Women	Men	Women	Men	Women
Year of birth	0.0287**	0.0188			0.0280**	0.0204	0.0282**	0.0232*
	-0.0129	-0.0122			-0.0133	-0.0126	-0.0133	-0.0126
bmi			0.0194	-0.0126	0.0162	-0.0139	0.0103	-0.0143
			-0.0257	-0.0206	-0.0257	-0.0206	-0.0259	-0.0206
bmi*yob							0.00673**	0.00729***
							-0.00343	-0.00246
Constant	1,708	2,172	1,556	1,997	1,556	1,997	1,556	1,997

7.4 Conclusions

In Research Paper IV, I join to the debate on population health changes and provide new evidence for the English setting over the past decade. The main contribution of the study consists in (i) considering multiple grades of disability; (ii) focusing on both proportional and absolute estimates of DFLE; (iii) exploring possible explanations for the observed changes in DFLE, modelling a corresponding outcome measurable at the individual level, YLD. To the best of my knowledge, this is the first time where health and mortality are combined in an individual level outcome and interactions of year of birth with a predictor (BMI) are estimated.

Results indicate a dynamic equilibrium of disability for women and expansion for men, although variations were very small. Men experienced a worse pattern than women, with increase in both mild and moderate disability, and severe disability roughly stayed stable. Among women only mild disability has increased and severe disability decreased in proportional terms. Evidence for an interactive role of year of birth and BMI on changes in YLD was also found, whereby high BMI was particularly detrimental for younger generations.

The results of this study indicated how distinguishing mild, moderate and severe forms of disability allows to capture specific patterns otherwise masked by averaging different trends. A recommendation deriving from the findings of Research Paper IV is that it is informative to distinguish milder and more severe levels of disability from moderate levels, because their trends seem to be divergent. The other recommendation is to always consider proportional changes in DFLE along with absolute changes, to fully understand the burden of disability on society and public health care. Finally, the evidence of interacting effects of BMI and year of birth on YLD can be taken as a warning for closely monitoring BMI in younger generations.

Research Paper IV has been presented at the international conference "Giornate di Studio sulla Popolazione 2017 - Population Days conference" in Florence in February 2017 and at the "International Health Policy Conference" hosted by the London School of Economics in London in February 2017. It has been accepted for oral presentation at the international conference of the Population Association of America which will take place in April 2017.

The slide of the most updated presentation are included in the appendix (figure 9.4).

Part III

Discussion

Chapter 8

Discussion

The relevance of population ageing to the development of health policies has motivated the direction taken by this thesis, as well as the urgency for understanding current and future trends in health inequalities. Within this framework, I focused on features and recent trends in survival rates in the English population aged 50 years and older and examined whether there was evidence of gender differences in the impact of disability on mortality. A systematic literature review of studies analysing socioeconomic and demographic inequalities in health expectancy among the older population has motivated and guided the direction of this work, shaping the main research questions and hence the thesis objectives. The work consists of two main bodies of research and is organised in three empirical research papers and the abovementioned systematic review. The first part investigated the association of disability with mortality, as empirically observed in England between 2002 and 2012, while addressing the question of whether it differs between men and women. The second part combined disability and mortality into a summary measure of population health and studied the observed changes in this indicator over the last decade, again assessing whether there is empirical evidence of differences between genders.

This chapter draws together the most important results from each contributing research paper. Strengths and limitations of the methods and data used are then reviewed, with avenues for future research discussed in the last section along with implications of these findings for public health provision.

8.1 Summary of findings of the thesis

This section synthesises the key findings of the four papers that constitute this thesis. These findings are discussed more extensively within each paper, but are linked up here to gain a general overview. They are presented according to the research objectives described in Chapter One and corresponding research questions.

8.1.1 Objective 1

- *What is known in literature about demographic and socioeconomic inequalities in health expectancy? What is unknown and what is controversial?*

To answer these questions, I performed a systematic literature review, focusing on the older population. I found evidence of inequalities in health expectancy associated with several factors, including gender, education, health-related behaviours, and race, in different countries including both high and low-middle income settings. All studies included in the review confirmed the existence of a gender paradox in health and mortality, consisting in the finding that women live longer than men but spend larger proportions of their life in poor health and with disability. On the other hand, it emerged that there was a lack of standardization in the definition of health and disability, leading to measures of Healthy Life Expectancy (HLE) and Disability-Free Life Expectancy (DFLE) based on inconsistent indicators, which limits and prevents comparability of results across studies. Moreover, where evidence of inequalities -of any sort- was found, hypotheses were suggested but without empirical support.

8.1.2 Objective 2

- *Is the WHO's International Classification of Functioning Disability and Health (ICF) a valid framework to capture disability among the older English population?*

The choice of adopting the WHO's ICF framework to measure disability was based on the fact that it is currently the predominant framework for disability [63, 84, 107]. The need of communicating research and speaking in a common language to depict the concept of disability has urged to explore the validity of the ICF as a theoretical framework for

accomplishing this goal. Although the ICF is quite good in differentiating concepts and categories within its framework -but some concerns about the lack of a clear operational differentiation between domains have been expressed [84, 95]-, it is not possible to rule out entirely the arbitrariness implied in the measurement of disability based on this model. This included the choice of the variables needed to capture each domain, their classification across impairment, activity limitations and participation restrictions and the approach to the modelling of disability.

The analyses performed in Research Paper II tested the ICF validity via a measurement model which confirmed the applicability of the WHO's ICF framework to capture disability among the older population in England. The ICF terminology had already been integrated specifically in a dataset targeted on older populations [108], but to the best of my knowledge, a construct validation was never attempted before for a specific older population. The applicability of the ICF framework to older populations and the fact that the three ICF domains could be detected using the questions asked in the English Longitudinal Study of Ageing (ELSA) represent the principal finding of research objective 2.

Some complementary and more specific aspects arose from the results produced in Research Paper II. Fitting the disability model allowed the assessment of which items were more important to capture each specific domain of disability and general disability. It emerged that items pertaining to "activity limitations", such as Activity of Daily Living (ADL) or mobility functions, were all highly correlated with the activity specific factor and the disability general factor. This suggests that these observed items are particularly informative to represent not only "activity limitations", but also general disability. ADLs and mobility functions are in fact the variables most commonly used to measure disability in the literature [79, 80, 81]. The systematic review indicated that of the 96 studies estimating health expectancy included, 42 used ADLs as indicator of health or disability. Interpreting disability only in terms of activity and mobility limitations, however, implies adopting a specific approach to its conceptualization, different from the ICF's idea that disability is an overlay concept that refers to the whole disablement process. Measuring disability using these variables, therefore, must be an aware decision based on the agreement upon a precise theoretical model. Nevertheless, often surveys do not offer a set of disability-related items

as broad as in ELSA, while ADLs are more commonly available. Therefore, although the measurement of disability cannot purely depend on availability of items, it is important to acknowledge the high relevance of ADLs and mobility functions in defining the disability factor's dimensionality found in this thesis.

8.1.3 Objective 3

- *How did disability measured at baseline affect mortality over the course of ten years?*

Along with objective 5, objective 3 is at the core of the general aim of this thesis.

The objective was to estimate the short and long term effect of disability, as measured at the beginning of the study. While it is commonly accepted that disability affects mortality, very little is known about this association over time (i.e. calendar time), and even less whether this is characterised by gender differences. Surprisingly, the body of studies investigating the relationship of disability with mortality is modest when it comes to older populations. This has been examined in a few papers [109, 110], very few of which set in the UK [111], and none using a multi-dimensional measure of disability.

In Research Paper II, I found evidence of a strong relationship between disability at baseline (2002/2003) and mortality observed over the course of a decade. Adjusting for demographic, socioeconomic and behavioural factors attenuated the association, but overall the effects remained strong and statistically significant.

- *Does the association of disability with mortality differ between men and women?*

The second question within research objective 3 relates to gender differences in the relationship of disability and mortality. From the Discrete-Time Survival Analysis (DTSA) performed in Research Paper II, it emerged that the significance of the association of disability with mortality held for men as well as for women, but disability at wave 1 was found to have a time-varying effect among men, and a smaller and constant effect for women. In particular, the effect among men was very strong in the short term and converged to the level observed in women with increasing survival time.

These results confirmed the existence of the gender paradox in health and mortality and helped to further understand it by showing that women spend a higher proportion of

their life in disability because they survive longer with disability, suggesting that higher prevalence among women may be a function of longer survivorship with disability rather than higher incidence of disability.

The sensitivity analyses performed in Research Paper II also helped to explain why the gender paradox may occur, and suggested that the observed differences in the association between disability and mortality in men and women were not driven only by gender-specific health conditions and body structures. This is in accordance with previous research that maintains that gender differences in mortality are not only dependent on purely biological factors, but also on psycho-social and behavioural aspects [112, 113]. This opens doors to further investigation of different mechanisms that underlie the association between disability and mortality and make women survive with disability better than men. I discuss it further in section 8.4, proposing possible mechanisms to be investigated in future research.

8.1.4 Objective 4

- *How can we optimally classify disability by severity level? Is a binary classification informative enough or would a more refined measure be preferable for health evaluation?*

In Research Paper III, I addressed these questions by applying sophisticated statistical methods of analysis to identify a categorization of disability that was best supported by the data, and then assessing it against its association with alternative specifications of disability, as well as with various health outcomes. This offered an innovative approach in at least two ways: (i) to the best of my knowledge, this was the first study attempting to identify categories of disability based on empirical evidence, rather than upon a priori theoretical classifications, for the English population; (ii) the usefulness and informativeness of multi-categorical disability, compared to its binary classification, was assessed by relating the alternative specifications of disability levels with health status and mortality.

The results indicated that the best classification of disability consisted of four classes, according to two of the adopted criteria (i.e. the estimated probabilities of disability items conditional on class membership and external validations based on the association of disability classes with health and mortality). Disability classes were labelled "non-disability"; "mild disability", which presented characteristics more similar to the "non-disability" group

than the other disability levels, and differed from the "non-disability" class especially in the impairment domain; "moderate disability", which presented intermediate level of limitations in the activity domain compared to the other groups, and was more similar to the "mild disability" group in terms of participation restrictions and closer to the "severe disability" group in the impairment domain; and "severe disability", which presented the highest limitations in all domains. Results also showed how binary categorizations of disability -compared to each of the more refined scales of severity- masked the very strong effect of severe disability because of their implicit averaging the burden of disability, and thus were not informative on the impact of low levels of disability.

Identifying adequate grades of severity of disability is therefore fundamental in order both to capture the impact on health and mortality of low disability, and equally to assess the strong burden of severe disability, which are relevant for the successful planning and implementation of health policies.

8.1.5 Objective 5

- *How has DFLE changed from 2002 to 2012 in England?*

As stated before, objective 5 specifies one of the core questions of this thesis. By assessing whether DFLE has changed over the past decade, both in absolute and relative terms, I intended to contribute to the debate on compression and expansion of disability. Overall, I found that the expected years of life without any level of disability have increased, for both men and women. However, the increase in number of years free of disability alone did not permit to establish whether disability had compressed or not. When the finding was interpreted in combination with results of changes in life expectancy, and assessed in absolute and proportional terms, the conclusions of my study were different for men and women, and depended on the level of disability. An exhaustive description of these findings is contained in the results and discussion sessions of Research Paper IV, here I report the most relevant results.

- *Were changes in DFLE different between men and women? Were they different depending on the level of disability?*

Briefly, women in England, overall, experienced greater improvements in DFLE than men, with reduced expected years of life with severe and moderate disability, and increasing expected years of life with mild disability, both in absolute and proportional terms. Men experienced nearly the opposite during the same period. Specifically, their life expectancy with moderate disability has increased, while years with severe disability remained stable, indicating a dynamic equilibrium of their disability over the past decade. Trends appeared to be quite stable across age groups for both genders. Only the older group (aged 75+ years) seemed to have experienced worse changes: for women aged 75 years and older life expectancy with moderate disability increased both in absolute term and especially as proportion of life expectancy; for men, the direction of changes was the same at all ages, but the older old experienced a larger reduction in the proportion of life expectancy without moderate disability compared to younger age groups.

8.1.6 Objective 6

- *Are there any factors likely to affect differently younger and older cohorts that may explain changes in DFLE observed over a decade?*

The research question corresponding to objective 6 was addressed by focusing on the impact of overweight and obesity (measured by Body Mass Index (BMI)) on Years Lost to Disability (YLD). It was an exploratory part of the analyses performed in Research Paper IV, which, as I will highlight later in this chapter requires further investigation, both to study in more depth the effect of BMI on mortality and disability across older cohorts, and to consider other behavioural or social factors likely to have changed in recent times, with their change affecting younger and older generations to different extents.

BMI was specifically chosen as a risk factor of interest for a combination of two key reasons: (1) overweight and obesity are known to be associated with higher risk of becoming and remaining disabled, but have little or no effect on life expectancy [90, 91, 92, 93]; (2) the prevalence of overweight or obesity varies across cohorts, and it has been found to be greater among younger cohorts [94]. When studying changes in DFLE at a certain age in different time periods, periods and cohorts are implicitly compared, and therefore the role of a factor that affects cohorts differently and is known to be associated with disability is

of particular interest. For this reason, the analyses adopted to address objective 6 focused on an individual level outcomes (YLD), replacing DFLE, which is an aggregate measure, in order to explore the role of individual BMI at baseline (which in this case was wave 2 because data was collected during the nurse visit). The main finding showed an interaction between BMI and year of birth, which pointed to the fact that BMI was found to be detrimental to disability, especially for younger individuals.

8.2 Main contributions of the thesis

This dissertation contributes to the understanding of the relationship between disability and mortality and how survival with and without disability has changed over the past decade in England among individuals aged 50 years and older. This work also improves the understanding of gender inequalities in disability and mortality.

Specific contributions were made to (1) research devoted to the conceptualization, measurement and classification of disability; (2) the debate on the gender paradox in health and mortality; (3) the debate on compression, expansion and dynamic equilibrium of disability and mortality.

8.2.1 Conceptualization, measurement and classification of disability

What was previously known:

In the last half century, several schools of thought have defined disability and related concepts. Among the most influential conceptual schemes there are the "Disablement Model" proposed in 1965 by Nagi [52], the WHO's International Classification of Impairments Disabilities and Handicaps (ICIDH) issued in 1980 [53], and the "Disablement Process" proposed in 1994 by Verbrugge and Jette [51]. These perspectives differ in their emphases and applications, but share the same focus on the process of disablement and expand on the medical model of disease, where only the characteristics of the individual -and not the environment- are salient to disability and its prediction [114]. Along with the literature on the conceptualization of disability, reviews aiming at drawing the link between different models and measurement methods have been produced [115], and studies of disability measurement and survey methodology issues figure prominently in the literature [61, 116].

A great deal of methodological research has been conducted to tackle challenges related to measuring disability. These include, for example, problems with the wording of the questions, time period references for which disability has to be reported, administration of surveys (see for example Freedman [117], Jette [118]), as recently presented in the Handbook of Aging and the Social Sciences. In chapter 4 of this handbook [61], the challenge of coding survey data for use in studies of disability, and therefore of choosing between binary or more refined measures is also acknowledged. The problems of categorizing disability and the context-dependence of the classification have received remarkably less attention compared to the problems of conceptualizing and measuring disability. Nevertheless, the relevance of identifying meaningful classes of disability is clear. In the context of understanding the dynamics of successful (i.e. healthy) ageing, the importance of considering the severity of conditions has been already acknowledged [20, 119], and recent studies [40, 76] have confirmed that trends in health expectancy differs according to the severity of one's disability. However, while the importance of distinguishing multiple grades of disability to understand the real burden of disability is becoming more established, there are very few reports of optimal classification, or related guidelines on how to perform the classification.

What this thesis adds:

The contribution of this thesis to this strand of research consists in:

1. Promoting the use of the ICF, which offers a comprehensive interpretation of disability encompassing both biological and social aspects of disability, for assessing disability in the older population. Joining the debate among those supporting [63, 84, 107] or critical [95] of the ICF as an appropriate framework to capture disability in the older population, I produced evidence to justify and enforce its use as a common language to delineate the concept of disability and implement its measurement. First, I showed via a systematic review (Research Paper I) the lack of agreement over the measurement of disability to estimate DFLE and remarked on the need for using a common measure across studies. Then, in Research Paper II, I tested and confirmed the construct validity of the WHO's ICF framework of disability for individuals aged 50 years and older living in England in the 2000s, using a set of variables available in the ELSA (but also generally available in other similar surveys) and I offered evidence that those

variables were suitable. The measurement model that I used provided a continuous measure of disability, in line with Wolf [61]'s statement that *"Continuous measures of disability severity, possibly based on weighted counts or factor loading scores, would have obvious utility in applied research, but few such disability scales appear to exist."*. This continuous score was subsequently used as an independent variable to explore the association of disability with mortality (Research Paper II), and then to derive a binary indicator that was compared with the national statistics of disabled and non-disabled population (Research Paper III).

2. The other contribution was probably also the most innovative, and consisted in identifying a severity gradient of disability distinguishing four classes (non-disability, mild, moderate and severe disability), based on empirical evidence rather than a-priori classification (Research Paper III). The advantages of distinguishing milder and more severe classes of disability were then evidenced in Research Paper IV, where DFLE was estimated for each severity level over a decade, and divergent patterns of changes were observed depending on the severity of disability and between genders. Such differences would have been ignored if considering only disabled and non-disabled individuals in terms of a binary classification.

8.2.2 Explanation of gender inequalities in disability and mortality

What was previously known:

The gender paradox in health and mortality has been well documented since the 1970s [67, 68], and has been mentioned repeatedly in this dissertation. The debate around why it occurs is lively, and multiple causes have been considered to explain the mechanisms which drive its occurrence. Poposed explanations include biological, social and psychological interpretations [120]. Methodological issues such as selection and information bias descending from gender differences in participation in surveys or reporting health problems have been considered, but the contribution of such biases to the gender paradox is considered to be small [121]. Gender inequalities in socioeconomic position and social support would explain the gender paradox in that women, especially at older ages, tend to have a lower education and socio-economic position and this may affect their health [122, 123, 124, 125]. Differences

in health-related behaviours and lifestyle have also been explored [126], as well as gender differences in disease profile, because men are known to suffer more than women from fatal conditions, and women are more likely to be affected by non-lethal and disabling diseases [47]. So far, however, there has not been a prevailing theory to justify why women live longer than men but spend larger proportion of their life in poor health and with disability.

