VOLUME 3 | ISSUE 3 | **OCTOBER 2012** ISSN 1757-9597

Longitudinal and Life Course Studies: International Journal

Published by

SLLS Society for Longitudinal and Life Course Studies

Inside this issue

Longview
Promoting Longitudinal and Life Course Studies

Co-sponsors:

VU university medical center

ISER INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

EDITORIAL BOARD

Open Journal System

The ***LLCS*** journal is produced using the Open Journal System, which is part of the Public Knowledge Project. OJS is open source software made freely available to journals worldwide, for the purpose of making open access publishing a viable option for more journals. Already, more than 8,000 journals around the world use this software.

Copyright Notice

AUTHOR GUIDELINES SUMMARY

Submission of Papers

All papers, written in the English language, should be submitted via the *LLCS* website as a Microsoft Word file. If there is a good reason why this is not possible, authors are requested to contact the Journal Manager **before** submitting the paper. All subsequent processes involving the author are carried out electronically via the website.

Preparation of Texts

Length. The paper should normally be **approximately 5000 words**, with longer papers also accepted up to a maximum of 7000 words. This word count excludes tables, figures and bibliography.

Font and Line Spacing. Please use Calibri (or similar sans serif) font, 12pt, with 1.5 line spacing in all main text, single space in figure or table captions.

Page Layout. All text should be justified to the left hand margin (all margins of at least 2.5cm) with no indents at the start of paragraphs and a line space separating paragraphs. All headings and sub-headings used within the body of the paper should also be left justified. Please do **NOT** use automatic text formatting.

Weblinks. To help our readers to look up cited references or other information available on the web, authors should ensure that all such references are activated.

Ensuring an Anonymous (Blind) Review. Please submit papers with a full detailed title page. Once a paper has been submitted to *LLCS* via the website, it will be 'anonymised' by the removal of all author name(s) and institution names on the title page, and any identifying electronic document properties will also be removed. Authors do not need to remove their name(s) from the main text or references but any reference to their work or themselves should be expressed in the third person.

Abstract. The abstract (no subheads or paragraphs) should be no more than 250 words (not part of the main word count).

Keywords. These should be included just below the author list (minimum 3 maximum 10).

Abbreviations. Words to be abbreviated should be spelt out in full the first time they appear in the text with the abbreviations in brackets. Thereafter the abbreviation should be used.

References. Please use ***APA version 6*** and refer to examples in the full Guidelines

Authors not complying with these reference guidelines will be asked to make all the necessary alterations themselves, if their paper is accepted for publication.

Notes. As a general rule, supplementary notes should be avoided, but if thought to be essential, they should *not* appear on the page as 'Footnotes', but instead, be included as Endnotes.

Supplementary material. Supplementary material may be uploaded with the submission, and if the paper is published, will be visible to the reader via a link on the RHS of the screen.

Submissions containing graphs, tables, illustrations or mathematics. All graphs, tables and illustrations should be embedded in the submitted text, and have clear, self-explanatory titles and captions. Mathematical expressions should be created in Word 2003 (or a compatible package) with equation editor 3.0, unless the author has good reason to use other software, in which case please contact the Journal Manager. All biological measures should be reported in SI units, as appropriate, followed, in the text, by traditional units in parentheses.

Author Citation. If the paper is accepted for publication, a version of the paper with all authors cited in full on the title page will be used. Only individuals who have contributed substantially to the production of the paper should be included.

Copy Editing. All accepted manuscripts are subject to copy editing, with reference back to author with suggested edits and queries for response. Particular care should be taken by the author to check references.

Proofs. The corresponding author will be asked to view a layout proof of the article on the website and respond with final amendments within 3 days.

(full Author Guidelines at: http://www.journal.longviewuk.com/index.php/llcs/about/submissions#authorGuidelines)

CONTENTS

Published by the Society for Longitudinal and Life Course Studies
Institute of Education, 20 Bedford Way, London, WC1H 0AL
Email: info@slls.org.uk
http://www.slls.org.uk

Editorial: finance, access and impact

John Bynner Executive Editor

To establish the journal, the first publisher, Longview, was awarded a development grant from the Nuffield Foundation sufficient to meet infrastructure and running costs for its first three years. When the development grant ended in March 2011 continuation depended on moving responsibility for publication and accompanying costs to the newly established *Society for Longitudinal and Life Course studies* www.slls.org.uk

The Society's main source of income and assets is membership fees and as membership grows, the costs of such activities as the Journal – currently £14,000 for three issues per annum - are increasingly likely to be met. The journal's Business Plan suggested that three years at least were needed for membership to reach such a 'steady state'. In the meantime an appeal to research centres affiliated to the society to join with SLLS in co-sponsoring the journal generated grants of £10-15,000, scheduled to taper down to zero by the end of the period. The Department of Psychiatry, Free University (VU), Amsterdam, led by Jan Smit and the Institute for Social and Economic research at the University of Essex led by Heather Laurie took up the offer.

We are most grateful to our co-sponsors for the support they have been giving the journal, which has been vital to sustaining it during the post-development phase. But because the Society has built more slowly towards self-sufficiency than the business plan optimistically predicted, there has been, and will continue to be for some time, a funding gap to make up. Acting on the advice of the journal's Editorial Committee, the Society's solution is to introduce the very modest charge for all readers, other than Editorial Board and Society members, of £20 per annum (US dollar and Euro equivalents) for registration to read the journal, starting in December 2012. If you are an author or reviewer for the journal you will of course continue to have free access to the paper(s) for which you are responsible and there are no publication charges. All papers will become fully accessible 12 months after publication.

The other route to free access is joining the Society itself – currently £65 per annum, and £25 or equivalent for doctoral students. Apart from access to current and past issues SLLS membership brings many additional benefits. These include a much reduced registration rate for the Society's annual conference - this year in Paris, next year in Amsterdam, a bimonthly newsletter containing articles and updates on the progress of the Society and the journal, notice of conferences and training activities and opportunities for networking and communications locally and internationally in between conferences through the Society's 13 chapters led by 'Global Representatives' in different world regions. 'Corporate membership' by organisations and societies is also an option offering the opportunity for reduced fees for five nominated individuals within the organisation and other collective benefits.

Another matter raised in the May 2012 issue, which will be a major topic at this year's SLLS annual conference 29-31 October, is that of impact ratings for the journal. The view is growing in some quarters that the virtual monopoly in this area of academic evaluation by one commercial company, Thomson Reuters, is not the ideal model that we should be signing up to. There are other opportunities for impact ratings such as Google Analytics and Scopus Elsevier (SNIP) that some argue will increasingly take over. Others maintain that publishing in journals with the T-R rating, while it exists, is vital for young scientists' careers. We look forward to a lively debate. If you are not able to attend the conference, please email me with your views.

This issue of the journal displays the variety of its content and longitudinal resources deployed to achieve scientific aims and the journal's different modes of presentation. We start with three Research Papers devoted respectively to 'parental education and adult health outcomes' (in Sweden), 'effects of deleterious ageing' (in Scotland) and the 'effects of marital status transitions on depression' (in the USA). Two Research Notes follow respectively on 'life course influences on quality of life at age 50 years' and 'parental separation and adult psychological distress' (both in the UK). Following in the footsteps of the Michigan *Panel Study of Income Dynamics (PSID)* profiled in the last issue, we then have a Study Profile of the *Household, Income and Labour Dynamics in Australia* (*HILDA*). Finally there is a review of the first twenty years 'Life History' of the pioneering UK Avon Longitudinal Survey of Parents and Children (ALSPAC) from the perspectives of 20 of the key actors who were involved at various stages. The source is an edited record of the Wellcome Trust's fortieth *History of Twentieth Century Medicine Witness Seminar.*

Longitudinal and Life Course Studies 2012 Volume 3 Issue 3 Pp 298 – 311 *ISSN 1757-9597*

Parental education and adult health outcomes: a cohort study examining disease-specific effects of education levels using Swedish nationwide registries across two generations

Hendrik Koffijberg
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht
h.koffijberg@umcutrecht.nl
Johanna Adami
Clinical epidemiology unit, Department of Medicine Solna, Karolinska Institutet, Stockholm
Erik Buskens
Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen
Mårten Palme
Department of Economics, Stockholm University, Stockholm

(Received August 2011 Revised April 2012)

Abstract

It is well known that children with less educated parents have inferior health status in later life. There are two competing hypotheses explaining the association found: the pathway hypothesis – suggesting that children from low educated households are more likely to obtain less education themselves, which, in turn, is associated with inferior health outcomes – and the life course hypothesis – suggesting that living conditions during childhood, as such affected by parental education level, is important for the formation of adult health status. We obtained data from National Swedish registries comprising health outcomes of individuals born between 1940 and 1949, and the Swedish Multi-generation Register. We assessed the differences in risk of hospital admission for individuals with low and high parental education as well as low and high own education. We found that for higher educated individuals, high parental education is associated with even better health outcomes: having a high versus low educated mother or high versus low educated father was associated with an overall decrease in the risk (hazard rate) of hospital admission by 5% (95% CI 0.91-0.98) and 3% (95% CI 0.95-0.99), respectively. This indicates that children from a relatively disadvantaged background, signaled by lower parental educational attainment, are more likely to continue accumulating risk throughout life. Even if they have higher qualifications they may still have a greater accumulation of risk, compared with other highly qualified children from a less disadvantaged background. We found that this effect is primarily attributed to circulatory diseases, and would appear to support the life course hypothesis. We conclude that parental education and ensuing early childhood or even fetal living conditions have a persistent effect on adult health.

Keywords: inter-generational; education; life course study; pathway hypothesis

Introduction

It has been established that children with less educated parents have inferior health status in their childhood and later in life compared to children with high educated parents (Case, Lubotsky & Paxson, 2002; Gakidou et al., 2010). Parental education may influence adult health only indirectly, i.e. as a result of the fact that children with low educated parents are themselves likely to obtain lower final educational levels, in turn associated with inferior health outcomes. This indirect effect is called the *pathway hypothesis. (*Marmot et al., 2001). Conversely, the inferior health may also be explained through a lasting, or accumulated, effect of inferior living conditions during childhood and adolescence. This direct effect is called the *life course hypothesis (*Kuh et al., 2003). The life course hypothesis can be further divided into (i) *the accumulation*, (ii) *the critical period* and (iii) *social mobility* hypotheses (Mishra et al., 2009).

Although evidence in support of the *life course hypothesis* exists, this evidence is mainly based on observations of child health, or adolescent health (Kuh et al., 2003). In studies on health outcomes at relatively young ages, a higher prevalence of diseases among children in families with lower socio-economic status was found (Case et al., 2002). Also, disease severity was worse among these families. Given that children born in families with low socio-economic position (SEP) have inferior health, it is then assumed that health problems at a young age are also related to poor health outcomes later in life. However, directly linking parental education to actual health outcomes later in life remains difficult, due to the long latency period in both cardiovascular disease and cancer (Gluckman et al., 2008).

In this study we evaluated the relation between parental education and subsequent health in later life as actually observed, i.e. hospital admission and death occurring in individuals older than 50 years. To this end, we used extensive registry data on a large population-based cohort with long term follow-up and information about potential confounders. Our data includes International Classification of Diseases (ICD) 9 or 10 codes for all hospital admissions in Sweden for the period 1987-2006 and data from the cause of death register. Thus, the usual extrapolation from child health to adult health could be avoided.

We studied the interaction between own education and paternal as well as maternal education on the health outcomes. This framework allowed us to test the pathway against the life course hypotheses. The life course hypothesis would predict an independent effect of parental education conditional on own education. We also discussed to what extent the results may be attributed to each of the three life course hypotheses mentioned above.

Since the available data did not allow us to identify the causal effect of education on health outcomes – i.e. we were unable to separate out the effect of education itself from all possible confounding factors – we limited the study to view education as a marker of socio-economic position.

Method

Data selection

In our study, we included all individuals born in Sweden between 1940 and 1949 and obtained the identities of these individuals from the population census. Information on date and region of birth as well as gender for the index persons and their parents, was collected from the population census (Statistics Sweden, 2009a). The linkage between the two generations was achieved through the Swedish Multi-generation Register (Statistics Sweden, 2009b).

Health

Data on all hospital admissions between 1987 and 2005 were collected from the Swedish National Patient Register (Socialstyrelsen, 2009a). This register contains administrative information such as date of admission, number of days in hospital care, as well as discharge diagnoses classified according to ICD 9 or ICD 10 codes. This data did not include out-patient visits and day surgery. Data on the occurrence of co-morbidities was limited and likely to be less accurate than data on the main disease of interest. Therefore, we did not use co-morbidity data in our analysis.

Information on the cause of death, for deceased index persons, was obtained from the National Cause of Death Registry for the period 1987-2006 (ICD 9 and ICD 10 codes) (Socialstyrelsen, 2009b). Data on all hospital admissions due to cancer were obtained from the Swedish Cancer Registry (Socialstyrelsen, 2009c). This register covers the whole population of Sweden and comprises all cancer diagnoses recorded, among others, as ICD 7

codes. The completeness of the registries has previously been shown to be very high (Barlow et al., 2009; Socialstyrelsen, 2009a; Socialstyrelsen, 2009b; Socialstyrelsen, 2009c).

Education

Data on highest educational level for the index persons was collected from the 1990 records in the Swedish Register of Education (Statistics Sweden, 2004). For the parents, the corresponding information was obtained from the 1970 census (Statistics Sweden, 2009a). The original education levels were re-coded into low education level, i.e. compulsory education up to vocational schooling, and high education level, i.e. upper secondary education up to doctoral education, for both index persons and parents. This categorization was required as the use of multiple education levels would result in a too large number of combinations of education of the index person, the father, and the mother. Moreover, only very few of the parents attained high education levels, which would result in many combinations with very few index persons and parents if more than two education levels were used.

Income and immigration

Data on taxable income from labour in 1990 was obtained for the index persons from the National Tax Register (Statistics Sweden, 2009c). The population census was used to collect information on migration status.

Definition of the cohort

In order to have accumulated a substantial medical history, the selected index persons should be as old as possible, e.g. around the age of 70 in 1990. However, parental education levels were only available for parents born from 1912 onwards. Therefore, we selected individuals born in the period 1940-1949, who had parents born from 1912 onwards, as index persons. Subsequently we assessed the medical history of the index persons starting from the age of 50 years. As a result, any disease histories resulting in death before the age of 50 were not taken into account. For each index person we retrieved all relevant hospital admissions and, if the index person had died, the cause of death. The risk of hospital admission related to cancer was ascertained twice, once using the Swedish National Patient Register and once using the Swedish Cancer Registry.

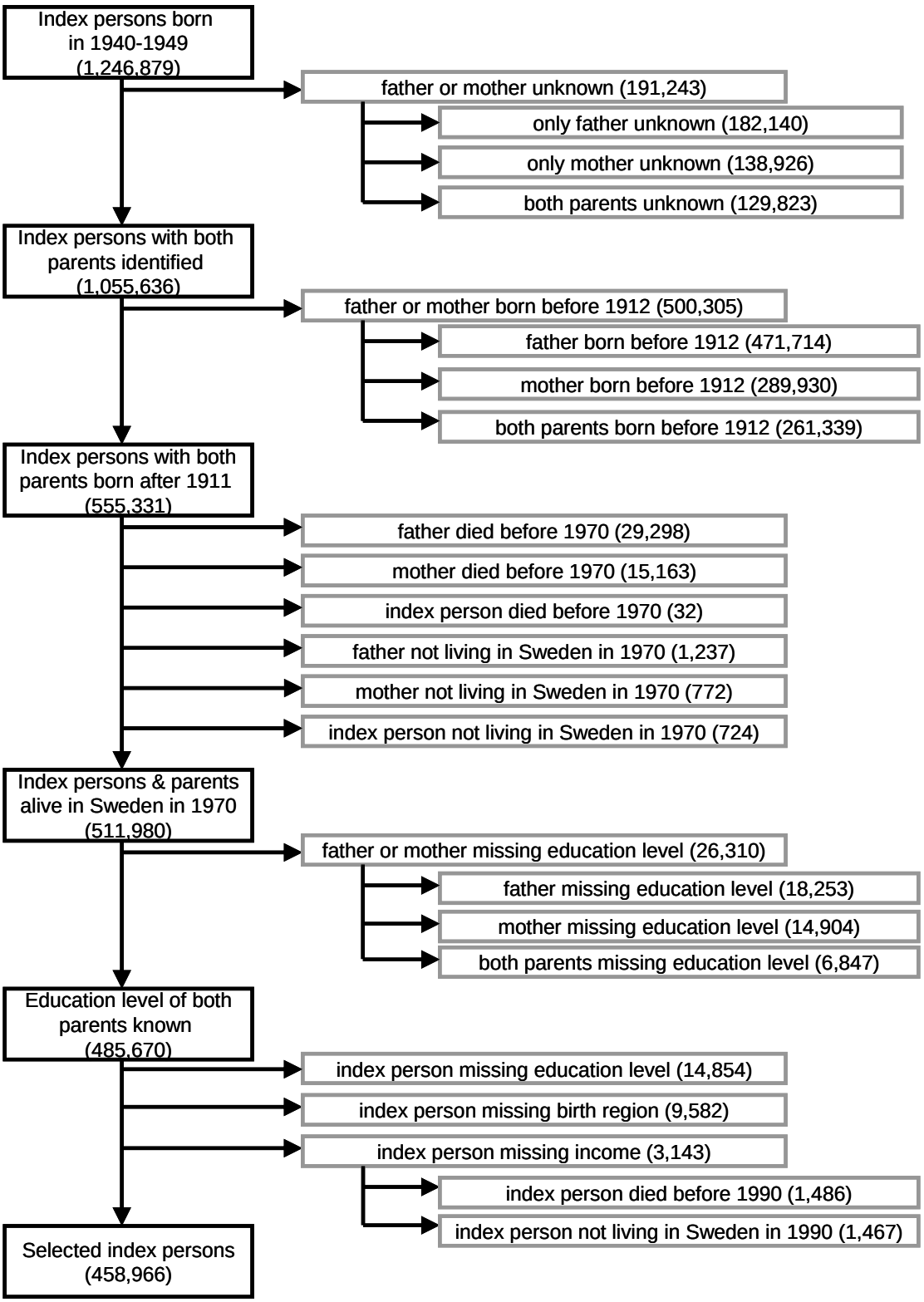
Statistical analyses

First, we estimated differences in incidence rates of hospital admission for index persons with low and with high education. Second, we estimated incidence rate ratios for the effects of high versus low maternal education level, and high versus low paternal education level, separately for index persons with low and high education level. Homogeneity of these groups of index persons was assessed with the Mantel–Haenszel test (Kirkwood and Sterne, 2003). Finally, we used Cox proportional hazard competing risk models (Kirkwood and Sterne, 2003). The dependent variable was time to the event under study, with the event defined as the primary diagnosis for hospital admission or death of the index person. For hospital admission, time to first hospital visit beyond the age of 50 years, caused by the indicated diagnosis was used. For mortality, the time to death by diagnosis group for index persons dying after the age of 50 years was used. For the analysis of mortality this means that individuals dying from another diagnosis than the considered type are treated as right censored at the date of death. Variables in the Cox model explaining the time to event were: education of the mother (high/low), education of the father (high/low), income of the index person (tertiles), sex, year of birth of the index person (1940/41, 1942/43, …, 1948/49), period of birth of the mother (1912/21, 1922/31), period of birth of the father (1912/21, 1922/31), adoption (yes/no), and region of birth in Sweden (northern, central, southern). Index persons with low and high education were analyzed separately. All analyses were performed with STATA 9.2 SE (StataCorp, Texas, USA).

Results

Figure 1 shows the flowchart for the data retrieval and selection process. Data was available on all relevant variables for 458,966 index persons and their parents after data selection. A large number of index persons had to be excluded, due to missing data on parental education levels. This selection accounted for 97% of all excluded index persons, and was mainly caused by one parent or both parents dying before the 1970 census in which data on education levels was collected.

Figure 1. Flowchart of the data retrieval and selection process



***Notes**. The majority of missing education levels for parents or index persons may be caused by death of the individuals prior to 1952. Education levels were determined in the 1970 census which was mandatory for all Swedish residents below the age of 65, i.e. born after 1911. However, death before 1970 cannot always be ascertained since the cause of death register was started on 1/1/1952. Parents and index persons dying before 1952 have no date of death, and are therefore assumed to be still alive in 1970 and 1990. These individuals will cause missing education levels and missing income data.*

Table 1. Characteristics of the selected index persons and their parents

		Index persons (n = 458,966)	
Gender			
	Male	225,057	(49.0%)
	Female	233,909	(51.0%)
Adopted			
	Yes	448,770	(97.8%)
	No	10,196	(2.2%)
Year of birth			
	1940-1941	36,973	(8.1%)
	1942-1943	64,425	(14.0%)
	1944-1945	97,999	(21.4%)
	1946-1947	122,306	(26.6%)
	1948-1949	137,263	(29.9%)
Region of birth			
	Northern	65,675	(14.3%)
	Central	175,886	(38.3%)
	Southern	217,405	(47.4%)
Monthly income (SEK) from the 1990 census			
	5th percentile	0	
	25th percentile	947	
	Median	1,448	
	75th percentile	1,908	
	95th percentile	3,000	
	Mean	1,469	
	SD	982	
Follow up (in years after the age of 50)			
	5th percentile	7.2	
	25th percentile	8.5	
	Median	10.4	
	75th percentile	12.5	
	95th percentile	15.5	
	Mean	10.7	
	SD	2.7	
	Sum	4,895,446	

		Index persons (n = 458,966)	
Education of the index person			
	Low	337,064	(73.4%)
	High	121,902	(26.6%)
Education of the mother			
	Low	444,691	(96.9%)
	High	14,275	(3.1%)
Education of the father			
	Low	396,090	(86.3%)
	High	62,876	(13.7%)
Lowest parental education level			
	Low	448,531	(97.7%)
	High	10,435	(2.3%)
Highest parental education level			
	Low	392,250	(85.5%)
	High	66,716	(14.5%)
Hospital admissions during follow up*			
	None	369,110	(80.4%)
	1	76,316	(16.6%)
	2	11,711	(2.6%)
	3	1,673	(0.4%)
	4+	156	(0.0%)
Admitted to the hospital during follow up*			
	No	369,110	(80.4%)
	Yes	89,856	(19.6%)
Admitted to the hospital during follow up #			
	No	250,705	(54.6%)
	Yes	208,261	(45.4%)
Death during follow up*			
	No	446,364	(97.3%)
	Yes	12,602	(2.7%)
Death during follow up #			
	No	443,734	(96.7%)
	Yes	15,232	(3.3%)

* Due to selected disease categories only
\# Due to any cause

Notes. *Low education level was defined as compulsory education up to vocational education. High education level was defined as upper secondary education up to doctoral education.*

Table 1 shows descriptive statistics of the final dataset. In general, education of the father was higher than that of the mother. Furthermore, education of the index-persons was in general higher than parental education levels. Because the average age at the end of follow up was around 61 years, more than half of the index persons (54.6%) had not been hospitalized, and only 3.3% had died during follow up. In total 4.9 million years of follow up were available. The effect of education of the index persons on the risk of hospital admission is shown in Table 2. In this table, all rate differences are positive; indicating that a low education level of the index persons increased their risk of hospital admission, for all disease categories considered.

Table 2. Differences in incidence rates of hospital admission for low and high educated index persons

	Index persons with low education (n = 337,064)		**Index persons with high education** (n = 121,902)		**Rate difference low vs high education** (x 10,000)		**Rate ratio low vs high education**	
	Events	***PY***	***Events***	***PY***	***Estimate***	***95% CI***	***Estimate***	***95% CI***
All hospital admissions combined	159,769	2,608,416	48,460	993,770	125	(120 - 130)	1.26	(1.24 - 1.27)
Cancer [140-208 / C00-C97]	19,911	3,543,764	6,447	1,258,696	5	(3 - 6)	1.10	(1.07 - 1.13)
Respiratory diseases [460-519 / J00-J99]	13,712	3,541,744	3,193	1,265,559	13	(12 - 15)	1.53	(1.48 - 1.60)
Diabetes [250 / E10-E14]	3,912	3,591,958	781	1,277,846	5	(4 - 5)	1.78	(1.65 - 1.93)
Chronic liver disease [571 / K70,K73-K74]	1,093	3,609,462	228	1,281,152	1	(1 - 2)	1.70	(1.47 - 1.97)
Mental disorders [303,305 / F10]	5,987	3,577,571	1,128	1,275,655	8	(7 - 9)	1.89	(1.78 - 2.02)
Circulatory diseases [390-459 / I00-I99]	37,986	3,415,613	11,005	1,226,142	21	(19 - 23)	1.24	(1.21 - 1.27)
- Cerebrovascular disease [430-438 / I60-I69]	7,409	3,580,231	2,013	1,273,204	5	(4 - 6)	1.31	(1.25 - 1.38)
- Ischemic heart disease [410-414 / I20-I25]	15,667	3,532,206	4,092	1,261,839	12	(11 - 13)	1.37	(1.32 - 1.42)
- Cerebrovascular or ischemic heart disease	22,204	3,496,604	5,915	1,252,279	16	(15 - 18)	1.34	(1.31 - 1.38)
- Other circulatory diseases	15,782	3,532,380	5,090	1,255,937	4	(3 - 5)	1.10	(1.07 - 1.14)

***Notes**. The observed number of hospital admissions and person-years for index persons with low and high education level. For each disease category the rate difference and rate ratio for low versus high educated index persons is given, along with the corresponding 95% confidence intervals.*
PY = person-years; CI = confidence interval.

To investigate whether the effects of maternal and paternal education level are similar in index persons with low and with high education, crude effects were estimated, as ratios of incidence rates, and are shown in Tables 3A-B. From these tables we see that, in general, high education of the mother and high education of the father reduced the risk of hospital admission, in both index persons with low and with high education themselves. However, for many disease categories, this reduction is more pronounced in individuals with high education level. Also, high maternal education appeared to be slightly more beneficial than high paternal education in most disease categories. The Mantel-Haenzel test indicated that the two subgroups of index persons were not homogeneous for several disease categories. In our subsequent, adjusted analyses, based on the Cox model, we therefore stratified by education level of the index persons.

Table 3. Crude effects of parental education level for index persons with low and high education

[A]

High vs low education of the mother	**Index persons with low education** (n = 337,064)		**Index persons with high education** (n = 121,902)		**M-H test of homogeneity** (p-value)	
	IRR	***95% CI***	***IRR***	***95% CI***		
All hospital admissions combined	0.92	(0.87 - 0.96)	0.91	(0.88 - 0.94)	0.74	
Cancer [140-208 / C00-C97]	0.99	(0.85 - 1.13)	1.04	(0.95 - 1.14)	0.51	
Respiratory diseases [460-519 / J00-J99]	0.79	(0.65 - 0.95)	0.80	(0.70 - 0.92)	0.92	
Diabetes [250 / E10-E14]	0.85	(0.59 - 1.20)	0.69	(0.50 - 0.93)	0.34	
Chronic liver disease [571 / K70,K73-K74]	0.72	(0.31 - 1.42)	0.72	(0.39 - 1.23)	1.00	
Mental disorders [303,305 / F10]	1.06	(0.81 - 1.35)	0.69	(0.53 - 0.89)	0.02	*
Circulatory diseases [390-459 / I00-I99]	0.84	(0.75 - 0.93)	0.73	(0.67 - 0.78)	0.03	*
- Cerebrovascular disease [430-438 / I60-I69]	1.01	(0.80 - 1.27)	0.74	(0.61 - 0.88)	0.03	*
- Ischemic heart disease [410-414 / I20-I25]	0.71	(0.58 - 0.85)	0.60	(0.52 - 0.69)	0.17	
- Cerebrovascular or ischemic heart disease	0.80	(0.68 - 0.92)	0.65	(0.58 - 0.73)	0.03	*
- Other circulatory diseases	0.91	(0.77 - 1.07)	0.82	(0.74 - 0.92)	0.32	

[B]

High vs low education of the father	**Index persons with low education** (n = 337,064)		**Index persons with high education** (n = 121,902)		**M-H test of homogeneity** (p-value)	
	IRR	***95% CI***	***IRR***	***95% CI***		
All hospital admissions combined	0.97	(0.96 - 0.99)	0.94	(0.92 - 0.96)	0.02	*
Cancer [140-208 / C00-C97]	1.06	(1.01 - 1.12)	1.00	(0.95 - 1.06)	0.12	
Respiratory diseases [460-519 / J00-J99]	0.96	(0.90 - 1.02)	0.89	(0.82 - 0.96)	0.14	
Diabetes [250 / E10-E14]	0.92	(0.81 - 1.03)	0.82	(0.69 - 0.96)	0.27	
Chronic liver disease [571 / K70,K73-K74]	1.02	(0.81 - 1.28)	0.97	(0.72 - 1.31)	0.78	
Mental disorders [303,305 / F10]	1.05	(0.96 - 1.16)	0.88	(0.77 - 1.01)	0.03	*
Circulatory diseases [390-459 / I00-I99]	0.89	(0.86 - 0.93)	0.85	(0.82 - 0.89)	0.11	
- Cerebrovascular disease [430-438 / I60-I69]	0.93	(0.85 - 1.02)	0.81	(0.73 - 0.90)	0.04	*
- Ischemic heart disease [410-414 / I20-I25]	0.82	(0.77 - 0.88)	0.80	(0.74 - 0.86)	0.48	
- Cerebrovascular or ischemic heart disease	0.86	(0.82 - 0.91)	0.81	(0.76 - 0.86)	0.10	
- Other circulatory diseases	0.95	(0.89 - 1.00)	0.91	(0.86 - 0.97)	0.44	

Notes. *The incidence rate ratio of hospital admissions corresponding to high versus low education of the mother [A], and the father [B], separately for low and high educated index persons, along with the corresponding 95% confidence intervals. Also, the p-value of the Mantel-Haenzel test for homogeneity between subgroups of individuals with low and high education is given. IRR = incidence rate ratio; CI = confidence interval; M-H = Mantel-Haenzel; * = p < 0.05*

Tables 4A-B show the hazard rates from the Cox models for high versus low parental education level, separately for hospital admission and death. An interaction term for education of the mother and education of the father was initially included in the Cox model. However, this term was not significant in our analyses and subsequently dropped from the model.