What this thesis adds:

This work contributes empirical evidence on the possible mechanisms leading men to live shorter lives but in better health compared to women, by looking at gender differences in the association between disability and mortality. A similar idea lied behind the work by Van Oyen et al. [127] where an aggregate-level analysis was performed to evaluate gender difference in Healthy Life Years (HLY) in Europe by estimating the contribution of women's mortality advantage versus women's disability disadvantage. The results reported in this thesis showed that women live longer with disability because disability has a smaller impact on their mortality than on males' mortality. The second contribution was to explore, through a number of sensitivity analyses, the potential confounding role of physical conditions, which are known to differ between men and women. Physical conditions (such as hypertension, arthritis, dementia, Parkinson, psychological problems and depression) did not seem to play a major role in confounding this association, and sensitivity analyses showed that gender did not affect the measurement of disability. The finding that in men, disability was strongly associated to mortality in the short term and then that the association converged to the level observed among women may be combined with the finding of expansion of moderate and severe disability observed among men, as reported in Research Paper IV. Given their stronger short-term effect of disability, men who survived disability may become more resilient the longer they survive, and therefore their life expectancy with disability would increase (i.e. expand) relatively more when compared to women.

8.2.3 New evidence for dynamics of population health changes in England over the past decade

What was previously known:

There is agreement in the literature with regards to the direction of population changes in England and the UK overall when one considers trends in life expectancy, but less so with regards to health expectancy. Between 1960 and 1990 the expansion of morbidity theory was dominant in most countries, including the UK [38]. However, when levels of disability were distinguished the dynamic equilibrium theory seemed prevailing [119, 128]. Recent estimates of DFLE for the UK in the first decade of the 21^{st} century, based on prevalence of persistent activity limiting illness (with no grading of disability identified) provided by the Office for National Statistics (ONS) [129], suggested that compression of morbidity took place in the UK. Data for England alone, however, looked less optimistic, and, life expectancy, in particular at age 65 years, was found to grow more than DFLE, both for English men and English women, with an absolute and proportional expansion of disability. Another study, set in England that distinguished less and more severe disability [40], pointed at decline in severe forms of disability and increase in milder levels, especially among women. Evidence supporting the theory of expansion, for the period going from 1990 to 2013, came from the work of Newton et al. [130], that extracted data from the Global Burden of Disease Study 2013 to compare mortality, causes of death (i.e. Years of Life Lost (YLL)), years lived with a disability (i.e. YLD), and Disability-Adjusted Life Years (DALY) in England, and found that levels of YLDs have declined much less (1.4% (0.1-2.8)) than YLLs (41.1% (38.3-43.6)).

To summarize, it seems that in the last quarter of the 20^{th} century expansion of disability prevailed in England, and over the past two decades this was observed especially at older ages (i.e. 65 years and older). When severity levels were differentiated, however, the dynamic equilibrium appeared to take precedence. It also appeared that results differed depending on how health and disability were measured and across age groups; gender-specific patterns were also unclear. These the context and gaps from which the research questions of Research Paper IV have been framed.

What this thesis adds:

The study presented in Research Paper IV joins an existing body of literature assessing trends in DFLE and sheds new light on the changes occurred over the past decade in England. The main contribution of this paper lies not within the novelty of its theme, but its corroboration of previous evidence. The innovative aspects of the analysis were to (i) interpret disability according to the ICF framework, and therefore adopting a comprehensive approach, including impairments, activity limitations and participation restrictions, rather than focusing on specific domains separately; (ii) distinguish four severity levels of disability (including non-disability) to better understand changes in DFLE; (iii) complement the analysis on DFLE, measuring also an individual level outcome, YLD. We confirmed previous findings of compression of severe forms of disability and expansion of milder levels for women (hence overall dynamic equilibrium of disability). Men, differently from women, experienced an increase in moderate disability and severe disability stayed constant rather than decrease, and mild disability slightly expanded. Such a difference between males and females would have not been captured distinguishing only less severe and more severe levels of disability. The consolidation of these trends is key to depict correctly the pattern of disability and put in place adequate resources and efforts to improve conditions of ageing. Finally, and from a different angle, an additional contribution of this work was to stress the importance of considering both proportional and absolute estimates of change in DFLE in parallel, as already pointed ouy by Robine et al. [73], and interpret results accordingly. The combination provides a clearer understanding of the dynamics of ageing, and also quantifies the actual burden of disability on society and health care system, which cannot be directly derived from relative measures alone.

8.3 Evaluation of data and methods

In this section, I discuss some challenges posed by the nature of the data and methods used in this thesis and outline the limitations descending from them. Some issues are common to all research using longitudinal data extracted from survey studies; others are more specific and technical aspects of the analyses performed in this thesis.

8.3.1 Replicability

"The replication of scientific findings using independent investigators, methods, data, equipment, and protocols has long been, and will continue to be, the standard by which scientific claims are evaluated. However, in many fields of study there are examples of scientific investigations that cannot be fully replicated because of a lack of time or resources. In such a situation, there is a need for a minimum standard that can fill the void between full replication and nothing. One candidate for this minimum standard is "reproducible research", which requires that data sets and computer code be made available to others for verifying published results and conducting alternative analyses." [131, p. 405]

This extract from an editorial of the journal *Biostatistics* highlights a critical issue which affects in particular two of the findings reported in this dissertation: the measurement of disability and the classification of the severity of disability levels.

As already stated, one of the contributions of this thesis was to promote the use of a broad measure of disability and overcome the lack of agreement and inconsistency of definitions, using the comprehensive interpretation proposed by the WHO. Disability was measured accordingly, using a wide set of variables combined in a measurement model, and the ICF was found an appropriate framework to capture disability among the older population. To make this finding useful and enforce the use of the WHO's ICF, the proposed measurement of disability needs to be *replicable* also in other settings. The rigour of the protocol adopted to select the variables and model them via a general-specific model, would certainly allow other investigators to *reproduce* the results shown in this thesis (see materials in Chapter Five and appendix B.1). What remains unexplored and needs to be assessed is the *validity* of findings in other settings and with different data. A possible first step in this direction would be to run the disability model developed in Research Paper II using the ELSA's sister studies, such as the Health Retirement Study (HRS), the Survey of Health, Ageing and Retirement (SHARE), and others, which offer a substantially similar set of variables and are targeted on the same type of population of ELSA (adults aged 50+) but in other countries.

Concerning the classification of severity levels of disability, the optimal number of classes was identified adopting a data-driven approach, as remarked in Research Paper III.

This overcame the criticality of using arbitrary a priori theoretical classifications -which is the approach most commonly used- but brought along the problem of replicability, raising the issue that the optimal level that was found could apply only to the specific study in which the analysis was carried out. The fact that the validity of the finding pertained to a certain setting and specific population limits the generalizability of my results. However, this limitation did not affect the rest of this thesis, because all the analyses referred to the same sample. Nevertheless, the relevance of the topic of identifying adequate severity levels of disability calls for further exploration of the general validity of my findings by replicating the analysis in other settings and populations, and comparing the conclusions across studies.

8.3.2 Generalizability

Another key aspect pertains to the representativeness of the general English population of the sample(s) used in the analyses presented in this dissertation. ELSA is meant to be representative of the non-institutionalised English population, but if respondents move into a residential care home or similar establishments during the course of the study, an "institution interview" is sought and they still remain in the study sample. Therefore, over time, ELSA's representativeness has changed slightly. In Chapter Four, a descriptive investigation of ELSA's representativeness, in particular with regards to mortality and disability, was offered. It emerged that, in terms of mortality, increasingly during the follow up, the ELSA sample became closer to the general English population aged 50+. And this has partly been explained by the fact that the representativeness of the study over time has slightly changed.

In Research Paper II, all variables were measured at wave 1 and therefore they pertained to the baseline sample of the ELSA, which, as said, consisted in individuals living in private households. Mortality information (i.e. the outcome variable), however, was available for all respondents who gave their consent to link their data to the relevant national registry and this included also those who moved to a residential care home during the follow-up. Therefore, results pertained to the effect of disability, measured among the non-institutionalised population, on mortality, whether it occurred when respondents were

still living in a private household or in residential care.

In Research Paper IV, the use of both longitudinal and cross-sectional respondents allowed us to see that the estimates of disability prevalence and DFLE were overall similar across the two groups of respondents, and to some extent, this indicated that attrition bias did not affect the measure of disability. In this paper, differently from Research Paper II, information on mortality came from the ONS and covered the total population, while disability prevalence referred to ELSA's respondents, and therefore did not include institutionalised individuals. The systematic literature review, presented in Chapter Two, clearly revealed the problem of combining national data on the general population, with survey-based information often targeted on non-institutionalised populations, in order to calculate health expectancy. Some studies addressed this limitation combining different sources of data to cover the total population (see for example Crimmins et al. [132]), others included hypotheses on the health condition of institutionalised population, assuming for example that they were either as healthy as non-institutionalised individuals or were all impaired (see for example Jagger et al. [133]). In this study, no assumptions were made, and the target population was English people aged 50 years and older at baseline and living in private household. From the results presented in Chapter Four, where the ELSA and English population mortality were compared, we know that the ELSA mortality converged to that of the general population over time. Making use of this finding, I suggest that the limitation of combining ELSA information, targeted on individuals living in private households, with mortality rates for the total population affected the measures of DFLE differently depending on the time period considered, i.e. wave 1 (2002) or wave 6 (2012), and between cross-sectional and longitudinal samples. In particular, the ELSA mortality rates were closer to those of the total English population for deaths occurred in 2012. Therefore the DFLE calculated at wave 6 (especially for the longitudinal sample) was based on mortality rates likely to be closer to those of the English population compared to wave 1.

A study by Cambois et al. [89] investigated in depth the impact of accounting or not for institutionalised population and warned that for populations where higher proportions of individuals live in communal establishments, such as older populations, the hypothesis that data from surveys on populations living in private households is representative for

the institutionalised populations is particularly critical. A commonly applied option was proposed by Sullivan [72] and consists in adjusting estimates of disability by assuming that the entire population of health-related institutions have disability. This implicitly requires knowing the distribution of institutionalised population by type of communal establishment. Cambois et al. [89] tested a number of hypotheses for including institutionalised individuals in disability prevalence estimates, among which the Sullivan's hypothesis, to assess the magnitude of the bias. It was found that the method suggested by Sullivan was most effective, but less so for advanced ages. For advanced age groups, the overestimation resulting from the Sullivan's hypothesis was greater than the underestimation descending from the assumption of same prevalence in institutionalised and household populations. These conclusions provide additional evidence of the complexity of the problem, and suggest that the limitations affecting Research Paper IV may be limited, but do not guarantee that the approach adopted was the most suitable.

8.3.3 Self-Report of disability

The presence or absence of disability was defined using self-reported measures, such that each disability domain (i.e. body functions and structure, activity limitations and participation restrictions) was captured by self-assessed variables asked in the Computer-Assisted Personal Interview (CAPI) questionnaires. Self-reported measures, in theory, may not merely reflect the objective status of disability, but also to some extent the subjective experience of health and disability. This may have introduced some bias to my analyses and driven, at least partly, the observed gender paradox, although sex differences in reporting health problems are considered to be small [121]. Various studies have shown that self-report of disability correlates sufficiently well with objective measures of disability, suggesting that the bias arising from self-reports may be limited [134, 135, 136]. Conversely, it has been remarked that for domains of functioning, such as sensory, cognitive, psychological and mental health functioning, a potential problem is posed by the fact that the presence of a deficit may itself undermine or hinder the attempt to measure it, and this may be less pronounced for the other dimension of disability, such as activities [61]. All functioning domains are "intrinsic" to the person, with the added disadvantage that their definition

disregards the interaction between person and environment widely understood to characterise disability, as presently conceptualised. This would imply that the self-reporting bias may affect the three domains of disability to different extents.

In Research Paper II, observer-measured health indicators (i.e. inflammation, blood clotting, cholesterol and a measure of respiratory functioning) were introduced into the analyses of the disability-mortality association, not to validate the disability measurement, but to control for their potential confounding effect. The results indicated that controlling for observer-measured health indicators slightly decreased the effect of disability among women, while for men adjusting or not adjusting for biomarkers did not change the results. Because biomarkers are informative of health conditions -which are not part of disability according to the ICF model- and adjusting for them did not change considerably the results, this suggested that the measure of disability that was used did not mask health domains. In other words, these domains would be external to the ICF-defined disability.

8.3.4 Causality

The aim of this thesis was to investigate gender inequalities in healthy ageing, and was achieved through a number of objectives. Most of the research objectives implied examining the relationship between an exposure and outcome and drawing conclusions about its strength. In this setting, ideally, the intention is to find a causal association, so that it is possible to identify modiable causes of the adverse event, and put in place effective interventions. There are several analytical methods that can help to draw reliable causal conclusions in observational studies, albeit under some unverifiable assumptions [137]. In this thesis, I used traditional regression methods with multivariable adjustment to estimate the association between my exposures and outcomes. Using this approach, however, the risk of unmeasured confounders is almost unavoidable, resulting in biased results, if such confounding is substantial.

In this section, I acknowledge this limitation and advocate for a careful interpretation of the results before drawing any causal conclusions from the observed effect. At the same time, I explain and justify the rationale behind the choice of not adopting alternative estimation methods to investigate these putative causal relationships. In the following, I

will focus on discussing the association between disability and mortality studied in Research Paper II.

In Research Paper II, I studied whether disability measured at baseline was associated with mortality and whether the effect was different for men and women. Given that disability is not assigned at random, the different relationship found in men and women may be due to unmeasured confounders that have sex-specific relationships with disability and with mortality. To control for these unmeasured confounders, one could use fixed effect modelling or instrumental variable methods. However, in this study, a fixed effect model was not applicable because disability was measured only at baseline. Similarly, the adoption of instrumental variable methods was problematic because of the difficulty of finding a convincing instrument, in particular one that satisfies the exclusion restriction; for example genes, that are sometimes used as instruments, are likely to affect mortality through other pathways not involving disability. Therefore, the nature of the analysis and the type of data used led me to choose traditional regression based multivariable adjustment to assess whether disability at baseline had an impact on mortality observed over ten years, where the adjustment involved a rich set of variables.

These variables, known to be associated with both mortality and disability, included demographic characteristics, socioeconomic status, health-related behaviours and illness, and indicators of social class from childhood. Furthermore, in sensitivity analyses, objective measures of health were introduced as additional confounders to assess whether gender differences in the association between disability and mortality were driven by gender differences in prevalence of specific disabling diseases.

Unmeasured confounders might still bias the association under study, and I acknowledge this limitation, but I believe the most relevant confounders were taken into account, therefore limiting the extent of this potential bias.

8.4 Implications and future research

Overall the work presented in this dissertation produced evidence that potentially has important implications for public health interventions and for future research priorities.

8.4.1 Implications for policies

- *Need for conceiving health expectancy as a comparable and repeatable measure in order to be used as an informative instrument for policy makers.*

The heterogeneity of measures of health and disability to estimate health expectancy has been highlighted in the systematic review presented in Chapter Two and commented in previous sections of this chapter. In this section, the need for a common indicator is expanded upon with the purpose of advocating action towards a harmonization of the measure of health expectancy, in particular DFLE, which would enable cross-study comparisons. The facts that the ICF is currently the most agreed upon model of disability [63, 84, 107] along with the finding produced in this dissertation that it is a measurable and valid indicator of disability for the older population, can be used to advocate its adoption as a measure of disability.

Among others, a specific addressee of this recommendation would be the Réseau Espérance de Vie en Santé (REVES). REVES is an international organization set up in 1989 by International Institutes, and promotes the use of health expectancy as a population health indicator. One of the activities of REVES is expressly *to propose a conceptual framework to organise the domain and classify all the health expectancy estimations.* The finding of this dissertation, confirming the validity of the ICF as a measure of disability among the older population, offers ground for promoting the use of this conceptual framework as a reference for measuring disability and calculating DFLE accordingly.

- *Need to identify and distinguish milder and more severe levels of disability.*

The main finding, reported in Research Paper III, and whose relevance was shown in Research Paper IV, consisted of the identification of a multi-categorical measure of disability that distinguished four classes, from non-disability to severe disability. The main conclusion and consequent recommendation is that milder and more severe levels of disability should always be identified when quantifying their impact on health and mortality, because these relationships changed across different levels of disability (Research Paper III), and trends in DFLE over the the first decade of the 21^{st} century in England also appeared to be different depending on the level of disability (Research Paper IV). This finding reinforces previous

evidence on the implications of differentiating levels of disability to interpret trends in DFLE [74, 75, 128]. Furthermore, a binary classification risks to average the effect of milder and more severe forms of disability together with those of moderate disability, masking the separation of the upward and downward trends seen for different levels of disability. A note of caution is deemed however with regards to these results. Their validity only applies to the context in which the analysis was performed, and thus the recommendation for a multi-categorical measure of disability does not necessarily point at four as the preferred number of categories. In Research Paper IV, where DFLE was estimated for three levels of disability, by identifying an intermediate level of disability it was possible to appreciate, for example, gender differences to a great extent.

- *Need to consider both proportional and absolute changes in DFLE to fully understand the burden of disability.*

In Research Paper IV, we empirically tested changes in DFLE over a decade in England in light of the theories of population health change [16, 17, 18, 20], and showed that absolute and relative changes were most commonly concordant but, in some cases, differed. In particular, when we found proportional compression of disability (for any level), we also observed that the expected number of years with disability expanded in absolute terms, with the exception of severe disability for women, which compressed both proportionally and absolutely. Number of years with disability can increase even if DFLE in relation to life expectancy compresses. Absolute changes in DFLE provide information on how long people with disability will survive and will need some level of assistance and care, and therefore inform on whether the overall burden of disability on health systems and families will increase, and to what extent. Hence, the consequent message is that proportional changes must be considered along with absolute changes. This consideration, however, is not new and joins previous recommendations already exposed in the 1990s [73], but that have not been always taken into consideration.

8.4.2 Implications for future research

- *Replication studies*

As reported in the previous paragraph, one of the limitations affecting some of the findings

of this thesis is the fact that their validity has been tested only in the ELSA. Given the relevance of the topic and findings -namely, the validity of the ICF to capture disability using a broad battery of items in a general-specific model, and the multi-categorical classification of disability composed by four classes- replicating the analyses in other contexts would be extremely informative. The availability of survey studies that are similar to ELSA and set in different countries (see Chapter Four for the list of studies) represents a great opportunity to test the validity of the results of this thesis for other settings. If confirmed, a common framework and measurement model to estimate disability among the older population in many countries would become available.

- *Mechanisms explaining the gender paradox in health and mortality*

One of the contributions of the thesis, as highlighted above, was to provide additional evidence to the gender paradox and understand why it occurs. Sensitivity analyses presented in Research Paper II suggested a minor role of physical conditions in explaining the smaller effect of disability on mortality observed among women compared to men, and therefore their longer expected life with disability did not seem to depend much on gender differences in diseases. This suggests further explorations of factors not related to physical conditions that affect disability and mortality, and that are likely to be different in men and women. Possible mechanisms that can be of particular interest to explain why women have larger proportions of disability than men but lower mortality rates, include health seeking behaviours and social networks. Women and men are known to have different attitudes towards these factors. Males have consistently been shown to have less positive attitudes toward seeking health services than females [138, 139, 140], and women to have a wider range of sources of emotional support and more close persons in their primary networks than men [141]. Moreover, studies have found these risk factors to affect mortality and health to different extent -and sometimes in opposite directions- between genders [138, 142]. Considering these findings, when assessing gender differences in the association between disability and mortality, can add further insight to understand why disability is more fatal for men than women. In another jargon, this would mean to include into the disability-mortality pathway factors known to be associated with both disability and mortality, and acting differently in men and women. For example, the fact that men tend to use health services less than

women would lead them to later diagnosis of diseases and disabilities, and therefore to underreport milder levels of disability and consequently to underestimate their survival to low levels of disability. A late treatment would as well increase the probability of dying due to disability. Therefore, evidence supporting the existence of any of these mechanisms would reveal a concrete avenue for intervention to promote higher gender equality in healthy ageing, ideally decreasing female disability prevalence and lowering male mortality rates.

- *Cohort differences in BMI affecting changes in DFLE*

Some risk factors are likely to explain changes in disability and mortality to different extents across different cohorts. This is of particular interest when comparing over time summary measures of population health, which implicitly compare cohorts in different periods relying on different assumptions depending on the methods applied. In this thesis, I looked at overweight and obesity as one of these factors and found a modifying effect with year of birth, suggesting that high BMI was more detrimental among younger generations. Nevertheless, it was not possible to disentangle the period and cohort effects represented by year of birth [104], with additional limitations also affecting the analysis (see Research Paper IV in Chapter Seven). The epidemics of obesity deserves greater exploration than offered in the concluding part of Research Paper IV, considering, for example, the timing of onset of overweight and obesity as an explanatory factor when performing gender and cohort comparisons. Cohort studies, such as the 1946 National Survey of Health and Development Cohort National Survey of Health and Development Cohort (NSHD), the 1958 National Child Development Study (NCDS), the 1970 British Cohort Study (BCS) and the Millennium Cohort Study (MCS) are excellent sources to pursue this type of investigation. The focus of this thesis was on the older population, however, with only NSHD and NCDS being of a comparative age range. Nevertheless, given that the onset of overweight and obesity keeps shifting to younger ages and that it affects new generations more than older ones [94], extending these analyses to younger samples would be informative, and would highlight whether interventions to delay the onset of overweight and obesity would be beneficial in terms of the relationship between disability and mortality.

- *Disentangling disability domains*

"The concept that disability should not be a specific point on the pathway but rather

an overlying concept that refers to the whole process has value, although it creates some confusion because of the very long usage of disability to reflect more specific issues.(..) In reading the ICF "bible", the study of the disablement process is really never clearly mentioned and in fact the much publicized figure of the model (figure 1 in Jette commentary) seeks to show that all aspects of disability are related to all other aspects, with all arrows going in both directions. One can only assume from this that a decision was made to not impose the concept of a pathway at all on the model and in fact the terms "disablement" or "disablement process" are not used." [95, p. 1169-1170]

The editorial of Guralnik and Ferrucci [95] published in 2009 in the *Journal of Gerontology* questioned the validity of the ICF framework to study disability in the older population, criticising in particular the ICF failure to differentiate concepts and categories between the activity and participation domains. Concerns about the lack of operational differentiation between these two spheres were expressed already by the Institute of Medicine [84] and other researchers [143, 144] and by Jette himself [63] in the editorial where he promoted the ICF as common language for disability. The essence of the discussion, however, underlies a more complex and substantive debate on the conceptualization of disability. Since disability has been interpreted not only as a medical, but also as a psychological and social condition, two different interpretations have emerged, which are not necessarily incompatible: one intends disability as a comprehensive concept composed by different and equally contributing domains, and the other as a hierarchical pathway consisting of separate subsequent domains. The ICF would correspond to the first view and frames disability as a single broad measure including multiple domains equally contributing to the overlaying concept. The Nagi's "Disablement model" would belong to the second approach and conceives disability through a pathway, such that diseases predicts impairments which lead to functional limitations which finally conduce to disability.

In this thesis, I shared the first view adopting the ICF's framework, and thus giving to disability an inclusive interpretation encompassing both biological and social aspects. When disability develops over time, the framework proposed by Nagi presents the advantage of allowing to distinguish and identify groups suffering from specific domains (functional limitations, impairments, etc.) and plan ad hoc interventions. A comprehensive approach,

on the other hand, would study trajectories from a level to another -with the intrinsic challenge of identifying severity levels. It was beyond the scope of this thesis to develop both the approaches and gain insight from the comparison of results obtained disentangling and not disentangling disability domains. If there is a preferable approach and which this is, is still an open debate among researchers who study disability. The relevance of the issue and the potential implications for policy planning encourage the development of this work towards adopting the hierarchical approach to disability and hence disentangling its domains and study their evolution and whether it differs in men and women.

8.5 Discussion in a nutshell

Table 8.1: Discussion in a nutshell

n	Objectives	Research Paper	Main findings	Implications
1	To systematically review research that investigates socioeconomic and demographic inequalities in health and mortality using summary measures of population health.	Research Paper I	- Existence of inequalities in health expectancy associated with several factors, such as gender (see the "gender paradox in health and mortality" as found in all studies). - Lack of a standardised definition of health expectancy which limits cross-study comparisons.	Need for conceiving health expectancy as a comparable and repeatable measure to make it an informative instrument for policy makers.
2	To theoretically conceptualise and empirically measure disability.	Research Paper II	Confirmation of the validity of the ICF framework to capture disability among the older population.	- Promote the use of the ICF framework to capture disability among the older populations. - Need for replication studies.
3	To explore gender inequalities in the effect of disability measured at baseline on mortality observed over a decade.	Research Paper II	- Strong association of disability with mortality, with a time-varying effect for men, and smaller constant effect for women. - Confirmation of the existence of the gender paradox showing that women survived longer to disability than men, although if men survived the first years the effect of disability converged to women's level. - Limited role of health conditions to explain the gender paradox.	- Gender paradox verified where women have differential survival rates by disability. - Need to explore other mechanisms, apart from health conditions, such as social networks and relations and health-seeking behaviours, to explain the gender paradox.
4	To identify an optimal number of severity levels of disability based on empirical evidence.	Research Paper III	- Preferable class model had 4 classes. - Binary indicators of disability averaged the burden of disability and masked the very strong effect experienced by individuals having severe disability, and were not informative on the effect of low levels of disability.	Need to identify and distinguish milder and more severe levels of disability.
5	To combine each level of disability with mortality in DFLE, and use it to estimate how expectancies of healthy life have changed over a decade across the two genders and across different levels of disability.	Research Paper IV	- Changes in DFLE observed in the last decade differed across gender and disability levels. - Severe and moderate disability compressed for women and stayed constant for men, while mild disability expanded for women and for men only in absolute terms.	Need to consider both proportional and absolute changes in DFLE to fully understand the burden of disability.
6	To explore possible demographic and behavioural factors explaining changes in DFLE observed over a decade.	Research Paper IV	Evidence of an effect modification of year of birth and BMI in changes in years lost due to disability (YLD), such that high BMI resulted particularly detrimental to younger cohorts.	- Warning for closely monitoring BMI in young generations. - Need to expand the analyses by capturing data on time of onset of overweight and obesity.

Chapter 9

PhD portfolio

Presentations:

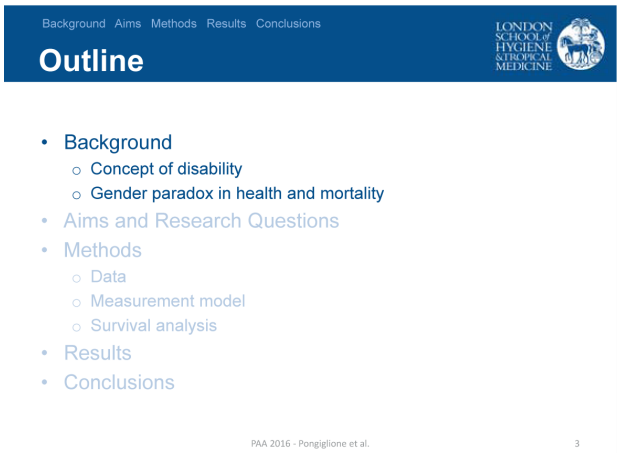
- **Research Paper II** "Disability and all-cause mortality in the older population: evidence from the English Longitudinal Study of Ageing"
  - Oral presentation *Population Association of America*, Washington. **April 2016**.
  - Oral presentation *59th Annual Scientific Meeting of the Society for Social Medicine*, University College Dublin, Ireland. **April 2016**.
  - Poster presentation **Best poster award**, *Research Degree Students' Poster Day*, London School of Hygiene and Tropical Medicine. **February 2016**.
  - Oral presentation *Economic and Social Research Council DTC Conference*, London. **May 2015**
- **Research Paper III** "Levels of disability in elderly people in England: comparing binary and ordinal classifications"
  - Oral presentation *Medical Research Council Unit for Lifelong Health and Ageing*, University College of London, London. **June 2016**.
  - Oral presentation *Lunch Seminar of the Medical Statistics Department*, London School of Hygiene and Tropical Medicine, London. **May 2016**.
- **Research Paper IV** "Disability-free life expectancy between 2002 and 2012 in England: trends differ across genders and levels of disability"

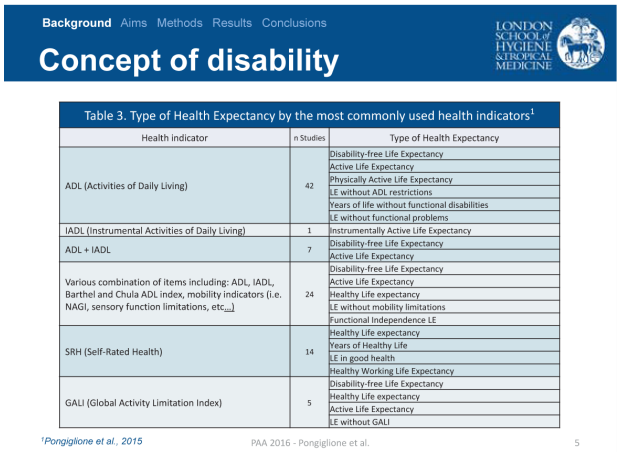
- Oral presentation *Population Association of America*, Chicago. **April 2017**.
- Oral presentation *International Health Policy Conference*, London School of Economics, London. **February 2017**.
- Oral presentation *Giornate di Studio sulla Popolazione 2017 - Population Days conference"*, Florence. **February 2017**.

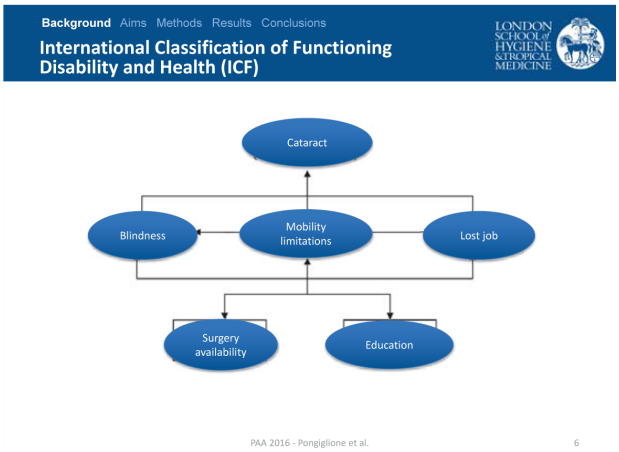
Publications:

- **Research Paper I**
  - Pongiglione B, De Stavola BL, Ploubidis GB. A Systematic Literature Review of Studies Analyzing Inequalities in Health Expectancy among the Older Population. PLoS One. 2015 Jun 26;10(6):e0130747.
- **Research Paper II**
  - Pongiglione B, De Stavola BL, Kuper H, Ploubidis GB. Disability and all-cause mortality in the older population: evidence from the English Longitudinal Study of Ageing. Eur J Epidemiol. 2016 Aug;31(8):735-46.
- **Research Paper III**
  - Pongiglione, B, Ploubidis GB, and De Stavola BL. Levels of disability in the older population of England: Comparing binary and ordinal classifications." Disability and Health Journal (2017).

Figure 9.1: Slides of Research Paper II







Background Aims Methods Results Conclusions

Gender paradox in health and mortality

- Women live longer than men, but spend a larger proportion of their life with disability/poor health.
- Some proposed explanations:
  - **Biological factors**: gender differences in disease profile: women may have higher prevalence of nonfatal but disabling diseases and men have higher prevalence of fatal diseases and chronic diseases strongly related to mortality.
  - **Socio-behavioural factors:** gender differences in risk behaviour patterns (smoking, alcohol, etc.); gender inequalities in socioeconomic status and social support.
  - **Reporting bias:** gender differences in self-reporting: i.e. women are more aware than men about their health conditions.

Background Aims Methods Results Conclusions

Outline

- Background
  - Concept of disability
  - Gender paradox in health and mortality
- Aims and Research Questions
- Methods
  - Data
  - Measurement model
  - Survival analysis
- Results
- Conclusions

Background **Aims** Methods Results Conclusions

Aims

The aim of this work is to explain the gender paradox in health and mortality by analysing whether the association of disability with mortality differs between women and men.

Background **Aims** Methods Results Conclusions

Research Questions

- Is the WHO's ICF a valid framework to capture disability among the older UK population?
- How disability measured at baseline affects mortality over the course of ten years?
- Does the association of disability with mortality differ between men and women?

Background **Aims** Methods Results Conclusions

Research Questions

- Is the WHO's ICF a valid framework to capture disability among the older UK population?
- How disability measured at baseline affects mortality over the course of ten years?
- Does the association of disability with mortality differ between men and women?

Background **Aims** Methods Results Conclusions

Research Questions

- Is the WHO's ICF a valid framework to capture disability among the older UK population?
- How disability measured at baseline affects mortality over the course of ten years?
- Does the association of disability with mortality differ between men and women?

Background Aims Methods Results Conclusions

Outline

- Background
  - Concept of disability
  - Gender paradox in health and mortality
- Aims and Research Questions
- Methods
  - Data
  - Measurement model
  - Survival analysis
- Results
- Conclusions

Background Aims **Methods** Results Conclusions

Data

English Longitudinal Study of Ageing (ELSA)

- Data from a representative sample of individuals aged 50 years and over resident in the household sector in England.
- Wave 1 (2002/2003) and mortality data from national mortality register until 2011.
- Sample size: 9,715 respondents interviewed at wave 1 who had complete data on all disability items.

Background Aims **Methods** Results Conclusions

Measures

- **Deaths:**
  Time of death available **by year** for respondents who consented to link their data to mortality records of the National Health Service Central Register.
- **Disability:**
  - Impairment: 19 items including self-rated eyesight and hearing, chronic conditions such as high blood pressure and arthritis, questions about pain
  - Activity: 20 items including ADLs, mobility functions, e.g. climbing flights of stairs or walking 100 yards.
  - Participation: 11 items including IADLs, and various limitations due to health problems, such as using public transports.
- **Covariates:**
  - Demographic variables: gender, age, ethnicity, marital and parental status
  - Socioeconomic status: education, wealth, income, social class, father's occupation
  - Health related behaviours: smoking, alcohol consumption, physical activity
  - Comorbidity: chronic illness

Background Aims **Methods** Results Conclusions

Measurement Model

- 50 binary or categorical items
- 3 domain-specific factors (impairment, activity, participation)
- 1 general factor (disability)

→ Approached with a 2-parameter logistic model

Categorical/binary observed indicators (y_{ij}) are related to continuous latent variables (η_j), such that:

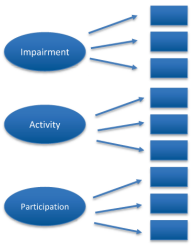
$$\Pr(y_{ij} = 1|\eta_j) = \frac{\exp(\beta_i + \lambda_i \eta_j)}{1+\exp(\beta_i + \lambda_i \eta_j)}$$

Rabe-Hesketh and Skrondal, 2008

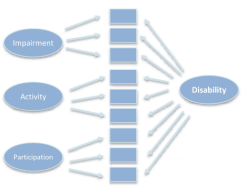
Background Aims **Methods** Results Conclusions

Measurement Model

1st step: first order model



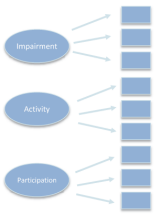
2nd step: general-specific model



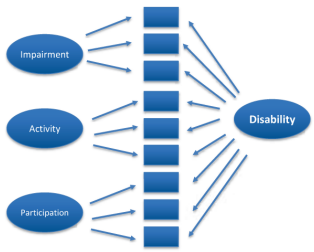
Background Aims **Methods** Results Conclusions

Measurement Model

1st step: first order model



2nd step: general-specific model



Background Aims **Methods** Results Conclusions

Survival Analysis

- **Discrete time survival analysis** in a general latent variable framework (Muthen and Masyn, 2005).
- 10 binary time-specific event indicators of death.
- All covariates are measured at baseline.
- The probability of each event is related to covariates through a logit link function. For a covariate x, its effect on the probability of event occurrence is expressed in terms of the odds ratio (OR).
- Likelihood Ratio Test (LRT) to establish if covariates have time-varying or invariant effect on mortality

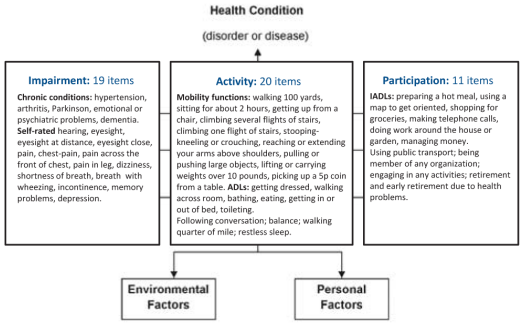
Background Aims Methods Results Conclusions

Outline

- Background
  - Concept of disability
  - Gender paradox in health and mortality
- Aims and Research Questions
- Methods
  - Data
  - Measurement model
  - Survival analysis
- Results
- Conclusions