Hospital admission and death for low educated index persons

From Table 4A we see that high maternal education level was associated with a 5% lower risk of hospital admission in general, but not with a lower risk of death for low educated individuals. Although the risk of hospital admission due to ischemic heart disease was lower, this was not the case for the risk of death due to ischaemic heart disease. Furthermore, the risk of hospital admission due to cerebrovascular disease was increased by 12%, although not significantly, and the risk of death due to cerebrovascular disease was increased by 91% as well (significantly).

High paternal education level was significantly associated with a 7% increased risk of hospital admission due to cancer and a 15% increased risk of admission due to mental disorders, and with decreased hospital admission due to all types of circulatory diseases, most notably a 10% reduced risk of hospital admission due to ischaemic heart disease. Furthermore, high paternal education level was significantly associated with a 13% increased risk of death in general, specifically a 10% increased risk of death due to cancer and a 59% increased risk of death due to mental disorders.

Table 4. Effects of parental education level for index persons with low and high education, obtained from the Cox proportional hazard competing risk model, and adjusted for confounders. [A]

Index persons with low education ***(n = 337,064)***		**High vs low education of the mother**		**High vs low education of the father**	
	Events	***HR***	***95% CI***	***HR***	***95% CI***
All hospital admissions combined	159,769	0.95	(0.90 - 0.99)	1.00	(0.98 - 1.02)
Cancer [140-208 / C00-C97]	19,911	0.99	(0.86 - 1.14)	1.07	(1.01 - 1.12)
Respiratory diseases [460-519 / J00-J99]	13,712	0.84	(0.70 - 1.02)	1.02	(0.95 - 1.09)
Diabetes [250 / E10-E14]	3,912	0.91	(0.65 - 1.28)	1.00	(0.88 - 1.13)
Chronic liver disease [571 / K70,K73-K74]	1,093	0.72	(0.35 - 1.45)	1.11	(0.89 - 1.39)
Mental disorders [303,305 / F10]	5,987	0.98	(0.76 - 1.26)	1.15	(1.04 - 1.26)
Circulatory diseases [390-459 / I00-I99]	37,986	0.92	(0.83 - 1.03)	0.95	(0.92 - 0.99)
- Cerebrovascular disease [430-438 / I60-I69]	7,409	1.12	(0.89 - 1.41)	0.99	(0.91 - 1.08)
- Ischemic heart disease [410-414 / I20-I25]	15,667	0.80	(0.67 - 0.97)	0.90	(0.85 - 0.96)
- Cerebrovascular or ischemic heart disease	22,204	0.89	(0.77 - 1.03)	0.93	(0.88 - 0.98)
- Other circulatory diseases	15,782	0.96	(0.82 - 1.14)	0.99	(0.93 - 1.05)
Above hospital admissions combined	69,826	0.92	(0.85 - 0.99)	1.02	(0.99 - 1.05)
Death due to all causes	12,453	1.12	(0.94 - 1.33)	1.13	(1.05 - 1.20)
Cancer [140-208 / C00-C97]	6,215	1.01	(0.79 - 1.31)	1.10	(1.01 - 1.21)
Respiratory diseases [460-519 / J00-J99]	438	0.52	(0.13 - 2.13)	0.93	(0.63 - 1.37)
Diabetes [250 / E10-E14]	268	0.83	(0.46 - 1.50)	0.94	(0.58 - 1.54)
Chronic liver disease [571 / K70,K73-K74]	388	1.43	(0.62 - 3.28)	1.12	(0.77 - 1.61)
Mental disorders [303,305 / F10]	287	1.57	(0.63 - 3.92)	1.59	(1.07 - 2.35)
Circulatory diseases [390-459 / I00-I99]	2,695	1.21	(0.83 - 1.76)	1.08	(0.93 - 1.25)
- Cerebrovascular disease [430-438 / I60-I69]	620	1.91	(1.01 - 3.64)	0.86	(0.62 - 1.20)
- Ischemic heart disease [410-414 / I20-I25]	1,446	0.78	(0.41 - 1.46)	1.17	(0.96 - 1.42)
- Cerebrovascular or ischemic heart disease	2,066	1.10	(0.70 - 1.73)	1.07	(0.90 - 1.27)
- Other circulatory diseases	629	1.53	(0.78 - 3.01)	1.12	(0.83 - 1.51)
Above causes of death combined	10,291	1.06	(0.87 - 1.29)	1.10	(1.03 - 1.19)

[B]

Index persons with high education (n = 121,902)		High vs low education of the mother		High vs low education of the father	
	Events	*HR*	*95% CI*	*HR*	*95% CI*
All hospital admissions combined	48,460	0.95	(0.91 - 0.98)	0.97	(0.95 - 0.99)
Cancer [140-208 / C00-C97]	6,447	1.05	(0.96 - 1.15)	0.99	(0.94 - 1.05)
Respiratory diseases [460-519 / J00-J99]	3,193	0.88	(0.76 - 1.01)	0.94	(0.87 - 1.02)
Diabetes [250 / E10-E14]	781	0.83	(0.60 - 1.13)	0.91	(0.76 - 1.08)
Chronic liver disease [571 / K70,K73-K74]	228	0.76	(0.43 - 1.34)	1.10	(0.81 - 1.49)
Mental disorders [303,305 / F10]	1,128	0.78	(0.60 - 1.02)	1.00	(0.87 - 1.15)
Circulatory diseases [390-459 / I00-I99]	11,005	0.83	(0.77 - 0.91)	0.95	(0.90 - 0.99)
- Cerebrovascular disease [430-438 / I60-I69]	2,013	0.88	(0.73 - 1.07)	0.90	(0.81 - 1.00)
- Ischemic heart disease [410-414 / I20-I25]	4,092	0.73	(0.63 - 0.85)	0.93	(0.86 - 1.00)
- Cerebrovascular or ischemic heart disease	5,915	0.78	(0.70 - 0.88)	0.92	(0.86 - 0.98)
- Other circulatory diseases	5,090	0.90	(0.80 - 1.01)	0.98	(0.92 - 1.05)
Above hospital admissions combined	20,026	0.90	(0.85 - 0.96)	0.96	(0.92 - 0.99)
Death due to all causes	2,782	0.97	(0.84 - 1.13)	0.97	(0.89 - 1.06)
Cancer [140-208 / C00-C97]	1,565	0.96	(0.78 - 1.17)	0.89	(0.79 - 1.01)
Respiratory diseases [460-519 / J00-J99]	53	0.56	(0.13 - 2.45)	0.87	(0.44 - 1.70)
Diabetes [250 / E10-E14]	42	1.18	(0.34 - 4.14)	0.91	(0.43 - 1.95)
Chronic liver disease [571 / K70,K73-K74]	60	0.17	(0.02 - 1.29)	1.49	(0.86 - 2.60)
Mental disorders [303,305 / F10]	48	0.88	(0.30 - 2.59)	1.75	(0.94 - 3.24)
Circulatory diseases [390-459 / I00-I99]	543	1.02	(0.71 - 1.45)	0.90	(0.73 - 1.11)
- Cerebrovascular disease [430-438 / I60-I69]	131	0.79	(0.37 - 1.68)	1.11	(0.74 - 1.66)
- Ischemic heart disease [410-414 / I20-I25]	268	1.61	(1.02 - 2.54)	0.76	(0.56 - 1.04)
- Cerebrovascular or ischemic heart disease	399	1.29	(0.87 - 1.90)	0.87	(0.68 - 1.11)
- Other circulatory diseases	144	0.37	(0.13 - 1.04)	0.99	(0.67 - 1.47)
Above causes of death combined	2,311	0.94	(0.79 - 1.11)	0.92	(0.84 - 1.02)

Notes. *Hazard rates and their 95% confidence intervals for hospital admission or death due to selected disease categories, for subgroups of index persons with a low [A] or high education level [B], with parents with low or high education level. For each disease category the corresponding ICD 9 and ICD 10 codes are given, separated with a slash. Note that for hospital admissions the events in the separate disease categories do not add up to the events in 'combined' categories because patients may be admitted for multiple diseases sequentially and the 'combined' categories only account for the first admission of any patient.*

Hospital admission and death for high educated index persons

Table 4B shows that a high level of education for either the mother or the father was associated with a 3-5% significantly decreased risk of hospital admission in general for index persons with high education. In particular, there was an association between high parental education and a reduced risk of hospital admission for circulatory diseases. However, this association was no longer observable for the corresponding risk of death. Moreover, high maternal education was significantly associated with a 61% increased risk of death due to ischaemic heart disease. Finally, high paternal education level was no longer significantly associated with an increased risk of admission or death due to mental disorders, as it was for index persons with low education.

Cancer related hospital admission

The hazard rates for hospital admission due to cancer, and the association with high education levels of the mother and father of the index persons, are shown in Tables 5A-B. Table 5A shows the results for low educated index persons and

Table 5B shows the results for high educated index persons. In general, parental education level was not associated with substantially lower or higher risk of admission due to cancer. However, high paternal education level was associated with a marginally significant 5% increased risk of admission for any cancer type, and a 20% decreased risk of admission due to haematopoietic cancer for index persons with low education. Also, a high paternal education level may be linked to a 16% increased risk of admission for breast cancer in high educated women.

Table 5. Risk of hospital admission due to cancer related to parental education levels for index persons with low and high education, obtained from the Cox proportional hazard competing risk model, and adjusted for confounders

[A]

Index persons with low education (n = 337,064)		**High vs low education of the mother**		**High vs low education of the father**	
	Events	***HR***	***95% CI***	***HR***	***95% CI***
All admissions for cancer	25,437	1.06	(0.94 - 1.20)	1.05	(1.00 - 1.10)
Cancer - digestive system [150-159]	3,821	1.07	(0.77 - 1.48)	0.97	(0.85 - 1.09)
Cancer - respiratory system [160-164]	2,090	0.99	(0.63 - 1.55)	1.06	(0.90 - 1.24)
Breast cancer [170]	5,481	1.22	(0.96 - 1.55)	1.06	(0.97 - 1.17)
Prostate cancer [177]	3,070	1.00	(0.69 - 1.46)	1.13	(0.99 - 1.29)
Kidney cancer [180]	582	0.18	(0.03 - 1.32)	0.97	(0.71 - 1.34)
Cancer - urinary organs [181]	860	0.58	(0.24 - 1.40)	1.21	(0.95 - 1.54)
Cancer - nervous system [193]	850	0.70	(0.31 - 1.57)	1.11	(0.87 - 1.42)
Hematopoietic cancer [200-209]	1,550	1.47	(0.94 - 2.31)	0.80	(0.65 - 0.98)
Other cancer types [194-199]	1,547	1.08	(0.66 - 1.76)	1.13	(0.94 - 1.35)
Above cancers combined	19,355	1.04	(0.91 - 1.20)	1.04	(0.99 - 1.10)

[B]

Index persons with high education (n = 121,902)		**High vs low education of the mother**		**High vs low education of the father**	
	Events	*HR*	*95% CI*	*HR*	*95% CI*
All admissions for cancer	8,609	0.99	(0.91 - 1.07)	1.03	(0.98 - 1.08)
Cancer - digestive system [150-159]	1,189	1.03	(0.83 - 1.29)	1.01	(0.88 - 1.15)
Cancer - respiratory system [160-164]	392	0.89	(0.59 - 1.36)	0.93	(0.73 - 1.19)
Breast cancer [170]	1,668	0.98	(0.83 - 1.16)	1.16	(1.04 - 1.29)
Prostate cancer [177]	1,757	0.96	(0.79 - 1.17)	1.00	(0.90 - 1.12)
Kidney cancer [180]	160	0.94	(0.49 - 1.79)	0.98	(0.67 - 1.42)
Cancer - urinary organs [181]	271	0.96	(0.59 - 1.56)	1.06	(0.80 - 1.41)
Cancer - nervous system [193]	329	1.29	(0.89 - 1.89)	1.02	(0.79 - 1.32)
Hematopoietic cancer [200-209]	585	1.05	(0.77 - 1.42)	1.03	(0.85 - 1.25)
Other cancer types [194-199]	445	0.63	(0.42 - 0.97)	1.06	(0.85 - 1.32)
Above cancers combined	6,609	0.98	(0.89 - 1.07)	1.04	(0.99 - 1.11)

Notes. *Hazard rates and their 95% confidence intervals for hospital admission of death due to cancer, for subgroups of index persons with a low [A] or high education level [B], with parents with low or high education level. For each cancer type the corresponding ICD 7 codes are given. The risk of breast cancer was assessed for women only, whereas the risk of prostate cancer was assessed for men only. Note that for hospital admissions, the events in the separate categories do not add up to the events in the 'combined' category because patients may be admitted for multiple types of cancer sequentially and the combined category only accounts for the first admission of any patient.*

Discussion

Addressing the research questions

The main objective of the paper was to test the pathway hypothesis against the life course hypothesis. To reject the pathway hypothesis, it is sufficient to show that there is an independent effect of parental education on adult health status, not just an indirect effect through the individual's own educational attainments. In Table 4 we showed that there is an effect of parental education on hospitalization also after stratifying on own educational attainments. This result is shown for the group of low educated individuals in Table 4A and for high educated individuals in Table 4B. For both groups, the results suggest that the effect mainly can be attributed to circulatory diseases. Our interpretation of these results is that we can reject the pathway hypothesis and maintain the life course hypothesis.

As described in the Introduction, the life course hypothesis can be further divided into (i) *the accumulation*, (ii) *the critical period* and (iii) *social mobility* hypotheses (Mishra et al., 2009). Under the *accumulation hypothesis,* the effect of staying in a particular SEP increases with the time spent in the SEP, while according to the *critical period hypothesis,* health or vulnerability to disease is to a large extent determined during a particular period of an individual's life course. Several previous studies have pointed out the very early period in one's life as of particular interest (Barker, 1995; Barker, 1999).

The support for the life course hypothesis obtained in this study: the independent effect of parental SEP on health later in life is in line with the *critical period* hypothesis, since parental educational attainments are closely associated with the individual's SEP during childhood and adolescence. To what extent the effect of SEP on health is accumulated over the life cycle is beyond the scope of this study and cannot be inferred from our results.

Under the *social mobility hypothesis,* the social mobility in itself, has an effect on health outcomes. The result that high paternal education is associated with increased mortality in low educated individuals, as shown in Table 4A, supports this hypothesis. Our interpretation of this result is that individuals originating from homes with high educated fathers who do not acquire more than compulsory schooling themselves, are on average adversely selected. They are more likely to have initial characteristics, or characteristics acquired early on, that are associated with both lower probability to obtain more than compulsory schooling and higher mortality, compared to the group originating from homes with low educated fathers. The estimated effect is thus not causal, but rather due to selection (Haas & Fosse, 2008; Blumenshine *et al.*, 2008; Jackson, 2009).

Previous findings

A systematic review of 29 previous studies on the relation between childhood SEP and health outcomes later in life has showed an independent relation between childhood SEP and the outcomes of mortality in stomach cancer and haemorrhagic stroke (Galobardes, Lynch & Davey Smith, 2004). The reviewed studies also found a relation between childhood SEP together with SEP in adult life and mortality in coronary heart disease and lung cancer, as well as in respiratory-related deaths. An update of the review including 11 additional studies confirmed the inverse relation between childhood SEP and mortality for all but one of the reviewed studies (Galobardes, Lynch, & Davey Smith, 2008). Adjusting for adult SEP accounted to a varying degree for the association.

An inverse effect of SEP measured in childhood and adolescence was found on lifestyle factors related to mortality in coronary heart disease, all-cause mortality and cardiovascular mortality for a cohort of 49,323 Swedish males born 1949-1951 (Hemmingsson and Lundberg, 2005). The study found that a substantial part of socio-economic differences in cardiovascular decrease among middle-aged men can be attributed to SEP in early life. A persistent adverse effect of low childhood SEP on adult health was also found in a 1,000 cohort of 26 year old individuals born in New Zealand in 1972-1973, when investigating health outcomes, including dental health and substance abuse (Poulton et al., 2002). A long-lasting effect of parental SEP on several diseases, in particular mortality from cardiovascular disease, was found in a sample of about 5,500 men aged 35-64 years drawn from 27 workplaces in Scotland (Smith et al., 1997). Moreover, in a study on differences in mortality by cause of death between individuals originating from different SEP groups, an adverse effect of childhood SEP on mortality in stroke and stomach cancer in adulthood was found, when controlling for adult SEP (Smith *et al.*, 1998). Finally,

an excess all-cause mortality as well as mortality due to ischaemic heart disease was demonstrated in a cohort of Swedish men born in 1946-1955 for individuals with non-employed and manual labour parents (Vagero and Leon, 1994). These findings provide support for the cumulative life course hypothesis.

In contrast, support for the pathway hypothesis was found in a study on coronary heart disease incidence, chronic bronchitis and depression in a sample of about 10,000 UK civil servants aged 35-55 years (Marmot et al., 2001). Here, childhood SEP did not have an effect on adult health outcomes when controlling for the SEP of individual. A study on a sample of 2,636 Finnish men yielded similar results: individuals with low income were found to have an increased risk of all cause and cardiovascular mortality relative to high income men (Lynch et al., 1994). Finally, evidence from a schooling reform did not indicate a substantial, direct effect of parental education on child health (Lindeboom, Llena-Nozal, & van der Klaauw, 2009).

Strengths

In this study we used data on parental education, income and health outcomes for all individuals born in Sweden between 1940 and 1949. Our study extends several aspects of the previous literature, as our data allowed us to study actual endpoints for all diseases, rather than a specific subset of diseases occurring either later in life or during adulthood. Compared to other registry-based studies, our analyses are more comprehensive, as we used data on health outcomes for a general population, i.e. not restricted to any occupational group or group of workplaces, and were able to adjust for important and potentially confounding factors. Also, we studied the direct relation between health and education, rather than a relation based on social class or income. Finally, given our large cohort, we were able to effectively analyze health outcomes separately for children who move up and children who move down in educational levels. The sheer numbers allowed subgroup analyses for samples based on own educational attainments.

Limitations

Our study has certain limitations. First, we used the primary diagnosis codes for hospital admissions to detect hospital admission due to diseases. However, from our data it was not possible to distinguish admission for follow up examinations from admission due to the event itself. As a result, admissions found just after the age of 50 might have occurred for either a new first admission, e.g. for a myocardial infarction at the age of 51, or a follow up check after a myocardial infarction at the age of 49. However, in both cases, the event of interest had occurred and only the time to event might have been recorded incorrectly. Moreover, this can only occur during the first 1 or 2 years of follow up. Second, we used income level measured at a single point in time as a proxy for SEP. Although such a proxy may be imprecise, we obtained the income of the individuals when they were 40-50 years old, i.e. generally by the time their career would plateau and when their income should be relatively stable (Haider & Solon, 2006). Third, we performed a large number of tests on our data, which may lead to chance findings. For example, high maternal education was associated with an 82% lower risk (not significant) of admission for kidney cancer for index persons with low education level, and high paternal education level was linked to a 16% statistically significant increased risk of admission for breast cancer for high educated women. These may be chance findings. Alternatively, the first finding may be explained by a study showing that conditions in utero, reflected by birth weight, may affect the risk of renal cell cancer in adulthood (Bergstrom, Lindblad, & Wolk, 2001) whereas the second finding may be explained by a higher age at first childbirth and fewer children for high educated women, as well as a more stringent self-diagnosis and compliance with screening compared with low educated women (Heck and Pamuk, 1997; Hussain et al., 2008). Fourth, given our selection process, in which index persons with parents who died relatively young had to be excluded, the least healthy families may have been excluded from our analysis. This may have resulted in an underestimation of the effect of parental education on adult health. Hence, the risk of disease, and associated mortality, may be larger than indicated in our results for index persons with low versus high maternal or paternal education level. Finally, we assumed that within the period of our analysis, factors affecting illness behaviour and referral practices in Sweden did not change substantially. This may not hold for all diseases, but is more likely to hold for many of the severe and life-threatening diseases on which we focus in our analysis.

Policy implications

Our findings have implications for the evaluation of programmes aimed at improving education or reducing health inequalities, and for projections of future health care needs of populations. Improvements in population education levels may have health effects beyond the cohort attaining more education. Our results show that it may be misleading to use cross-sectional data to establish the relation between educational attainments and health when education is seen and used as marker of a life course disadvantage. At least some of the health differences are acquired early on and educational attainments are correlated across generations. This, in turn, implies that the estimated health effects of educational policies designed to increase the educational level of the population are likely to be biased when the persistent effects of parental education are ignored. This bias should be added to the bias originating from unobserved confounders that may plague observational studies on the effects of educational attainments on various individual outcomes.

Acknowledgements

The research has been funded by The Swedish Council for Social Research (FAS) and The Institute for Evaluation of Labour Market and Education Policy (IFAU).

References

Barker, D. J. (1995). Fetal origins of coronary heart disease. British Medical Journal, 311(6998), 171-74. Retrieved from: *http://www.ncbi.nlm.nih.gov/pubmed/7613432*

Barlow, L., Westergren, K., Holmberg, L., & Talback, M. (2009). The completeness of the Swedish Cancer Register: a sample survey for year 1998. *Acta Oncologica, 48*(1), 27-33. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/18767000

Bergstrom, A., Lindblad, P., & Wolk, A. (2001). Birth weight and risk of renal cell cancer. *Kidney International, 59*(3), 1110-1113. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11231367

Blumenshine, S. L., Vann, W. F. Jr., Gizlice, Z., & Lee, J. Y. (2008). Children's school performance: impact of general and oral health. *Journal of public health dentistry, 68*(2), 82-87. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/18221320

Case, A., Lubotsky, D., & Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review, 92*(5), 1308-34. Retrieved from: http://apps.webofknowledge.com/InboundService.do?SID=S2G2il9AD4bnBKh8cLh&product=WOS&UT=000180306100003&Init=Yes&action=retrieve&mode=FullRecord

Gakidou, E., Cowling, K., Lozano, R., & Murray, C. J. (2010). Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. *Lancet, 376*(9745), 959-74. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/20851260

Galobardes, B., Lynch, J. W., & Davey Smith, G. (2008). Is the association between childhood socio-economic circumstances and cause-specific mortality established? Update of a systematic review. *Journal of Epidemiology and Community Health, 62*(5), 387-390. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/18413449

Galobardes, B., Lynch, J. W., & Davey Smith, G. (2004). Childhood socio-economic circumstances and cause-specific mortality in adulthood: systematic review and interpretation. *Epidemiologic reviews, 26*, 7-21. Retrieved *from:* http://www.ncbi.nlm.nih.gov/pubmed/15234944

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *The New England Journal of Medicine, (1),* 61-73. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/18596274

Haas, S. A., & Fosse, N. E. (2008). Health and the educational attainment of adolescents: evidence from the NLSY97. Jou*rnal of Health and Social Behavior, 49*(2), 178-92. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/18649501

Haider, S., & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. American Economic Review, 96(4), 1308-20. Retrieved from: http://apps.webofknowledge.com/InboundService.do?SID=S2G2il9AD4bnBKh8cLh&product=WOS&UT=000240789200019&Init=Yes&action=retrieve&mode=FullRecord

Heck, K. E., & Pamuk, E. R. (1997). Explaining the relation between education and postmenopausal breast cancer. *American Journal of Epidemiology, 145*(4), 366-72. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9054241

Hemmingsson, T., & Lundberg, I. (2005). How far are socio-economic differences in coronary heart disease hospitalization, all-cause mortality and cardiovascular mortality among adult Swedish males attributable to negative childhood circumstances and behaviour in adolescence? *International Journal of Epidemiology*, *34*(2), 260-267. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/15333622

Hussain, S. K., Altieri, A., Sundquist, J., & Hemminki, K. (2008). Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. *International Journal of Cancer, 122*(1), 165-69. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/17708572

Jackson, M. I. (2009). Understanding links between adolescent health and educational attainment. Demography, *46*(4), 671-94. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/20084824

Kirkwood, B. R. & Sterne, J. A. C. (2003). *Essential Medical Statistics (2nd ed.).* York: Blackwell Science Ltd.

Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health, 57*(10), 778-83. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/14573579

Lindeboom, M., Llena-Nozal, A., & van der Klaauw, B. (2009). Parental education and child health: evidence from a schooling reform. *Journal of Health Economics, 28*(1), 109-31. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/18952306

Lynch, J. W., Kaplan, G. A., Cohen, R. D., Kauhanen, J., Wilson, T. W., Smith, N. L., & Salonen, J. T. (1994). Childhood and adult socio-economic status as predictors of mortality in Finland. *Lancet, 343*(8896), 524-27. Retrieved *from:* http://www.ncbi.nlm.nih.gov/pubmed/7906766

Marmot, M., Shipley, M., Brunner, E., & Hemingway, H. 2001. Relative contribution of early life and adult socio-economic factors to adult morbidity in the Whitehall II study. *Journal of Epidemiology and Community Health, 55*(5), 301-7. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/11297647

Mishra, G., Nitsch, D., Black, S., De Stavola, B., Kuh, D., & Hardy, R. 2009. A structured approach to modelling the effects of binary exposure variables over the life course. *International Journal of Epidemiology, 38*(2), 528-37. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/19028777

Poulton, R., Caspi, A., Milne, B. J., Thomson, W. M., Taylor, A., Sears, M. R., & Moffitt, T. E. (2002). Association between children's experience of socio-economic disadvantage and adult health: a life-course study. *Lancet, 360*(9346), 1640-1645. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/12457787

Smith, G. D., Hart, C., Blane, D., Gillis, C., & Hawthorne, V. (1997). Lifetime socio-economic position and mortality: prospective observational study. *British Medical Journal, 314*(7080), 547-52. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9055712

Smith, G. D., Hart, C., Blane, D., & Hole, D. (1998). Adverse socio-economic conditions in childhood and cause specific adult mortality: prospective observational study. *British Medical Journal, 316*(7145), 631-35. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9603744

Socialstyrelsen. (2009)a. *The Swedish National Patient Register. Inpatient diseases in Sweden 1987-2007*. Stockholm, Sweden,. Retrieved from: http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/inenglish

Socialstyrelsen. (2009b). *The Swedish Cause of Death Registry*. *Causes of death. Stockholm, Sweden*. Retrieved from: http://www.socialstyrelsen.se/register/dodsorsaksregistret

Socialstyrelsen. (2009)c. *The Swedish Cancer Registry. Cancer incidence in Sweden.* Stockholm, Sweden. Retrieved from: http://www.socialstyrelsen.se/register/halsodataregister/cancerregistret/inenglish

Statistics Sweden. (2004). *The Swedish Register of Education*. Stockholm, Sweden.
Retrieved from: http://www.scb.se/statistik/UF/UF0506/Produktbeskrivning_short_English_UF0506_20040101r.doc

Statistics Sweden. (2009a). *Population and Housing Census 1970.* Stockholm, Sweden. Retrieved from: http://www.scb.se/default____2154.aspx

Statistics Sweden. (2009b). *The Swedish Multigeneration Register*. Stockholm, Sweden. Retrieved from: http://www.scb.se/default____2154.aspx

Statistics Sweden. (2009c). *Longitudinal integration database for health insurance and labour market studies.* Stockholm, Sweden. Retrieved from: http://www.scb.se/Pages/List____257743.aspx

Vagero, D., & Leon, D. (1994). Effect of social class in childhood and adulthood on adult mortality. *Lancet, 343*(8907), 1224-25. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/7909885

Can we spot deleterious ageing in two waves of data? The Lothian Birth Cohort 1936 from ages 70 to 73

Wendy Johnson,[1,2,4] Alan J Gow,[1,4] Janie Corley,[4] Paul Redmond,[4] Ross Henderson[5], Catherine Murray,[4] John Starr,[1,3,4] and Ian J Deary[1,4]

[1]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, UK

[2]Department of Psychology, University of Minnesota – Twin Cities, USA

[3]Department of Geriatric Medicine, University of Edinburgh, UK

[4]Department of Psychology, University of Edinburgh, UK

[5]Edinburgh Medical School, Edinburgh, UK

wendy.johnson@ed.ac.uk

(Received April 2012 Revised August 2012)

Abstract

'Younger' old age (the late 60s through early 70s) is, for many, a period of stability of lifestyle and considerable freedom to pursue leisure activities. Despite the stability that many enjoy, the mortality rate is about 2% per year in western nations. This increases to about 5% by age 80. It would be useful to know if those most vulnerable can be identified through patterns of deleterious ageing, and especially if this could be accomplished with just two waves of data. The Lothian Birth Cohort 1936 was surveyed on a host of individual difference variables including cognition, personality, biomarkers of physical health, and activities at ages 70 and 73 years. Overall, the group showed the expected basic stability in mean levels for these variables, but some individuals had died and others did show substantial changes that could be considered statistically reliable. These presumably reliable changes were at least as likely to be positive (reflecting improved condition/ability) as negative (reflecting decline/ageing). Moreover, limitations in the estimated reliabilities of the measures meant that most of the observed changes could not be considered reliable. The changes clustered only weakly around general health to predict death over the next approximately two years. We concluded that two waves of longitudinal data were not sufficient to assess meaningful patterns of ageing, despite often being used to do so.