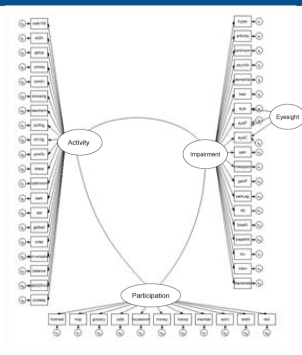
Background Aims Methods **Results** Conclusions

Selected disability items



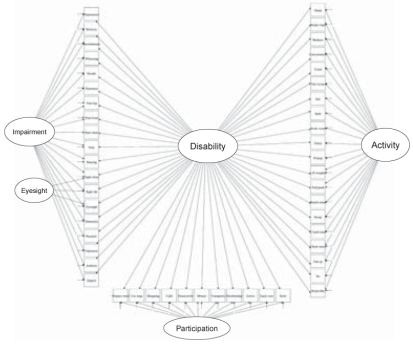
Background Aims Methods **Results** Conclusions

First order model



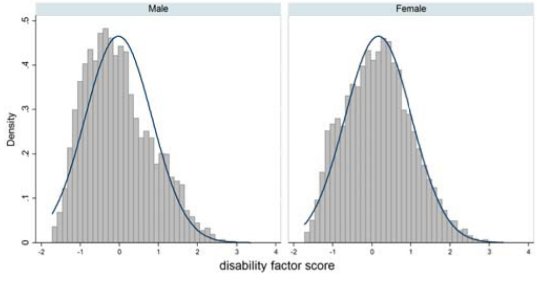
Background Aims Methods **Results** Conclusions

General-specific model



Background Aims Methods **Results** Conclusions

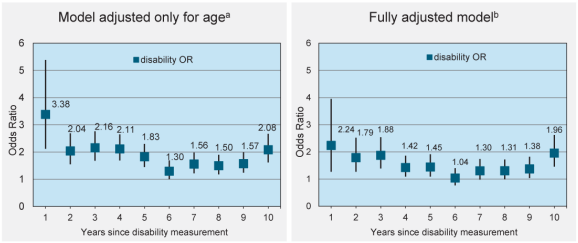
Disability factor score by gender



Background Aims Methods **Results** Conclusions

Discrete-time Survival Analysis

Men



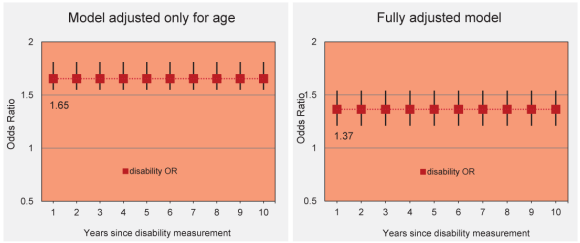
[a] test for linear trend chi2(8) = 17.54, p-value = 0.025

[b] test for linear trend chi2(8) = 17.54, p-value = 0.025

Background Aims Methods **Results** Conclusions

Discrete-Time Survival Analysis

Women



Background Aims Methods Results Conclusions

Outline

- Background
  - Concept of disability
  - Gender paradox in health and mortality
- Aims and Research Questions
- Methods
  - Data
  - Measurement model
  - Survival analysis
- Results
- Conclusions

Background Aims Methods Results **Conclusions**

Conclusion

RQ 1: Is the WHO's ICF a valid framework to capture disability among the older UK population?

- Our work confirmed the construct validity of the WHO's ICF framework to capture disability among the older population and identified the more relevant factors for each component.

Background Aims Methods Results **Conclusions**

Conclusion - cont'd

RQ 2 : How disability measured at baseline affects mortality over the course of ten years?

- Increasing odds of dying as the baseline disability score increased, with the association being stronger among men;
- Time-varying association for men: the impact of baseline disability on their mortality decreased with longer survival;
- Constant effect of disability over the ten-year period of observation for women.

Background Aims Methods Results **Conclusions**

Conclusion - cont'd

RQ 3: Does the association of disability with mortality differ between men and women?

- **Evidence of 'gender paradox'** in health and mortality: women spend a higher proportion of their life in disability because they survive longer to disability.

Background Aims Methods Results **Conclusions**

Conclusion - cont'd

- Help to explain gender paradox:
  - Is it due to gender differences in disease profile? Sensitivity analysis dropping some impairment items that describe health functions led to same results.
  - Is it due to gender inequalities in socio-behavioural factors? Controlling for SEP and health-related behaviours attenuated disability impact, but it did for both men and women.
  - Is it due to gender differences in self-reporting health measures? Sensitivity analysis using biomarkers led to similar results.

Background Aims Methods Results **Conclusions**

Future work

- Gender paradox in health mortality is not driven only by gender-specific health conditions and body structures
  → there must be some other mechanisms acting within the pathway between disability and mortality that make women cope with disability better than men.
- Exploring some of these mechanisms should be the next step to explain further the gender paradox in health and mortality.

Thank you for your attention

E-mail: benedetta.pongiglione@lshtm.ac.uk
Twitter: @benedettapongi

Figure 9.2: Poster of Research Paper II

Disability and All-Cause Mortality in the Older Population: Evidence from the English Longitudinal Study of Ageing

B. Pongiglione[1], B.L. De Stavola[1], H. Kuper[2], G.B. Ploubidis[3]

[1] Dept of Medical Statistics, London School of Hygiene and Tropical Medicine
[2] International Centre for Evidence in Disability, London School of Hygiene and Tropical Medicine
[3] Centre for Longitudinal Studies, Institute of Education

1 Background

- Concept of disability: the **International Classification of Functioning, Disability and Health (ICF)**, developed by the WHO, conceives disability as composed by three interconnected areas (see Fig.1): i) impairments that are problems in body function; ii) difficulties in executing activities; iii) participation restrictions that are problems with involvement in any area of life [1].
- Disability & Mortality: a key aspect in studying disability and mortality is related to gender differences. The **gender paradox in health and mortality** [2] consists in the finding that women live longer than men, but tend to have more disability than males.

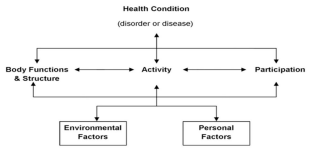


Fig. 1 Representation of the ICF. Source: World Health Organization, Geneva 2002

2 Aim & Research Questions

The aim of this work is to assess the validity of the ICF framework and explore the role of disability in explaining mortality and assess differences between men and women.

Research Questions

- Is the WHO's ICF a valid framework to capture disability among the older UK population?
- How disability measured at baseline affects mortality over the course of ten years?
- Does the association of disability with mortality differ between men and women?

3 Methods

Data & Measures

- **Data:** English Longitudinal Study of Ageing (ELSA), Wave 1 (2002/2003); mortality recorded in the national mortality register from 2002 to 2011
- **Sample:** 9,715 respondents aged 50+ who have complete data on all disability items
- **Measures:** 50 items measuring disability
- **Controls:** demographic characteristics; socioeconomic position (SEP); health-related behaviors; limiting long-lasting illness

Analysis

- **Measurement model:** general-specific model

50 binary or categorical items
3 domain-specific factors (impairment, activity, participation)
1 general factor (disability)

- **Discrete-time survival analysis**

10 binary time-specific event indicators of death
All covariates are measured at baseline
Likelihood Ratio Test (LRT) to establish if covariates have time-varying or invariant effect on mortality

4 Results

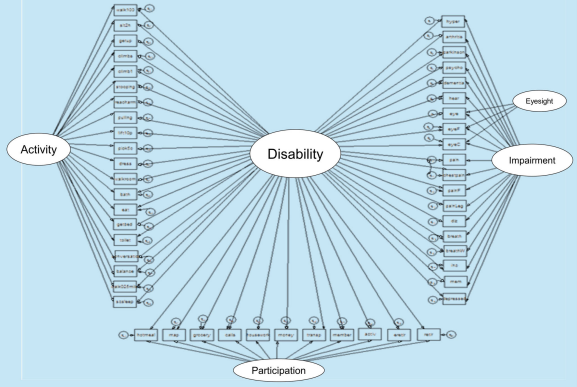


Fig. 2 General-specific measurement model

4 Results

Measurement Model

The distribution of the disability factor score, predicted by the general-specific model, is shown in Fig.3 by gender. Male distribution is slightly more right-skewed than that for females, meaning that, compared to women, fewer men had high disability score.

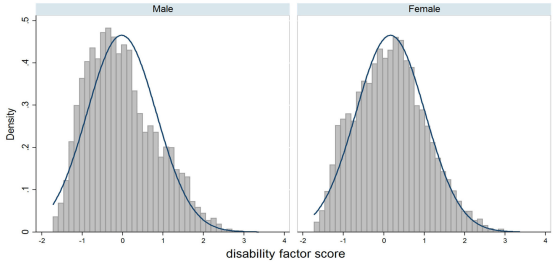


Fig. 3 Disability factor score predicted by the general-specific model, by gender

Discrete-time survival analysis

Fig. 4 shows the effect of disability on mortality by gender. Following LRT, disability measured at baseline resulted to have time-varying effect for men and time-invariant effect for women. Mortality odds ratio dramatically declines from period one to period two among men, indicating a strong effect of disability on mortality in the very short term.

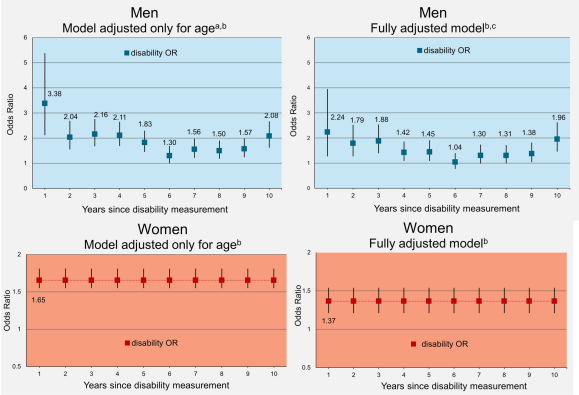


Fig. 4 Disability odds ratios for mortality

[a] test for linear trend chi2(8) = 17.54, p-value = 0.025 [b] SE estimated from pooled logistic regression
[c] test for linear trend chi2(8) = 15.96, p-value = 0.043

5 Conclusions

- The study confirms the **validity of the WHO's ICF** framework to capture disability among the older population.
- **Evidence of 'gender paradox'** in health and mortality: women spend a higher proportion of their life in disability because they survive longer to disability.
- Help to explain gender paradox:
  - **Is it due to gender differences in disease profile?** Sensitivity analysis dropping some impairment items that describe health functions led to same results.
  - **Is it due to gender inequalities in SEP and social support?** Controlling for SEP attenuated disability impact, but it did for both men and women.
  - **Is it due to gender differences in self-reporting health measures?** Sensitivity analysis using biomarkers led to similar results.

6 References

1. World Health Organization, World Bank. World report on disability. Geneva, Switzerland: World Health Organization; 2011.
2. Luy M, Minagawa Y. Gender gaps-Life expectancy and proportion of life in poor health. Health Rep. 2014 Dec 17;25(12):12-9.

Figure 9.3: Slides of Research Paper III

Beyond binary definitions of disability: how many levels are needed?

Benedetta Pongiglione

Department of Medical Statistics
London School of Hygiene and Tropical Medicine

MRC Unit for Lifelong Health and Ageing
June 22, 2016

Outline

Outline

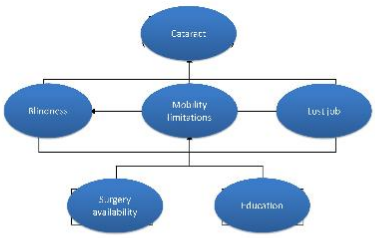
Aim

The aim of this work is to foster and advance the debate on the usefulness and relevance of adopting a finer categorization of disability, and discuss why it is important to go beyond a binary definition.

Research Questions

- How severe does a continuous measure of disability need to be for a person to be classified as disabled?
- Would a finer classification of disability better capture clustering of disability dimensions and of heterogeneities in later disease and mortality rates?.
- What is the number of disability levels that is most useful for health evaluation?

International Classification of Functioning Disability and Health (ICF)



Data

English Longitudinal Study of Ageing (ELSA)

- Data from a representative sample of individuals aged 50 years and over resident in the household sector in England.
- 6 waves issued from 2002/2003 to 2012/2013.
- Sample: respondents interviewed at wave 1 who had complete data on all disability items. 9,715 observations at baseline.

Measures

Disability:

- **Impairment:** 19 items including self-rated eyesight and hearing, chronic conditions such as high blood pressure and arthritis, questions about pain.
- **Activity:** 20 items including ADLs, mobility functions, e.g. climbing flights of stairs or walking 100 yards.
- **Participation:** 11 items including IADLs, and various limitations due to health problems, such as using public transports.

Measures

Health outcomes:

- **Mental health:** measured at every wave using eight items of the CES-D scale and treated as a continuous variable ranging from 0 to 8.
- **Grip strength:** grip strength in the dominant hand, average of three measurements done using the Smedley dynamometer. Measured during the nurse visits, which took place every two waves, i.e. wave 2, 4 and 6.
- **Coagulation:** fibrinogen, measured during the nurse visits.

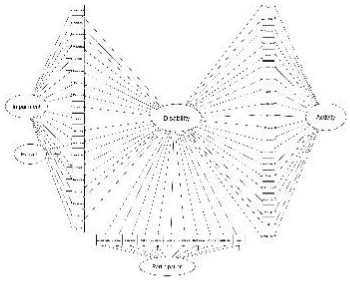
Outline

Latent variable framework

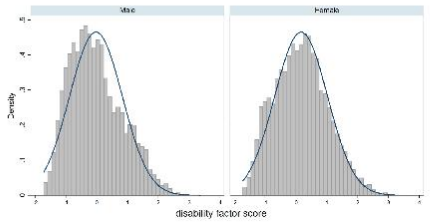
	Continuous Observed Variables	*Discrete Observed Variables*
Continuous Latent Variables	Factor Analysis	Item Response Theory
Discrete Latent Variables	Latent Profile Analysis	Latent Class Analysis

Cf. Bartholomew, Knott and Moustaki (2011)

Continuous Disability



Continuous Disability



Binary Disability

1. **WHO's approach**
   Threshold set according to *the average disability score observed among those reporting extreme difficulties in any of the domain functioning used to measure disability.*

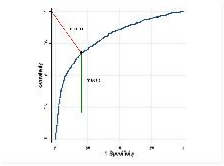
   → We replicated the WHOs approach looking at *the average disability score observed among those reporting at least one limitation in any of the six disability questions selected by the Washington Group* (WG).

Binary Disability

2. **ROC curve**
   Agreement of disability score with an external gold standard, consisting in *receiving health or disability benefits.*

   The cut-off in the continuous disability score is chosen using two alternative criteria (Kumar and Indrayan, 2011):

- the point on the curve closest to the (0, 1) (where specificity=1 and sensitivity=1), i.e. the minimal distance r.
- the Youden index that maximizes the vertical distance (j) from line of equality to the point [x, y] on the curve.



Results

1. **WHO's approach**
   62.4 percent of men and **60.7** percent of women had at least one limitation in the six WG activities.

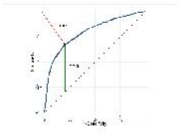
   The mean disability score among the respondents belonging to this group was **0.44** and **0.47** for men and women respectively, on a range going from -1.71 to 3.42.

   Setting the cut-off point at these values led to **31.5** percent of male respondents and **31.7** percent of female respondents having disability.

Results

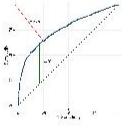
2. **ROC curve**

Men
Cut-off based on minimum r= Cut-off based on Youden index= **0.51**



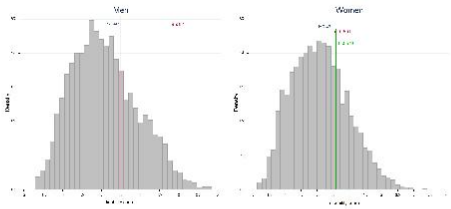
Cut-off point corresponds to **29.2** percent of disabled men.

Women
Cut-off based on minimum r=**0.5**
Cut-off based on Youden index= **0.58**



Cut-off points correspond to either **30.3** percent (Youden index) or **27.9** percent (min r) of disabled women.

Results



Results

Table: Binary Classifications

Approach	Proportion	
	Men	Women
WHO	31.5	31.7
ROC curve -min(r)	29.2	30.3
ROC curve - max(j)	29.2	27.9
ONS	32.8	32.9

Outline

Latent variable framework

Continuous Latent Variables	Factor Analysis	Item Response Theory
Discrete Latent Variables	Latent Profile Analysis	Latent Class Analysis
	Continuous Observed Variables	*Discrete Observed Variables*

Cf. Bartholomew, Knott and Moustaki (2011)

Latent variable framework

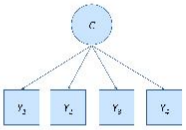
Continuous Latent Variables	Factor Analysis	Item Response Theory
Discrete Latent Variables	Latent Profile Analysis	Latent Class Analysis
	Continuous Observed Variables	*Discrete Observed Variables*

Cf. Bartholomew, Knott and Moustaki (2011)

Benedetta Pongiglione (LSHTM) MRCLHA Seminar / 42

Latent Class Analysis

We assume that the patterns of responses are due to the influence of a discrete latent variable C. The discrete latent variable C can have 1 to k classes, with k specified by us. k can be as many as the number of possible response patterns, but hopefully much fewer.



Latent Class Analysis

For binary items and a categorical latent variable C_k with K classes ($C = k; K = 2, \ldots 6$), the marginal probability of observed item u_j (with j=1,2 . . . ,50) being equal to 1 is

$$Pr(u_j = 1) = \sum_{k=1}^{K} Pr(C_k = k) Pr(u_j = 1 | C_k = k)$$

Where $Pr(u_j = 1 | C_k = k)$ denotes the conditional probability of the item being equal to 1 the class is equal to k and $Pr(C_k{=}k)$ is the probability of class k.