Keywords: cognitive ability, health, longitudinal data, ageing, mortality

Introduction

'Younger' old age (the late 60s through early 70s) is, for many in developed economies, a period of stability of lifestyle and general condition. Though far from universally true, many are in good health, experience few restrictions in activities, enjoy the newfound freedoms of recent occupational retirement and absence of financial responsibility for offspring, and are only beginning to tap their old-age financial resources. Despite this stability and generally favourable picture, however, at age 70 the mortality rate in western nations is about 2% per year, with male mortality running about 50% higher than female (http://www.mortality-trends.org/, 16 January, 2012, based on World Health Organisation data). The rate increases to about 5% per year by age 80. This means that about 22% of the age-70 population dies between ages 70 and 80.

Unfortunately, few in western nations are fortunate enough to experience good health until just prior to death; many go through some period of illness and/or increasing disability in the years prior to death, a period that often involves substantial reduction in quality of life, inability to live independently, and extensive medical care with its associated costs. Given these facts, the ability to identify those most likely to die within the next 5-10 years in an apparently healthy group in this younger range of old age could be very useful in developing methods to minimize these periods of illness and disability and perhaps even to extend longevity.

This idea has long been discussed with respect to cognitive ageing, or the normative declines in many cognitive functions that accompany old age. Beginning with observations in the late 1950s and early 1960s, that those who survived and agreed to be retested at some later point in time had generally showed higher original cognitive test scores than those who either refused testing or died in the interim, gerontologists have speculated that perhaps the observed declines in mean function with age result primarily from the presence in elderly samples of individuals who do not survive long beyond assessment (Jarvik & Falek, 1963; Lieberman, 1966; Rabbitt et al., 2011; Riegel & Riegel, 1972). That is, it is possible that decline in function is minimal until some overt disruption takes place, after which decline is quite rapid and ends in death. Even samples screened for clinical manifestations of impairment would inevitably include some who had entered this period but remained undiagnosed, and their numbers could be expected to increase with age, which could create the declines in mean function with age that studies consistently show. As longitudinal data samples have proliferated and statistical analytical techniques have improved, research efforts have been directed toward describing individual trajectories of decline (e.g., Finkel et al., 2005; McGue & Christensen, 2002) and/or the specific interval before death at which some period of 'terminal decline' begins (e.g., Rabbitt et al., 2011; Sliwinski et al., 2006; Terrera et al., 2011; Wilson et al., 2003). These studies have produced widely varying results about the extent of change, its rate of acceleration if considered, and the length of any terminal decline interval.

There are likely many features of study design that contribute to variation in results, including differences in sample selectivity, differences in measures assessed and the rates of normative rates of change to which they may be subject, unacknowledged constraints on results imposed by the models used, and the large age ranges sampled in many studies. In addition, three inter-related realities complicate even the most optimal study design. These apply to all measures of aging, whether cognitive, physiological, or behavioral. First, at any age, there are large individual differences in function. They stem from diverse sources, but many of the most salient ones show considerable rank-order stability throughout the lifespan. For example, those within an ageing cohort who perform relatively well on cognitive tasks or show relatively high levels of physical fitness, would have tended to do so in youth as well, had they been assessed then. This is borne out by data from the few studies for which such information is available. For example, the correlation between IQ scores at ages 11 and 79 in the Lothian Birth Cohort 1921 was .66; with correction for restriction of range in the sample, it rose to .73 (Deary et al., 2004). Without clear recognition in study design of the stability of inter-individual variability, this stability can easily be mistaken for intra-individual change, especially in cross-sectional studies or longitudinal studies with large sample age ranges in relation to the follow-up periods (Sliwinski, Hoffman, & Hofer, 2010). The second reality is that, over the timespans of most longitudinal studies, average rates of decline in function are small in relation to lifespan-stable individual differences in level of function. And third, individual differences in rates of decline are also small in relation to individual differences in lifespan-stable level. These latter realities act to keep statistical power low to detect rates of change accurately.

The problem of estimating rates of change or intervals of terminal decline is further complicated by the strong likelihood that, in addition to individual differences in linear rates of decline, there is also meaningful variability in rates of acceleration of decline with age and/or intervals of terminal decline. Most modeling of terminal decline intervals has been based on the assumption that there is one uniform interval of decline for all (Sliwinski et al., 2006), and violations of this assumption may especially bias estimates of individual differences in rates of change before and after the beginning of the supposedly uniform terminal decline interval. One very likely

possibility is that any terminal decline intervals vary at least with the pathology finally involved in death. Evidence for this was provided recently (Rabbitt et al., 2011).

Links among physical, cognitive, and psychosocial characteristics in ageing

Gerontological research over the past 10 years or so has increasingly suggested that ageing is a rather general systemic process. That is, degrees of well-being in many aspects of life; including physiological processes and biomarkers, health as measured both by self-assessment, and more objective criteria such as clinical diagnoses and medication, cognitive function, and psychosocial affect; all tend to be correlated, in level and possibly also rate of change (e.g., Backman & MacDonald, 2006; Deary et al., 2011; Dixon, 2011; Dolcos et al., 2012; Johnson et al., 2009; Li & Lindenberger, 2002). This repeated observation alone suggests that, examination of a wide range of potential risk and protective factors in the same group of individuals, may be worthwhile in helping to identify individuals entering final stages of ageing that portend some period of acute disability or impairment ending in death.

Theoretical considerations also point to the value of examining a wide range of potential factors involved in ageing. There are several reasons that many otherwise disparate aspects of function may be linked, particularly in old age. First, many chronic physical illnesses common in old age, including cardiovascular disease and diabetes (e.g., Cosway et al., 2001; Hassing et al., 2004; Rafnsson et al., 2007; Schram et al., 2007) also undermine cognitive function, possibly because they increase proinflammatory and oxidative stress markers and impede vascular function. These conditions are often aggravated by failure to adhere to somewhat detailed treatment regimens, and such failure is more common when cognitive function is impaired (Deary et al., 2009). As well, the chronic nature of these conditions and associated disabilities can contribute to reduction in quality of life, leading to depression (Fiske, Wetherell, & Gatz, 2009; Kendler et al., 2009). Second, lifetime-stable cognitive ability can contribute to the development and maintenance of lifestyle and habits such as nutrition, exercise, smoking, and drinking that support or undermine physical health. Evidence for this comes from studies (e.g. Adler & Snibbe, 2003; Hart et al., 2003) that have found that lower IQ scores in childhood were associated with smoking and other unhealthy habits in later life as well as with greater morbidity (Batty, Deary & Macintyre, 2007; Batty et al., 2007; Deary, Weiss & Batty, 2010). Also, lower IQ scores have also been robustly associated with poorer living circumstances that can contribute to poor psychological wellbeing and depression (Adler & Snibbe, 2003; Gallo & Matthews, 2003). Finally, there may be individual differences in some form of biological or constitutional 'integrity' and/or ageing processes that contribute to both physical and cognitive function as well as ability to sustain psychological wellbeing (Christensen et al., 2001; Gale et al., 2009; Li & Lindenberger., 2002). Clearly, these possibilities are not mutually exclusive.

Measuring change in evaluating ageing processes in two data waves

The simplest way to evaluate ageing processes is through analysis of differences among individuals of different ages. Early studies of ageing made it clear that this is of limited value, however, due to the large likelihood of sampling differences and cohort effects among the ages (e.g., Riegel & Riegel, 1972). This realisation led to the development of longitudinal studies. Efficiency and convenience in data collection is always a consideration, so many studies have been designed to sample individuals from a wide age range, for example 20 years, only twice, for example 3 years apart. This makes it possible in principle to address change over the whole 23-year period. Though many samples with this design remain in use in recent publications (e.g., Dolcos et al., 2012; Gayman, Turner & Cui, 2008; Gerstorf, Rocke & Lachman, 2011; Hanson et al., 2011; Kooij & Van De Voork, 2011; Lapi et al., 2009; Mather et al., 2010, Menezes et al., 2011; Ramsden et al., 2011; Schelleman-Offermans, Kuntsche & Knibbe, 2011; Whitehead et al., 2011), the limitations of two waves of data for understanding change have been well documented (e.g., Rogosa, 1995; Rogosa, Brandt, & Zimoski, 1982), as have the additional complications introduced by large age ranges within samples (e.g., Sliwinski, Hoftman, & Hofer, 2010). Essentially, the limitations of two waves of data surround the fact that, with only two waves of data, it is not possible to distinguish true between-individual differences in overall level from error of measurement in the estimation of indvidual change trajectories. Additional complications are introduced in samples with wide age ranges because it is necessary to assume that the individual change trajectories depend only on

the ages of the indviduals, and not on when the individuals attained those ages. There is long-standing and substantial evidence that this assumption is often violated (e.g., Kuhlen, 1940; Schaie, 1965). In addition, it is necessary to assume that each age group within the sample represents the underlying population to the same degree. This assumption too is generally violated due to the underlying associations between survival and ability to participate at any given age and overall function noted above.

The purpose of this study

Despite these limitations and complications, pressure from funding agencies, publication goals, and sheer scientific curiousity provide strong temptation to wring some information from two waves of data, while study administrators await the availability of, and/or justify requests for, funding of additional data waves. Though limited, some information *can* be gleaned from such an approach, and the greater the volume of data available about each participant, the greater the amount of information it should be possible to extract, especially if the complications associated with wide sample age ranges can be avoided. At 2% mortality rate per year, some proportion of participants should be either in or entering such periods, though the specific proportion would depend on the length of the terminal decline period. The purpose of this study was to explore the potential capacity to use two data waves to identify individuals who, in the period from ages 70 to 73 years, might be experiencing negative changes in many areas consistent with terminal decline. We did this through examination of three questions: 1) How and to what extent did individuals change during this period? 2) Were there correlates or predictors of these changes? and 3) Did changes tend to cluster in ways that could distinguish healthy ageing from terminal decline? Our analysis made use of the Lothian Birth Cohort 1936 (LBC1936; Deary et al., 2007; Deary et al., in press), a sample of 1,091 initially healthy 70-year-olds living in the Edinburgh area of Scotland, all of whom were born in 1936 and who completed a broad assessment of both cognitive and physical function. We thus avoided the analytical complications associated with samples with large age ranges and had access to a wealth of information about these individuals.

Method

Participants

The LBC1936 study was designed to take advantage of the Scottish Mental Survey 1947 (Scottish Council for Research in Education 1949). On 4 June 1947, almost all children born in 1936 and attending school in Scotland on that day completed a valid cognitive ability test. LBC1936 recruited 1091 of these individuals who were living independently in the area of Edinburgh, Scotland when they were mean age 70 between 2004 and 2007, with the intent of following them through old age. Recruitment was limited to the Edinburgh area for practical reasons of ease of access to the clinical research facility; within the recruitment catchment area the goal was to recruit at least 1,000. The only eligibility criteria were birth in 1936, enrolment in school in Scotland at age 11, and current ability to get to the clinic to participate in about 4 hours of psychological and medical testing. Taxi transportation was provided if needed. Recruitment was accomplished with the assistance of the Lothian Health Board and through advertisements. The Lothian Health Board wrote to 3,810 individuals on the Lothian Community Health Index who were born in 1936 and thus might have taken part in the Scottish Mental Survey 1947, of whom 3,686 were invited to hear about the study. Of these, 1,703 (46.2%) responded, 1,226 (72.0% of 1,703) were interested and considered themselves eligible, and 1,091 (89.0% of 1,226) participated, with some small supplementation from advertisements. See Deary et al., (2007 in press) for further details on participant recruitment and assessment. Participants included 548 males and 543 females, aged 67.7 to 71.3 years at time of first assessment in old age (mean=69.6, SD =0.80). The presence in the sample of fewer females than males suggests that the female participants may have been less representative of the overall population in this age group than the males, as the population sex ratio favours females in this age group due to longer female longevity. Ethical approval for the study was granted by the Multi-Centre Research Ethics Committee for Scotland and by Lothian Research Ethics Committee. The study was carried out in compliance with the Helsinki Declaration.

The sample was assessed in essentially the same way a second time, approximately 3 years later (mean=3.0, SD=.3). Of the original sample, 866 (79%) returned (448 males, 82% of original; 418

females, 77% of original). The primary reason for failure to return was death, or self-assessment of inability to participate. Compared to returning participants, non-returners were poorer at the first assessment in immune and inflammation indicators; lung function; walk and visual search speed; memory span; and performance on logical memory, matrix reasoning, block design, and reaction time task performance, with standardised mean score differences (Cohen's *d*s) ranging from .25 to .35. Non-returning participants also had lower cognitive ability scores at age 11 from the Scottish Mental Survey (*d*=.22), fewer years of education (*d*=.20), earlier retirement age (*d*=.19), and lower current social class (*d*=.20), suggesting that their lower level of at least cognitive function may have been long-term rather than some indication of greater failing health or proximity to death.

Measures

Participants were interviewed and tested individually during a single session in each testing wave, by a trained psychology researcher and a research nurse at the Wellcome Trust Clinical Research Facility at the Western General Hospital in Edinburgh. The assessment was broken by two periods of at least 15 minutes for rest and refreshments. It began with orientation to the study and collection of informed written consent to participate, followed by provision of basic demographic and medical information through structured interview. This included educational attainment, primary occupation during working life (and that of spouse for married women), age of retirement, history of medical diagnoses and current prescription medications, smoking history and current status, and current pattern of alcohol consumption. For this study, we made use of the current demographic variables and numbers of current medical diagnoses and prescription medications from this interview (see Deary et al., 2007 for further details of the assessment).

Hospital anxiety and depression scale (Zigmond & Snaith 1983). Participants completed this written questionnaire, which consists of 14 items, half of which reflect anxiety and half depression. Maximum score on each scale is 21, with probable clinical levels at scores of at least 11.

Tests of current cognitive function. Participants completed a battery of cognitive tasks intended to measure various aspects of cognitive function. The tests used here, their content, and sources are listed in Table 1.

Test of childhood cognitive function. Most participants in LBC1936 had participated in the Scottish Mental Survey 1947, which took place on 4 June 1947, when participants were age about 11 years. The primary focus of this survey was administration of Moray House Test No. 12, a well-validated predominantly verbal reasoning test. Scores on this test were obtained from the survey records for the purposes of LBC1936. LBC1936 participants completed the same test again at age 70.

Physical examination and interview. This included measurement of height and weight, time in seconds to walk 6 metres, demi-span in cm, responses to a 9-item activities of daily living scale (Townsend, 1979), sitting and standing systolic and diastolic blood pressure, forced expiratory volume from lungs in 1 sec. (FEV1; best of 3), grip strength in the right and left hands, and corrected and uncorrected vision in right and left eyes. Participants provided blood samples used to assess haemoglobin, white cell and platelet counts, prothrombin time, activated partial thromboplastin time (APTT), fibrinogen, serum folate, albumin, calcium, cholesterol, HDL cholesterol, glycated haemoglobin (HbA1C), C-reactive protein levels, and estimated glomerular filtration rate. To avoid distortions, we did not make use of prothrombin time for any participant taking the medication warfarin.

LBC1936 Study questionnaires. Participants were distributed questionnaires and stamped return envelopes, with instructions for completing the questionnaires at home and returning them. For this study, we made use of the personality questionnaire that was measured at both time points. This was the International Personality Item Pool inventory of 50 items, 10 measuring each of the so-called Big Five personality traits of Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect**,** which can be freely downloaded at http://ipip.ori.org/. The word 'I' was added to each of the fragments making up these items, to make them more closely match wording in other questionnaires used. Participants rated how well they believed the items described them on a 5-point scale (very accurate to very inaccurate). At Time 1, 87% of participants returned these questionnaires. At Time 2, 99% did so.

Table 1. Cognitive tests administered at ages 70 and 73 to Lothian Birth Cohort 1936 participants

Test	General Description	Source (Citations below)
Logical Memory	Verbal declarative memory	Wechsler Memory Scale-III^{UK}
Spatial Span	Non-verbal memory	Wechsler Memory Scale-III^{UK}
Verbal Paired Associates	Verbal learning and memory	Wechsler Memory Scale-III^{UK}
Symbol Search	Speed of information processing	Wechsler Adult Intelligence Scale-III^{UK}
Digit Symbol	Speed of information processing	Wechsler Adult Intelligence Scale-III^{UK}
Matrix Reasoning	Pictorial pattern inference	Wechsler Adult Intelligence Scale-III^{UK}
Letter-Number Sequencing	Working memory	Wechsler Adult Intelligence Scale-III^{UK}
Digit Span Backwards	Manipulation of memory	Wechsler Memory Scale-III^{UK}
Block Design	Constructional ability	Wechsler Adult Intelligence Scale-III^{UK}
Simple Reaction Time	Mean response to simple stimulus	
Choice Reaction Time	Mean defined response to specific stimulus among 4 alternatives	
Inspection Time	Visual discrimination of briefly presented stimulus	
Verbal Fluency	Attention focus; association flexibility	
National Adult Reading Test	Reading vocabulary	
Wechsler Test of Adult Reading	Reading vocabulary	
Mini-Mental State Exam	Space/time orientation; dementia screen	

Notes. Wechsler Memory Scale-III^{UK}: Wechsler, D. (1998). WMS-IIIUK Administration and Scoring Manual. London: Psychological Corporation. Wechsler Adult Intelligence Scale-III^{UK}: Wechsler, D. (1998). WAIS-IIIUK Admini stration and Scoring Manual. London: Psychological Corporation. Reaction times: Deary, I. J., Der, G., & Ford, G.(2001). Reaction times and intelligence differences: A population-based cohort study. Intelligence, 29, 389-399. Inspection Time: Deary, I. J., Simonotto, E., Meyer, M., Marshall, A., Marshall, I., Goddard, N., & Wardlaw, J. M. 2004). The functional anatomy of inspection time: an event-related fMRI study. NeuroImage, 22, 1466-1479. Verbal fluency: Lezak, M. (2004). Neuropsychological Testing. Oxford: Oxford University Press. National Adult Reading Test: Nelson, H. E., & Willison, J. R. (1991). National Adult Reading Test (NART) Test Manual (Part II). Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

Death information. The study receives monthly reports of participant deaths from the General Register Office for Scotland, part of the National Records for Scotland. We made use of reports through 31 December, 2011. At that date, there had been 30 deaths among participants who had completed Time 2 assessments, the majority in 2009.

Data treatment

We followed several steps in preparing the data for analysis. First, we regressed the effects of sex and height from the measures of 6-metre walk time, demi-span, FEV, and grip strength. We then standardised all variables to place them on the same scale, trimming isolated outliers by observation, separately by variable. We consolidated height and weight by calculating body mass index (BMI) as weight in kg/height in metres^2. Several other variables tapped similar constructs, suggesting the formation of summary or composite variables. We combined the systolic and diastolic blood pressure readings to estimate mean arterial pressure, using the formula (mean sitting diastolic reading) + (mean sitting systolic reading – mean sitting diastolic reading)/3. We averaged grip strength readings for right and left hands; scores from the Digits Backwards and Letter-Number Sequencing tests to form a Memory Span variable; scores from the two scores from the two reading tests to form a Word Reading variable; corrected vision in right and left eyes; standardised prothrombin time and APTT to form an indicator of blood clotting function; standardised folate, albumin, and cholesterol levels (reversed as appropriate) to form a variable indicating nutritive status; standardised platelet counts, fibrinogen levels, and C-reactive

protein levels to measure inflammation; and standardised HDL cholesterol and glycated haemoglobin (reversed as appropriate) to form an indicator of metabolic vascular risk. This resulted in 36 variables for study. We adjusted Time 2 standardised variables to reflect their mean standardised differences from Time 1 standardised variables, while preserving variance differences from Time 1. Finally, we subtracted Time 1 values from Time 2 values to produce change scores.

The decision to make use of difference scores as the measure of change requires some explanation, as statistical advice against this is not uncommon (e.g., Campbell & Kenny, 1999; Cronbach & Furby, 1970). Two common observations have led to this advice. First, difference scores are often negatively correlated with Time 1 scores, leading to a perception that the difference score is a negatively biased estimate of change. The reality is the opposite: the difference score is an unbiased estimate of true change, but the observed correlation is a negatively biased estimate of the correlation between true initial status and true change (Rogosa, Brandt, & Zimowski, 1982), due to the presence of the error of measurement associated with the observation of initial status in both that observation and the observed change. Second, it has been common to assume that variance remains constant from one measurement occasion to another. Under this assumption, as the correlation between the two measurement occasions increases, the reliability of the difference score decreases, yielding the impression that greater reliability of measures leads to lower reliability of difference scores. In reality, however, it is very common in developmental situations for variances to change over time (e.g., McCardle & Woodcock, 1997). The reliability of the difference score is sensitive to these changes (Nesselroade & Cable, 1974). The greater the changes are, the more reliable is the difference score. Because variance changes are common, the difference score is commonly quite reliable. The most commonly used alternative, the residual from regressing Time 2 values on Time 1 values, is imprecise and often considerably biased (Rogosa, Brandt, & Zimoski, 1982). Perhaps most important, however, the regression residual does not address the simple question of change. Instead, it addresses the question of what the expected change for an individual would have been, had that individual been at the mean level at Time 1. Appropriate interpretation of any answer to this question is far from clear.

Results

Basic change statistics

Table 2 shows descriptive statistics for the raw study variables at the two assessments. Most variables showed large variation among individuals, but the means of some gave clear indications of the overall health and wellbeing of participants. For example, at Time 1, participants had on average (SD) 3.8 (1.9) diagnosed medical conditions, for which they took an average of 3.0 (2.5) prescription medications. By Time 2, these averages had grown to 4.4 (2.0) and 4.0 (2.3). Despite this, average difficulties with activities of daily living were very small at both time points, with most participants reporting no difficulties at all. Average BMIs were 27.82 (4.56) and 27.92 (4.43), respectively at Times 1 and 2, which would generally be considered overweight but not obese. Hospital Anxiety and Depression averages were low at both time points, indicating generally good levels of well-being. Mini-Mental State Exam average scores were 28.8 (1.4) and 28.8 (1.4) at Times 1 and 2 respectively. At Time 1, 21 had scores less than 25; at Time 2, 18, and the lowest score at both time points was 22. Thus, most participants suffered some health impairments, but not sufficiently to have undermined basic cognitive function and overall well-being.

The variables all showed both substantial correlations between Time 1 and Time 2 indicating considerable stability and substantial individual differences in extent of change. The smallest correlation was .47 for the MMSE, likely because of restriction of range due to a strong ceiling effect. The largest correlation was .96 for Word Reading, and the overall mean correlation was .71. Across all the variables, the mean largest individual participant increase was 3.49 SDs, and the mean largest decrease was 3.91 SDs. The distributions of change variables were generally close to normal, with average skewness at -.002. This suggested two things: first, positive change was basically as likely as negative change. Second, it would likely be difficult to identify participants who were experiencing terminal decline, which would be evidenced by substantially skewed distributions in which most values hovered around 0. The mean change, adjusted to reflect the direction of each measure that indicated decline in function, was -.01 SD, indicating small overall absolute decline. Acknowledging the multiple tests run, 26 of the 36 variables showed no significant mean change. There was considerable heterogeneity in direction even among those variables showing significant mean

change, and cognitive functions were largely stable at the mean level. Lung function, walk speed, reading and search abilities, nutritive status, and grip strength declined significantly and participants were taking significantly more medications. Inflammatory and metabolic vascular risk markers, and corrected vision improved significantly. The variables that showed significant declines more likely indicated ageing, while the variables that showed overall improvement more likely reflected better health care. Full details of this information are shown in Table 3.

Reliability of changes

Another way to evaluate changes is to measure the extent to which the changes observed could be considered reliable. To do this, we made use of the Reliable Change Index (Christensen & Mendoza, 1986; Hsu, 1989), which indicates the magnitude of change that can be considered reliable after accounting for measurement error and regression to the mean. The full formula is

$$\frac{x_2 - x_1}{\sqrt{2SD_1^2(1-r)}}$$

where the subscripts refer to the time points, *x* to a data point, *SD* to standard deviation, and *r* to test-retest reliability. The denominator is the standard error of the difference between the two test scores, and describes the expected variability in change scores if no actual change occurred. If the variables are normally distributed, the index will be too, and there will be 95% probability that change did occur if the index is greater than 1.96. Conceptually, the situation is analogous to inferring that mean differences are significant when their 95% confidence intervals do not overlap: here the differences between measures at two time points are significantly different when the intervals reflecting their standard errors of measurement do not overlap.

Implementing this formula involved some judgment, as the formula requires short-term test-retest reliability of the measures and this information was not available for most of our measures. Psychometricians tend to think of test reliabilities in the range of .75-.85 (often inappropriately assessed with a measure of internal consistency rather than short-term test-retest correlations) as strong. Medical practitioners, however, tend to think of a clinical measure as reliable when a short-term reassessment would likely generate a deviation of no more than 10% from the first observation, and many of our variables were clinical/medical. To understand the medical perspective in psychometric terms, consider IQ scores and the T-scale. Most IQ tests are scaled with mean 100 and SD 15, and the T-scale has mean 50 and SD 10. Medical practitioners, then, might expect a short-term retest of an IQ-scale observation of 100 to generate a score between 90 and 110, and a T-scale observation of 50 to generate a score between 45 and 55. Translation to psychometric perspective can be modeled by adding uniformly-distributed random values within .67 SD to the IQ-scale scores and within .5 SD to the T-scale scores from any sample, and correlating these scores with the original scores. This generates correlations in the range of .92-.97 for .67 SD and .96-.98 for .5 SD. To produce correlations in the range generally considered to indicate reliability by psychometricians, it is necessary to add random uniformly-distributed values within 1.0-1.5 SD, or 15-22 points on the IQ scale and 10-15 points on the T-scale. The reliability guidelines used by psychometricians thus allow very considerable individual variation, perhaps more than most researchers realise.

This was reflected in application of the Reliable Change Index to our data. We lacked short-term test-retest reliability data for our measures in general, but could reasonably assume that average reliability lay in the .75-.85 range. We thus estimated the numbers of variables on which each participant showed reliable changes if test-retest reliability was .75 and if test-retest reliability was .85. Assuming .85 reliability, on average, participants showed reliable changes in 10.4 (SD=6.0) of the 36 variables, with a range of 0-24. The distribution of reliable changes was slightly skewed at .70. Assuming .75 relability, on average participants showed reliable changes in 7.9 (SD=4.8, skewness=.43) variables, with a range of 0-22. Regardless of level of test reliability, about half the reliable changes represented improvements rather than declines in function. We also separated the reliable changes for each participant into those that represented decline and improvement. Numbers of reliable changes indicating improvements were correlated .53 (p<.001) with numbers of reliable changes indicating declines in function assuming .85 reliablity, and .52 (p<.001) assuming .75 reliability, indicating a substantial tendency for the same participants to show changes reflecting both improvements and declines in function. This is summarised in Table 4.