Outline

1. Aims and background
2. From continuous score to binary classification
3. Categorical classification
4. Criteria for selecting the number of classes
5. Discussion and future work

Criteria for selecting the number of classes

i Model fit indicators
ii Estimated probabilities
iii Gradients in associations with health outcomes and mortality

Criteria for selecting the number of classes

i Model fit indicators
- Entropy
- Bayesian Information Criterion (BIC)
- Bootstrap likelihood ratio test (BLRT) (Nylund, Asparouhov, Muthn, 2007)

ii Estimated probabilities
iii Gradients in associations with health outcomes and mortality

Criteria for selecting the number of classes

i Model fit indicators
ii Estimated probabilities
- Item profiles for each class-model. Descriptive way to assess differences across classes. It is a support to theoretical considerations.

iii Gradients in associations with health outcomes and mortality

Criteria for selecting the number of classes

i Model fit indicators
ii Estimated probabilities
iii Gradients in associations with health outcomes and mortality

Association of disability with mortality and health

Latent growth model

With y_{it} representing the health outcome being observed in individual i at occasion t:

$$y_{it} = \lambda_{0t} intercept_i + \lambda_{1t} slope_i + e_{it} \quad \text{(1) first level equation}$$

$$intercept_i = \beta_{00} + \mathbf{Z}_i^{\top}\beta_{01} + u_{0i} \quad \text{(2) second level equations}$$
$$slope_i = \beta_{10} + \mathbf{Z}_i^{\top}\beta_{11} + u_{1i}$$

Association of disability with mortality and health

Latent growth model

Figure: Typical latent linear growth model for an outcome measured on six occasions with one manifest explanatory variable

Results

i. Fit indicators

Men

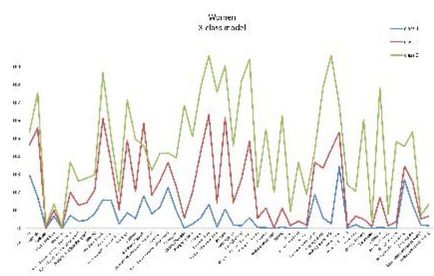
n classes	BIC	entropy	free parameters
3	184507.246	0.914	242
4	181752.864	0.904	323
5	180330.003	0.901	404
6	179795.197	0.898	485

Women

n classes	BIC	entropy	free parameters
3	235433.017	0.906	242
4	232979.961	0.889	323
5	231568.245	0.892	404
6	230449.939	0.885	485

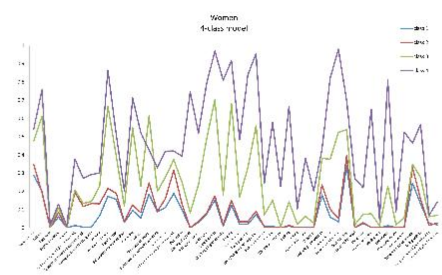
Results

ii. Estimated probabilities



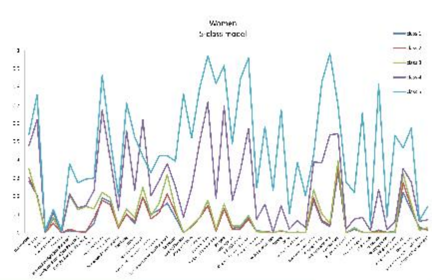
Results

ii. Estimated probabilities



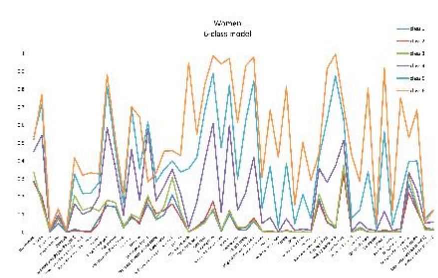
Results

ii. Estimated probabilities



Results

ii. Estimated probabilities



Results

iii. Association of alternative specifications of disability levels with health

Table: -Multicategorical classification-

Women

Specification of disability levels		Grip strength		Coagulation		Mental health	
		intercept	slope	intercept	slope	intercept	slope
3 classes - LCA	group 2	-2.092***	0.053	0.106***	-0.003	0.827***	0.014
	group 3	-5.398***	0.363***	0.071*	0.017	1.971***	-0.072***
4 classes - LCA	group 2	-0.594**	0.003	-0.031	0.006	0.191***	-0.002
	group 3	-2.604***	0.118*	0.092**	-0.006	0.94**	0.003
	group 4	-5.829***	0.32***	0.045	0.015	1.994***	-0.082***
5 classes - LCA	group 2	-0.006	-0.078	0.026	-0.012	0.085	-0.014
	group 3	-0.316	-0.105	-0.034	0	0.2**	0
	group 4	-2.376***	0.027	0.089**	-0.011	0.99***	0
	group 5	-5.658***	0.255**	0.05	0.005	2.007***	-0.08**
6 classes - LCA	group 2	0.091	-0.058	0.02	-0.013	0.055	-0.016
	group 3	-0.177	-0.092	-0.041	0.001	0.13	0.005
	group 4	-1.751***	-0.051	0.09***	-0.012	0.792***	-0.007
	group 5	-4.02***	0.169	0.126***	0.013	1.577***	-0.006
	group 6	-7.366***	0.644**	0.092	-0.011	2.421***	-0.173***
Observations		3,817		3,083		4,962	

Results

iii. Association of alternative specifications of disability levels with health

Table: -Binary classification-

Women

Specification of disability levels		Grip strength		Coagulation		Mental health	
		intercept	slope	intercept	slope	intercept	slope
2 classes - WHO	disabled	-3.387***	0.259***	0.082**	0.009	1.099***	-0.006
2 classes - ROC curve (min r)	disabled	-3.439***	0.236***	0.098**	0.011	1.108***	-0.008
2 classes - ROC curve (max j)	disabled	-3.702***	0.285***	0.082**	0.017*	1.115***	-0.016
2 classes - LCA	disabled	-3.66***	0.281***	0.099**	0.011	1.223***	-0.025

Outline

Discussion

- Importance of going beyond a binary definition of disability stressed in recent studies (Jagger et al., 2015; Wolf et al. 2015).
- The balance between the need of identifying detailed disability groups and the threat of subgrouping the population in repeated categories is not easy to achieve.

Discussion

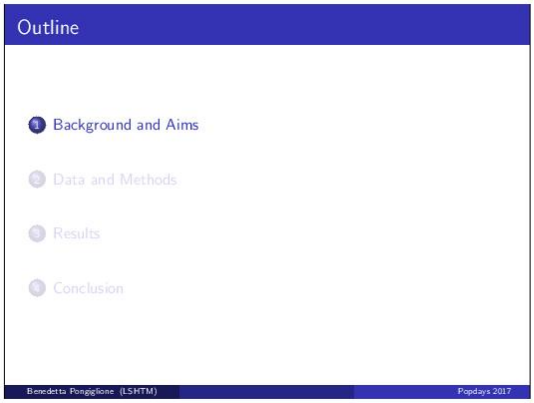
- Existence of a disability gradient, but only for the higher grades of disability.
- The binary indicator averages the burden of disability and masks the very strong effect experienced by individuals having severe disability, and is not informative for low level of disability.

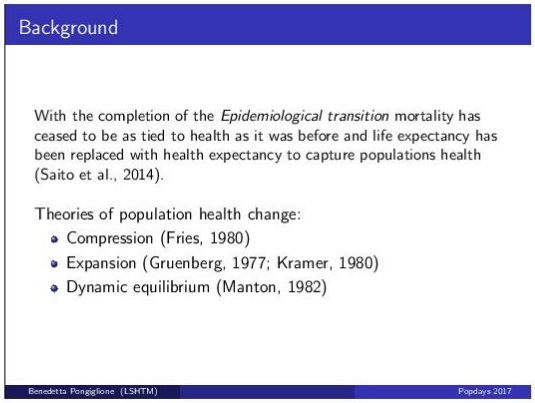
→ Our results indicated that the best classification of disability has **four classes**.

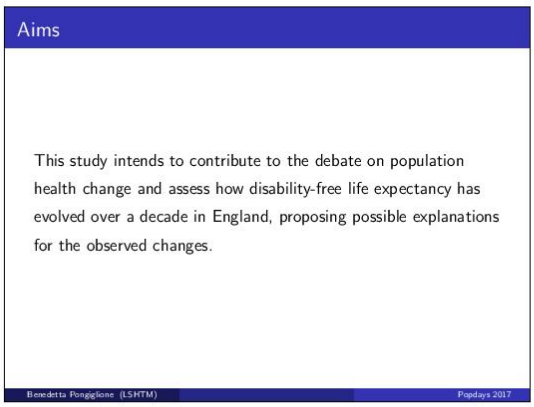
Implication and future work

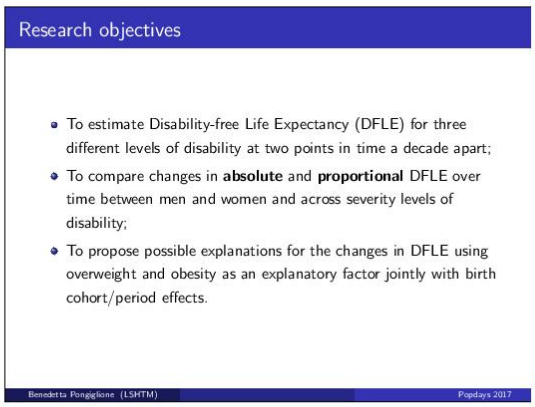
- Use the 4-class classification of disability to build a summary measure of population health: disability-free life expectancy and disability life expectancy for each disability level
- Assess trends in these measures over time

Figure 9.4: Slides of Research Paper IV









Outline

1. Background and Aims
2. Data and Methods
3. Results
4. Conclusion

Data

English Longitudinal Study of Ageing (ELSA)

- Data from a representative sample of individuals aged 50 years and over resident in the household sector in England.
- 6 waves issued from 2002/2003 to 2012/2013.
- Sample: core members interviewed who had complete data on all disability items interviewed at wave 1 (n=9,731) and wave 6 including refreshment samples (n=7,507).

Methods

Measurement of disability

- Latent class analysis (LCA): the patterns of responses are due to the influence of a discrete latent variable C that we specified having 4 classes.
  For binary items and a categorical latent variable C with 4 classes (k=1,...4), the marginal probability of observed item u_j (with j=1,2...,42) being equal to 1 is

$$Pr(u_j = 1) = \sum_{k=1}^{4} Pr(C_k = k) Pr(u_j = 1 | C_k = k)$$

Disability-free life expectancy

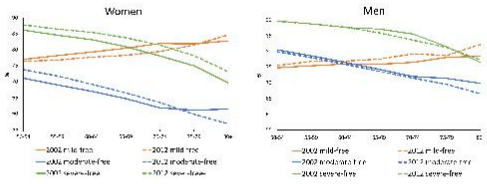
- Estimated using the Sullivan method (Sullivan, 1971).
- DFLE estimated for each grade of disability (i.e. mild-DFLE, moderate-DFLE, severe-DFLE).

Outline

1. Background and Aims
2. Data and Methods
3. Results
4. Conclusion

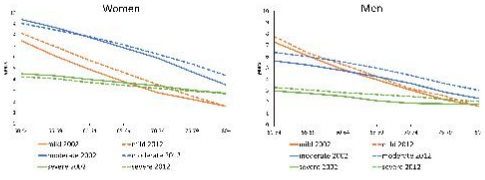
Results

Proportion of disability-free life expectancy on total life expectancy by age and gender



Results

Absolute disability life expectancy by age and gender



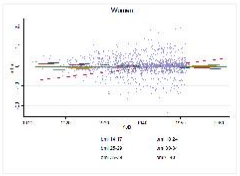
Results

Table: Linear regressions on the difference between age-adjusted YLD residuals at wave 6 and wave 2 ($\hat{\epsilon}_6 - \hat{\epsilon}_2$).

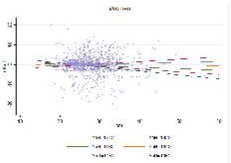
	(1)		(2)		(3)		(4)	
	Men	Women	Men	Women	Men	Women	Men	Women
YoB	0.031**	0.013			0.03**	0.014	-0.185*	-0.164**
(s.e.)	(-0.013)	(-0.013)			(-0.013)	(-0.013)	(-0.096)	(-0.075)
BMI			0.018	-0.013	0.015	-0.014	-15.00**	-12.54**
(s.e.)			-0.026	-0.021	(-0.026)	(-0.021)	(-6.621)	(-5.176)
Bmi×YoB							0.008**	0.007**
(s.e.)							(-0.003)	(-0.003)

Results

Is there a cohort effect moderated by BMI?

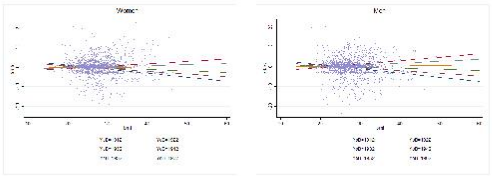


Is there a BMI moderated by birth cohort?



Results

Is there a BMI effect moderated by birth cohort?



Outline

Conclusions

- Results differ depending on level of disability → crucial identifying appropriate disability levels.
- Preliminary evidence for an interactive role of cohort and BMI in changes in YLD, such that high BMI resulted particularly detrimental to younger cohorts.
- **Policy implications:**
  - Absolute expansion of (almost) all levels of disability will have dramatic implications on society and health system.
  - Need to monitor BMI and age at onset of overweight and obesity, especially among the younger cohorts.

Thank you for your attention

E-mail: benedetta.pongiglione@lshtm.ac.uk
Twitter:@benedettapongi

Appendix A

Supporting Information to the Systematic Literature Review

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	2
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			

Table A.1: PRISMA statement for the systematic review

Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	2
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	3
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	3,4
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	NA
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis.	4
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	NR
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	4

Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	S1 table
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	NA
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	5-9
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	10
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	10
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	10
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

Table A.3: Descriptive characteristics of included studies

#	Year	Authors	Where	Who (participants)	Why (exposure)	What (outcome)	Measure for outcome	Method	Data
1	2007	Al Snih et al.	US	aged 65+	sex, age, BMI	TLE, DFLE	ADL	Multistate life table	Epidemiologic Studies of the Elderly (EPESE)
2	2011	Andrade et al.	Brazil	aged 60+	sex, age	TLE, DFLE, DLE	ADL	Multistate life table	Survey on Health, Well-Being, and Aging in Latin America and the Caribbean (SABE)
3	2014	Belon, Lima, and Barros	Brazil	aged 60+	sex, age	TLE, HLE, UHE	SRH; GALI	Sullivan's method	City of Campinas Health Survey (ISACAMP)
4	2013	Beltran-Sanchez and Andrade	Brazil and Mexico	aged 60+	sex, age, education	TLE, DFLE, DLE	Nagi (1976) functional limitations	Micro-simulation methods	(1) (SABE) (2) Mexican Health and Aging Study (MHAS)
5	1991	Branch et al.	US	aged 65+	sex, age, state, disability status	TLE, ALE	ADL	Multistate life table	EPESE
6	2001	Brayne et al.	England and Wales	aged 64+	sex, age	TLE, -LE w/ physical illness, LE w/ functional problems, LE w/ cognitive problems	ADL; MMSE score; physical health: cancer, non-cancer morbidity, hearing and sight problems	Sullivan's method	Medical Research Council Cognitive Function and Ageing Study (MRC CFAS)
7	2005	Bronnum-Hansen	Denmark	aged 20 aged 65	sex, age	TLE, LE in good SRH, LE w/o LLI	SRH; LLI	Sullivan's method	Danish Health Interview Surveys
8	2006	Bronnum-Hansen, Juel and Davidsen	Denmark	aged 65+	sex, disease	TLE,-LE w/o LLI	LLI	Sullivan's method	Danish Health Interview Survey
9	2007	Cai, and Lubitz,	US	aged 65+	sex, disability status	TLE, ALE, DLE	ADL; IADL	Multistate life table	Medicare Current Beneficiary Survey (MCBS)

10	2001	Cambois, Robine, and Hayward	France	men aged 35+	occupational class, year, age	TLE and partial TLE; DFLE and partial DFLE; DLE and partial DLE	GALI	Sullivan's method	French Health and Medical Care Survey
11	2013	Cambois, Blachier, and Robine	France	aged 50+	sex, age	TLE, DFLE, DLE	Functional Limitations in: (i) hearing and seeing that are not corrected by devices for sensory; (ii) walking, bending down, grabbing and reaching; (iii) remembering, understanding, orientating, behaving inappropriately. Activity Restriction: (i) ADL; (ii) IADL; (iii) GALI	Sullivan's method	(1)Survey on health and health consumption (2) Household survey on disability and health (3); SILC, EU-Statistics on income and living conditions (4) Survey of health, aging and retirement in Europe (SHARE) (5) Survey on health and social protection
12	2011	Cambois et al.	France	aged 50 aged 65	sex, age, occupational class	TLE, 5 measures of HE	SRH; self-rated chronic diseases; physical and sensory functional limitations; GALI;ADL	Sullivan's method	French health survey (DREES 2008)
13	2013	Campolina et al.	Brazil	aged 60+	sex, age, disease	TLE, DFLE, DLE	ADL	Sullivan's method	SABE
14	2011	Chan, Zimmer, and Saito	Japan	aged 65+	sex, age (controlling for education, occupational class, income, chronic conditio)	TLE, DFLE	ADL	Multistate life table	Nihon University Japanese Longitudinal Study of Aging (NUJLSOA).
15	2010	Cheung and Yip	Hong Kong	aged 60+	sex, age	TLE, DFLE, chronic morbidity-free life expectanc (CMFLE)	ADL, questions on chronic diseases	Sullivan's method	Thematic Household Survey (THS)
16	1989	Crimmins, Saito, and Ingegneri	US	-birth -aged 65 -aged 85	sex, age, race, country	TLE, -DFLE (distinguishing with short and long term disability, institutionalised)	normal activities; secondary or nonmajor activities	Sullivan's method	National Health Interview Survey (NHIS)

17	1996	Crimmins, Hayward and Saito	US	aged 70+	sex, education, race	TLE, ALE	ADL, IADL, functions indicating stamina and mobility	Multistate life table	Longitudinal Study of Aging (LSOA)
18	2009	Crimmins et al.	US	aged 70+	age, year, disability status	TLE, DFLE, DLE	ADL, IADL	Multistate life table	LSOA
19	2001	Crimmins and Saito,	US	aged 30+	sex, race, education	TLE, HLE	GALI	Sullivan's Method	NHIS
20	1992	Crimmins, Saito, and Ingegneri	Book chapter, review						
21	2002	Deeg, Portrait and Lindeboom	The Netherldands	aged 65+	sex, age	LE in 6 health profiles (cancer; "other" chronic diseases; cognitive impairment; frailty or multimorbidity; cardiovascular diseases; good health)	self-reported functional limitations, test of physical performance, self-reported chronic diseases, vision, hearing, depressive symptoms (20-item scale), cognitive impairment (MMSE)	Life table	Longitudinal Aging Study Amsterdam
22	2002	Diehr et al.	US	aged 65+	sex, BMI	YOL, YHL	SRH	NA	Cardiovascular Health Study (CHS)
23	2008	Diehr et al.	US	aged 65+	sex, BMI, ADL, SRH	YOL, YHL, ALE	ADL, SRH	Multistate life table	CHS
24	1999	Ferrucci et al.	US	aged 65+	sex, smoking, physical activity	TLE, ALE, DLE	ADL	Markov chain	EPESE
25	2009	Gu et al.	China	aged 65+	sex, age, year	TLE, ALE, DFLE, HLE	ADL , chronic conditions, SRH	Sullivan's method	-Old-Age Support Survey of the Chinese Elderly (OSSCE) -Chinese Longitudinal Healthy Longevity Survey (CLHLS)
26	1993	Guralnik et al.	US	aged 65+	sex, education, race	TLE, ALE, DLE	ADL	Multistate life table	Piedmont Health Survey of the Elderly (PHSE)
27	1998	Hayward, Crimmins and Saito	US	aged 70+	sex, age, cause of death	TLE, ALE, DLE	ADL, IADL	Multistate life table	Longitudinal Study of Aging