Associations with pervasiveness of change

Number of reliable changes was not significantly correlated with age-11 IQ (.00 [-.04], assuming .85 [.75] reliability, nor with age-70 IQ (.04 [.01]) (all p>.15). Years of education and social class status were similarly uncorrelated with reliable changes. With no adjustment for multiple testing and assuming .75 reliability (thus a liberal reading), numbers of reliable changes were correlated with changes in numbers of drugs (.12), BMI (-.08), six-metre walk time (.24), ADLs (.10), mean arterial pressure (-.08), FEV (-.07), Logical Memory (-.13), Search Speed (-.13), simple reaction time (.07), inspection time (-.08), haemoglobin (-.10), and metabolic vascular risk (-.11), generally indicating that greater decline in function was associated with greater numbers of reliable changes. This is summarised in Table 4. Most Time 1 variables were not correlated with number of reliable changes, but there were more significant correlations between better function and number of changes indicating declines in function, than changes indicating improvements in function. These observations thus likely reflected the negative bias in correlations between observed Time 1 scores and observed difference scores, which would make them largely artefactual; one hint that this might be the case was that the strongest such associations with number of reliable changes were -.15 with IPIP Conscientiousness and -.12 with IPIP Agreeableness.

Clustering of change variables

Significance of the correlation between numbers of changes indicating improvements and declines in function suggested some tendency for changes to cluster, though it also suggested that the changes may be random. We nevertheless ran a factor analysis of the change variables. Parallel analysis indicated 4 factors, but most variables did not show substantial loadings on any of them. Table 4 shows the results, which did support some clustering of sources of change. The first factor appeared to indicate declines in memory and speed of information processing; the second suggested declines in emotional well-being and positive personality function; the third grouped increases in numbers of diseases and numbers of drugs taken with decreases in BMI, mean arterial pressure, and haemoglobin (which can be negative health indicators in this age group, and are increasingly so with greater age beyond 73); and the fourth grouped increases in numbers of diseases with decreases in, haemoglobin, white cell count, and inflammation. The factors were basically independent, with all correlations being between .05 and .06. We labelled the factors Memory/Speed, Personality, Metabolism, and Physical Robustness. Full results are shown in Table 5.

Deaths since Time 2

As noted, we observed 30 deaths between the Time 2 assessment and 31 December, 2011. It is not possible to make precise calculations of how much power we had to detect whether participants were in or entering periods of terminal decline, without indications of the expected effect sizes and lengths of periods over which terminal decline might operate. Based on the overall 2% per year death rate for this age group, however, we would expect about 200 deaths from the full sample over the 10-year period from ages 70 to 80. Given a not-uncommon estimate in the literature of periods of terminal decline on the order of 5-7 years, we would expect at least that number to have been in or entered such a period between ages 70 and 73. Based on this, we had over 80% power to detect changes on the order of .2 standard deviations, noticeably smaller than we could actually measure with any reliability.

We regressed death status as of 31 December, 2011 on the factor scores to assess their potential as markers of terminal decline. Together, the 4 variables explained 11.3% of variance, but the overall regression did not reach significance (p=.103). Entered singly, declines in Metabolism and Memory/Speed were each significant predictors of mortality, but Memory/Speed was not significant when entered with Metabolism. Their odds ratios were 1.58 (95% confidence interval 1.09-2.29) for Metabolism and 1.66 (95% confidence interval 1.10-2.54) for Memory/Speed when entered singly. At the same time, numbers of reliable change indicating declines in function at 75% reliability predicted death with an odds ratio of 1.27 (95% confidence interval 1.14-1.41, p<.001). But so did numbers of reliable change indicating improvement in function at 75% reliability (odds ratio = 1.13, 95% confidence interval 1.03-1.32, p=.031). This is summarised in Table 6.

Table 2. Descriptive statistics of raw study variables

	All age 70					Age 70, returning at 73					Age 73			
	N	Min.	Max.	Mean	SD	N	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD
Number of diseases	1091	0	11	3.8	1.9	866	0	11	3.7	1.9	0	13	4.4	1.9
Number of drugs taken	1091	0	8	3.0	2.5	762	0	8	2.9	2.5	1	8	4.0	2.3
HADS anxiety score	1089	0	17	4.9	3.2	865	0	16	4.8	3.1	0	18	4.5	3.1
HADS depression score	1086	0	13	2.8	2.2	865	0	13	2.7	2.1	0	13	2.6	2.2
BMI	1089	16.02	72.00	27.82	4.56	866	16.02	48.52	27.80	4.37	16.67	48.50	27.92	4.43
6 metre walk time	1085	2.0	11.0	3.85	1.13	860	2.0	11.0	3.78	1.05	2.30	12.30	4.34	1.24
Demi-span in cm	1088	65	91	77.8	4.8	864	65	91	78.1	4.8	66	91	78.2	4.7
Activities of daily living	1089	0	14	1.0	2.0	865	0	13	.9	1.9	0	14	1.0	2.1
Mean arterial pressure	1088	72.00	145.00	104.15	11.87	866	72.00	145.00	104.02	11.88	63.78	140.00	101.64	11.33
Forced expiratory vol.	1085	.74	4.34	2.36	.68	856	.74	4.34	2.41	.68	.78	4.25	2.30	0.67
Average grip strength	1086	5.00	55.50	28.03	9.93	865	5.00	55.50	28.49	9.78	5.00	53.25	28.11	9.26
Logical memory	1087	16	117	71.4	17.9	864	16	117	72.56	17.26	20	116	74.3	17.8
Spatial span	1084	5	24	14.7	2.8	861	6	24	14.8	2.8	7	23	14.7	2.8
Verbal paired assoc.	1050	0	40	26.4	9.1	843	1	40	26.9	9.0	0	16	9.1	3.8
Symbol search	1086	2	49	24.7	6.4	862	2	49	25.0	6.4	3	45	24.6	6.2
Digit symbol	1086	25	98	56.6	12.9	862	25	98	57.5	12.7	22	94	56.4	12.3
Matrix reasoning	1086	4	24	13.5	5.1	863	4	24	13.9	5.1	4	25	13.2	5.0
Letter-number seq.	1079	1	21	10.9	3.2	863	1	21	11.1	3.1	1	20	10.9	3.1
Digit Span Backwards	1090	2	14	7.7	2.3	866	2	14	7.8	2.3	2	14	7.8	2.3
Block design	1085	10	65	33.8	10.3	864	11	65	34.5	10.1	10	66	33.6	10.1
Log simple RT	1085	.16	.51	.24	.04	865	.16	.47	.24	.04	.17	.44	.24	.04
Choice reaction time	1084	.45	1.13	.64	.09	865	.45	1.13	.64	.08	.46	.97	.65	.09
Inspection time	1041	70	140	112.2	10.9	838	70	140	112.7	10.7	67	137	111.3	11.4
Verbal fluency	1087	10	83	42.4	12.5	865	10	83	42.8	12.7	8	85	43.2	12.9

(Table 2 cont'd)

Natl adult reading test	1089	10	50	34.5	8.1	864	10	50	34.8	8.1	9	50	34.4	8.1
Wechsler T Adult Reading	1089	14	50	41.0	7.2	864	14	50	41.3	7.1	16	50	41.0	6.9
Mini-mental state exam	1090	22	30	28.8	1.4	865	22	30	28.8	1.4	22	30	28.8	1.4
Average corr. vision	775	-.10	.80	.10	.16	607	-.10	.80	.10	.16	-.10	.90	.15	.17
Haemoglobin	1063	101	181	145.2	13.0	825	101	181	145.1	13.1	101	180	140.1	13.3
White cell count	1062	2.5	27.0	7.05	2.22	824	3.0	27.0	6.96	2.24	2.6	27.0	6.95	2.20
Platelet count	1061	105	508	274.1	64.6	820	105	508	271.1	61.8	73	460	245.0	58.2
Prothrombin time	1051	9	13	9.7	.6	820	9	13	9.7	.6	9	22	11.6	1.0
APTT	1051	21	41	28.6	3.1	820	21	41	28.6	3.1	21	50	31.0	3.9
Fibrinogen	1051	1.6	5.5	3.3	.6	819	1.6	5.5	3.25	.62	1.8	5.3	3.32	0.59
Serum folate	911	3.3	25.0	12.84	6.29	716	3.6	25.0	13.03	6.30	2.5	25.0	11.66	6.11
Albumin	1058	37	54	44.7	3.0	831	37	54	44.7	3.0	35	51	43.8	2.9
Cholesterol	1054	2.7	8.9	5.45	1.15	832	2.7	8.9	5.45	1.14	2.4	9.3	5.15	1.15
HDL cholesterol	969	.53	3.32	1.52	.44	832	.68	3.32	1.53	0.44	.57	3.00	1.46	0.43
HbA1C	1061	4.5	9.5	5.93	.71	826	4.5	9.5	5.93	0.74	4.4	8.9	5.75	0.65
C-reactive protein	1053	1.5	45.0	5.19	6.03	830	1.5	45.0	4.98	5.69	1.5	45.0	4.80	5.80
Glomerular filtration rate	1060	41	160	81.4	18.2	833	41	160	81.5	18.1	30	157	77.8	19.2
IPIP extraversion	954	1	40	21.3	7.1	854	1	40	21.3	7.0	2	40	21.6	7.2
IPIP agreeableness	952	11	40	31.1	5.4	854	12	40	30.9	5.5	10	40	30.8	5.6
IPIP conscientiousness	952	9	40	28.2	6.0	854	9	40	28.1	6.1	5	40	27.7	6.1
IPIP emotional stability	950	1	40	24.6	7.7	853	1	40	24.9	7.7	2	40	25.0	7.7
IPIP intellect	948	5	40	23.8	5.7	852	5	40	23.9	5.7	5	40	23.7	5.9

Table 3. Descriptive statistics of individual-level changes from ages 70 to 73

	Mean paired difference	Standard deviation	Standard error of mean	Skewness	Kurtosis	t	df	Prob. (2-tailed)	T1-T2 correlation
Number of diseases	.040	.655	.022	.36	.60	1.78	865	.08	.78
Number of drugs Ttaken	-.096	.684	.025	.35	2.29	-3.88	761	<.001*	.75
HADS anxiety	.021	.768	.026	.21	1.90	.79	862	.43	.70
HADS depression	.040	.820	.028	.48	3.28	1.42	860	.16	.65
BMI	.006	.323	.011	-.45	2.71	.52	864	.60	.95
6-Metre walk time	.510	.875	.030	1.49	9.59	17.05	856	<.0001*	.63
Demi-span	.062	.962	.033	-.32	3.04	1.89	863	.06	.53
Activities of daily living	.042	.699	.024	.89	9.25	1.78	864	.08	.75
Mean arterial pressure	.011	1.014	.034	-.30	1.46	.31	865	.76	.49
Forced expiratory vol.	-.152	.514	.018	.29	5.60	-8.63	852	<.001*	.86
Average grip strength	-.056	.404	.014	-.44	2.22	-4.07	863	<.001*	.93
Logical memory	-.062	.758	.026	-.09	.41	-2.41	861	.02	.70
Memory span	-.058	.751	.026	.17	.48	-2.26	865	.02	.72
Verbal paired assoc.	-.053	.773	.027	.02	.42	-1.98	827	.05	.69
Search speed	-.073	.575	.020	-.58	3.48	-3.75	862	.000*	.83
Matrix Reasoning	-.070	.836	.028	-.05	.26	-2.45	861	.01	.65
Block design	-.062	.689	.023	-.15	.84	-2.64	860	.01	.76
Verbal fluency	-.028	.619	.021	.02	.08	-1.32	863	.19	.81
Log simple reaction time	.070	.906	.031	-.25	5.00	2.26	863	.02	.56
Choice reaction time	.066	.682	.023	-.30	4.42	2.86	863	.00	.76
Inspection time	-.053	.895	.031	-.20	3.00	-1.69	821	.09	.59
Word reading	-.044	.282	.010	-.32	1.13	-4.58	863	<.001*	.96
Mini-mental state exam	-.035	1.015	.035	.16	1.58	-1.03	864	.30	.47
Average corr. vision	.054	.148	.006	-.11	1.38	8.61	551	<.001*	.58

(Table 3 cont'd)

White cell count	-.046	.629	.026	.06	2.85	-2.07	810	.04	.73
Haemoglobin	-.063	.711	.025	-.02	2.28	-2.52	812	.01	.76
Clot	-.062	.531	.019	-.03	.29	-3.23	754	.001	.50
Nutritive status	-.461	.896	.031	.29	.21	-14.70	816	<.001*	.57
Inflammation	-.348	.917	.032	.09	5.72	-10.90	823	<.001*	.52
Metabolic vascular risk	-.569	.709	.025	-.29	4.61	-23.01	822	<.001*	.74
Glomerular filtration rate	-.004	.579	.020	-.08	1.19	-.21	817	.83	.83
IPIP extraversion	.005	.576	.021	-.02	.31	.22	776	.82	.84
IPIP agreeableness	.039	.748	.027	-.47	2.01	1.46	775	.14	.72
IPIP conscientiousness	.057	.676	.024	.03	.33	2.36	774	.02	.77
IPIP emotional stability	-.030	.701	.025	-.12	.81	1.20	771	.23	.76
IPIP intellect	.017	.699	.025	-.22	.86	.67	771	.51	.76

Note. Variables were standardised to Time 1 level, so the mean difference was effect size relative to that level. With adjustment for multiple testing, only probability levels of .001 or less should be considered significant (*). Change was Time 2 less Time 1, so negative differences indicate declines in scores.

Table 4. Observations regarding numbers of reliable changes

	85% Reliability	75% Reliability
Number of reliable changes - mean (sd)	10.4(6.0)	7.9(4.8)
Correlation between reliable improvements and declines	.53	.52
Correlation of number of reliable changes with:		
age-11 IQ	ns	ns
age-70 IQ	ns	ns
years of education	ns	ns
social class	ns	ns
number of drugs taken	---	.12
BMI	---	-.08
6-metre walk time	---	.24
activities of daily living	---	.10
mean arterial pressure	---	-.08
forced expiratory volume		-.07
logical memory	---	-.13
search speed	---	-.13
simple reaction time	---	.07
inspection time	---	-.08
haemoglobin	---	-.10
metabolic vascular risk	---	-.11

Note. Only significant (with no adjustment for multiple testing) correlations with change are shown. 'ns' is 'not significant.' '---' is 'not calculated.' Number of reliable changes refers to number of changes per person that were reliable among 36 variables assessed at about ages 70 and 73. 85/75% reliability refer to assumed test-retest reliability of the measures. See text for further explanation.

Table 5. Factor analysis of change variables

	Memory/ Speed	Personality	Metabolism	Physical robustness
Number of diseases	-.02	.05	**.18**	**.17**
Number of drugs taken	-.08	.05	**.32**	.06
HADS anxiety	-.06	**.33**	.10	.00
HADS depression	-.03	**.32**	.07	.04
BMI	.03	-.07	**-.37**	.08
6-metre walk time	-.14	.03	.06	.02
Demi-span	-.02	-.01	-.08	.02
Activities of daily living	-.05	.01	.04	.08
Mean arterial pressure	-.06	.02	**-.65**	-.01
Forced expiratory vol.	-.01	-.04	.07	.04
Average grip strength	.08	-.01	-.13	-.08
Logical memory	**.57**	.01	-.01	.01
Memory span	**.18**	.00	-.06	.01
Verbal paired assoc.	**.40**	.04	-.02	.09
Search speed	.14	-.03	-.07	.07
Matrix reasoning	.06	.04	-.01	-.01
Block design	.13	-.02	.09	-.04
Verbal fluency	**.21**	-.03	.03	.07
Simple reaction time	**-.33**	.02	.00	.04
Choice reaction time	**-.39**	.11	-.01	.01
Inspection time	.07	.06	-.04	-.05
Word reading	**.18**	.05	.02	.02
Mini-mental state exam	**.18**	.02	.05	.07
Average corr. vision	.03	-.04	-.07	-.01
White cell count	.01	.05	-.01	**-.67**
Haemoglobin	.01	.05	**-.27**	**-.19**
Clot	-.07	.03	.03	-.01
Nutritive status	.00	.02	.14	-.11
Inflammation	-.04	-.06	.04	**-.49**
Metabolic vascular risk	.05	-.10	-.10	-.10
Glomerular filtration rate	.03	.06	.14	.14
IPIP extraversion	-.09	**-.37**	-.01	-.01
IPIP agreeableness	.05	**-.49**	-.04	-.04
IPIP conscientiousness	-.02	**-.33**	.04	.04
IPIP emotional stability	-.05	**-.34**	.09	.10
IPIP intellect	-.02	**-.38**	.03	.03

Note. Factor loadings greater than .15 in absolute value are in bold.

Table 6. Change variables as individual predictors of death since Time 2

	Odds ratio	95% Confidence interval
Decline in memory/speed	1.67	1.10-2.54
Decline in personality traits	ns	---
Decline in metabolism	1.58	1.09-2.29
Decline in physical robustness	ns	---
Number of reliable declines	1.27	1.14-1.41
Number of reliable improvements	1.13	1.03-1.32

Note. 'ns' is 'not significant.' Logistic regression including all factors was not significant.

Discussion

In this study, we explored the potential capacity to use two longitudinal data waves to distinguish 'normal' from disadvantageous and even terminal ageing patterns in the LBC1936 between ages 70 and 73. To do this, we examined 36 variables indicating aspects of cognitive, emotional, and physical function. In the process, we addressed 3 questions: 1) How and to what extent did individuals change during this period? 2) Were there correlates or predictors of these changes? and 3) Did changes tend to cluster in ways that could distinguish healthy ageing from terminal decline? Overall, our measures tended to show mean changes indicating declines in function, as would be expected. Most of these mean differences were not significant after adjustment for multiple testing, however, despite our good-sized sample. The lack of statistical significance of most of the mean changes indicated the gradual nature of the overall ageing process, especially since some of the significant mean differences indicated improvement in average function, including number of drugs taken, average corrected vision, and inflammation.

Within each variable, we observed substantial individual differences in change, but most could not be considered reliable. Moreover, the changes that could be considered reliable were at least as likely to indicate improvements in function as declines and changes indicating improvements and declines in function were substantially correlated. This suggests that ageing is far from a uniform process, but it also suggests that increasing variability in 'measurability' may be an important indicator of its progress, a topic receiving increasing attention in the ageing literature (e.g., Ram et al., 2011). Of course practice effects on some variables could have accounted for improvements as well. In sum, we observed substantial change, in the aggregate indicating ageing, but most of the individual observations could not be considered reliable. The lack of reliability of measures of change based on two waves of data is well known, and certainly results from the inability to distinguish error of measurement of individual level from error of measurement of individual change. It is rare, however, to see it so clearly documented as was possible here, given the large number of variables available. Even if more measurement occasions made it possible to minimise error of measurement, it is possible that a proportion of individual variation in ageing trajectories reflects a random walk process, wherein differences between adjacent measurement occasions are at least partially completely random. Such random walk processes are common throughout nature and society, and their presence and importance in understanding developmental progressions is increasingly being recognised, for example, in economics (Kuljanin, Braun, & DeShon, 2011).

Study limitations

Before discussing our observations in further detail, we note the primary limitations of our study. The most important of these was the relatively select nature of our sample, which was of somewhat higher childhood mental ability than the overall population. The average age-11 IQ score in the LBC1936 was 0.78 standard deviation higher

than the overall average for the full Scottish Mental Survey 1947, and the variance was restricted by 44%. In general, range restriction tends to reduce the magnitudes of associations involving the variable on which range has been restricted, but this is not always the case. Any such reduction would have been small in this case (.01-.03 at most), due to the small magnitudes of the correlations observed. The sample selectivity at least partly reflected general mortality patterns, as IQ is associated with greater longevity (Batty, Deary, & Gottfredson, 2007; Calvin et al., 2011). Our sample was likely of higher educational attainment and social class than the overall population as well, though we could not quantify the degree to which this was the case. We did not have short-term test-retest reliability data for most of our measures, and thus were forced to make assumptions about their likely values. We did this somewhat crudely, making two overall assumptions for all variables. Within this, however, our conclusions were very similar for the two levels assumed.

Correlates or predictors of change and clustering of change variables

We found no associations between number of reliable changes and age-11 IQ, age-70 IQ, number of years of education, or current social class, suggesting that change was relatively evenly distributed throughout the sample. Personality may have contributed to lower reliability of measurement, as lower IPIP Conscientiousness and Agreeableness at age 70 had the largest negative correlations with number of reliable changes. This question deserves greater research attention. To the extent we were able to pick up leading indicators of ageing, they appeared to be increases in numbers of drugs, six-metre walk time, ADLs, and clot function, and decreases in BMI, mean arterial pressure, FEV, Logical Memory, Search Speed, haemoglobin, and metabolic vascular risk.

One intriguing observation was that the individual age-70 variable with the strongest correlation with number of reliable changes was IPIP Conscientiousness, with IPIP Agreeableness second. Both correlations were negative (-.15 and -.12, respectively), indicating that those with lower Conscientiousness and Agreeableness at age 70 tended to show larger changes that could be considered statistically reliable. At the same time, however, given the variables in question, it seems likely that those changes, though large enough to be considered reliable statistically, were in fact not particularly reliable at all. That is, participants scoring lower in Conscientiousness and/or Agreeableness may have used less care in completing all measures over which they had some overt control, and may have varied more even on the physiological measures due to less routine in daily dietary, sleep, and other habits. This should be pursued in future research.

Evidence of clustering among the change variables was weak. The average absolute value of correlation was less than .05, and most variables did not load on any of the four factors that might reasonably be considered substantial within the data. In general, however, the variables that did load on these factors clustered around constructs of memory and attention, personality, perhaps appetite status, and overall health. These factors basically reflected the possible leading-indicator variables just noted. Indicators of failing overall health, particularly failing appetite in the form of increases in BMI, mean arterial pressure, glomerular filtration rate, and total number of reliable declines in function appeared to be the strongest candidates as markers of terminal decline, in the sense that they predicted rather imminent death. Many would consider these to be obvious, and effects were sufficiently weak that recovery from any particular state was clearly possible.

Conclusions

The limitations of two waves of longitudinal data to explore change have been well documented, but studies claiming to have uncovered important associations with change based on only two waves continue to be published regularly. We undertook this study both to provide an empirical demonstration of the degree of uncertainty of such estimates of change and to explore whether such change estimates might aggregate in more meaningful ways. Within levels of short-term test-retest reliability considered acceptable in the field of psychology, this study showed that most change observed over a 3-year period was not reliable. Moreover, levels of correlation considered by psychologists to indicate substantial stability actually allow a level of movement, whether random or systematic, that many psychologists will likely find surprising. Even change that could be considered reliable often represented improvement (some of which might

have been test familiarity) rather than decline in function in this ageing sample in which overall decline might be expected. This suggests strongly that ageing proceeds in fits and starts, and it is possible that individual trajectories of any one variable include a considerable component best described as a random walk. Though it increases the assessment burden considerably, we urge researchers planning longitudinal projects to anticipate the need for more than two waves of data in order to draw meaningful conclusions about change, and we strongly suggest that journals publishing two-wave studies require inclusion of information about the extent to which the change measures could be considered reliable.

Acknowledgements

We thank the LBC1936 participants; Caroline Brett, Alison Pattie, Michelle Taylor and Caroline Cameron for data collection; the LBC1936 Study Secretary, Paula Davies; and the nurses and other staff at the Wellcome Trust Clinical Research Facility, Edinburgh (http://www.wtcrf.ed.ac.uk). The LBC1936 sample was funded by Research Into Ageing (Wave 1), and continues to be funded by Age UK as part of the Disconnected Mind project (Wave 2). Wendy Johnson, Alan Gow, John Starr, and Ian Deary are members of the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (http://www.ccace.ed.ac.uk), which is supported by the BBSRC, EPSRC, ESRC, and MRC as part of the cross-council Lifelong Health and Wellbeing Initiative (Grant #G0700704/84698).

References

Adler, N. E., & Snibbe, A. C. (2003) The role of psychosocial processes in explaining the gradient between socioeconomic status and health. *Current Directions in Psychological Science,* 12, 119-123.

Backman, L., & MacDonald, S. W. (2006). Death and cognition: Synthesis and outlook. *European Psychologist,* 11, 224-235.

Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (early life) IQ and later mortality risk: Systematic review. *Annals of Epidemiology,* 17, 278-288.

Batty, G. D., Deary, I. J. & Macintyre S. (2007). Childhood IQ in relation to physiological and behavioural risk factors for premature mortality in middle-aged persons: the Aberdeen Children of the 1950s Study. Journal of Epidemiology and Community Health*, 61, 241-247.*

Batty G.D., Deary, I.J., Schoon, I., & Gale, C. R. (2007). Childhood mental ability in relation to food intake and physical activity in adulthood: the 1970 British Cohort Study. Pediatrics*, 119, e38-e45.*

Calvin, C. M., Deary, I, J., Fenton, C., Roberts, B., Der, G., Leckenby, N., & Batty, G. D. (2011). Intelligence in youth and all-cause mortality: systematic review with meta-analysis. *International Journal of Epidemiology*, 40, 626-644.

Campbell, D. T., & Kenny, D. A. (1999). *A Primer on Regression Artifacts.* New York: Guilford Press.

Christensen, H., Mackinnon, A. J., Korten, A., & Jorm, A. F. (2001). The 'common cause' hypothesis of cognitive ageing: Evidence not only a common factor but also specific associations of age with vision and grip strength in a cross-sectional analysis. *Psychology and Aging,* 16, 588-599.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: an alteration of the RC index. *Behavior Therapy,* 17, 305-308.

Cosway, R., Strachan, M. W., Dougall, A., Frier, B. M., & Deary, I. J. (2001). Cognitive function and information processing in Type 2 diabetes. *Diabetes Medicine,* 18, 803-810.

Cronbach, L. J., & Furby, L. (1970). How should we measure change? Or should we? *Psychological Bulletin,* 74, 68-80.

Deary, I. J., Gale, C. R., Stewart, M. C. W., Fowkes, F. G. R., Murray, G. D., Batty, G. D.,& Price, J. F. (2009). Intelligence and persisting with medication for two years: analysis in a randomised controlled trial. *Intelligence*, 37, 607-612.

Deary. I. J., Gow, A. J., Pattie, A., & Starr, J.M. (in press). Cohort profile: The Lothian Birth Cohorts of 1921 and 1936. *International Journal of Epidemiology*.

Deary, I. J., Gow, A. J., Taylor, M. D., Corley, J., Brett, C., Wilson, V., Campbell, C., Whalley, L. J., Visscher, P. M., Porteus, D. J. & Starr, J. M. (2007). The Lothian birth Cohort 1936: A study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatrics,* 7, doi:10.1186/1471-2318-7-28.

Deary, I. J., Johnson, W., Gow, A. J., Pattie, A., Brett, C. E., Bates, T. C., & Starr, J. M. (2011). Losing one's grip: A bivariate growth curve model of grip strength and nonverbal reasoning from age 79 to 87 years in the Lothian Birth Cohort 1921. *Journals of Gerontology, Series B: Psychological Sciences,* 66, 699-707.

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest*, 11, 53-79.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J. & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology,* 86, 130-147.

Dixon, R. A. (2011). Enduring theoretical themes in psychological aging: Derivations, functions, perspectives, and opportunities. In K. W.Schaie amd S. L. Willis (eds.), *Handbook of the Psychology of Aging, 7th Ed.* (Pp 3-23). San Diego, CA: Academic Press.

Dolcos, S., MacDonald, S. W., Braslavsky, A., Camicioli, R., & Dixon, R. A. (2012). Mild cognitive impairment is associated with selected functional markers: Integrating concurrent, longitudinal, and stability effects. *Neuropsychology,* 26, doi:10.1037/a0026760.

Finkel, D., Reynolds, C. A., McCardle, J. J., & Pedersen, N. L. (2005). The longitudinal relationship between processing speed and cognitive ability: Genetic and environmental influences. *Behavior Genetics,* 35, 535-549.

Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in older adults. *Annual Review of Clinical Psychology,* 5, 363-389.

Gale, C. R., Batty, G. D., Cooper, C., & Deary, I. J. (2009). Psychomotor coordination and intelligence in childhood and heatlh in adulthood: Testing the 'system integrity' hypothesis. *Psychosomatic Medicine,* 71, 675-681.

Gallo, L. C., & Matthews, K. A. (2003). Understanding the association between socioeconomic status and physical health: Do negative emotions play a role. *Psychological Bulletin,* 129, 10-51.

Gayman, M. D., Turner, R. J., & Cui, M. (2008). Physical limitations and depressive symptoms: Exploring the nature of the association. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences,* 63, S219-S228.