28	2014	Hayward et al.	US	aged 50+	sex, race/ethnicity	TLE, DFLE, DLE	ADL	Multistate life table	Health and Retirement Study (HRS)
29	2007	Hidajat, Hayward and Saito	Indonesia	aged 50+	sex, age, education	TLE, ALE, ILE	Nagi	Multistate life table	Indonesian Family Life Survey
30	2002	Ishizaki et al.	Japan	aged 65+	sex, age	TLE, ALE in terms of: (1) physically active life expectancy (2) instrumentally active life expectancy	ADL, IADL	Multistate life table	Saku Longitudinal Study on Aging
31	2000	Izmirlian et al.	US	aged 65+	sex, age, smoking, education	TLE, ALE, DLE	ADL	Markov chain	EPESE
32	2002	Jagger and Matthews	England	aged 65+	sex, age	TLE, HE by number of impairments	Functional limitations: Townsend disability scale, cognitive impairment: MMSE, physical impairment: presence of chronic conditions	Sullivan's method	MRC CFAS
33	2007	Jagger et al.	UK	aged 65+	sex, age, disease status	TLE, DFLE (free of any disability and free of moderate or Severe Disability)	ADL, IADL	Multistate life table	MRC CFAS
34	2011	Jagger et al.	13 European countries	aged 50+	sex, country	TLE, LE free of chronic morbidity, LE w/o physical functional limitations, HLY, LE w/o IADL, LE w/o ADL, LE in good SRH	chronic diseases, functional limitations, ADL, IADL, SRH	Sullivan's method	SHARE
35	2007	Jagger et al.	England and Wales	aged 65+	sex, age, education	TLE, DFLE, DLE	ADL, mobility disability (get up and down stairs)	Multistate life table	MRC CFAS

36	2008	Jeune and Bronnum-Hansen	Denmark	aged 65+	sex, year	TLE, exp. lifetime w and w/o long-standing illness, exp. lifetime w and w/o LLI, exp. lifetime w and w/o mobility limitations, exp. lifetime w and w/o communication limitations, exp. lifetime in good, fair or poor SRH	SRH, LLI, mobility restrictions, communication restrictions	Sullivan's method	Danish Health Interview Surveys
37	2012	Jiawiwatkul et al.	Thailand	aged 60+	sex, age	TLE, ALE	ADL	Sullivan's method	National Health Examination Surveys
38	2000	Jitapunkul and Chayovan	Thailand	aged 60+	sex, age, year	TLE, HLE	SRH	Sullivan's method	(1) Socioeconomic Consequences of the Ageing of the Population (SECAPT) (2) Survey of the the Welfar of Elderly in Thailand (SWET)
39	2001	Jitapunkul et al.	Thailand	aged 60+	sex, age	TLE, ALE	Barthel ADL index, Chula ADL index	Sullivan's method	A cross-sectional mulitistage random sampling survey
40	2003	Jitapunkul et al.	Thailand	aged 60+	sex, age	TLE, DFLE, Self-care LE	GALI, modified Barthel ADL index	Sullivan's method	National Health Examination Survey
41	1993	Kai	Japan	aged 63+	sex, age	TLE, ALE	ADL	Abridged life table	Cohort in a rural community in Japan
42	2005	Kaneda, Zimmer and Tang	China	aged 55+	sex, age, SES: (education, income, occupation, household possession)	TLE, ALE, IALE	ADL, Nagi (1976) functional limitations	Multistate Life Table	Beijing Multidimensional Longitudinal Study on Aging
43	2013	Karcharnubarn Rees and Gould	Thailand	aged 60+	sex, age, year	TLE, SRHLE,-Mobility Disability Free Healthy LE, Self-care DFLE	ADL, SRH	Sullivan's method	Surveys of Elderly in Thailand
44	1983	Katz et al.	US	aged 65+	sex, age, income	TLE, ALE	ADL	Abridged life table	Massachusetts Health Care Panel Study

45	2011	Klijs, Mackenbach and Kunst	The Netherlands	aged 55+	smoking, weight, drinking controlling for age sex and marital status	TLE, DFLE, DLE	ADL	Sullivan's method	Dutch Permanent Survey of the Living Situation (POLS)
46	2004	Konno et al.	Japan	aged 65+	sex, age	TLE, ALE	ADL, IADL	Modified Branch's method	Annual surveys on preventive care for senior citizens
47	2002	Laditka and Laditka	Critical Review						
48	2014	Laditka and Laditka	US	aged 55+	sex, ethnicity, education, cognitive impairment	TLE, ALE, DLE	ADL	Markov chain	Panel Study of Income Dynamics
49	2009	Lai	China	all age groups	sex, age, year	TLE, expected years of life lived w/ and w/o handicap	6 categories of handicap: ocular, aural (including language), intellectual, skeletal, psychological and multiple handicap	Sullivan's method	National sampling survey
50	2000	Lai, Lee and Lee	China	all age groups	sex, age, residence	TLE, handicap-Free LE, expected Years of Life with Handicap	Ocular, aural, intellectual, skeletal, psychological and multiple handicaps	Sullivan's method	Survey on the handicapped
51	1994	Land, Guralnik and Blazer	US	aged 65+	sex, age, race, education	TLE, ALE, DLE	ADL	Multistate life table	EPESE
52	2007	Lievre et al.	12 European countries	aged from 50 to 70	sex, age	Healthy Working Life Expectancy	SRH	Multistate life table	European Community Household Panel (ECHP)
53	2009	Liu et al.	China	aged 60+	sex, age	TLE, DFLE, DLE	Visual, hearing and speech, physical, intellectual, and mental disabilities assessed by clinicians	Sullivan's method	China National Sample Survey on Handicap/Disability

54	1995	Liu et al.	Japan	aged 60+	sex, education, social network	TLE, functional independence LE, functional disability LE	Functioning: (1) bathing oneself, (2) climbing two or three flights of stairs, and (3) walking about 200-300 meters or a few blocks	Multistate life table	Longitudinal study of adult Japanese aged 60 and over.
55	2010	Majer et al.	10 Western - European countries	aged 50 aged 65	sex, education, country	TLE, DFLE	GALI	Multistate life table	ECHP
56	2008	Manton, Gu and Lowrimore	US	aged 65+	sex, cohort, disability status	TLE, ALE	ADL, IADL, Nagi functional limitations, sensory function (vision)	Stochastic process life tables	National Long-Term Care Survey
57	1997	Manton, Stallard, and Corder	US	aged 65+	sex, age, education	TLE, DFLE	Functioning	Stochastic process life tables	National Long Term Care Surveys
58	2006	Matthews et al.	UK	aged 65+	sex, age, town	TLE, LE free of cognitive impairment, ALE, LE in good health	SRH, ADL, risk factors for dementia, cognitive function and medication (MMSE)	Sullivan's method	MRC CFAS
59	2006	Matthews, Jagger and Hancock	UK	aged 75+	sex, socio-economic conditions	TLE, DFLE	ADL	Multistate life table	Melton Mowbray Health Check data
60	2000	Melzer et al.	England	aged 65+	sex, social class	TLE, DFLE, DLE	mentally impairment: Automated Geriatric Examination Computer Assisted Taxonomy; phisically disability: Modified Townsend Disability Scale; need for constant care: functional incapacity and/or mental state (MMSE)	Sullivan's method	MRC CFAS

61	2005	Minicuci and Noale	Italy	aged 65-84	sex, age, education	TLE, DFLE, DLE (mild and severe)	ADL, IADL	Multistate life table	Italian Longitudinal Study on Aging (ILSA)
62	2005	Minicuci and Noale	Italy	aged 65-84	sex, age, geographical area	TLE, DFLE	ADL, IADL, physical performance test (PPT)	Sullivan's method	ILSA
63	2011	Minicuci et al.	Italy, Bulgaria and Latin America	aged 65-84	sex, age, country	TLE, DFLE	ADL	Sullivan's method	(1) NHIS (Bulageria) (2) Multidisciplinary survey among Italian families: health status and health services access (IMF-S) in Italy. (3) SABE
64	2014	Montez and Hayward	US	aged 50+	sex, childhood SES, childhood health, education	TLE, ALE, ILE	ADL, IALD, functional limitations	Multistate life table	HRS
65	2011	Muangpaisan et al.	Thailand	aged 50+	sex, age	TLE, cognitive impairment free LE, physical illness-free LE, psychological impairment-free LE, self-care LE, DFLE, HLE	Thai MMSE, psychological Quality of Life score, ADL, self reported physical illness and hearing and visual problems, self-care domains, SPGH (self-perceived global health)	Sullivan's method	Bangkok Longitudinal Study by Siriraj Hospital for the Older Men and Women (BLOSSOM)
66	1997	Mutafova et al.	Bulgaria	aged 60+	sex, age	TLE, DFLE, handicap-free LE, handicap LE	ADL , SRH, handicap (confined home, chair, bed)	Sullivan's method	First pilot health interview survey and in Svishtov
67	2001	Newman and Brach	Review						
68	2004	Ofstedal et al.	6 Asian countries	aged 60+	sex, age, country	TLE, HLE, ULE	SRH	Sullivan's method	National surveys: (1) Indonesian Family Life Survey (2) Philippine Survey of the Near-Elderly and Elderly (3) National Survey of Senior Citizens in Singapore (4) Taiwan Survey of Health and Living Status of the Middle-Aged and Elderly

69	2008	Peres et al.	England and Wales	aged 65+	sex, age, emotional problem, multimorbidity	TLE, DFLE	ADL	Multistate life table	MRC CFAS
70	2005	Reyes-Beaman et al.	Mexico	aged 60+	sex, age, local regional area	TLE, ALE	ADL	Sullivan's method	National Survey on Ageing carried out in the Mexican Institute of Social Security (IMSS)
71	2008	Reynolds, Haley, and Kozlenko,	US	aged 70+	sex, age, depression and chronic diseases	TLE, ALE, DLE	ADL	Multistate life table	Asset and Health Dynamics Among the Oldest Old (AHEAD)
72	2009	Reynolds and McIlvane	US	aged 70+	sex, age, obesity status, arthritis	TLE, ALE, DLE	ADL	Multistate life table	AHEAD
73	2005	Reynolds, Saito, and Crimmins	US	aged 70+	sex, age, obesity status	TLE, ALE, DLE	ADL	Multistate life table	AHEAD
74	2002	Robine, Jagger, and Cambois	12 European countries	aged 65	sex, country	TLE , fully active LE, partly active LE	GALI	Sullivan's method	European Community Household Panel
75	1989	Rogers, Rogers and Belanger	US	aged 70+	sex, age	TLE, ALE	ADL	Multistate life table	LSOA
76	2014	Romero-Ortuno, Fouweather, and Jagger	Europe	aged 50+	sex, age, country	LE w/ and w/o frailty	Frailty Instrument, GALI	Sullivan's method	SHARE
77	2005	Sagardui-Villamor et al.	Spain	aged 65+	sex, age, year	TLE, DFLE, DLE	5 major types of disability (seeing, hearing, selfcare, walking, cognitive disabilities)	Sullivan's method	National Disability, Impairment and Handicap Surveys

78	2008	Santos Camargos, Machado, and Rodrigues	Brazil	aged 60+	sex, age (predictor for SRH: living arrangement, marital status, schooling and income, functional disability and chronic illnesses)	TLE, HLE, ULE	SRH	Sullivan's method	SABE
79	2008	Santos Camargos, Machado and Rodrigues	Brazil	aged 60+	sex, age	TLE, DFLE (Mild, Moderate, Severe)	ADL, mobility	Sullivan's method	National Sample Household Survey (PNAD)
80	2007	Santos Camargos, Machado and Rodrigues	Brazil	aged 60+	sex, age, education	TLE, DFLE, DLE	ADL	Sullivan's method	SABE
81	2001	Sauvaget, Jagger and Arthur	UK	aged 75+	sex, age	TLE, ALE, cognitive impairment free LE	ADL, Clifton Assessment Procedures of the Elderly	Multistate life table	Melton Mowbray Health Check data
82	1999	Sauvaget et al.	Japan	aged 65+	sex, age	TLE, DFLE	ADL, IADL, mobility	Multistate life table	A 2-year prospective study of older residents in a rural Japanese community.
83	2011	Szwarcwald, et al.	Brazil	aged 18+	sex, age, location (slums)	TLE, HLE, Healthy Life Lost	(1) daily activities (work or household activities), (2) mobility (moving around), (3) self-care (washing or dressing yourself), (4)sensation of pain (5) learning (6) vision	Sullivan's method	World Health Survey

84	2013	Tareque, Begum and and Saito	Bangladesh	aged 60+	sex, age	TLE, DFLE	(a) vision; (b) hearing; (c) walking and climbing; (d) difficulty in remembering and concentrating; (e) selfcare; (f) speaking and communicating.	Sullivan's method	Bangladesh's Household Income and Expenditure Survey (HIES)
85	2011	Tian et al.	China	aged 55+	sex, age, smoking status (controlling for marital status, education, district, alcohol consumption (in men only), srh, and physical disabilities)	TLE, ALE	ADL	Sullivan's method	Beijing Longitudinal Study of Ageing
86	1995	Tsuji et al.	Japan	aged 65+	sex, age	TLE, ALE	ADL	Multistate life table	Sendai Longitudinal Study of Aging
87	2002	Tsuji, Sauvaget, and Hisamichi	Japan	aged 65+	sex, age	TLE, ALE, DLE	ADL, IADL, locomotion	Multistate life table	Surveys conducted by the authors
88	1993	van den Bos and van der Maas	The Netherlands	aged 55-79	sex, age, index of income, eduation and occupation	TLE, morbidity-free Life Expectancy, DFLE	chronic conditions, disability	Sullivan's method	
89	2009	Walter et al.	The Netherlands	aged 55+	sex, BMI, waist circum (controlling for age edu income smoking alcohol living situatio)	TLE, DFLE, DLE	ADL	Multistate life table	Rotterdam Study
90	2009	Yardim et al.	Turkey	aged 60+	sex, rural urban national level	TLE, HLE	not mentioned	Preston-Coale (50%) Method corresponding to Horiuchi Method	(1) Turkish Demographic and Health Survey (2) Verbal Autopsy Survey

91	2004	Yi, Danan and Land	China	aged 80-105	sex, age, disability status	TLE, ALE, DLE	ADL	Multistate life table	CLHLS
92	2001	Yi et al.	China	aged 80+	sex, age, rural urban area	TLE, ALE, DLE (mild, severe)	ADL	Sullivan's method	Healthy Longevity Survey
93	2009	Yong and Saito	Japan	aged 25+	sex, age, year	TLE, HLE (good, average and poor health)	SRH	Sullivan's method	Kokumin Seikastu Kiso Chosa (Comprehensive Survey of Living Conditions of the People on Health and Welfare)
94	2011	Yong, Saito and Chan	Singapore	aged 55+	sex, age	TLE, years of life w/ and w/o diseases, years of life w/ and w/o impairments, years of life w/ and w/o functional disabilities	chronic diseases, impairments (bone/joint, lung/breathing, hearing, eye/vision, and mobility problems), ADL	Sullivan's method	National Survey of Senior Citizens (NSSC)
95	2012	Yong and Sayto	Japan	aged 65+	sex, age, education	TLE, ALE, IALE	ADL, IADL	Multistate life table	Nihon University Japanese Longitudinal Study of Aging
96	2010	Zimmer et al.	China	aged 55+	age, sex, rural urban area	TLE, ALE	ADL and general physical movement tasks necessary for independent living	Multistate life table	Beijing Multidimensional Longitudinal Study of Aging

Table A.5: Full population of studies screened for the systematic review

Prisma phases	Exclusion criteria	n	Author	Year	Title
Full-text articles assessed for eligibility	Not focused on outcome of interest	5	Di Gessa and Grundy	2014	The relationship between active ageing and health using longitudinal data from Denmark, France, Italy and England
			Leveille et al.	1999	Aging successfully until death in old age: Opportunities for increasing active life expectancy
			Avlund et al.	1999	Active life in old age - Combining measures of functional ability and social participation
			Hirai et al.	2012	Social Determinants of Active Aging: Differences in Mortality and the Loss of Healthy Life between Different Income Levels among Older Japanese in the AGES Cohort Study
			Yang et al.	2014	Socioeconomic status, comorbidity, activity limitation, and healthy life expectancy in older men and women: a 6-year follow-up study in Japan
	Not focused on exposure of interest	3	Laditka and Wolf	1998	New methods for analyzing active life expectancy
			Manton et al.	1993	Forecasts of active life expectancy: policy and fiscal implications
			Mathers et al.	2002	Global patterns of healthy life expectancy for older women
	Not focused on older population	6	Geronimus et al.	2001	Inequality in life expectancy, functional status, and active life expectancy across selected black and white populations in the United States
			Lau et al.	2012	Healthy Life Expectancy in the Context of Population Health and Ageing in India
			Molla, Centers for Disease and Prevention Control	2013	Expected years of life free of chronic condition-induced activity limitations - United States, 1999-2008
			Murray and Lopez	1997	Regional patterns of disability-free life expectancy and disability adjusted life expectancy: Global Burden of Disease Study
			Payne et al.	2013	Disability Transitions and Health Expectancies among Adults 45 Years and Older in Malawi: A Cohort-Based Model
			Wolfson	1996	Health-adjusted life expectancy
	Methods	6	Fukuda et al.	2005	Municipal health expectancy in Japan: decreased healthy longevity of older people in socioeconomically disadvantaged areas
			Noale et al.	2012	Longevity and health expectancy in an ageing society: implications for public health in Italy
			Salomon et al.	2012	Healthy life expectancy for 187 countries, 1990-2010: a systematic analysis for the Global Burden Disease Study 2010
			Jagger et al.	2008	Inequalities in healthy life years in the 25 countries of the European Union in 2005: a cross-national meta-regression analysis
			Kondo et al.	2005	Factors explaining disability-free life expectancy in Japan: the proportion of older workers, self-reported health status, and the number of public health nurses
			Van Oyen et al.	2010	Gender gaps in life expectancy and expected years with activity limitations at age 50 in the European Union: associations with macro-level structural indicators