Gerstorf, D., Rocke, C., & Lachman, M. E. (2011). Antecedent-Consequent Relations of Perceived Control to Health and Social Support: Longitudinal Evidence for Between-Domain Associations Across Adulthood. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences,* 66, 61-71.

Hanson, L. L. M., Akerstedt, T., Naswall, K., Leineweber, C., Theorell, T., & Westerlund, H. (2011). Cross-Lagged Relationships Between Workplace Demands, Control, Support, and Sleep Problems. *Sleep,* 34, 1403-U147.

Hart, C. L., Taylor, M. D., Davey-Smith, G., Whalley, L. J., Starr, J. M., Hole, D. J., & Deary, I. J. (2003). Childhood IQ, social class, deprivation, and their relationships with mortality and morbidity in later life: Prospective observational study linking the Scottish Mental Survey 1932 and the Midspan studies. *Psychosomatic Medicine,* 65, 877-883.

Hassing, L. B., Grant, M. D., Hofer, S. M., Pedersen, N. L., Nilsson, S. E., & Berg, S. (2004). Type 2 diabetes mellitus contributes to cognitive decline in old age: A longitudinal, population-based study. *Journal of the International Neurological Society,* 10, 599-607.

Hsu, L. M. (1989). Reliable changes in psychotherapy: Taking into account regression toward the mean. *Behavioral Assessment,* 11, 459-467.

Jarvik, L. F., & Falek, A. (1963). Intellectual stability and survival in the aged. *Journal of Gerontology,* 18, 173-176.

Johnson, W., Deary, I. J., McGue, M., & Christensen, K. (2009). Genetic and environmental transactions linking cognitive ability, physical fitness, and education lin late life. *Psychology and Aging,* 24, 48-62.

Kendler, K. S., Gardner, C., Fiske, A., & Gatz, M. (2009). Major depression and coronary artery disease in the Swedish Twin Registry: Phenotypic, genetic, and environmental sources of comorbidity. *Archives of General Psychiatry,* 66, 857-863.

Kuhlen, R. G. (1940). Social change: A neglected factor in psychological studies of the life span. *School and Society,* 52, 14-16.

King, L. A., King, D. W., McArdle, J. J., Saxe, G. N., Doron-LaMarca, S., & Orazem, R. J. (2006). Latent difference score approach to longitudinal traums research. *Journal of Traumatic Stress,* 19, 771-785.

Kooij, D., & Van De Voorde. (2011). How changes in subjective general health predict future time perspective, and development and generativity motives over the lifespan. *Journal of Occupational and Organizational Psychology,* 84, 228-247.

Kuljanin, G., Braun, M. T. & DeShon, R. P. (2011). A cautionary note on modeling growth trends in longitudinal data. *Psychological Methods,* 16, 249-264.

Lapi, F., Pozzi, C., Mazzaglia, G., Ungar, A., Fumagalli, S., Marchionni, N., Geppetti, P., Mugelli, A., & Di Bari, M. (2009). Epidemiology of Suboptimal Prescribing in Older, Community Dwellers A Two-Wave, Population-Based Survey in Dicomano, Italy. *Drugs & Aging*, 26, 1029-1038.

Li, K. Z., & Lindenberger, U. (2002). Relations between ageing sensory/sensorimotor and cognitive functions. *Neuroscience Biobehavioural Review,* 26, 777-783.
Lieberman, M. A. (1966). Observations on death and dying. *Journal of Gerontology,* 6, 70-72.
Mather, K. A., Jorm, A. F., Anstey, K. J., Milburn, P. J., Eastel, S., & Christensen, H. (2010). Cognitive performance and leukocyte telomere length in two narrow age-range cohorts: a population study. *BMC Geriatrics*, 10, 62.
McCardle, J. J., & Woodcock, R. W. (1997). Expanding test-retest designs to include developmental time-lag components. *Psychological Methods,* 2, 403-435.
McGue, M., & Christensen, K. (2002). The heritability of level and rate-of-change in cognitive function in Danish twins age 70 years and older. *Experimental Aging Research,* 28, 435-451.
Menezes Costa, L. D., Maher, C. G., McAuley, J. H., Hancock, M. J., & Smeets, R. J. E. M. (2011). Self-efficacy is more important than fear of movement in mediating the relationship between pain and disability in chronic low back pain. *European Journal of Pain,* 15, 213-219.
Nesselroade, J. R., & Cable, D. E. (1974). "Sometimes it's okay to factor difference scores" -- The separation of state and trait anxiety. *Multivariate Behavioral Research,* 9, 273-284.
Rabbitt, P., Lunn, M., Pendleton, N., & Yardafagar, G. (2011). Terminal pathologies affect rates fo decline to different extents and age accelerates the effects of terminal pathology on cognitive decline. *Journals of Gerontology Series B - Psychological Sciences and Social Sciences,* 66, 325-334.
Rafnsson, S. B., Deary, I. J., Smith, F. B., Whiteman, M. C., Rumley, A., & Lowe, G. D. (2007). Cognitive decline and markers of inflammation and hemostasis: The Edinburgh Artery Study. *Journal of the American Geriatric Society,* 55, 700-707.
Ram, N., Gerstorf, D., Lindenberger, U., & Smith, J. (2011). Developmental change and intraindividual variability: Relating cognitive aging to cognitive plasticity, cardiovascular lability, and emotional diversity. *Psychology and Aging*, 26, 363-371.
Ramsden, S., Richardson, F. M., Josse, G., Thomas, M. S. C., Ellis, C., Shakesshaft, C., Seghier, M. L. & Price, C. J. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature***,** 479, 113-116.
Riegel, K. F., & Riegel, R. M. (1972). Development, drop, and death. *Developmental Psychology,* 6, 306-319.
Rogosa, D. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J. L. Gottman(Ed.), *The Analysis of Change* (pp 1-55). Mahwah, NJ: Lawrence Erlbaum Associates.
Rogosa, D., Brandt, D., & Zimoski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin,* 92, 726-748.
Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin,* 64, 92-107.
Schelleman-Offermans, K., Kuntsche, E., & Knibbe, R. A. (2011). Associations between drinking motives and changes in adolescents' alcohol consumption: a full cross-lagged panel study. (2011) *Addiction,* 106, 1270-1278.
Schram, M. T., Euser, S. M., de Craen, A. J., Witteman, J. C., Frolich, M., & Hofman, A. (2007). Systemic markers of inflammation and cognitive decline in old age. *Journal of the American Geriatric Society,* 55, 708-716.
Scottish Council for Research in Education. (1949). *The Trend of Scottish Intelligence: A Comparison of the 1947 and 1932 Surveys of the intelligence of Eleven-Year-Old Pupils.* London:University of London Press.
Sliwinski, M. J., Stawski, R. S., Hall, C. B., Katz, M., Verghese, J., & Lipton, R. (2006). Distinguishing preterminal and terminal cognitive decline. *European Psychologist,* 11, 172-181.
Sliwinski, M., Hoftman, L., & Hofer, S. M. (2010). Evaluating convergence of within-person change and between-person age differences in age-heterogeneous longitudinal studies. *Research in Human Development,* 7, 45-60.
Terrera, G. M., van den Hout, A., & Matthews, F. E. (2011). Random change point models: Investigating cognitive decline in the presence of missing data. *Journal of Applied Statistics,* 38, 705-716.
Townsend, P. (1979). *Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living.* Harmondsworth: Penguin.
Whitehead, B. P., Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. S. (2011). Are neurocognitive speed and inconsistency similarly affected in type 2 diabetes? *Journal of Clinical and Experimental Neuropsychology,* 33, 647-657.
Wilson, R. S., Beckett, L. A., Bienias, J. L., Evans, D. A., & Bennett, D. A. (2003). Terminal decline in cognitve function. *Neurology,* 60, 1782-1787.
Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavia,* 67, 361-370.

Longitudinal and Life Course Studies 2012 Volume 3 Issue 3 Pp 332 – 345 *ISSN 1757-9597*

The effects of marital status transitions on alcohol use trajectories

Hui Liew
Mississippi State University
hpliew104@gmail.com

(Received November 2011 Revised May 2012)

Abstract

This study aims to exploit the longitudinal nature of the US Health and Retirement Study (HRS), using a semi-parametric mixture modeling (SPMM) approach to examine whether the presence of marital transitions has an impact on alcohol use trajectories among the elderly. The empirical work of this study is based on the 1994-2008 Health and Retirement Study (HRS). Findings provide support to the marriage protection effect: for both elderly men and women, remaining divorced or widowed had detrimental effects on alcohol consumption. Findings also concur with the potential roles of assortive mating / marital homophily: having a drinking spouse increased alcohol consumption.

Introduction

For the past few decades, most researchers have discussed, but few have explicitly tested, the potential roles of marriage protection and assortive mating / marital homophily on subsequent alcohol consumption as an explanatory mechanism. The protective mechanism underscores the social integrative function of marriage and the role of social control over risk-taking behavior. As such, being married is a way of gaining increased protection against adverse health outcomes, and marriage has a protective effect in reducing the risk of substance abuse. The assortive mating / marital homophily mechanism implies that marriage can offer individuals a shared lifestyle (Pienta & Franks, 2006). As such, alcohol consumption may also decrease following the loss of a spouse, especially in the case of the loss of a drinking partner.

Findings pertaining to the marriage protection scenario found inconclusive results. On one hand, proponents of the marriage protection effect have found that transition to marriage reduces alcohol consumption, while transition out of marriage results in increased alcohol consumption (e.g. Bachman, Wadsworth, O'Malley, Schulenberg, & Johnston, 1997; Hajema & Knibbe, 1998; Chilcoat & Breslau, 1996; Leonard and Rothbard, 1999; Miller-Tutzauer, Leonard, &Windle, 1991; Power & Estaugh, 1990). Previous studies have found that married individuals have decreased alcohol use (Harford, Hanna, & Faden, 1994) and steeper rates of decline in alcohol use over time (Curran, Muthen, & Harford, 1998). More recent research by Bogart and her colleagues (2005) found that women who married as young adults were less likely than singles to engage in any alcohol use, heavy episodic drinking, or experience negative consequences, and reported less alcohol use at age 29. Women who married in adolescence (if they had not divorced) consumed less alcohol and were either less likely to engage in heavy episodic drinking (Bogart, Collins, Ellickson, Martino, & Klein, 2005). The impact of remarriage and widowhood has also been highlighted. Divorced women who remarried, drank less than those who did not remarry (Prescott & Kendler, 2001). Relative to men who stayed married over four years, men who became widowed increased their alcohol consumption (Eng, Kawachi, Fitzmaurice, & Rimm, 2005). Indeed, according to Leonard and Rothbard (1999), marriage

ushers in a change in social and recreational activities that can lead to a reduction in alcohol consumption, particularly heavy drinking. Some of these studies found that the protection effect of marriage is stronger among women (Bachman et al., 1997) and Whites (Curran et al., 1998).

On the other hand, other researchers have challenged the marriage protective effect. Early opponents of this effect have provided strong support for increased alcohol consumption following marital dissolution. Still, other researchers as early as the mid-1980s (e.g. Jacob & Bremer, 1986; Yamaguchi & Kandel, 1993) have acknowledged the potential roles of assertive mating / marital homophily, where individuals tended to select spouses or partners with similar drinking patterns. Indeed, earlier research (Wilsnack, Klassen, Schur, & Wilsnack, 1991), found that post-dissolution alcohol consumption was lower among women who transitioned out of marriage. They attributed their finding to the possibility that these women were problem/heavy drinkers and they were removed from their heavy-drinking spouses following marital dissolution (Wilsnack et al., 1991). Evidence in France revealed that divorcees and widows drank less than married women, as measured by fewer glasses per day and fewer days per week drinking wine (Zins, Gueguen, Leclerc, & Goldberg, 1991). Likewise, getting married was accompanied by an increased level of drinking, especially of wine, beginning a year before the wedding and lasting until 4 years after it (Zins et al., 1991). Unfortunately, as shown in Table 1, even though research on the effect of marital transition on alcohol consumption began as early as the 1980s, most existing research tends to focus on young adults (e.g. Bachman et al., 1997; Chilcoat & Breslau, 1996; Power & Estaugh, 1990; Power, Rodgers, & Hope, 1999). Other studies were region-specific (e.g. Bogart et al., 2005; Hajema & Knibbe, 1998; Matzger, Delucchi, Weisner, & Ammon, 2004). To the best of my knowledge, only four studies have examined the marriage effect later in life using repeated measure approaches (e.g. Eng et al., 2005; Miller-Tutzauer et al., 1991) and latent growth models (e.g. Curran et al., 1998; Prescott & Kendler, 2001). On one hand, the risk of alcohol consumption may increase following marital transition, as separated / divorced or widowed individuals may view drinking as a way of to cope with their loss and to reduce stress, especially in the initial months of separation and widowhood. On the other hand, alcohol consumption may also decrease following the divorce, separation, or widowhood, especially if the spouse was a drinking partner, as marriage can offer individuals a shared lifestyle (Pienta & Franks, 2006). This is especially true for females who were married to a male alcoholic (Brennan, Moos, & Kim, 1993).

Alcohol use among the elderly has increasingly become a common but unrecognized problem among health practitioners and policy makers (Dar, 2006) but there is still a scarcity of knowledge regarding the impact of marital transition on alcohol consumption later in life. To fill this research gap, this study aims to exploit the longitudinal nature of the Health and Retirement Study (HRS) to examine whether the presence of marital transitions has an impact on alcohol use trajectories among the elderly.

Table 1. Findings on the impact of marital status on alcohol consumption

Author	Age	Setting	Statistical Method	Dataset / Subjects
Bogart et al., 2005	1,138 women aged 18-29.	U.S. (middle schools in California and Oregon)	Not provided	Data were collected at age 13 and respondents were followed from ages 18 to 29.
Brennan et al., (1993)	Late middle-age women (N=183) and men (N=476).	U.S.	Not provided	Not provided
Curran et al., 1998	4,052 men and women at least 21 of age in 1982.	U.S.	Random coefficient latent curve models	National Longitudinal Survey of Youth 1982-1985.
Dar (2006)	Not provided	England	Not provided	Not provided
Eng et al., 2005	38,865 men aged 40-75	U.S.	Multivariate linear regression performed on repeated measures for marital status and health behaviors.	Longitudinal data with collected at 4-year intervals (1986-90; 1990-94).
Hajema & Knibbe, 1998	1,327 men and women aged 16-69.	Limburg, Netherlands	Not provided	Not provided
Matzger et al., 2004	600 dependent and 992 problem drinkers (men and women).	North Carolina	Nested models using maximum likelihood estimation	Participants were interviewed at baseline and again 1, 3, 5, and 7 years later
Miller-Tutzauer et al., 1991	80,944 women aged 46–71.	U.S.	Repeated-measures MANCOVA (Multivariate Analysis of Variance)	Youth Cohort of the National Longitudinal Survey of Labor Market Experience.
Author	**Age**	**Setting**	**Statistical Method**	**Dataset / Subjects**
Power et al. 1999	35 men and 20 women aged 23-33.	England	Not provided	1958 British birth cohort.
Prescott & Kendler 2001	1,986 female twins aged between 17-61	Not provided	Latent growth models	Data were collected between 1-3 waves over 8 years.
Wilsnack et al., 1991	Not provided	Not provided	Not provided	1981 national survey of women's drinking
Zins et al. 2003	4,782 women	France	Not provided	Data were collected and followed up over a 5-year period.

Method

Subjects

The empirical work of this study is based on seven waves (1994-2008) of the Health and Retirement Study (HRS). The HRS, which is a collaborative effort between the United States (US) Social Security Administration (SSA) and the National Institute on Aging (NIA) at the National Institutes of Health (NIH), has been the principal source of health data on the advanced age population (50 and over) in America. The HRS is an ongoing longitudinal survey with a high re-interview rate. In addition to providing current information on gender, age, race, self-reported health status, marital status, personal income, and education, individuals participated in in-depth interviews about health behaviors, health care utilization and costs, housing, income, assets, employment status and history, physical and mental health, insurance coverage, financial status, family structure, family support systems, labor market status, and retirement planning. The analysis is limited to elderly men and women whose information on alcohol use, marital history, drinking behavior of the spouse, age, sex, ethnicity, education, depression, and chronic illness is available. The final analysis sample consists of 11,742 elderly men and 15,482 elderly women.

Measures

The outcome variable *alcohol consumption*, is computed from responses to questions on the number of drinks consumed on a typical day. I used information from *marital histories* to create dummy variables for always married (reference), continuously divorced or separated, continuously widowed, never married, and transitioned to marriage from singlehood, divorce or widowhood. These variables are constructed by following respondents at each wave of the survey.

Both time-constant and time-varying control variables were included as control variables. Time-constant covariates included age, gender, education, spouse was ever a drinker, and pattern of attrition. *Age* is treated as a continuous variable. *Gender* is formulated as a dichotomy for males and females (reference). Respondent's race is indicated by dummy variables for White/Caucasian (reference), Black / African-American, and Hispanics. *Education* is indicated by dummy variables for less than high school, high school or GED (General Education Development) (reference), and greater than high school. *Spouse was ever a drinker* is a dummy variable, coded as 1 if the spouse reported having consumed alcohol.

Like any other longitudinal studies of older adults, HRS faces the possibility of attrition (primarily due to death and loss at follow-up). Following Vandecasteele and Debels (2007), *pattern of attrition* will be classified into Always Participating, Monotone Participation, and Variable Participation. Always Participating refers to non-attritors. Individuals who were deceased or loss at follow up are considered monotone attritors. Individuals who are in some waves but not other waves are considered variable participants. The implications of attrition might differ across non-attritors, monotone attritors, and variable participants.

Time varying covariates included *comorbidity* (i.e. the number of chronic health conditions) and *depression*. The former is a continuous variable measuring the co-occurrence (sum of) of multiple diagnosed health conditions in an individual within a year. The illnesses included are hypertension, diabetes, cancer, lung disease, heart disease stroke, psychiatric problems, and arthritis. The latter is derived from a mental health index using a score based on the Center for Epidemiologic Studies Depression (CESD) scale. The CESD score, originally developed by Lenore Radloff of Utah State University, is a self-report scale designed to help an individual to determine his or her depressive feelings and behaviors (Radloff, 1977). In the Health and Retirement Study, it is calculated by summing five "negative" indicators and subtracting two "positive" indicators (St. Clair, Blake, Bugliari, Chien, Hayden, Hurd, Ilchuk, Kung, Miu, Panis, Pantoja, Rastegar, Rohwedder, Roth, Carroll, & Zissimopoulos, 2010). The negative indicators measure whether the respondent has experienced depressive sentiments all or most of the time (St. Clair et al., 2010). The five negative indicators ask respondents if they 1) feel depressed, 2) feel that everything they did is an effort, 3) slept restlessly, 4) feel alone, 5) feel sad, or 6) feel like they could not get going (St. Clair et al., 2010). The positive indicators measure whether the respondent felt happy and enjoyed life, all or most of the time (St. Clair et al., 2010).

Method of Analysis

Group-based modeling is a type of latent class growth analysis, and was pioneered by Daniel Nagin and Kenneth Land in 1993. It is used to describe and identify the trajectories, patterns, and changes over time in alcohol use among the elderly men and women. Unlike the other longitudinal data analysis techniques such as the hierarchical or growth curve models, it does not assume a monotonic and regularly varying growth in the overall population (Jones, Nagin, & Roeder, 2001; Nagin, 2005). Instead, it uses a probability function and a semi-parametric mixture modeling (SPMM) approach to classify and assign individuals to their distinctive drinking trajectories based on the patterns of their longitudinal trajectories (Jones et al., 2001). It assumes a number of discrete classes with varying "growths" or trajectories of change, each having a fixed intercept and slope and an estimate of population prevalence (Jones et al., 2001). Under this approach, clusters of individuals with similar developmental trajectories of alcohol consumption over time were identified using the PROC TRAJ procedure in SAS procedure. The likelihood functions of the finite mixture model are formulated by summing *J* finite number of group-specific, or conditional functions, $p^j(Y_i)$ that compose the population, Y_i:

$p(Y_i) = \sum_j^J \pi_j p^j(Y_i)$, where $p(Y_i)$ is the unconditional probability of observing individual *i*'s longitudinal alcohol use trajectories, Y_i, $p^j(Y_i)$ is the probability of Y_i given membership in group J. It is formulated by $p^i(Y_i) = \prod^T p^j(y_{it})$, where $p^j(y_{it})$ is the probability distribution function of y_{it} for an individual *i* in group *j*. π_j is the probability of a randomly chosen individual belonging to group *j*. It is formulated by $\pi_j = \frac{e^{\theta_j}}{\sum_{j=1}^{j} e^{\theta_j}}$. Since there are no predictors for group membership included, θ_j is referred to as a set of base parameters estimated from the multinomial logit model.

The analyses are conducted using SAS 9.2. Respondents are assigned to the trajectory for which they have the largest posterior probability estimate — the group that best conforms to the individual's observed alcohol use severity. Subjects with a high probability (i.e., p≥0.7) of a specific trajectory group were included from further analyses because they were considered correctly assigned to a specific alcohol use groups. Bayesian Information Criterion (BIC)[i] is used to determine the optimal number of trajectory groups that describes the data (Jones et al., 2001) (see Table 2). The analyses examined models for three through seven trajectory groups, with consideration of cubic, quadratic, and linear forms. The procedure initially specified cubic functions for the trajectories, the most general and default option, and the Quadratic or linear forms were utilized in instances where an improvement in fit was noted. Trajectory groups with less than 10% of the sample were considered problematic and rejected (Jones et al., 2001). The approach then uses a multinomial modeling strategy (estimating the parameters of finite mixture models by maximum likelihood) to compare the groups on various predictor variables (Jones et al., 2001). The likelihood for the entire sample is formulated as $L = \prod^N p(Y_i)$, it is the product of the individual likelihood functions of the N individuals who make up the sample, $p(Y_i)$. With the addition of covariates in the model, π_j is now formulated as $\pi_j(x_i) = \frac{e^{x_i\theta_j}}{\sum_j e^{x_i\theta_j}}$, where x_i is the vector for the potential predictors that are associated with a specific group *j* in the model and $\pi_j(x_i)$ is the probability of belonging to group j conditional on x_i. It is much more flexible than hierarchical linear modeling because it allows cross-group differences in the level and shape of trajectories (Jones et al., 2001). Group-based modeling is appropriate for this study because it seeks to understand how multiple distinct patterns of alcohol use change over time and to predict who is likely to belong to a particular trajectory class.

Table 2. Model fit indices, HRS 1994-2008

Number of trajectory groups	Males (N= 11,742)	Females (N = 15,482)
2	-79,240.62	-75,440.25
3	-73,209.68	-68,731.48*
4	-72,974.69*	-67,429.17

Analysis Plan

Two regression models were estimated for the dependent variable. Since the purpose of this study is to examine the impact of marital status transition on alcohol consumption, the first model only included marital history. The second model built upon the first model to add age, race, education, and marital history. It has been well established in every society that men drink more than women (Wilsnack et al., 2000; Holmila & Raitasal, 2005; Emslie, Lewars, Batty, & Hunt, 2009) even though the adverse consequences of alcohol abuse are greater in women, because they are more sensitive to raised blood alcohol levels after a shorter period of drinking smaller amounts of alcohol than for men (Brady & Randall, 1999). Therefore, two separate but identical regression models - one for males and the other for females - will be estimated.

Results

Descriptive Statistics

In both the male and female samples (see Table 3), more than 80 percent of the sample was white and more than 70 percent of the sample had at least a high school education. Most respondents (85 percent for men and 74 for women) remained married. On average, respondents were 75 years of age. The average posterior assignment probabilities for each group (provided upon request) indicate that 91–96% of elderly men and 96-98% of elderly women were correctly assigned to different depression groups.

Table 3. Descriptive Statistics

	Mean (SD) / %	
	Men (N=11,742)	**Women (N=15.482)**
No marital transition		
Always married/partnered*	85.33	74.11
Never married	0.46	0.51
Always divorced*	3.69	5.50
Always widowed*	0.13	1.32
Experienced transition		
Marriage to divorce*	0.65	0.50
Marriage to widowed*	7.45	15.90
Single/Divorced/Widowed to married*	2.28	2.16
Age in 1994	75.21 (12.29)	75.10 (13.81)
Race		
White*	82.76	80.08
Black*	14.80	17.39
Hispanic	2.43	2.53
Education		
Less than high school	26.44	26.30
High School or GED*	32.26	37.71
Greater than high school*	41.30	35.99

* *statistically significant at $p \leq .05$*

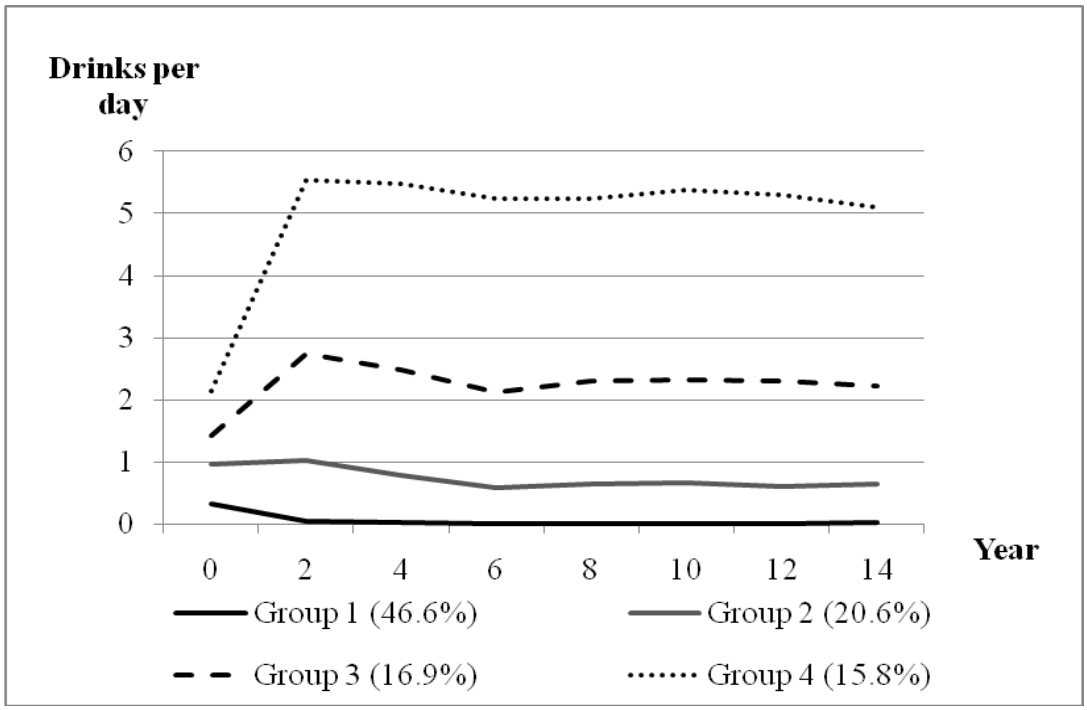
Multivariate analysis

With respect to alcohol use, four trajectory models were retained as the final model for men and three trajectory group models were retained for women. For men, alcohol use trajectories of Low and Decreasing, Moderate and Fluctuating, and High and Chronic groups were contrasted with those of non-drinkers. For women, alcohol use trajectories of Low and High and Chronic groups were contrasted with those of non-drinkers. Thus, the interpretations for all the parameter estimates are relative to the reference (i.e. non-drinking) group.

Elderly men

Figure 1 showed that the first trajectory (i.e. Group 1), representing nearly half of the population, may be depicted as a non-drinking group. This trajectory is essentially flat. The second trajectory (i.e. Group 2), constituting about 21 percent of the population, may be depicted as a Low and Decreasing alcohol use group. It showed a very slight increase in alcohol use from an immediate to the second year, a decline from the second to the sixth year, and tended to stabilize thereafter. The third trajectory (i.e. Group 3), constituting nearly 17 percent of the population, may be depicted as a Moderate and Fluctuating alcohol use group. Individuals in this group experienced a drastic increase in alcohol use from an immediate to the second year, followed by a decline from the second to sixth year, and tended to stabilize thereafter. The fourth trajectory (i.e. Group 4), constituting about 16 percent of the population, is a High and Chronic alcohol use group. Individuals in this group experienced a drastic increase in alcohol use from in the first two years, and tended to slightly fluctuate thereafter.

(**Figure 1. Group trajectories for alcohol consumption, Elderly Men, HRS 1994-2008**



The interpretations for all the parameter estimates are relative to the reference (i.e. non-drinking) group. Elderly men who remained divorced were more likely than their counterparts who remained married to belong to the Low and Decreasing and Moderate and Fluctuating alcohol use groups. Results in models 1 and 2 of Table 4 also suggested that the effect of marital statuses on alcohol use is conditioned by the socio-demographic characteristics of the individuals (age, race, education, the drinking behavior of the deceased spouse) and the pattern of attrition (monotone attrition or variable participation). Results in Model 2

suggested that elderly men who remained widowed were more likely than their counterparts who remained married to belong to the High and Chronic alcohol use group when these variables were taken into account. Controlling for individual socio-demographic characteristics and pattern of attrition, elderly men who transitioned from marriage to divorced, were less likely to than their counterparts who remained married, to belong to the Low and Decreasing, Moderate and Fluctuating, and High and Chronic alcohol use groups. In contrast, elderly men who became widowed, were more likely than their counterparts who remained married, to belong to the High and Chronic alcohol use group.

As reported in Model 2 of Table 4, the estimates of comorbidity were negative and significant for all alcohol use groups except the High and Chronic group, suggesting that in general, increased number of chronic illnesses is associated with a reduction in alcohol use over time. Results reported in Model 2 also revealed that the likelihood of belonging to the Low and Decreasing, Moderate and Fluctuating, and High and Chronic alcohol use groups decreased with the age. When compared to White men, Black men were more likely to belong to the Moderate and Fluctuating alcohol use groups, but were less likely to belong to the High and Chronic alcohol use group. Hispanic men were more likely than White men to belong to the Low and Decreasing and Moderate and Fluctuating alcohol use groups.

Elderly men with less than a high school education were less likely than their counterparts with a high school or GED education to belong to the Low and Decreasing, Moderate and Fluctuating, and High and Chronic alcohol use groups, while the opposite is observed among their counterparts with education beyond high school. Elderly men who were married to a drinker were more likely to belong to the Low and Decreasing, Moderate and Fluctuating, and High and Chronic alcohol use groups. Monotone (permanent) attritors were less likely than non-attritors to belong to the High and Chronic alcohol use group. Temporary attritors (variable participation) were more likely than non-attritors to belong to the Moderate and Fluctuating alcohol use group.

Table 4. Parameter estimates for depression, elderly men, HRS 1994-2008

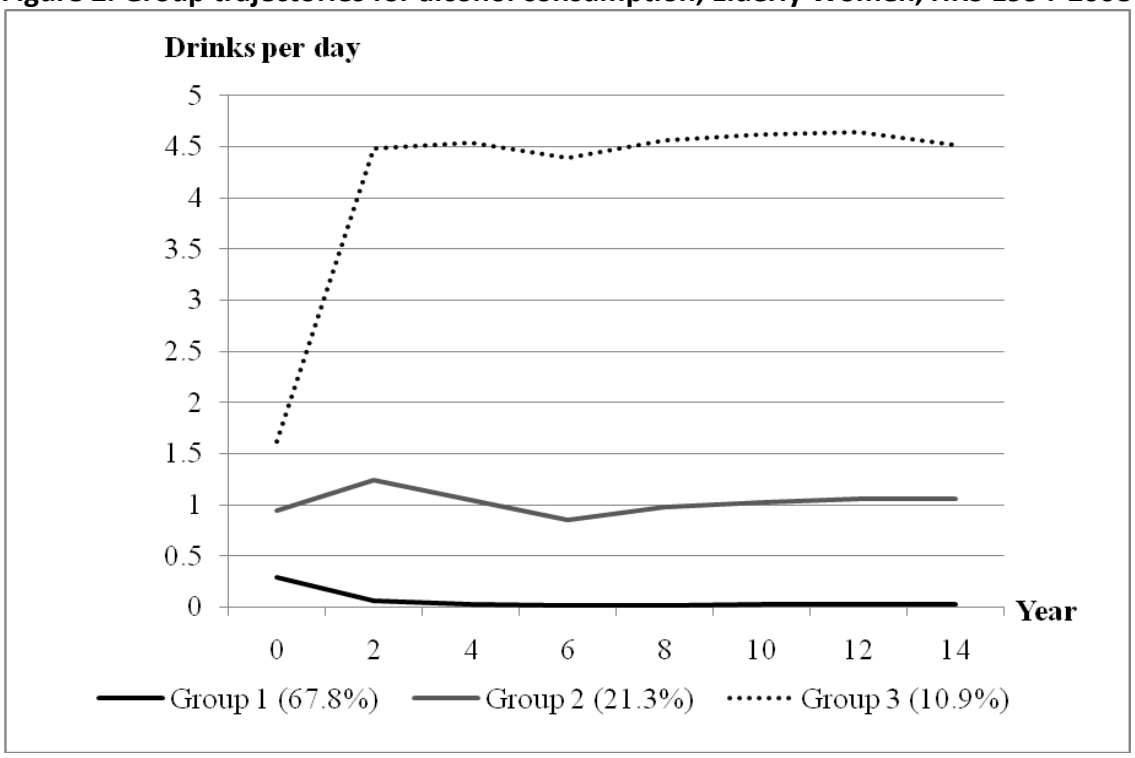
Group	Parameter	Model 1	Model 2
Non-Drinker (1)	Intercept	-0.160***	-0.934***
	Linear	-0.899***	-0.967***
	Quadratic	0.053***	0.056***
	Comorbidity		-0.110**
	Depressive Symptoms		-0.035
Low and Decreasing (2)	Intercept	0.040***	0.149***
	Linear	-0.090***	-0.108***
	Quadratic	0.004***	0.006***
	Comorbidity		-0.111***
	Depressive Symptoms		0.002
Moderate and Fluctuating (3)	Intercept	0.806***	0.804***
	Comorbidity		-0.018*
	Depressive Symptoms		0.003
High and Chronic (4)	Intercept	1.604***	1.590***
	Comorbidity		0.021***
	Depressive Symptoms		-0.010**

*** $P<0.001$; ** $P<0.01$; * $P<0.05$

Elderly women

Figure 2 showed that the first trajectory (i.e. Group 1), representing nearly 70 percent of the population, may be depicted as a non-drinking group. This trajectory is essentially flat. The second trajectory (i.e. Group 2), constituting about 21 percent of the population, may be depicted as a Low alcohol use group. It showed a very slight increase in alcohol use from an immediate to the second year, a decline from the second to the sixth year, and tended to stabilize thereafter. The third trajectory (i.e. Group 3), constituting nearly 11 percent of the population, may be depicted as a High alcohol use group. Individuals in this group experienced a drastic increase in alcohol use from an immediate to the second year, followed by stabilization thereafter.

Figure 2. Group trajectories for alcohol consumption, Elderly Women, HRS 1994-2008



Results in Model 1 of Table 5 revealed that elderly women who remained divorced and elderly women who transitioned from marriage to widowhood were more likely than their counterparts who remained married to belong to the Low and High and Chronic alcohol use groups. Elderly women who remained widowed were more likely than their counterparts who remained married to belong to the Low alcohol use group. Elderly women who transitioned from singlehood, divorce, and widowhood to marriage were more likely than their married counterparts to belong to the Low alcohol use group. However, the effect of transitioning into marriage became insignificant when individual socio-demographic and pattern of attrition were added in Model 2.

As reported in Model 2 of Table 5, the estimates of depression were negative and significant for all alcohol use groups among elderly women. This suggests that an increased number of chronic illnesses is associated with a reduction in alcohol use over time. The estimates for comorbidity were negative and significant for all alcohol use groups except the High and Chronic alcohol use group. This suggests that in general, increased depression is associated with decreased alcohol use for elderly women.

As reported in Model 2 of Table 5, the likelihood of belonging to the Low alcohol use groups decreased with the age. However, the likelihood of belonging to the High and Chronic alcohol use group increased with age. Black and Hispanic women were less likely than their White counterparts to belong to the Low and High and Chronic alcohol use groups. Elderly women with less than a high school education were less likely than their counterparts with a high school or GED education to belong to the Low and High and Chronic alcohol use groups, while the opposite is observed among their counterparts with education beyond high school. Elderly women who were married to a drinker were more likely to belong to the Low and High and Chronic alcohol use groups. Monotone (permanent) attritors were less likely than non-attritors to belong to the Low and High and Chronic alcohol use group.

Table 5. Parameter estimates for depression, elderly women, HRS 1994-2008

Group	Parameter	Model 1	Model 2
Non-Drinker (1)	Intercept	-1.279***	-0.916***
	Linear	-0.729***	-0.704***
	Quadratic	0.043***	0.041***
	Comorbidity		-0.215***
	Depressive Symptoms		-0.052**
Low (2)	Intercept	0.060**	0.196***
	Linear	-0.020**	-0.016*
	Quadratic	0.001**	0.002**
	Comorbidity		-0.118***
	Depressive Symptoms		-0.011*
High and Chronic (3)	Intercept	0.940***	0.972***
	Linear	0.145***	0.145***
	Quadratic	-0.008***	-0.008***
	Comorbidity		-0.009
	Depressive Symptoms		-0.015***

**** P<0.001; ** P<0.01; * P<0.05*

Discussion

Patterns of alcohol consumption across time were relatively stable among the elderly. A similar result was found by Penny Brennan and her colleagues (Brennan, P., Schutte, K., & Moos, 2010). Findings provide support to the marriage protection effect. Remaining divorced or widowed had detrimental effects on alcohol consumption for elderly men and women. Moderate and high alcohol consumption over time was observed among elderly men who remained divorced. High alcohol consumption over time was observed among elderly men who remained divorced or widowed and among elderly women who remained divorced. This resonates with previous research (i.e. Ekerdt, deLabry, Glynn, & Davis, 1989; Eng et al., 2005; Power et al., 1999). Indeed, earlier research has established that heavy drinking is related to high levels of stress (Kleinke, Staneski, & Meeker, 1983) and low levels of social support (Berkman & Syme, 1979). More recent research also revealed that more frequent visits with family members and friends was predictive of a lower initial level of, and slower rate of change in, alcohol consumption over time (Brennan et al., 2010). A related finding by Stolzenberg (2001) also revealed that spouses, especially wives, tend to play a greater role in the social control of health behaviors and

maintaining contact with friends and family. Thus, it is possible that these elderly men and women who remained divorce or separated experienced strains associated with household management and changes in the patterns of kin network and social support.

Elderly men who transitioned from marriage to divorce drank significantly less than their counterparts who remained married. Elderly women who transitioned from marriage to widowhood were less likely to become moderate or heavy drinkers compared to the married counterparts. Likewise, having a drinking spouse increased alcohol consumption for both men and women. These findings concurred with those of the earlier studies that acknowledged the potential roles of assortive mating / marital homophily in reducing alcohol consumption following divorce or separation (e.g. Jacob & Bremer, 1986; Wilsnack et al., 1991; Yamaguchi & Kandel, 1993). Since marriage can offer individuals a shared lifestyle, alcohol consumption may decrease following the loss of a spouse, especially if the spouse was a drinking partner (Pienta & Franks, 2006) because these elderly men were removed from their drinking spouses following divorce or separation (Wilsnack et al., 1991) or widowhood.

Age had different effects on drinking trajectories for elderly men and elderly women. Among elderly men, the likelihood of belonging to the Low and Decreasing, Moderate and Fluctuating, and High and Chronic alcohol use groups decreased with the age. However, among elderly women, the likelihood of belonging to the High and Chronic alcohol use group increased with age. Similarly, an earlier study in France also found evidence that women in the oldest generation drank more than the younger women (Zins et al., 1991).

Black and Hispanic men were more likely than Whites to be moderate drinkers. However, Black men were less likely than White men to be heavy drinkers. For women, ethnic minorities were less likely than Whites to be moderate and heavy drinkers. The results corroborate the findings by Stephen Gilman and his colleagues (2008), where Blacks, Asians, and Hispanics have significantly lower odds of alcohol dependence than their white counterparts. Another related finding by Platt, Sloan, and Costanzo (2010) found that Whites were more likely than Blacks to increase alcohol consumption over time. Earlier findings also found that Blacks tended to have lower and later initiation, as well as a slower rate of acceleration into regular alcohol use than Whites (Johnston et al., 1995).

Elderly men and women with education beyond high school have greater likelihood of belonging to the Moderate and High alcohol use groups than their counterparts with a high school or GED education, while the reverse is true for elderly men and women with less than a high school education. This finding resonated with that of Ruchlin (1997), who found a positive association between educational levels and binge drinking. A more recent study by Platt, Sloan, and Costanzo (2010) found an increase in alcohol consumption over time among highly educated individuals.

Implications

The HRS, however, does not provide information on marital quality (e.g. happiness, satisfaction, sharing of activities, conflict, communication, etc.), family support (e.g. tangible help, emotional attachment, etc.), age of drinking onset, the duration of alcohol consumption, and personality traits. Once more recent, detailed and sufficient data become available, one of the important tasks for future research is to examine the interplay between family support and subsequent alcohol consumption following the death of a spouse, to explore how age of drinking onset and the duration of alcohol consumption affects alcohol use trajectories, and to focus on the ways in which personality traits and marital quality may impact subsequent drinking trajectories among the elderly population.

Information about dementia is not included in the list of number of chronic health conditions in the HRS. Even though it is widely acknowledged that dementia can be a significant problem among the elderly and is related to alcohol consumption, we do not know whether an elderly respondent who participated in the survey has dementia. Once more recent, detailed and sufficient data become available, future researchers should focus on how dementia may impact subsequent drinking trajectories among the elderly population.

The HRS does not provide information on clinically diagnose depression, and CESD is the only proxy for the presence of depression symptoms in the HRS. The scoring system of the CESD can only be used as an indicator of symptoms relating to depression but not to clinically diagnose depression. Higher scores may indicate that further clinical tests/screenings are warranted, and lower scores do not suggest the absence of clinical depression.

Despite the limitations of the data, this study is a step towards describing and identifying the trajectories, patterns, and changes over time in alcohol consumption, and subsequently identifying the potential predictors of these trajectories. This is more essential than ever, as late-life alcohol use and misuse has become a significant problem among the older adult population, and many elderly people are more vulnerable to the detrimental effects of alcohol (Salmon & Forester, 2012). Indeed, the number of adults over 50 with any substance abuse problems (alcohol, prescription/ illicit drugs) is projected to triple and the number of older adults needing substance abuse treatment is projected to double in 2020 (Gfroerer, Penne, Pemberton, & Folsom, 2003). As the United States continues to experience rapid social, economic, demographic, and health changes as baby boomers reach old age, this unique understanding can inform policy makers dealing with health and retirement policies on issues that are pertinent to the elderly widowed population.

Older adults with alcohol abuse and dependence problems are far less likely to be identified as having substance use disorders than younger persons (Blazer and Wu, 2009). As the needs of the elderly are different from those of younger individuals, the tools used for screening younger populations, such as standard questionnaires and other instruments, may not be adequate for use among the elderly (Beullens & Aertgeerts, 2004; Nemes et al., 2004). Therefore, few treatment facilities are designed with older adults in mind and are often ill-equipped to meet the specific needs of the elderly, such as the ability to address the limited mobility and access to transportation among patients (Schultz, Arndt, & Liesveld, 2003). Therefore, health practitioners need to be equipped with the necessary clinical skills to prevent, diagnose, and treat alcohol problems among the elderly. Targeted screening that considers current alcohol consumption, the history of drinking problems, and help-seeking, could help identify older adults at higher risk for excessive or problematic drinking (Moos et al., 2010). Elderly men, Whites, respondents who remained divorced or widowed, and respondents who were married to a drinker, should be the main target of policy interventions designed to reduce alcohol use disorders.

References

Bachman, J. G., Wadsworth, K. N., O'Malley, P. M., Schulenberg, J., Johnston, L. D. (1997). Marriage, divorce, and parenthood during the transition to young adulthood: Impacts on drug use and abuse. In J. Schulenberg, J. L. Maggs, & K. Hurrelmann. (Eds.), *Health Risks and Developmental Transitions During Adolescence* (pp. 246-279). New York: Cambridge University Press.

Berkman, L. F., & Syme, S. L. (1979). Social networks, host resistance and mortality: a nine year follow-up study of Alameda county residents. *American Journal of Epidemiology 109*(2),186-204.

Beullens, J., & Aertgeerts, B. (2004). Screening for alcohol abuse and dependence in older people using DSM criteria: a review. *Aging and Mental Health 8*(1), 76-82.

Blazer, D. G., & Wu, L. T. (2009). The epidemiology of substance use and disorders among middle aged and elderly community adults: National survey on drug use and health. *American Journal of Geriatric Psychiatry 17*(3), 237-245.

Bogart, L. M., Collins, R. L., Ellickson, P. L., Martino, S. C., & Klein, D. J. (2005). Effects of early and later marriage on women's alcohol use in young adulthood: a prospective analysis. *Journal of Studies on Alcohol and Drugs 66*(6), 729-737.

Brady, K.T., & Randall, C. L. (1999). Gender differences in substance use disorders. *The Psychiatric Clinic of North America 22*(2), 241-252.

Brennan, P. L., Moos, R. H., & Kim, J. Y. (1993). Gender Differences in the Individual: Characteristics and Life Contexts of Late-Middle-Aged and Older Problem Drinkers. *Addiction,* 88(6), 781-787.

Brennan, P., Schutte, K., & Moos, R. (2010). Patterns and predictors of late-life drinking trajectories: A 10-year longitudinal study. *Psychology of Addictive Behaviors*, *24*(2), 254-264.

Chilcoat, H. D., & Breslau, N. (1996). Alcohol disorders in young adulthood: effects of transitions into adult roles. *Journal of Health and Social Behavior,* 37(4)*,* 339-349.

Curran, P. J., Muthen, B. O., & Harford, T. C. (1998). The Influence of Changes in Marital Status on Developmental Trajectories of Alcohol Use in Young Adults. *Journal of Studies on Alcohol and Drugs, 59*(6), 647-658.

Dar, K. (2006). Alcohol use disorders in elderly people: fact or fiction? *Advances in Psychiatric Treatment,* 12, 173-181.

Ekerdt, D. J., deLabry, L. O., Glynn, R. J., & Davis, R. (1989). Change in drinking behaviours with retirement: findings from the normative ageing study. *Journal of Studies on Alcohol and Drugs, 50*(4), 347-353.

Eng, P. M., Kawachi, I., Fitzmaurice, G., & Rimm, E. B. (2005). Effects of marital transitions on changes in dietary and other health behaviours in US male health professionals. *Journal of Epidemiology and Community Health, 59*(1), 56-62.

Emslie, C., Lewars, H., Batty, G. D., & Hunt, K. (2009). Are there gender differences in levels of heavy, binge and problem drinking? Evidence from three generations in the west of Scotland. *Public Health, 123*(1), 12-14.

Gfroerer, J., Penne, M., Pemberton, M., & Folsom, R. (2003). Substance abuse treatment need among older adults in 2020: the impact of the aging baby-boom cohort. *Drug and Alcohol Dependence, 69*(2), 127-135.

Hajema, K. J., & Knibbe, R.A. (1998). Changes in social roles as predictors of changes in drinking behavior. *Addiction,* 93(11), 1717-1727.

Harford, T. C., Hanna, E. Z., & Faden, V. B. (1994). The long-and short-term effects of marriage on drinking. *Journal of Substance Abuse, 6*(2), 209-217.

Holmila, M., & Raitasalo, K. (2005). Gender differences in drinking: why do they still exist? *Addiction, 100*(12), 1763-1769.

Jacob, T., & Bremer, D. A. (1986). Assortative mating among men and women alcoholics. *Journal of Studies on Alcohol and Drugs*, *47*(3), 219-222.

Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (1995). *National Survey Results on Drug Use from the Monitoring the Future Study, 1975-1994*. NIH Publication No. 95-4026, Washington: Government Printing Office.

Jones, B. L., Nagin, D., & Roeder, K. (2001). A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. *Sociological Methods & Research, 29*(3), 374-393.

Kleinke, C. L., Staneski, R. A., & Meeker, F. B. (1983). Attributions for smoking behavior: comparing smokers with nonsmokers and predicting smokers' cigarette consumption. *Journal of Research in Personality, 17*(2), 242-255.

Leonard, K. E., & Rothbard, J. C. (1999). Alcohol and the marriage effect. *Journal of Studies on Alcohol and Drugs*, *Suppl.* 13, 139-146.

Matzger, H., Delucchi, K., Weisner, C., & Ammon, L. (2004). Does marital status predict long-term drinking? Five-year observations of dependent and problem drinkers. *Journal of Studies on Alcohol and Drugs, 65*(2), 255-265.

Miller-Tutzauer, C., Leonard, K. E., & Windle, M. (1991). Marriage and alcohol use: a longitudinal study of ''maturing out''. *Journal of Studies on Alcohol and Drugs, 52*(5), 434-440.

Moos, R. H., Schutte, K. K., Brennan, P. L., & Moos, B. S. (2010). Late-life and life history predictors of older adults' high-risk alcohol consumption and drinking problems. *Drug Alcohol Dependence*, *108*(1–2), 13–20.

Nagin, D. S. (2005). *Group-based modeling of development*. Cambridge, Massachusetts: Harvard University Press.

Nemes, S., Rao, P. A., Zeiler, C., Munly, K., Holtz, K. D., & Hoffman, J.H. (2004). Computerized screening of substance abuse problems in a primary care setting: Older vs. younger adults. *American Journal of Drug and Alcohol Abuse, 30*(3), 627-642.

Pienta, A. M., & Franks, M. M. (2006). A closer look at health and widowhood: do health behaviors change after the loss of a spouse? In D. Carr., R. M. Nesse., C. B. Wortman, (eds.), *Spousal bereavement in late life* (pp. 117-142). New York: Springer.

Platt, A., Sloan F. A., & Costanzo, P. (2010). Alcohol-consumption trajectories and associated characteristics among adults older than age 50. *Journal of Studies on Alcohol and Drugs, 71*(2), 169-179.

Power, C., & Estaugh, V. (1990). The role of family formation and dissolution in shaping drinking behavior in early adulthood. *British Journal of Addiction, 85*(4), 521-530.

Power, C., Rodgers, B., & Hope, S. (1999). Heavy alcohol consumption and marital status: disentangling the relationship in a national study of young adults. *Addiction, 94*(10), 1477–1487.

Prescott, C. A., & Kendler, K. S. (2001). Associations between marital status and alcohol consumption in a longitudinal study of female twins. *Journal of Studies on Alcohol and Drugs 62*(5), 589-604.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Ruchlin, H.S. (1997). Prevalence and correlates of alcohol use among older adults. *Preventive Medicine, 26*(5, Pt.1), 651-657.

St. Clair, P., Blake, D., Bugliari, D., Chien. S., Hayden, O., Hurd, M., Ilchuk, S., Kung, F. Y., Miu, A., Panis, C., Pantoja, P., Rastegar, A., Rohwedder, S., Roth, E., Carroll, J., & Zissimopoulos, J. (2010). *RAND HRS Data Documentation,* Version J. Santa Monica, California: RAND Corporation, RAND Center for the Study of Aging.

Salmon, J. M. & Forester, B. (2012). Substance Abuse and Co-occurring Psychiatric Disorders in Older Adults: A Clinical Case and Review of the Relevant Literature. *Journal of Dual Diagnosis*, *8*(1), 74-84.

Schultz, S. K., Arndt, S., & Liesveld, J. (2003). Locations of facilities with special programs for older substance abuse clients in the US. *International Journal of Geriatric Psychiatry, 18*(9), 839–843.

Stolzenberg, R. (2001). It's about time and gender: Spousal employment and health. *American Journal of Sociology, 107*(1), 61-100.

Vandecasteele, L., & Debels, A. (2007). Modelling attrition in panel data: the effectiveness of weighting. *European Sociological Review, 23*(1), 81-97.

Wilsnack, S. C., Klassen, A. D., Schur, B. E., & Wilsnack, R. W. (1991). Predicting onset and chronicity of women's problem drinking: a five-year longitudinal analysis. *American Journal of Public Health, 81*(3), 305-318.

Wilsnack, R.W., Vogeltanz, N. D., Wilsnack, S. C., Harris, T. R., Ahlström, S., Bondy, S., Csémy, L., Ferrence, R., Ferris, J., Fleming, J., Graham, K., Greenfield, T., Guyon, L., Haavio-Mannila, E., Kellner, F., Knibbe, R., Kubicka, L., Loukomskaia, M., Mustonen, H., Nadeau, L., Narusk, A., Neve, R., Rahav, G., Spak, F., Teichman, M., Trocki, K., Webster, I., & Weiss, S. (2000). Gender differences in alcohol consumption and adverse drinking consequences: cross-cultural patterns. *Addiction 95*(2), 251-265.

Yamaguchi, K., & Kandel, D. B. (1993). Marital homophily on illicit drug use among young adults: assortative mating or marital influence? *Social Forces, 72*(2), 505-528.

Zins, M., Gueguen, A., Leclerc, A., & Goldberg, M. (1991). Alcohol Consumption and Marital Status of French Women in the GAZEL Cohort: A Longitudinal Analysis between 1992 and 1996. *Journal of Studies Alcohol and Drugs, 64*(6), 784-789.

Endnote

[i] A criterion for model selection with the purpose of providing a measure of the weight of evidence to select the model that describes the data adequately. It is defined as BIC = log(L) – 0.5$k$log(N), where L is the value of the model's maximized likelihood, N is the sample size, and k is the number of parameters in the model. The model with the smallest absolute BIC is chosen.

RESEARCH NOTE

Life course influences on quality of life at age 50 years: evidence from the National Child Development Study (1958 British birth cohort study)

D Blane, [1,2] M Wahrendorf, [3,2] E Webb, [1,2] G Netuveli [1,2]

1. Imperial College London
2. ESRC *International Centre for Life Course Studies in Society and Health*
3. University of Duesseldorf

d.blane@imperial.ac.uk

Abstract

The objectives of this study were to investigate whether prospective data reveal life course influences on quality of life at older ages; to establish a baseline for the evolution of quality of life through the Third Age; and to estimate the relative importance of direct and indirect effects in these life course relationships. We used the age 50 years sweep of the National Child Development Study (1958 British birth cohort study) included the CASP measure of positive quality of life at older ages, allowing prospective path analysis of life course influences on quality of life at the start of the Third Age. We found that material (social class; deprivation) and psycho-social (family conflict; family fracture) circumstances in childhood and adulthood were linked using path analysis to CASP scores at age 50 years. The strength of these relationships was modest; and their influence was primarily indirect via well-recognised contemporaneous factors. Prospective data revealed life course influences on quality of life at the start of the Third Age. We conclude that the influence of these longitudinal factors is weak in comparison with that of contemporaneous circumstances. In this respect quality of life differs from health.

Keywords: National Child Development Study; life course; CASP measure of positive quality of life at older ages; *Third Age*

Introduction

The presently reported study, which analyses prospective life course data up to age 50 years, is situated at the transition between mid-life and early old age; in Laslett's terms, the Second and Third Ages (Laslett, 1996). As well as being of substantive interest in its own right, the study is a baseline for future studies of the evolution of quality of life through the Third Age, including whether the balance between contemporaneous and life course influences shifts with age towards the former. Three issues are of particular interest: whether prospective data confirm previous findings based on retrospective data; whether prospective data reveal life course influences on quality of life which are akin to those on health, where there are important long-lasting direct effects; or whether life course influences on quality of life are primarily indirect via recognised contemporaneous factors.

Mid-life is characterised by responsibility for others, in particular: supporting and socialising offspring through their childhood and adolescence into early adulthood; satisfying employers, to maintain employment income, and cooperating with fellow employees. As well as being a stage of

life with high responsibilities, mid-life includes the experience of success or failure in achieving important goals and satisfying important non-material needs (Willis, Martin, & Rocke, 2010). Among the latter, continued experience of control and autonomy and regular experience of personal achievement are core elements of quality of life (Higgs et al., 2003), which also influence later health (Jopp & Schmitt, 2010).

Recent increases in life expectancy at middle age and the spread of second pensions, both private and occupational, which in some countries supplement the State Pension, have combined to create a new phase of the life course which Laslett has christened the *Third Age* (Laslett, 1996). This Third Age lies between the end of family formation and paid employment, on the one hand, and the onset of physical dependency on the other. Freed of responsibility for children and to employer and in reasonable health and financial circumstances, a person, perhaps for the first time in their life, can concentrate on self and self-realisation. Although the concept of the Third Age has its critics, such as Bury (1995) who wondered whether it is a middle class phenomenon, and Gilleard & Higgs (2000), who stress the diversity of responses to this stage of life, it undoubtedly points to the potentially positive opportunities opening up to a growing proportion of people in early old age. The Third Age does not have specified age limits, but most descriptions would include the ages 55-75 years.