Appendix B

ICF items classification

Domain	#	Questions	1th level	2nd level	3rd level	4th level
Impairment	1	High blood pressure or hypertension	4	20		
	2	Arthritis (including osteoarthritis , or rheumatism)	7	10		
	3	Parkinson's disease	7	35	6	
	4	Any emotional, nervous or psychiatric problems	1	26	3	
	5	Dementia, organic brain syndrome, senility or any other serious memory impairment	1	17		
	6	Is your eyesight (using glasses or corrective lens as usual) [excellent to poor]	2	10		
	7	How good is your eyesight for seeing things at a distance, like recognising a friend across the street (using glasses or corrective lens as usual)? [excellent to poor]	2	10	0	0
	8	How good is your eyesight for seeing things up close, like reading ordinary newspaper print (using glasses or corrective lens as usual)? [excellent to poor]	2	10	0	1 or 2
	9	Is your hearing (using a hearing aid as usual) [excellent to poor]	2	30		
	10	How often do you have problems with dizziness when you are walking on a level surface?	4	15	8	
	11	Have you ever had any pain or discomfort in your chest?	2	80	1	1
	12	Have you ever had a severe pain across the front of your chest lasting for half an hour or more?	2	80	1	1

Table B.1: ICF items classification

	13	Are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill?	4	60		
	14	Have you had attacks of wheezing or whistling in your chest at any time in the last 12 months?	4	60		
	15	This might not be easy to talk about, but we would like to ask you about incontinence. During the last 12 months, have you lost any amount of urine beyond your control?	6	20	2	
	16	(Much of the time during the past week), you felt depressed?	1	26		
	17	Are you often troubled with pain?	2	80		
	18	How would you rate your memory at the present time? [excellent to poor]	1	44		
	19	Do you get pain or discomfort in either of your legs which comes on when you walk?	2	80		
Activity	1	Walking 100 yards	4	50	0	
	2	Sitting for about two hours	4	10	3	
	3	Getting up from a chair after sitting for long periods	4	10	8	
	4	Climbing several flights of stairs without resting	4	55	1	
	5	Climbing one flight of stairs without resting	4	55	1	
	6	Stooping, kneeling, or crouching	4	10	2	
	7	Reaching or extending your arms above shoulder level	4	45	2	
	8	Pulling or pushing large objects like a living room chair	4	49		
	9	Lifting or carrying weights over 10 pounds, like a heavy bag of groceries	4	30		
	10	Picking up a 5p coin from a table	4	40	0	
	11	Dressing, including putting on shoes and socks	5	40		
	12	Walking across a room	4	50	0	
	13	Bathing or showering	5	10	1	
	14	Eating, such as cutting up your food	5	50		
	15	Getting in or out of bed	4	10		
	16	Using the toilet, including getting up or down	5	30	8	
	17	Do you find it difficult to follow a conversation if there is background noise, such as TV, radio or children playing (using a hearing aid as usual)?	2	30	8	
	18	How often do you have problems with keeping your balance when you are walking on a level surface?	4	15	8	
	19	(Much of the time during the past week), your sleep was restless?	1	34	3	
	20	By yourself and without using any special equipment, how much difficulty do you have walking for a quarter of a mile? [1 no difficulty, 2 some difficulty, 3 much difficulty? 4 or, are you unable to do this?]	4	50	0	

Participation	1	Preparing a hot meal	6	30		
	2	Using a map to figure out how to get around in a strange place	4	69		
	3	Shopping for groceries	6	20	0	
	4	Making telephone calls	3	60	0	
	5	Doing work around the house or garden	6	49		
	6	Managing money, such as paying bills and keeping track of expenses	8	60		
	7	Do you use public transport Why don't you use public transport more often?[My health prevents me]	4	70	2	
	8	What were your reasons for retiring? [2 Own ill health]	8	45	2	
	9	What were your reasons for taking early retirement? [2 Own ill health]	8	45	2	
	10	Are you a member of any of these organisations, clubs or societies? (Political party, trade union or environmental groups; Tenants or resident groups; Neighbourhood Watch; Church or other religious groups; Charitable associations; Education, arts or music groups or evening classes; Social and Sports clubs; Any other group)	9	50		
	11	Did you do any of these activities during the last month, that is since date a month ago? (Paid work; Self-employment; Voluntary work; Cared for a sick or disabled adult; Looked after home or family; Attended a formal educational or training course)	8	50		

Appendix C

Overview of the content of the ELSA questionnaire

Table C.1: Overview of the content of the ELSA questionnaire by modules

Module	Content	Respondent	W1	W2	W3	W4	W5	W6
Household Demographic (HD)	The HD module is used to collect basic demographic information about everyone living in the household.	Household respondents	x	x	x	x	x	x
Individual Demographics (ID)	The ID module is at the start of the ELSA interview. Each respondent is asked details about their legal marital status, closest relatives and their own circumstances in childhood.	All respondents	x	x	x	x	x	x
Health (HE)	The HE module concerns the respondent's state of health, functional limitations, and certain behavioural aspects of their daily life that are likely to influence health.	All respondents	x	x	x	x	x	x
Social Participation (SP)	The SP module covers the frequency with which respondents participated in certain social activities, whether they were limited from participating in these activities, their care-giving and use of public transport.	All respondents	x	x	x	x	x	x
Work and Pension WP)	The WP module collects employment details, job characteristics, earnings, occupational pension contributions or receipts, and retirement decisions. It includes information about job search, training and voluntary activities if relevant.	All respondents	x	x	x	x	x	x
Income and Assets (IA)	The IA module asks those identified as providers of financial information about their individual and joint income, assets and debts.	Financial respondents	x	x	x	x	x	x
Housing (HO)	The HO module collects information about current housing situation, housing-related expenses, ownership of durable goods and cars, and expenditure on food.	Housing respondents	x	x	x	x	x	x
Cognitive Function (CF)	The CF module is asked privately and covers different aspects of the respondent's cognitive function, including memory, speed, mental flexibility and numeracy.	All respondents	x	x	x	x	x	x

Expectations (EX)	The EX module is asked privately and measures people's expectations in a number of dimensions, the level of certainty respondents felt about the future, financial decision-making within households and the time frame they thought about when making financial decisions.	All respondents	x	x	x	x	x	x
Psychosocial Health (PS)	The PS module is asked privately and measures how respondents viewed their lives across a variety of dimensions and included a mental health scale and questions about respondents' attitudes towards ageing.	All respondents	x	x	x	x	x	x
Final Questions and consent (FQ)	The FQ module is asked privately and collects miscellaneous information that does not easily fit elsewhere, for example classificatory information on ethnic group, country of birth, education.	All respondents	x	x	x	x	x	x
Measurement - Time Walked (MM)	In the MM module the test of walking speed is conducted, recording the time taken by respondents to walk a distance of 8 feet at their usual walking pace.	All respondents aged 60+	x	x	x	x	x	x
Self-Completion Questionnaire (SC)	The SC questionnaire contains instruments that are standardly collected in this way to safeguard privacy or because they are straightforward to collect. Core items are the CASP quality of life questionnaire, views of relationships with family and friends, and perceived social status.	All respondents	x	x	x	x	x	x
Effort and Reward (ER)	The ER module is asked privately and assesses respondents' motivations for voluntary work, caring for others & the relationship between effort and reward.	All respondents		x	x	x	x	
Health and Work Self-completion (HW)	The HW SC module asks to rate health - situation on 5-point scale including mobility, pain, sleep, cognition and depression and work - rate hypothetical people in various scenarios on 5-point scale.	Sub-sample of W3 respondents (third received work, third received health, third received core)			x			

Risk Module (RS)	The RS module comprises 22 games respondents play on a laptop as a self-completion CASI (computer-assisted self-interviewing) interview. Games are designed to measure respondents' attitudes towards financial risk taking and willingness to accept delay in receiving money.	Pre-selected sub-sample				x	
Nurse	The nurse interview comprises a personal face-to-face CAPI interview, the taking of a number of different measures and an additional voluntary self-completion questionnaire.	Core members who had an interview in person		x	x		x
Life history data	The Life History Interview collects data in a number of different areas including fertility, relationships, housing and mobility, living situation at 10 years old, jobs and earnings, health, childhood health, smoking and gynaecology, relationship with parents in childhood, important and difficult events in life	Respondents interviewed at wave 3			x		

References

[1] United Nations. *World population ageing, 1950-2050.* United Nations, New York, 2002.

[2] J. R. Wilmoth. Demography of longevity: past, present, and future trends. *Experimental gerontology*, 35(9):1111–1129, 2000.

[3] United Nations. World population ageing 2015. *Department of Economic and Social Affairs, Population Division*, (ST/ESA/SER.A/390), 2015. doi: citeulike-article-id: 14015720.

[4] J. E. Bennett, G. Li, K. Foreman, N. Best, V. Kontis, C. Pearson, P. Hambly, and M. Ezzati. The future of life expectancy and life expectancy inequalities in england and wales: Bayesian spatiotemporal forecasting. *Lancet*, 386(9989):163–70, 2015.

[5] United Nations. Department of Economic. *World population ageing 2009*, volume 295. United Nations Publications, 2010.

[6] A. Sonnega. The future of human life expectancy: have we reached the ceiling or is the sky the limit. *Research highlights in the demography and economics of aging*, 8: 1–4, 2006.

[7] S. J. Olshansky. The law of mortality revisited: interspecies comparisons of mortality. *J Comp Pathol*, 142 Suppl 1:S4–9, 2010.

[8] S. J. Olshansky, B. A. Carnes, and C. Cassel. In search of methuselah: estimating the upper limits to human longevity. *Science*, 250(4981):634–40, 1990.

[9] S. J. Olshansky, B. A. Carnes, and A. Desesquelles. Demography. prospects for human longevity. *Science*, 291(5508):1491–2, 2001.

[10] K. G. Manton, E. Stallard, and H. D. Tolley. Limits to human life expectancy: Evidence, prospects, and implications. *Population and Development Review*, 17(4): 603–637, 1991.

[11] J. Oeppen and J. W. Vaupel. Demography. broken limits to life expectancy. *Science*, 296(5570):1029–31, 2002.

[12] J. W. Vaupel. Biodemography of human ageing. *Nature*, 464(7288):536–42, 2010.

[13] J. W. Vaupel, J. R. Carey, K. Christensen, T. E. Johnson, A. I. Yashin, N. V. Holm, I. A. Iachine, V. Kannisto, A. A. Khazaeli, P. Liedo, V. D. Longo, Y. Zeng, K. G. Manton, and J. W. Curtsinger. Biodemographic trajectories of longevity. *Science*, 280(5365):855–60, 1998.

[14] E. M. Crimmins and H. Beltrán-Sánchez. Mortality and morbidity trends: is there compression of morbidity? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(1):75–86, 2011.

[15] J. James. *The Health of Populations Beyond Medicine*. Elsevier, United States, 2016.

[16] E. M. Gruenberg. The failure of success. *Milbank Memorial Fund Quarterly Health and Society*, 55(1):3–24, 1977.

[17] M. Kramer. The rising pandemic of mental disorders and associated chronic diseases and disabilities. *Acta Psychiotrica Scandinaveca*, 62:282–297, 1980.

[18] J. F. Fries. Aging, natural death, and the compression of morbidity. *Gerontological Perspecta*, 1:5–8, 1980.

[19] J. F. Fries. Aging, natural death, and the compression of morbidity. 1980. *Bulletin of the World Health Organization*, 80(3):245–250, 2002.

[20] K. G. Manton. Changing concepts of morbidity and mortality in the elderly population. *Milbank Mem Fund Q Health Soc*, 60(2):183–244, 1982.

[21] S. M. McKinlay and J.B. McKinlay. Examining trends in the nation's health., 1979.

[22] J. B. McKinlay and et al. *Mortality, Morbidity and the Inverse Care Law*, pages 99–138. Sage, Beverly Hills, 1983.

[23] V. A. Freedman, E. Crimmins, R. F. Schoeni, Brenda C. Spillman, H. Aykan, E. Kramarow, K. Land, J. Lubitz, K. Manton, L. G. Martin, et al. Resolving inconsistencies in trends in old-age disability: report from a technical working group. *Demography*, 41(3):417–441, 2004.

[24] V. A. Freedman, L. G. Martin, and R. F. Schoeni. Recent trends in disability and functioning among older adults in the united states: a systematic review. *Jama*, 288 (24):3137–3146, 2002.

[25] E. Crimmins and Y. Saito. Change in the prevalence of diseases among older americans: 1984-1994. *Demographic Research*, 3(9), 2000. 10.4054/DemRes.2000.3.9.

[26] E. M. Crimmins, Y. Saito, and S. L. Reynolds. Further evidence on recent trends in the prevalence and incidence of disability among older americans from two sources: the lsoa and the nhis. *J Gerontol B Psychol Sci Soc Sci*, 52(2):S59–71, 1997.

[27] V. A. Freedman and L. G. Martin. Understanding trends in functional limitations among older americans. *Am J Public Health*, 88(10):1457–62, 1998.

[28] K. G. Manton, X. Gu, and V. L. Lamb. Change in chronic disability from 1982 to 2004/2005 as measured by long-term changes in function and health in the u.s. elderly population. *Proc Natl Acad Sci U S A*, 103(48):18374–9, 2006.

[29] R. F. Schoeni, V. A. Freedman, and R. B. Wallace. Persistent, consistent, widespread, and robust? another look at recent trends in old-age disability. *J Gerontol B Psychol Sci Soc Sci*, 56(4):S206–18, 2001.

[30] G. Doblhammer and J. Kytir. Compression or expansion of morbidity? trends in healthy-life expectancy in the elderly austrian population between 1978 and 1998. *Soc Sci Med*, 52(3):385–91, 2001.

[31] J. Sagardui-Villamor, P. Guallar-Castillon, M. Garcia-Ferruelo, J. R. Banegas, and F. Rodriguez-Artalejo. Trends in disability and disability-free life expectancy among elderly people in spain: 1986-1999. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences*, 60(8):1028–1034, 2005.

[32] P. Graham, T. Blakely, P. Davis, A. Sporle, and N. Pearce. Compression, expansion, or dynamic equilibrium? the evolution of health expectancy in new zealand. *Journal of Epidemiology and Community Health*, 58(8):659–666, 2004.

[33] E. Cambois, A. Clavel, I. Romieu, and J. M. Robine. Trends in disability-free life expectancy at age 65 in france: consistent and diverging patterns according to the underlying disability measure. *European Journal of Ageing*, 5(4):287, 2008.

[34] J.M. Robine and J.P. Michel. Looking forward to a general theory on population aging. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 59 (6):M590–M597, 2004.

[35] A. Bebbington. The expectation of life without disability in england and wales. *Social Science and Medicine*, 27:321–326, 1988.

[36] S. Kelly, A. Baker, and S. Gupta. Healthy life expectancy in great britain, 1980–96, and its use as an indicator in united kingdom government strategies. *Health Statistics Quarterly*, 7:32–37, 2000.

[37] A. Bebbington and R. Darton. Healthy life expectancy in england and wales: recent evidence. 1996.

[38] M.R. Bone, A.C. Bebbington, C. Jagger, K. Morgan, and G Nicolaas. Health expectancy and its uses. *HMSO London*, 1995.

[39] C. Jagger. Trends in life expectancy and healthy life expectancy. future of an ageing population project: evidence review. Report, Foresight, UK Government Office for Science, 2015.

[40] C. Jagger, F. E. Matthews, P. Wohland, T. Fouweather, B. C. M. Stephan, L. Robinson, A. Arthur, and C. Brayne. A comparison of health expectancies over two decades in england: results of the cognitive function and ageing study i and ii. *The Lancet*, 387(10020):779–786, 2015.

[41] B. Pongiglione, B. L. De Stavola, and G. B. Ploubidis. A systematic literature review of studies analyzing inequalities in health expectancy among the older population. *PLOS ONE*, 10(6):e0130747, 2015.

[42] B. Pongiglione, B. L. De Stavola, H. Kuper, and G. B. Ploubidis. Disability and all-cause mortality in the older population: evidence from the english longitudinal study of ageing. *European journal of epidemiology*, 31(8):735–746, 2016.

[43] B. Pongiglione, G. B. Ploubidis, and B. L. De Stavola. Levels of disability in the older population of england: Comparing binary and ordinal classifications. *Disability and Health Journal*, 2017.

[44] B Pongiglione, G. B. Ploubidis, B. L. De Stavola, and H Kuper. Op20using a latent disability structure to explain mortality among the older population. *Journal of Epidemiology and Community Health*, 69(Suppl 1):A17–A18, 2015.

[45] E.M. Crimmins, M.D. Hayward, A. Hagedorn, Y. Saito, and N. Brouard. Change in disability-free life expectancy for americans 70 years old and older. *Demography*, 46 (3):627–646, 2009.

[46] E. M. Crimmins, M. D. Hayward, and Y. Saito. Differentials in active life expectancy in the older population of the united states. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 51B(3):S111–S120, 1996.

[47] L. Cai and J. Lubitz. Was there compression of disability for older americans from 1992 to 2003? *Demography*, 44(3):479–495, 2007.

[48] World Health Organization and World Bank. *World report on disability*. World Health Organization, Geneva, Switzerland, 2011.

[49] A. Harris and S. Enfield. Disability, equality and human rights. *A training manual for development and humanitarian organisations, Oxfam*, 2003.

[50] T. Shakespeare and N. Watson. The social model of disability: an outdated ideology? In *Exploring theories and expanding methodologies: Where we are and where we need to go*, pages 9–28. Emerald Group Publishing Limited, 2001.

[51] L. M Verbrugge and A. M Jette. The disablement process. *Social science & medicine*, 38(1):1–14, 1994.

[52] S. Nagi. *Some conceptual issues in disability and rehabilitation*, pages 100–113. American Sociological Association, Washington, DC, 1965.

[53] World Health Organization. *International classification of impairments, disabilities, and handicaps : a manual of classification relating to the consequences of disease*. World Health Organization, Geneva, 1980.

[54] World Health Organization et al. *International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease; publ. for trial purposes in accordance with resolution WHA29. 35 for the Twenty-ninth World Health Assembly, May 1976*. na, 1980.