Previous research has identified the main predictors of positive quality of life in early old age as contemporaneous good physical functioning, financial adequacy, an absence of clinical depression, good quality personal relationships & frequent social contact, participation in social and voluntary activities and, negatively, being unable to move residence from a disliked neighbourhood (Bowling, Farquhar, & Browne, 1991; Farquhar, 1995; Bowling, 1995; Bowling, 1996; Wiggins et al., 2004: Netuveli et al., 2006; Wahrendorf et al., 2006; Knesebeck, Ovd *et al.,* 2007; Siegrist & Wahrendorf, 2009; Wahrendorf & Siegrist, 2010; Webb et al., 2011). The current evidence suggests that these same factors account for short term change in quality of life (Webb et al., 2011), but that any long-term life course effects are either non-existent or weak (Blane et al., 2004). In this respect, quality of life at older ages seems to differ from physical health, where life course effects can be long-term and powerful (Davey Smith et al., 1997; Montgomery et al., 2000; Langenberg et al., 2003; Kuh et al., 2006; Langenberg et al., 2006;).

In contrast to the evidence relating to life course influences on physical health at older ages, which have come from prospective data, previous investigations of potential life course effects on quality of life at older ages have had to depend on retrospective data (Blane et al., 2004; Siegrist & Wahrendorf, 2009). The present research note corrects this deficiency by using prospectively collected life course data to investigate whether circumstances during childhood and adulthood influence quality of life at the start of the Third Age. We hypothesise that such effects will appear prospectively, with some direct effects in addition to indirect effects via the recognised contemporaneous factors mentioned above.

Methods

Data

The National Child Development Study (NCDS), also known as the 1958 British birth cohort study, sampled all those born in England, Wales and Scotland during one week in 1958. Study participants have been re-visited periodically to produce a prospective longitudinal data set, which is deposited on open academic access at the UK Data Archive. The presently reported NCDS analyses use information collected at birth and from follow-up sweeps at ages 7, 11, 16, 23, 33 and 50 years.

Missing data, due to sample attrition, item non-response and so forth, are a significant problem in birth cohort studies such as NCDS, whose original sample of some 18,000 births reduced to, at best, around 7,000 with complete data in the present analyses at age 50 years. We have addressed this problem by means of multiple imputation. Imputation was conducted on all NCDS members interviewed in 2008 (N=9790), thereby excluding all who had left the study due to death and other reasons. Results from both data sets, complete and imputed, are shown.

Operationalisation

The logic of the investigation was to start from the main contemporaneous influences on quality of life at older ages, about which there is a secure consensus, to specify features of childhood and adult life which plausibly might feed into them and,

thereby, to identify any unexplained influences which appear as direct effects. Life course material disadvantage was operationalised in terms of social class and deprivation, which are seen as potential influences on financial adequacy and physical functioning in early old age. The parallel psycho-social disadvantages are family conflict and family fracture (divorce or separation of parents, or death of one or both parents), impacting later clinical depression and social participation (the term clinical depression is used to distinguish it from transient feelings of unhappiness, not to imply validation by clinical interview).

Measures during childhood and adulthood

(a) Social class was measured using the Registrar General's classification of occupations, aggregated into social classes V-IIIM (manual) and IIIN-I (non-manual), based on father's occupation in 1958, when the NCDS participant was born, and the NCDS participant's own occupation at age 33 years. (b) Accumulated deprivation during childhood was measured using the sum of four binary items, giving a scale 0-4. The items were: father being in social class V or IV when the NCDS participant was born; family having reported financial difficulties when the NCDS participant was aged 7 years; NCDS participant having received free school meals at age 11 years; family having reported financial difficulties when the NCDS participant was aged 16 years. Accumulated deprivation during adulthood was measured using the sum of three binary items at ages 23 and 33 years, giving a scale 0-6: occupying crowded residential accommodation; receiving welfare benefits; living in rented accommodation. (c) Family conflict was measured by whether or not the school health visitor reported parental discord when the NCDS participant was aged 7 years. The NCDS data set does not contain a comparable measure at 23 or 33 years, so family fracture at age 33 years was used instead. (d) Family fracture during childhood and at age 33 years were measured as divorce, separation or death of parent(s) before the NCDS participant was aged 7 years; and divorce, separation or death of spouse by age 33 years.

Measures at age 50 years

Financial situation was measured using a five category variable from the question whether the household of the NCDS participant experienced financial difficulties (1 = none; 5 = great difficulties). We considered limiting illness was indicated by the SF36 physical functioning score (0-100 re-scaled 0-10). Clinical depression was measured by the nine-item version of the Malaise Inventory (Rutter, Tizzard, & Whitmore, 1970) on a continuous scale 0-9, where a score of four or more is considered indicative of depression. Social participation was measured as a score based on the past, current and active membership of civil society organisations, such as parents' associations, religious groups, political parties and social and voluntary organisations. We used these variables either as continuous or as multiple categories to use the maximum information captured within them.

Quality of life was measured in its positive, *Third Age,* sense (Higgs et al*.,* 2003; Hyde et al*.,* 2003) by CASP-12 version 2 (Wiggins et al*.,* 2008), which consists of 12 Likert-scaled items, each scored 0-3, that access the dimensions of control, autonomy, self-realisation and pleasure. Although CASP was developed to study the third age, by focusing on the psychosocial needs of this age group, we assume that its domains are relevant also to the mid-life period where, as mentioned previously, the experience of control and autonomy and regular experiences of personal achievement are important for quality of life (Willis, Martin, & Rocke, 2010) and later health (Jopp & Schmitt, 2010). CASP-12v.2 has a range of 0-36. For comparing size effects, the difference in mean CASP-12v.2 scores between those with no current long-standing illness and those with a current limiting long-standing illness, in this NCDS sample, is 4.9 units on the CASP-12v.2 scale.

Analysis

Data were analysed in STATA version 11 for descriptive analyses and Mplus version 4 for path analyses, with SEM models of direct and indirect effects, using WLSMV, as binary variables in the paths (Muthen, 2011). The multiple imputation created five sets of data with missing values imputed, giving 97% efficiency based on a comparison of the confidence interval in the imputed data with the confidence interval of the results obtained from the data if there was no missing information (Rubin, 1987). Path analysis disaggregated the relationships between the material and psycho-social features of childhood and adulthood, on the one hand, and quality of life

at age 50 years, on the other, into direct effects and indirect effects acting via those factors at age 50 years which are known to influence quality of life. Given that gender did not interact with childhood exposures, gender was adjusted in all models.

Results

Moderate to severe financial difficulties were experienced by 7.5% of the sample. The average SF36 physical functioning score was 1.4 (SD 2.2), with 29.9% of the sample above this level. The mean Malaise Inventory score measuring depression was 1.5 (SD 1.9), with 14.7% recording scores of four or more (indicating clinical depression). The maximum value of the social participation score was 29 out of a possible 50; with a mean value of 26.4 (SD 3.3) and 58.9% above this value (results not shown).

Table 1 shows that mean quality of life at age 50 years, as measured by CASP-12v.2, varies between those who had been in non-manual and manual social classes at birth (difference: 0.95 CASP units; $p<0.001$) and at age 33 years (difference: 1.17 CASP units; $p<0.0001$), with non-manual having the higher quality of life.

It is related in a stepwise fashion also to the level of deprivation experienced during childhood (difference 0-3 items: 2.87 CASP units; test for trend $p<0.0001$) and adulthood (difference 0-6 items: 4.34 CASP units; test for trend $p<0.0001$), with the least deprived having the highest quality of life.

In terms of psychosocial factors, mean quality of life at age 50 years, as measured by CASP-12v.2, is associated with family conflict at age 7 years (difference: 1.54 CASP units; $p<0.0001$) and with family fracture during childhood (difference: 1.52 CASP units; $p<0.0001$), but not with family fracture at age 33 years (difference: 0.44 CASP units; p 0.132).

In six out of the seven instances, the difference between the complete case and imputed means is small; the modest exception is deprivation during adulthood.

Figures 1-4 show the results of the path analyses. Following the convention to present the most parsimonious model, a path is shown only when its coefficient is significant statistically at the 95% level or above. The path coefficients are standardised beta coefficients, allowing comparison of relative importance. The un-bracketed coefficients are derived from complete data, while the bracketed coefficients are derived from the imputed data.

Below each path diagram, a table reports the contribution of the various pathways to the relationship between childhood circumstances and quality of life at age 50 years, in both CASP and standard deviation units. The final column shows the share of each pathway. In Figure 1 for example, of the effect of father's social class at birth on quality of life at age 50 years : 26% goes through social class at 33 years and its impact on quality of life at 50 years ; 45% goes through social class at 33 years and its impact on financial situation at 50 years; and 29% goes through social class at 33 years and impact on limiting illness at 50 years.

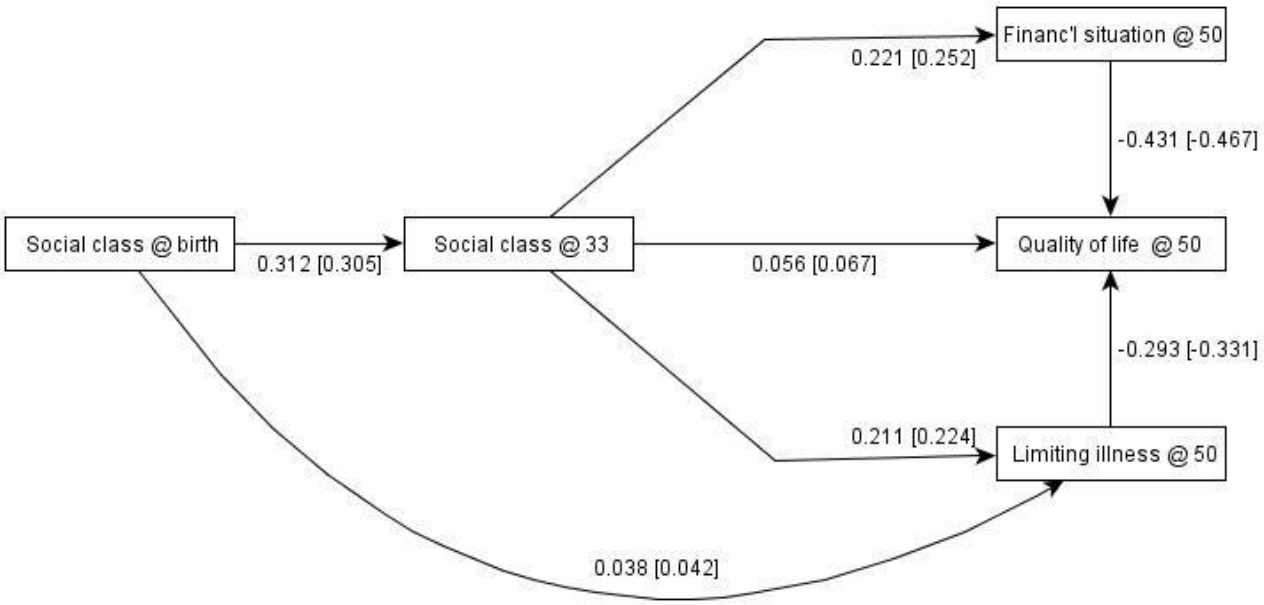
Figure 1 shows: (1) social class at birth is directly associated with the NCDS member's social class at age 33 years and with limiting illness at age 50 years; with the former having the stronger effect - so, if you are born into a working class family, you are more likely in early adulthood to be working in a manual occupation and to enter early old age with a limiting illness; (2) social class at 33 years is directly associated with financial situation and limiting illness at 50 years, as well as having a residual positive direct effect on quality of life at 50 – if you are a manual worker, you are more likely to enter early old age with financial and health problems although, in partial compensation, some aspect of working class life, perhaps its generosity, will enhance your later quality of life; (3) both financial situation and limiting illness at 50 years have inverse direct effects on contemporaneous quality of life – entering early old age with financial and health problems reduces quality of life. The variables included in the model explain 44% of the relationship between social class at birth and quality of life at 50 years. The combined direct and indirect effects of life course (childhood and 33 years) circumstances on quality of life at 50 years is modest compared with the influence of contemporaneous financial situation and limiting illness. The coefficients derived from the imputed data set, differ little from those based on complete cases.

Table 1. Quality of life (mean CASP-12v.2) at age 50 years by circumstances in childhood and adulthood; complete and imputed data

	Childhood						Adulthood					
	All available data			Imputed data			All available data			Imputed data		
	N	Mean	95%CI	N	Mean	95%CI	N	Mean	95%CI	N	Mean	95%CI
Social Class*												
Non-	2385	26.83	26.61, 27.05	2862	26.76	26.56, 26.96	4509	26.65	26.34, 26.62	6021	26.48	26.49, 26.82
Manual	5439	25.88	25.73, 26.04	6928	25.65	25.51, 25.79	2564	25.48	24.97, 25.36	3769	25.16	25.25, 25.70
Deprivation counts**												
0	2903	26.67	26.46, 26.87	2627	26.8	26.59, 27.01	2304	27.1	26.97, 27.41	2195	27.19	26.88, 27.31
1	4731	26.02	25.85, 26.18	5585	25.92	25.77, 26.07	3507	26.46	26.28, 26.62	4055	26.45	26.28, 26.65
2	751	25.33	24.91, 25.76	1096	25.35	25.00, 25.70	1615	24.99	24.97, 25.48	2051	25.22	24.69, 25.29
3	244	23.8	22.97, 24.63	482	23.49	22.87, 24.10	608	24.74	24.17, 24.94	1007	24.56	24.23, 25.24
4							196	23.78	22.36, 23.76	365	23.06	22.86, 24.69
5							46	22.76	19.74, 22.47	117	21.1	20.77, 24.75
Family fracture‡												
	8295	26.17	26.05, 26.30	9395	26.02	25.91, 26.14	4982	26.55	25.92, 26.16	8905	26.04	26.39, 26.70
	334	24.65	23.95, 25.35	395	24.8	24.17, 25.42	405	26.11	24.88, 25.70	885	25.29	25.54, 26.68
Family conflict#												
	6284	26.29	26.15, 26.43	9293	26.04	25.93, 26.16						
	324	24.75	24.06, 25.44	497	24.6	24.02, 25.17						

Notes.
** Childhood: at birth. Adulthood: at 33 years.*
*** Childhood: father's social class at birth, family financial difficulties at age 7, free school meals at age 11, and family financial difficulties at age 16. Adulthood: living in crowded accommodation, receipt of benefits, and living in rented accommodation at ages 23 and 33.*
‡Divorce, separation or death of parent(s) at age 7; that of spouse at age 33.
#Childhood: School health visitor reported parental discord at age 7.

Figure 1. Path analysis of father's social class at time of NCDS member's birth and NCDS member's quality of life (CASP-12v.2) at age 50 years; results for complete data (N=6552) and imputed data [N= 9790]

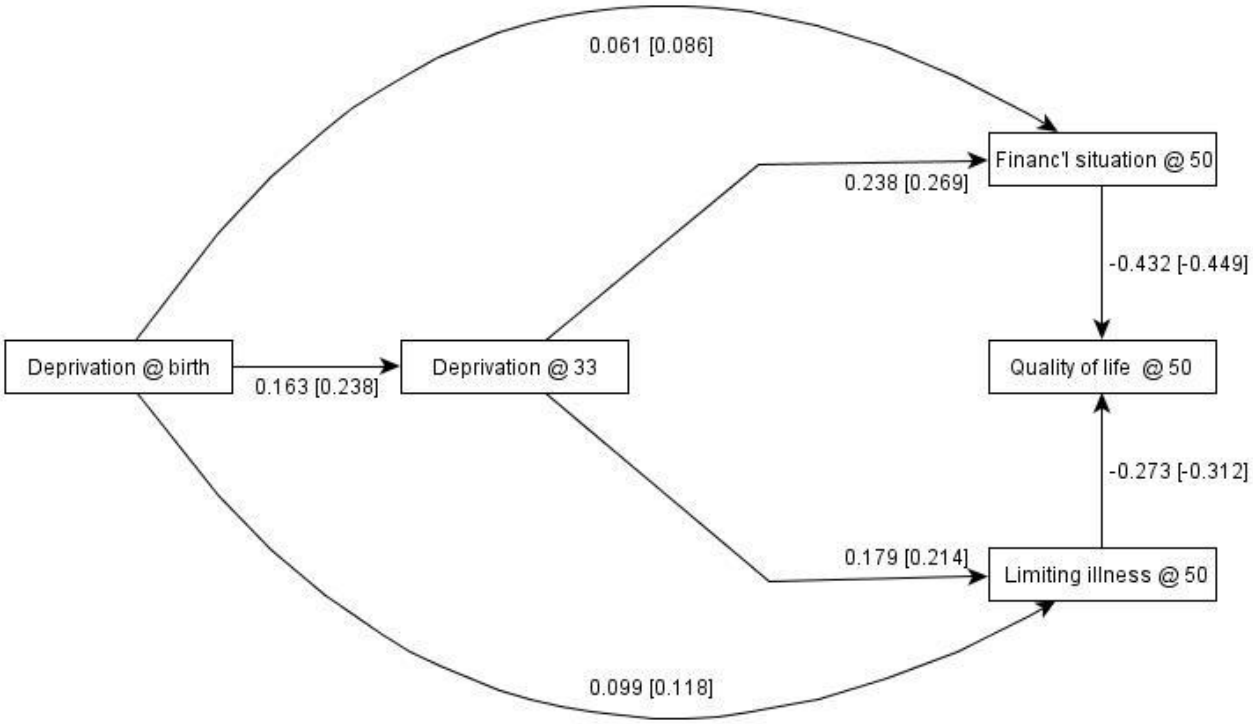


Effect of father's social class at birth on quality of life at 50 years

Paths	**Effect**		**Contribution**
	CASP units	SD units	%
Social class @33	0.217 [0.262]	0.017 [0.020]	26 [26]
Social class @33 > Financ'l situation @ 50	-0.372 [-0.461]	-0.030 [-0.036]	45 [45]
Social class @33 > Limiting illness @ 50	-0.241 [-0.290]	-0.019 [-0.022]	29 [29]
Total indirect effects	-0.396 [-0.489]	-0.032 [-0.038]	100 [100]
Proportion of total association explained by the model		44%	

Figure 2 shows that: (1) deprivation during childhood has direct effects on deprivation at age 33 years and on financial situation and limiting illness at 50 years, with the former having the strongest effect; (2) deprivation at 33 years has direct effects on financial situation and limiting illness at 50 years, although unlike social class at 33 there is no direct effect on quality of life; (3) as in Figure 1, financial situation and limiting illness at 50 years have inverse direct effects on contemporaneous quality of life.

Figure 2. Path analysis of accumulated material deprivation during NCDS member's childhood and adolescence and their quality of life (CASP-12v.2) at age 50 years; results for complete data (N=3147) and imputed data [N=9790]

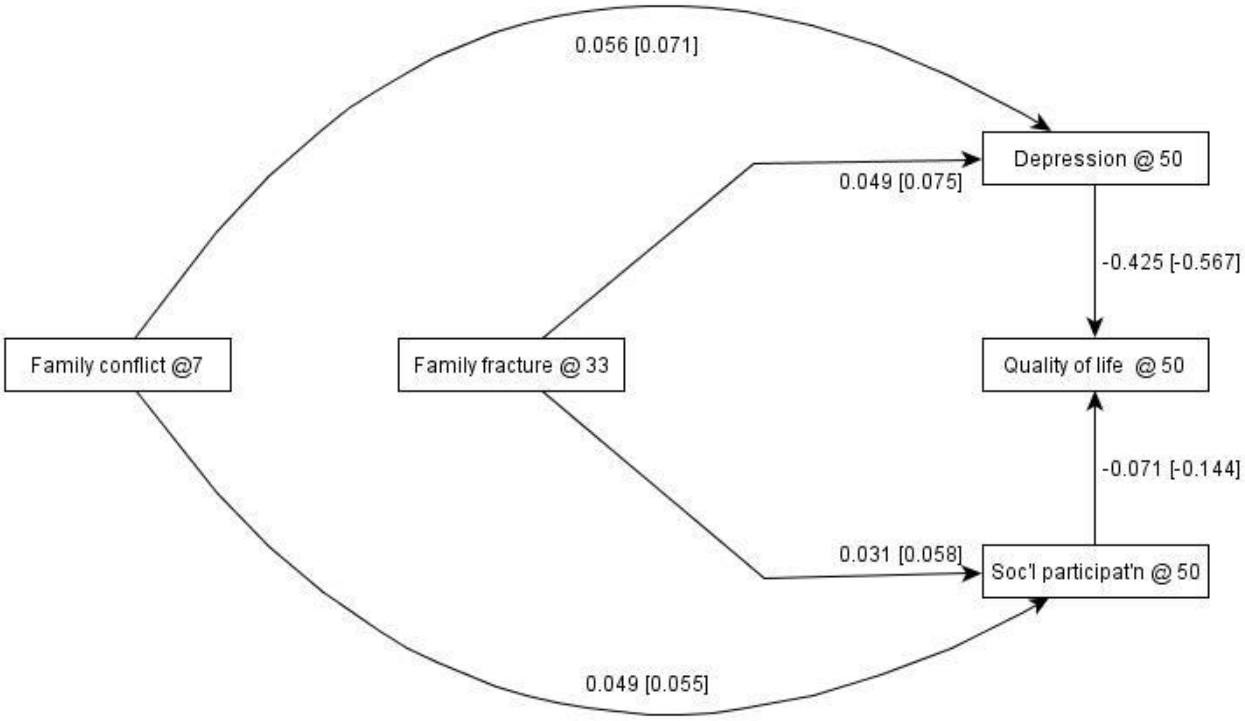


Effect of childhood and adolescent deprivation on quality of life at 50 years.

Path	Effect		Contribution
	CASP units	SD units	%
Deprivation @33 > Financ'l situation @ 50	-0.134 [-0.208]	-0.017 [-0.028]	68 [68]
Deprivation @33 > Limiting illness @ 50	-0.064 [-0.115]	-0.008 [-0.016]	32 [32]
Total indirect effects	-0.198 [-0.323]	-0.025 [-0.044]	100 [100]
Proportion of total association explained by the model		32%	

Figure 3 shows that: (1) family conflict at age 7 years has direct effects on depression and social participation at 50 years, but no direct effect on family fracture at 33 years; (2) family fracture at 33 years has direct effects on depression and social participation at 50 years; (3) both depression and social participation at 50 years have inverse direct effects on contemporaneous quality of life.

Figure 3. Path analysis of family conflict when NCDS member was aged seven years and NCDS member's quality of life (CASP-12v.2) at age 50 years; results for complete data (N=4619) and imputed data [N=9790]

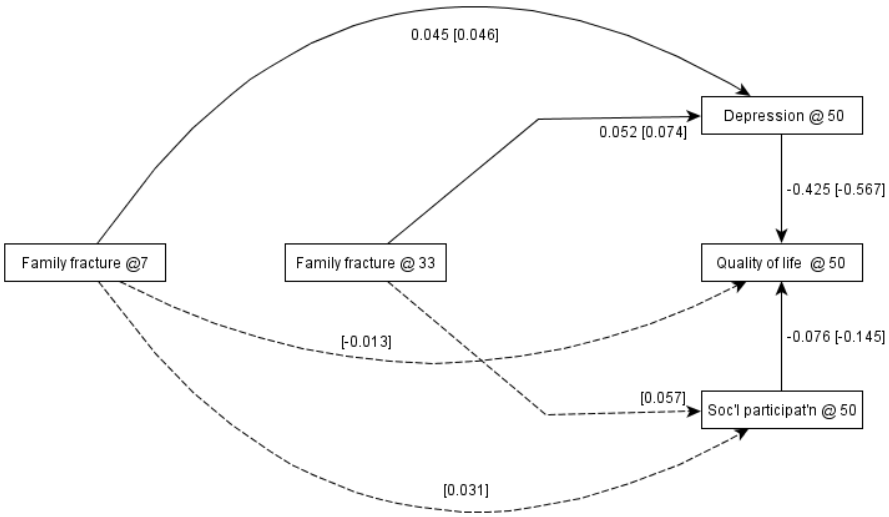


Effect of family conflict at seven on quality of life at 50 years.

Path	**Effect**		**Contribution**
	CASP units	SD units	%
Depression @50	-0.713 [-1.065]	-0.024 [0.003]	87 [89]
Social participation @ 50	-0.104 [-0.128]	-0.003 [-0.008]	13 [11]
Total indirect effects	-0.817 [-1.193]	-0.027 [-0.047]	100 [100]
Proportion of total association explained by the model		87%	

Figure 4 shows that: (1) family fracture before age 7 years has a direct effect on depression at age 50 years, but no direct effect on family fracture at 33 years; (2) family fracture at 33 years has a direct effect on depression at 50 years; (3) both depression and social participation at 50 years have inverse direct effects on contemporaneous quality of life; (4) three pathways are significant statistically only in the imputed data set: a direct path between family fracture before age 7 years and quality of life at 50 years; and paths to social participation at age 50 years from family fracture before age 7 and family fracture by age 33 years.

Figure 4. Path analysis of family fracture by the time the NCDS member was aged seven years and NCDS member's quality of life (CASP-12v.2) at age 50 years; results for complete data (N=5919) and imputed data[N=9790]



Note. Dotted line refer to paths significant only in imputed datasets

Effect of family fracture by age seven on quality of life at 50 years.

Path	Effect		Contribution
	CASP units	SD units	%
Depression @50	-0.609 [-0.776]	0.019 [0.026]	92 [60]
Social Participation @50	-0.050 [-0.135]	-0.002 [-0.004]	8 [10]
Total indirect effects	-0.659 [-0.911]	-0.021 [-0.030]	100 [70]
Direct effects	0 [-0.390]	0 [-0.013]	0 [30]
Total effects	-0.659 [1.301]	-0.021 [-0.043]	100 [100]
Proportion of total association explained by the model		99%	

Discussion

The use of prospective data has revealed life course influences on quality of life at age 50 years which were not detected when broadly similar analyses were conducted on life course data collected retrospectively. Four things are worth noting about these life course influences. First, their effect is modest; only the experience of deprivation in childhood and deprivation at age 33 years have an effect approaching in size the effect of having a limiting illness. Second, their effect is modest also in comparison with the influence on quality of life of well-recognised contemporaneous factors, such as financial adequacy, physical functioning, clinical depression and social participation. Third, with few exceptions, the life course affects quality of life at 50 years indirectly, via influence on these well recognised contemporaneous factors; only social class at age 33 years and, in the imputed data set only, family fracture before age seven years have a direct impact on quality of life at 50. Finally, the modest and largely indirect nature of life course influences on quality of life differs from the life course influences on physical health, where long-term effects are integral to the aetiology and pathology of, for example, cardiovascular and respiratory disease.

Three potential weaknesses need to be considered. First, an important assumption when using multiple imputation is that data are missing at random. In the case of the NCDS this may not be an always appropriate assumption, although grounds for cautious confidence come from the fact that in the present analyses multiple imputation mostly did not revise the substantive results derived from complete data. Comparison between the coefficients from the complete and imputed data models is difficult, because they come from different data sets, but the similarity in the contribution of the different paths suggests that our models are valid. The one exception is the path model for family fracture at age seven years, where the imputed data showed three paths which were not significant statistically in the complete cases model and contemporaneous depression and social participation had stronger associations with quality of life. This exception may be due to the involvement of family fracture in the imputation of all missing dependant variables (primary outcome and mediators) in the regression equations. As a counter-balance, it is worth remembering that loss to follow-up in longitudinal studies tends to be greatest at the extremes of the social structure, thereby giving a conservative bias to results.

Second, we had the modest aim of showing the direct and indirect influences of childhood factors on adult wellbeing, so, when specifying our path models, we were thrifty in our use of variables, thereby exposing ourselves to the danger of omitting variables that might confound our models (Ploubidis & DeStavola, 2011).

Third, it is reasonable to protest that the *Third Age* does not start until after age 50 years, the age at which the outcome data used here were collected. The present findings come from the transition between mid-life and the Third Age; and should be seen as the baseline for future studies of early old age. It will be interesting to see whether the impact of life course influences on quality of life become stronger and more direct as the NCDS cohort ages deeper into the *Third Age*.

Two substantive conclusions of the analyses here reported are interesting. The first is scientific: that different dimensions of life at older ages appear to vary in their sensitivity to life course influences. The second concerns policy: that material deprivation early in life can have a significant long-term effect, but that contemporaneous circumstances are nevertheless the key to *Third Age* quality of life. These are considered in greater detail below.

It is reasonable to assume that the various dimensions of life at older ages differ in their relationship to life course processes. Some dimensions may reflect primarily current circumstances and cultural context, while others involve changes that are written into the body's structure. Mapping such differences is part of life course research. Compare quality of life in this respect with diet and nutrition and physical health. The present paper has reported that life course influences on quality of life at the start of the *Third Age* are of a modest size, similar to those life course effects on nutrition at older ages found by Maynard and colleagues (Maynard et al., 2006). Using data from the Boyd Orr Cohort (Martin et al., 2005), namely dietary intake measured during childhood (household food inventory) and again in early old age (food frequency questionnaire), Maynard showed that the quality of a person's diet in early old age was strongly associated with their current

way of life (financial and marital circumstances; tobacco smoking) but that, even after allowing for these contemporaneous factors, vegetable consumption in childhood predicted dietary quality 60 years later.