[55] T. B. Ustun, S. Chatterji, N. Kostansjek, and J. Bickenbach. Who's icf and functional status information in health records. *Health Care Financing Review*, 24(3):77–88, 2003.

[56] WG Washington Group on Disability Statistics (WG). Development of an internationally comparable disability measure for censuses, 2008. URL `http://www.cdc.gov/nchs/data/washington_group/meeting8/NSO_report.pdf`.

[57] A. Cieza, T. Brockow, T. Ewert, E. Amman, B. Kollerits, S. Chatterji, T. B. Ustun, and G. Stucki. Linking health-status measurements to the international classification of functioning, disability and health. *J Rehabil Med*, 34(5):205–10, 2002.

[58] International Centre for Evidence in Disability (ICED). Measuring disability in surveys and programs: A summary, 2014. URL `http://disabilitycentre.lshtm.ac.uk`.

[59] I. Mactaggart, H. Kuper, G. V. S. Murthy, J. Oye, and S. Polack. Measuring disability in population based surveys: The interrelationship between clinical impairments and reported functional limitations in cameroon and india. *PLoS ONE*, 11(10):e0164470, 2016.

[60] D. Mont. Measuring health and disability. *The Lancet*, 369(9573):1658, 2007.

[61] D. Wolf. *Chapter 4 - Late-Life Disability Trends and Trajectories*, pages 77–99. Academic Press, San Diego, 2016.

[62] L. I. Iezzoni, E. P. McCarthy, R. B. Davis, and H. Siebens. Mobility problems and perceptions of disability by self-respondents and proxy respondents. *Med Care*, 38 (10):1051–7, 2000.

[63] A. M. Jette. Toward a common language of disablement. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 64(11):1165–1168, 2009.

[64] A. Sonnega, J. D. Faul, M. B. Ofstedal, K. M. Langa, J. W. R. Phillips, and D. R. Weir. Cohort profile: the health and retirement study (hrs). *International Journal of Epidemiology*, 43(2):576–585, 2014.

[65] R. Taylor, L. Conway, L. Calderwood, C. Lessof, H. Cheshire, K. Cox, and S. Scholes. Technical report (wave 1): health, wealth and lifestyles of the older population in england: the 2002 english longitudinal study of ageing. Report, National Centre for Social Research, 2007.

[66] S. Seaman, J. Galati, D. Jackson, and J. Carlin. What is meant by missing at random? *Statist. Sci.*, 28(2):257–268, 05 2013. doi: 10.1214/13-STS415. URL `http://dx.doi.org/10.1214/13-STS415`.

[67] C. A. Nathanson. Illness and the feminine role: a theoretical review. *Social Science and Medicine*, 9(2):57–62, 1975.

[68] L. M. Verbrugge. Sex differentials in health. *Public Health Reports*, 97(5):417–37, 1982.

[69] F. F. Chen, S. G. West, and K. H. Sousa. A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, 41(2):189–225, 2006.

[70] A. R. Omran. The epidemiologic transition: A theory of the epidemiology of population change. *The Milbank Quarterly*, 83(4):731–757, 2005.

[71] B. S. Sanders. Measuring community health levels. *American Journal of Public Health and the Nation's Health*, 54:1063–70, 1964.

[72] D. Sullivan. A single index of mortality and morbidity. *HSMHA Health Reports*, 86: 347354, 1971.

[73] J.M. Robine, C. Jagger, C. D. Mathers, E. M. Crimmins, and R. M Suzman. *Determining health expectancies*. John Wiley & Sons, 2003.

[74] R.J.M. Perenboom, H. van Oyen, and M. Mutafova. *Chapter 18 - Health Expectancies in European Countries*, pages 359–376. John Wiley & Sons, Ltd, 2003.

[75] J.M. Robine and I. Romieu. Health expectancies in the european union: Progress achieved. *REVES paper*, 319, 1998.

[76] V. A. Freedman, D. A. Wolf, and B. C. Spillman. Disability-free life expectancy over 30 years: A growing female disadvantage in the us population. *Am J Public Health*, 106(6):1079–85, 2016.

[77] E. M. Crimmins, Y. Zhang, and Y. Saito. Trends over 4 decades in disability-free life expectancy in the united states. *Am J Public Health*, 106(7):1287–93, 2016.

[78] L. Sundberg, N. Agahi, J. Fritzell, and S. Fors. Trends in health expectancies among the oldest old in sweden, 19922011. *The European Journal of Public Health*, 2016. doi: 10.1093/eurpub/ckw066.

[79] C. Jagger, A. J. Arthur, N. A. Spiers, and M. Clarke. Patterns of onset of disability in activities of daily living with age. *Journal of the American Geriatrics Society*, 49 (4):404–409, 2001.

[80] E. N. Lazaridis, M. A. Rudberg, S. E. Furner, and C. K. Cassel. Do activities of daily living have a hierarchical structure? an analysis using the longitudinal study of aging. *Journal of gerontology*, 49(2):M47–M51, 1994.

[81] D. D. Dunlop, S. L. Hughes, and L. M. Manheim. Disability in activities of daily living: patterns of change and a hierarchy of disability. *American journal of public health*, 87(3):378–383, 1997.

[82] W. D. Spector and J. A. Fleishman. Combining activities of daily living with instrumental activities of daily living to measure functional disability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 53(1):S46–S57, 1998.

[83] P. Barberger-Gateau, C. Rainville, L. Letenneur, and J.F. Dartigues. A hierarchical model of domains of disablement in the elderly: a longitudinal approach. *Disability and rehabilitation*, 22(7):308–317, 2000.

[84] M. J. Field, A. Jette, et al. *The future of disability in America*. National Academies Press, 2007.

[85] J. Ormel, F. V. Rijsdijk, M. Sullivan, E. Van Sonderen, and G. Kempen. Temporal and reciprocal relationship between iadl/adl disability and depressive symptoms in late life. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(4):P338–P347, 2002.

[86] C. Jagger, R. Matthews, F. Matthews, T. Robinson, J.M. Robine, C. Brayne, et al. The burden of diseases on disability-free life expectancy in later life. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(4):408–414, 2007.

[87] M. T. Puts, D. J. Deeg, N. Hoeymans, W. J. Nusselder, and F. G. Schellevis. Changes in the prevalence of chronic disease and the association with disability in the older dutch population between 1987 and 2001. *Age and ageing*, 37(2):187–193, 2008.

[88] K. Peres, C. Verret, A. Alioum, and P. Barberger-Gateau. The disablement process: factors associated with progression of disability and recovery in french elderly people. *Disability and rehabilitation*, 27(5):263–276, 2005.

[89] E. Cambois, C. Jagger, W. Nusselder, H. Van Oyen, and J.M. Robine. International comparisons of disability prevalence estimates: Impact of accounting or not accounting for the institutionalized population. In *International Measurement of Disability*, pages 207–229. Springer, 2016.

[90] S. L. Reynolds, Y. Saito, and E. M. Crimmins. The impact of obesity on active life expectancy in older american men and women. *The Gerontologist*, 45(4):438–444, 2005.

[91] S. Al Snih, K. J. Ottenbacher, K. S. Markides, Y.F. Kuo, K. Eschbach, and J. S. Goodwin. The effect of obesity on disability vs mortality in older americans. *Archives of internal medicine*, 167(8):774–780, 2007.

[92] S.a Walter, A. Kunst, J. Mackenbach, A. Hofman, and H. Tiemeier. Mortality and disability: the effect of overweight and obesity. *International Journal of Obesity*, 33 (12):1410–1418, 2009.

[93] S. L. Reynolds and J. M. McIlvane. The impact of obesity and arthritis on active life expectancy in older americans. *Obesity*, 17(2):363–369, 2009.

[94] W. Johnson, L. Li, D. Kuh, and R. Hardy. How has the age-related process of overweight or obesity development changed over time? co-ordinated analyses of individual participant data from five united kingdom birth cohorts. *PLoS Med*, 12(5):e1001828; discussion e1001828, 2015.

[95] J. M. Guralnik and L. Ferrucci. The challenge of understanding the disablement process in older persons commentary responding to Jette AM. toward a common language of disablement. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, page glp094, 2009.

[96] K. Imai and S. Soneji. On the estimation of disability-free life expectancy: Sullivan' method and its extension. *J Am Stat Assoc*, 102(480):1199–211, 2007.

[97] C. D. Mathers and J. M. Robine. How good is sullivan's method for monitoring changes in population health expectancies? *J Epidemiol Community Health*, 51(1): 80–6, 1997.

[98] D. A. Wolf and T. M. Gill. Modeling transition rates using panel current-status data: how serious is the bias? *Demography*, 46(2):371–86, 2009.

[99] D. A. Wolf, V. A. Freedman, J. I. Ondrich, C. L. Seplaki, and B. C. Spillman. Disability trajectories at the end of life: A "countdown" model. *J Gerontol B Psychol Sci Soc Sci*, 70(5):745–52, 2015.

[100] C. Jagger, B. Cox, Roy S.L., and THE EHEMU TEAM. Health expectancy calculation by the sullivan method: a practical guide, 3rd edition. *EUROPEAN HEALTH EXPECTANCY MONITORING UNIT, France*, 2007.

[101] J. A. F. Rushby and K. Hanson. Calculating and presenting disability adjusted life years (dalys) in cost-effectiveness analysis. *Health Policy and Planning*, 16(3):326–331, 2001.

[102] D. Donev, L. Zaletel-Kragelj, V. Bjegovic, and G. Burazeri. *Measuring the burden disease: Disability adjusted life year (DALY)*, volume 2, book section Measuring the burden disease: Disability adjusted life year (DALY)., pages 393–416. Jacobs Verlag, Lage, Germany, 2013. URL `www.mf.uni-lj.si`.

[103] G. B. Ploubidis, A. Sullivan, M. Brown, and A. Goodman. Psychological distress in mid-life: evidence from the 1958 and 1970 british birth cohorts. *Psychol Med*, pages 1–13, 2016.

[104] K. M. Keyes, R. L. Utz, W. Robinson, and G. Li. What is a cohort effect? comparison of three statistical methods for modeling cohort effects in obesity prevalence in the united states, 1971-2006. *Soc Sci Med*, 70(7):1100–8, 2010.

[105] L. Li, R. Hardy, D. Kuh, R. Lo Conte, and C. Power. Child-to-adult body mass index and height trajectories: a comparison of 2 british birth cohorts. *Am J Epidemiol*, 168 (9):1008–15, 2008.

[106] C. J. Murray. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull World Health Organ*, 72(3):429–45, 1994.

[107] M. Cerniauskaite, R. Quintas, C. Boldt, A. Raggi, A. Cieza, J. E. Bickenbach, and M. Leonardi. Systematic literature review on icf from 2001 to 2009: its use, implementation and operationalisation. *Disability and rehabilitation*, 33(4):281–309, 2011.

[108] P. De Vriendt, M. Lambert, and T. Mets. Integrating the international classification of functioning, disability and health (icf) in the geriatric minimum data set-25 (gmds-25) for intervention studies in older people. *Journal of Nutrition, Health and Aging*, 13(2):128–34, 2009.

[109] C. Bjrkenstam, K. Alexanderson, E. Bjrkenstam, C. Lindholm, and E. Mittendorfer-Rutz. Diagnosis-specific disability pension and risk of all-cause and cause-specific

mortality a cohort study of 4.9 million inhabitants in sweden. *BMC Public Health*, 14:1247, 2014.

[110] V. L. Forman-Hoffman, K. L. Ault, W. L. Anderson, J. M. Weiner, A. Stevens, V. A. Campbell, and B. S. Armour. Disability status, mortality, and leading causes of death in the united states community population. *Med Care*, 53(4):346–54, 2015.

[111] I. M. Majer, W. J. Nusselder, J. P. Mackenbach, and A. E. Kunst. Socioeconomic inequalities in life and health expectancies around official retirement age in 10 western-european countries. *Journal of Epidemiology and Community Health*, 65(11):972–979, 2011.

[112] A. Barford, D. Dorling, G. Davey Smith, and M. Shaw. Life expectancy: women now on top everywhere. *Bmj*, 332, 2006.

[113] A. Gjonça, C. Tomassini, B. Toson, and S. Smallwood. Sex differences in mortality, a comparison of the united kingdom and other developed countries. *Health Statistics Quarterly*, 26(2):6–16, 2005.

[114] S. M. Lynch, J. S. Brown, and M. G. Taylor. *Demography of Disability*, pages 567–582. Springer Netherlands, 2009.

[115] M. Palmer and D. Harley. Models and measurement in disability: an international review. *Health Policy and Planning*, 27(5):357–64, 2011.

[116] D. A. Wolf, K. Hunt, and J. Knickman. Perspectives on the recent decline in disability at older ages. *The Milbank Quarterly*, 83(3):365–395, 2005.

[117] V. A. Freedman. Implications of asking "ambiguous" difficulty questions: An analysis of the second wave of the asset and health dynamics of the oldest old study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55(5): S288–S297, 2000.

[118] A. M. Jette. How measurement techniques influence estimates of disability in older populations. *Social Science and Medicine*, 38(7):937–942, 1994.

[119] N. Spiers, C. Jagger, and M. Clarke. Physical function and perceived health: cohort differences and interrelationships in older people. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 51(5):S226–S233, 1996.

[120] A. Oksuzyan, K. Juel, J. W. Vaupel, and K. Christensen. Men: good health and high mortality. sex differences in health and aging. *Aging clinical and experimental research*, 20(2):91, 2008.

[121] A. Oksuzyan, I. Petersen, H. Stovring, P. Bingley, J. W. Vaupel, and K. Christensen. The male–female health–survival paradox: a survey and register study of the impact of sex-specific selection and information bias. *Annals of epidemiology*, 19(7):504–511, 2009.

[122] N. Bossuyt, S. Gadeyne, P. Deboosere, and H. Van Oyen. Socio-economic inequalities in health expectancy in belgium. *Public health*, 118(1):3–10, 2004.

[123] E. Cambois, J.M. Robine, and M. D. Hayward. Social inequalities in disability-free life expectancy in the french male population, 1980–1991. *Demography*, 38(4):513–524, 2001.

[124] E. M. Crimmins and Y. Saito. Trends in healthy life expectancy in the united states, 1970–1990: gender, racial, and educational differences. *Social science & medicine*, 52 (11):1629–1641, 2001.

[125] R. Perenboom, L. M. van Herten, H. C. Boshuizen, and G. van Den Bos. Life expectancy without chronic morbidity: trends in gender and socioeconomic disparities. *Public health reports*, 120(1):46–54, 2005.

[126] K. Juel, J. Sørensen, and H. Brønnum-Hansen. Supplement: Risk factors and public health in denmark. *Scandinavian journal of public health*, 36(1_suppl):1–227, 2008.

[127] H. Van Oyen, W. Nusselder, C. Jagger, P. Kolip, E. Cambois, and J.M. Robine. Gender differences in healthy life years within the eu: an exploration of the health–survival paradox. *International journal of public health*, 58(1):143–155, 2013.

[128] C. Jagger. Compression or expansion of morbidity–what does the future hold? *Age and ageing*, 29(2):93–94, 2000.

[129] Office for National Statistics. Statistical bulletin: Health expectancies at birth and age 65 in the united kingdom, 2009-11. Report, Office for National Statistics, 2014.

[130] J. N. Newton, A. D. Briggs, C. J. Murray, D. Dicker, K. J. Foreman, and et al. Wang, H. Changes in health in england, with analysis by english regions and areas of deprivation, 1990-2013: a systematic analysis for the global burden of disease study 2013. *Lancet*, 386(10010):2257–74, 2015.

[131] R. D. Peng. Reproducible research and biostatistics. *Biostatistics*, 10(3):405–8, 2009.

[132] E. M. Crimmins, Y. Saito, and D. Ingegneri. Changes in life expectancy and disability-free life expectancy in the united states. *Population and development review*, pages 235–267, 1989.

[133] C. Jagger, C. Weston, E. Cambois, H. Van Oyen, W. Nusselder, G. Doblhammer, J. Rychtarikova, J.M. Robine, et al. Inequalities in health expectancies at older ages in the european union: findings from the survey of health and retirement in europe (share). *Journal of Epidemiology and Community Health*, pages jech–2010, 2011.

[134] M. E. Bravell, S. H. Zarit, and B. Johansson. Self-reported activities of daily living and performance-based functional ability: a study of congruence among the oldest old. *European Journal of Ageing*, 8(3):199–209, 2011.

[135] P. Kivinen, R. Sulkava, P. Halonen, and A. Nissinen. Self-reported and performance-based functional status and associated factors among elderly men: the finnish cohorts of the seven countries study. *J Clin Epidemiol*, 51(12):1243–52, 1998.

[136] C. L. van den Brink, M. Tijhuis, S. Kalmijn, N. S. Klazinga, A. Nissinen, S. Giampaoli, P. Kivinen, D. Kromhout, and G. A. van den Bos. Self-reported disability and its association with performance-based limitation in elderly men: a comparison of three european countries. *J Am Geriatr Soc*, 51(6):782–8, 2003.

[137] A. Nichols et al. Causal inference with observational data: Regression discontinuity and related methods in stata. In *North American Stata Users' Group Meetings 2007*, number 2. Stata Users Group, 2007.

[138] C. S. Mackenzie, W. L. Gekoski, and V. J. Knox. Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. *Aging Ment Health*, 10(6):574–82, 2006.

[139] F. I. Matheson, K. L. W. Smith, G. S. Fazli, R. Moineddin, J. R. Dunn, and R. H. Glazier. Physical health and gender as risk factors for usage of services for mental illness. *Journal of Epidemiology and Community Health*, 2014.

[140] A. E. Thompson, Y. Anisimowicz, B. Miedema, W. Hogg, W. P. Wodchis, and K. Aubrey-Bassler. The influence of gender and other patient characteristics on health care-seeking behaviour: a qualicopc study. *BMC Family Practice*, 17:38, 2016.

[141] T. C. Antonucci. *A life-span view of womens social relations.*, pages 239–269. Sage Publications, Thousand Oaks, 1994.

[142] R. Fuhrer and S. A. Stansfeld. How gender affects patterns of social relations and their impact on health: a comparison of one or multiple sources of support from "close persons". *Soc Sci Med*, 54(5), 2002.

[143] A. M. Jette, S. M. Haley, and J. T. Kooyoomjian. Are the icf activity and participation dimensions distinct? *Journal of Rehabilitation Medicine*, 35(3):145–149, 2003.

[144] M. F. Schuntermann. The implementation of the international classification of functioning, disability and health in germany: experiences and problems. *International Journal of Rehabilitation Research*, 28(2):93–102, 2005.