Life course influences on physical health at older ages appear to be stronger than those on quality of life or nutrition. Using data from a sub-sample of the Boyd Orr Cohort who had completed a lifegrid (Blane et al.,1999), Montgomery and colleagues showed that slow pre-pubertal growth, interpreted as an indicator of psycho-social stress during childhood, interacted with later occupational stress during adulthood to raise blood pressure in early old age (Montgomery et al., 2000); findings which were replicated in the National Survey of Health and Development (Langenberg et al., 2005).

The variability between quality of life, nutrition and health, in the importance of life course influences, prompts the question of how other aspects of life at older ages might behave. Montgomery's work on resilience (Osika et al., 2006; Montgomery et al., 2007; Osika & Montgomery, 2008) suggests that more complex relationships are possible in which, for example, early life advantage can paradoxically be a disadvantage when faced with adversity at older ages (the disappointment paradox).

Of the life course exposures investigated here, material deprivation during childhood and adulthood had the most substantial effects on quality of life at age 50 years. This is a demonstration of a relationship found previously in cross-sectional data from the English Longitudinal Study of Ageing, where social class differences in quality of life at ages 50-74 years were largest in those unemployed or permanently sick (Blane et al., 2007). In the presently reported analyses, these relationships were indirect, via contemporaneous financial adequacy and physical functioning.

The work of Morris and colleagues on the Minimum Income for Healthy Living (MIHL) is relevant to both financial adequacy and physical functioning at older ages. Morris found that the UK State Pension in 2007 provided only two-thirds of the Minimum Income for Healthy Living for an older person without disability, and that even the higher UK Pension Credit Guarantee was some £12 pounds per person per week less than MIHL (Morris et al., 2007). The weekly shortfall in income will be greater for older people with disability and impaired physical functioning. With good reason, Goldblatt and Marmot's recent policy review of health inequalities in England and Wales emphasised the importance of the Minimum Income for Healthy Living (Marmot Review 2010).

Conclusions

Prospective data reveal life course influences on quality of life at the start of the *Third Age*. The influence of these longitudinal factors mainly is indirect and weak in comparison with that of contemporaneous circumstances. In this respect quality of life differs from health.

Acknowledgements

Professor Harvey Goldstein gave us the idea of including adult circumstances in our analyses. The comments of our anonymous reviewers improved greatly our original text. ESRC *International Centre for Life Course Studies in Society and Health* (ICLS) provided the intellectual context of this work. And UK Economic and Social Research Council funded ICLS: Award number: RES-596-28-0001.

References

Blane, D., Berney, L., Davey Smith, G., Gunnell, D., & Holland, P. (1999). Reconstructing the life course: a 60 year follow-up study based on the Boyd Orr cohort. *Public Health,* 113, 117-124.

Blane, D., Netuveli, G., & Bartley, M. (2007). Does quality of life at older ages vary with socio-economic position? *Sociology* 41, 717-726.

Blane, D., Higgs, P., Hyde, M., & Wiggins, R. (2004). Life course influences on quality of life in early old age. *Social Science and Medicine* 58, 2171-2179.

Bowling, A., Farquhar, M., & Browne, P. (1991). Life satisfaction and associations with social networks and support variables in three samples of elderly people. *International Journal of Geriatric Psychiatry* 6, 549-566.

Bowling, A. (1995). The most important things in life: comparisons between older and younger population age groups by gender; results from a national survey of the public's judgements. *International Journal of Health Sciences* 6, 169-175.

Bowling, A. (1996). The effects of illness on quality of life: findings from a survey of households in Great Britain. *Journal of Epidemiology and Community Health* 50, 149-155.

Bury, M. (1995). Ageing, gender and sociological theory. In S. Arber, & J. Ginn (Eds.), *Connecting Gender and Ageing* (pp.15-30). Philadelphia: Open University Press.

Davey Smith, G., Hart, C., Blane, D., Gillis, G., & Hawthorne, V. (1997). Lifetime socioeconomic position and mortality: prospective observational study. *British Medical Journal* 314, 547-552.

Farquhar, M. (1995). Elderly people's definitions of quality of life. *Social Science and Medicine* 41, 1439-1446.

Gilleard, C., & Higgs, P. (2000). *Cultures of Ageing.* London: Prentice Hall.

Higgs, P., Hyde, M., Wiggins, R., & Blane, D. (2003). Researching quality of life in early old age: the importance of the sociological dimension. *Social Policy and Administration* 37, 239-252.

Hyde, M., Wiggins, R., Higgs, P., & Blane, D. (2003). A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). *Aging and Mental Health* 7, 186-194.

Jopp, D., & Schmitt, M. (2010). Dealing with negative life events: differential effects of personal resources, coping strategies and control beliefs. *European Journal of Ageing* 7, 167-180. doi: 10.1007/s10433-010-0160-6.

Knesebeck, Ovd., Wahrendorf, M., Hyde, M., & Siegrist, J. (2007). Socio-economic position and quality of life among older people in 10 European countries: results of the SHARE study. *Ageing and Society* 27, 269-284.

Kuh, D., Hardy, R., Butterworth, S., Okell, L., Richards, M., Wadsworth, M., Cooper, C., & Sayer, A. (2006). Developmental origins of midlife physical performance: evidence from a British birth cohort. *American Journal of Epidemiology* 164, 110-121.

Langenberg, C., Hardy, R., Kuh, D., Brunner, E., & Wadsworth, M. (2003). Central and total obesity in middle aged men and women in relation to lifetime socioeconomic status: evidence from a national birth cohort. *Journal of Epidemiology and Community Health* 57, 816-822.

Langenberg, C., Hardy, R., Breeze, E., Kuh, D., & Wadsworth, M. (2005). Influence of short stature on the change in pulse pressure, systolic and diastolic blood pressure from age 36 to 53 years : an analysis using multilevel models. *International Journal of Epidemiology* 34, 905-913.

Langenberg, C., Kuh, D., Wadsworth, M., Brunner, E., & Hardy, R. (2006). Social circumstances and education: life course origins of social inequalities in metabolic risk in a prospective national birth cohort. *American Journal of Public Health* 96, 2216-2221.

Laslett, P. (1996). *A Fresh Map of Life* (2nd Edition). London: Macmillan Press.

Marmot Review. (2010). *Fair Society, Healthy Lives* : strategic review of health inequalities in England post-2010. London: Marmot Review.

Martin, R., Gunnell, D., Pemberton, J., Frankel, S., & Davey Smith, G. (2005). Cohort profile : The Boyd Orr Cohort – an historical cohort study based on the 65 year follow-up of the Carnegie Survey of Diet and Health (1937-39). *International Journal of Epidemiology* 34, 742-749.

Maynard, M., Gunnell, D., Abraham, L., Ness, A., Bates, C., & Blane, D. (2006). What influences diet at older ages? Prospective and cross-sectional analyses of the Boyd Orr cohort. *European Journal of Public Health* 16, 315-323.

Muthén, B. (2011). Applications of causally defined direct and indirect effects in mediation analysis using SEM in Mplus. Retrieved from http://www.statmodel.com/download/causalmediation.pdf.

Montgomery, S., Berney, L., & Blane, D. (2000). Pre-pubertal growth and blood pressure in early old age. *Archives of Disease in Childhood* 82, 358-363.

Montgomery, S., Netuveli, G., Hildon, Z., & Blane, D. (2007). Does financial disadvantage at older ages eliminate the potential for better health? *Journal of Epidemiology and Community Health* 61, 891-895.

Morris, J.N., Wilkinson, P., Dangour, A., Deeming, C., & Fletcher, A. (2007). Defining a minimum income for healthy living (MIHL) : older age, England. *International Journal of Epidemiology* 36, 1300-1307.

Netuveli, G., Hildon, Z., Montgomery, S., Wiggins, R., & Blane, D. (2006). Quality of life at older ages: evidence from English Longitudinal Study of Ageing. *Journal of Epidemiology and Community Health* 60, 357-363.

Osika, W., Ehlin, A., Montgomery, S. (2006). Does height modify the risk of angina associated with economic adversity? *Economics & Human Biology* 4, 398-411.

Osika, W., & Montgomery, S. (2008). Economic disadvantage modifies the association of height with low mood in the US, 2004: the disappointment paradox. *Economics & Human Biology* 6, 95-107.

Ploubidis, G., & DeStavola, B. (2011). Health differentials in the older population of England: an empirical comparison of the materialist, lifestyle and psychosocial hypotheses. *BMC Public Health* 11:390.

Rubin, D. (1987). *Multiple Imputation for Non-response in Surveys*. New York: J. Wiley & Sons.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, Health and Behaviour*. London: Longmans.

Siegrist, J., & Wahrendorf, M. (2009). Participation in socially productive activities and quality of life in early old age: findings from SHARE. *Journal of European Social Policy* 19, 317-326.

Wahrendorf, M., Knesebeck, O., & Siegrist, J. (2006). Social productivity and well-being of older people: baseline results from the SHARE study. *European Journal of Ageing* 3, 67-73.

Wahrendorf, M., & Siegrist, J. (2010). Are changes in productive activities of older people associated with changes in their well-being?: results from a longitudinal European study. *European Journal of Ageing* 7, 59-68.

Webb, E., Netuveli, G., McMunn, A., & Blane, D. (2011). Proximal predictors of change in quality of life at older ages. *Journal of Epidemiology and Community Health* 65, 542-547.

Wiggins, R., Higgs, P., Hyde, M., & Blane, D. (2004). Quality of life in the third age: key predictors of the CASP-19 measure. *Aging and Society* 24, 693-708.

Wiggins, R., Netuveli, G., Hyde, M., Higgs, P., & Blane, D. (2008). The evaluation of a self-enumerated scale of quality of life (CASP-19) in the context of research on ageing: a combination of exploratory and confirmatory approaches. *Social Indicators Research* 89, 61-77.

Willis, S., Martin, M., & Rocke, C. (2010). Longitudinal perspectives on midlife development : stability and change. *European Journal of Ageing* 7, 131-134. doi: 10.1007/s10433-010-0162-4.

Longitudinal and Life Course Studies 2012 Volume 3 Issue 3 Pp 359 – 368 ISSN 1757-9597

RESEARCH NOTE

Parental separation and adult psychological distress: evidence for the 'reduced effect' hypothesis?

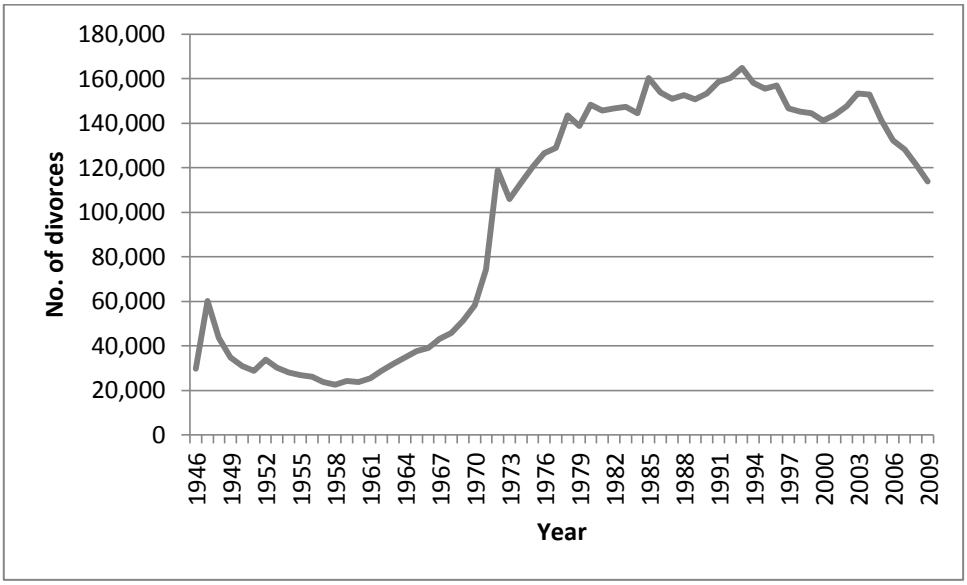
Rebecca Lacey
International Centre for Life Course Studies in Society and Health, Dept. Epidemiology & Public Health, University College London
Rebecca.Lacey@ucl.ac.uk
Mel Bartley
International Centre for Life Course Studies in Society and Health, Dept. Epidemiology & Public Health, University College London
Hynek Pikhart
Dept. Epidemiology & Public Health, University College London
Mai Stafford
MRC Unit for Lifelong Health and Ageing, London
Noriko Cable
International Centre for Life Course Studies in Society and Health, Dept. Epidemiology & Public Health, University College London
Lester Coleman
One Plus One Marriage and Relationship Research, London

Abstract

Parental separation has been linked to increased likelihood of reporting psychological distress in adulthood, but relatively little is known about how this association may have changed over time. One hypothesis is that as the experience of separation has become more common, the association with psychological distress will reduce (the 'reduced effect' hypothesis). Previous evidence using the British birth cohorts does not support this hypothesis, but past studies have several limitations which we aim to address. In this study we measure parental separation from birth, account for missing data and statistically test cohort differences using data from two British birth cohorts – the 1958 National Child Development Study (NCDS) and 1970 British Cohort Study (BCS). Despite these methodological improvements, we find little evidence to support the 'reduced effect' hypothesis: parental separation was associated with a similarly increased likelihood of reporting psychological distress in adulthood for men and women in both cohorts.

UK divorce rates have increased rapidly since the mid-20th century (fig. 1). This trend translates into an increasing probability that children will experience family disruption across this period. A longitudinal link between parental separation and adult psychological distress indicates that parental separation occurring during childhood is associated with increased reporting of psychological distress in adulthood (e.g. Chase-Lansdale, Cherlin & Kiernan, 1995; Kendler, Shether, Garner, & Prescott, 2002; Kuh & Maclean, 1990; Rodgers, 1990; Rodgers & Pryor, 1998). However the experience of parental separation may change

Figure 1 – Trend in the number of divorces, England and Wales, 1946-2009



Source: Data taken from 'Divorces 1858-2003' (Office for National Statistics 2006); 'Marriage, divorce and adoption statistics. FM2 No 35' (ONS 2010)

over time because social institutions and organisations, such as marriage and the family, have links to individuals through social trajectories which change with life experiences and historical time (Elder, 1994). In particular it is thought that the association between parental separation and adult psychological distress will lessen over time (termed the 'reduced effect' hypothesis) through reducing stigma surrounding divorce, increasing availability of services (e.g. family mediation) and greater awareness of the importance of maintaining contact between children and both parents following separation (Ely, Richards, Wadsworth, & Elliott, 1999).

There is some evidence to support this hypothesis, as the association between parental divorce and adult psychological health was found to be greater in studies conducted during the 1950s and 1960s (effect size=-0.31, $p<0.001$), compared to studies from the 1980s (effect size=-0.13, $p<0.001$) (Amato & Keith, 1991). However Ely and colleagues (1999) failed to support this hypothesis, not finding any change in the association between parental separation and children's educational attainment in their comparison of three British birth cohorts (born in 1946, 1958 and 1970). Also Sigle-Rushton, Hobcraft, and Kiernan (2005) reported that the associations between parental divorce and adult psychological distress were similar in the 1958 and 1970 birth cohorts. Despite the lack of evidence for the 'reduced effect' hypothesis in the British birth cohorts, these and other studies have methodological limitations which may have affected the results and conclusions: for example, cohort differences in the association between parental divorce and adult outcomes were not statistically tested in some studies; only separations and divorces from mid-childhood onwards were examined; and cases with missing values were dropped from the analyses. In this research note we aim to assess whether there is evidence for the 'reduced effect' hypothesis once these methodological problems have been addressed.

In addition it has been suggested that girls may be more affected by parental separation than boys, as girls are more likely to take on excessive responsibilities in the home, or emotionally support a parent (Hetherington & Stanley-Hagan, 1999). Parental divorce has been found to be associated with an increase in adult affective disorders for women but not for men in some samples (see e.g. Rodgers, 1990),

but knowledge regarding changes in gender differences in the association between separation and psychological distress over time is limited. We aim to extend current knowledge by including parental separation from birth in the study model; by statistically testing the 'reduced effect' hypothesis; by accounting for missing data; and by testing for possible gender differences in the association.

Method

Data

This study uses data from two of the British birth cohorts - the 1958 National Child Development Study (NCDS) and 1970 British Cohort Study (BCS) – which were also used by Ely et al. (1999) and Sigle-Rushton et al. (2005). The comparison of these two cohorts is particularly appropriate for this study: NCDS cohort members experienced relatively high levels of family stability as they grew up in the 1960s-1970s, sometimes referred to as the 'golden period' of marriage (Festy, 1980), when marriage occurred at young ages and amongst the vast majority of the population. Conversely, BCS cohort members grew up in a time of greater family instability, when societal divorce rates were increasing rapidly and the 'golden period' of marriage was in decline. These two studies are similarly designed, have large sample sizes, and contain directly comparable measures.

The NCDS was initially started as the Perinatal Mortality Study, recruiting 17,415 participants born in one week in 1958 in England, Wales, and Scotland (Shepherd, 1995). The participants' families were retraced in 1965 (age 7 years) and a longitudinal study initiated. Immigrants who were born in the same target week and who were identified by their schools were also added to the sample at ages 7, 11, and 16 years (total sample=18,558). Data has been collected from cohort member and parental interviews, medical examinations, teacher assessments, and school tests at birth, and at ages 7, 11, 16, 23, 33, 42, 46, and 50 years. The BCS was started as the British Births Survey recruiting 16,752 babies born in one week of 1970 (Elliott & Shepherd, 2006). These children were retraced at age 5. Again, immigrants born during the same week were added to the sample in 1975 (age 5) and 1980 (age 10) - total sample=18,732. By 2008, BCS cohort members have been followed up at eight points across their lives (birth, age 5, 10, 16, 26, 30, 34, and 38 years). Our study uses data from the first six sweeps of both studies.

Measures

Parental separation was measured between the ages of 0-16 years. As this study was interested in the experience of parental separation, mother's marital status at birth was used to restrict the samples to those participants whose mothers were married or in a relationship at the time of the cohort member's birth: 17,887 in the NCDS (96.4% of total sample including immigrants) and 17,467 in the BCS (93.2%). Parental separation was derived from items detailing change in parental figures for each inter-sweep period, and reasons for this. If the reason was 'separation' or 'divorce', participants were coded as having experienced parental separation. If a reason of 'separation' or 'divorce' was given but the response to the parental figures question was missing, these people were also coded as having experienced parental separation.

Psychological distress was measured at age 33 (NCDS) and age 30 (BCS) using the Malaise Inventory (Rutter, Tizard, & Whitmore, 1970). This is comprised of 24 items (yes/no responses) which cover emotional disturbance and somatic symptoms. Internal consistency of the scale is acceptable: Cronbach's α is 0.78 in the NCDS and 0.79 in the BCS. The total affirmative answers were summed giving a score out of 24. This was dichotomised using the frequently-used cut-off point of 7 to indicate 'caseness' (Hope, Rodgers, & Power, 1999; Sacker & Cable, 2005).

Data were pooled across cohorts, and gender was taken from the birth sweep. *A priori* confounding variables included in the analysis were mother's education, taken from

the birth sweep of both cohorts (0 = *stayed in education beyond minimum age*, 1 = *left education before or at minimum leaving age*), mother's psychological distress (mother's psychiatric illness at age 11 in the NCDS and mother's Malaise Inventory score at age 10 in the BCS), and mother's age at the birth of the cohort member. Father's social class at the birth of the cohort member, coded according to the Registrar-General's scheme (see below), was included to control for selection effects, to account for the increased likelihood of parents from more disadvantaged backgrounds to separate. As both cohorts were initiated as medical studies, the Registrar-General's scheme is the only marker of social disadvantage available in both cohorts. This social class measure is based upon occupation and has the following categories: I (professional), II (managerial and technical), IIINM (skilled non-manual), IIIM (skilled manual), IV (semi-skilled manual) and V (unskilled).

Statistical methods

In order to go some way towards accounting for missing data, multiple imputation by chained equations (MICE) was conducted. This method is particularly useful in studies where missing values are present in many variables (Carpenter & Plewis, 2011), as in these datasets. The performance of MICE is highly dependent upon the imputation model specified. The imputation model used here contained all analysis variables, variables predictive of missingness, and variables thought to be useful in providing information to fill in the gaps (indicators of health, such as reporting a limiting longstanding illness, and markers of social disadvantage, such as housing tenure and social class). Psychological distress was included in all imputations as its omission may result in bias (White, Royston, & Wood, 2011). Following Von Hippel's (2007) method of multiple imputation then deletion, values for missing psychological distress scores were imputed using MICE, but only those participants with complete data on distress were used in the analysis. The final sample sizes for each cohort were: 10,923 in the NCDS and 10,714 in the BCS, representing those with complete data on psychological distress. Twenty datasets were imputed for each cohort and estimates given in the results section are those obtained by combining estimates across all data sets using Rubin's rules (1987). Table 1 shows the proportion of missingness within each analysis variable and also provides a comparison of observed and imputed data. The proportions for those with observed and imputed data are very similar, suggesting that the imputation has been suitably implemented.

The association between parental separation and adult psychological distress was tested using logistic regression. Likelihood ratio tests are often conducted in post-estimation hypothesis tests following logistic regression, but these are not appropriate for use with multiply-imputed data, as the estimates do not come from a single model and therefore the true likelihood is unknown (Medeiros, 2008). Wald tests are therefore recommended for use with multiply-imputed data (White, et al., 2011), and these were used to test both separation-gender and separation-cohort interactions. Analyses were adjusted for father's social class as a pre-separation risk factor, and for other confounders (mother's education, age and psychological distress).

Table 1. % missingness, and comparison of observed and imputed data in the NCDS and BCS

	NCDS				BCS			
Variable	Missingness %	Observed data %	Imputed data %	Imputed data: range of participants (mean)*	Missingness %	Observed data %	Imputed data %	Imputed data: range of participants (mean)*
Parental separation (0-16 yrs)								
No	28	91.2	90.9	9,893-9,972 (9,926)	52.9	80.9	79.6	8,481-8,575 (8,529)
Yes		8.8	9.1	951-1,030 (997)		19.1	20.4	2,139-2,233 (2,185)
Psychological distress (30/33 yrs)								
No	0	93.2	93.2	10,183**	0	87.4	87.4	9,359**
Yes		6.8	6.8	740		12.7	12.7	1,355
Gender (birth)								
Male	0	49.4	49.4	5,396	8.7	48.6	48.9	5,207-5,267 (5,235)
Female		50.6	50.6	5,527		51.4	51.2	5,447-5,507 (5,479)
Confounders:								
Father's social class (birth)								
I	7.1	4.7	4.8	508-532 (519)	9.2	5.4	5.4	569-596 (580)
II		13.8	13.8	1,487-1,520 (1,506)		12.5	12.5	1,320-1,357 (1,337)
IIINM		10.2	10.2	1,102-1,135 (1,117)		13.1	13.0	1,374-1,418 (1,394)
IIIM		50.8	50.8	5,512-5,562 (5,545)		46.4	46.4	4,938-4,997 (4,969)
IV		12.1	12.0	1,296-1,332 (1,312)		14.1	14.1	1,485-1,517 (1,507)
V		8.4	8.5	900-936 (925)		8.5	8.7	915-949 (927)
Mother's education (birth)								
Stayed beyond minimum age	6.5	26.1	26.1	2,835-2,875 (2,853)	10.0	35.9	36.1	3,854-3,882 (3,868)
Left at or before minimum age		73.9	73.9	8,048-8,088 (8,070)		64.1	63.9	6,832-6,860 (6,846)
Mother's age (birth)								
Mean (range, standard deviation)	5.1	27.7 (16-47 yrs, 5.6)	27.7 (16-48 yrs, 5.6)	27.65-27.72***	9.2	26.2 (16-50 yrs, 5.3)	26.2 (16-50 yrs, 5.3)	26.16-26.21***
Mother's psychological distress (10/11 yrs)								
Not distressed	17.4	98.4	98.4	10,739-10,754 (10,746)	23.6	48.6	48.9	9,012-9,078 (9,049)
Distressed		1.6	1.6	169-184 (177)		51.4	51.2	1,636-1,702 (1,665)

Notes. **The exact number of participants within each category of each variable cannot be given, as MICE results in the creation of 20 imputed data sets, within which, the number of participants vary in each variable (with the exception of psychological distress). The table thus shows the range of participants for each variable along with an average number of participants for each category of every variable. **There is no variation in the number of participants for psychological distress, as only those with complete malaise scores were used in the analysis. ***Range of means of mother's age across imputed data sets.*

Results

The proportion of participants who experienced parental separation approximately doubled across the two cohorts; in the NCDS 9.1% of participants experienced separation, but 20.4% did so in the BCS (Table 1). We tested the association between cohort and parental separation using logistic regression, and found an increased likelihood of experiencing separation in the BCS: taking the NCDS as the reference category, those in the BCS were approximately 2.6 times more likely to experience separation (OR=2.58, 95% CI: 2.30, 2.88, $p<0.001$).

Table 1 also shows that reporting of psychological distress was higher in the BCS compared to the NCDS. The majority of NCDS participants were from manual social classes; in the BCS there were relatively higher proportions of children born into non-manual social classes, though the majority were still born into manual social classes, particularly class IIIM. The proportion of mothers who stayed in education beyond the minimum school leaving age was higher in the BCS than the NCDS. There was a slight decrease in mother's age between the two cohorts. This study does not look at birth order and therefore this apparent decrease does not necessarily translate into a reducing age of women at childbearing. In the NCDS just 1.6% of mothers reported having been given a diagnosis of psychiatric illness. In the BCS the Rutter Malaise Inventory was completed by mothers when the children were aged 10 years, and the reported levels of psychological distress were relatively high (15.5%). It should be noted that this figure would be expected to be higher than in the NCDS, as the Malaise Inventory does not just capture diagnosed psychiatric illness but also lower levels of psychological distress.

Table 2 shows the results of the adjusted logistic regression analysis. The interactions between parental separation and gender were not statistically significant (see results at bottom of Table 2), suggesting that the association between parental separation and psychological distress did not differ between men and women in either cohort. Consequently the rest of the analyses were conducted for both genders together. The results showed that despite the methodological issues we addressed (missing data, investigating parental separation between 0-16 years and statistically testing whether the association has changed over time), there was still no clear evidence that the association between parental separation and adult psychological distress had reduced over time (p=0.69 for the separation-cohort interaction). Therefore our results suggest that in both cohorts, experiencing parental separation during childhood is associated with increased odds of reporting psychological distress in adulthood, after accounting for mother's education, age and psychological distress, and father's social class.

Table 2. Results of logistic regression (OR 95% CIs) testing associations between parental separation and adult psychological distress, by cohort, adjusted for confounders (mother's education, age, psychological distress and father's social class)

Variable	**1958 NCDS**		**1970 BCS**	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Parental separation (0-16 yrs)				
No	Ref		Ref	
Yes	1.75 (1.35, 2.26)	<0.001	1.52 (1.26, 1.83)	<0.001
Confounders:				
Mother's education (birth)				
Stayed after min. age	Ref		Ref	
Left at/before min. age	1.56 (1.25, 1.94)	<0.001	1.16 (1.01, 1.33)	0.037
Mother's age (birth)				
Per 1 yr increase	1.01 (1.00, 1.02)	0.113	1.00 (0.98, 1.01)	0.524
Mother's psychological distress (10/11 yrs)				
No	Ref		Ref	
Yes	1.19 (0.70, 2.04)	0.524	1.63 (1.39, 1.91)	<0.001
Father's social class (birth)				
I	0.47 (0.27, 0.82)	<0.001*	0.68 (0.48, 0.95)	<0.001*
II	0.58 (0.42, 0.78)		0.71 (0.57, 0.89)	
IIINM	0.79 (0.59, 1.05)		0.81 (0.66, 0.99)	
IIIM	Ref		Ref	
IV	1.15 (0.91, 1.44)		1.08 (0.91, 1.28)	
V	1.51 (1.18, 1.92)		1.13 (0.92, 1.39)	
Total	***N*=10,923**		***N*=10,714**	
Interaction between gender and separation	0.71 (0.44, 1.14)	**0.13**	0.92 (0.68, 1.24)	**0.51**
Interaction between cohort and separation	OR=0.89 (0.64, 1.25)			**0.69**

Note. **P value for trend given*

Discussion

In this study we found that parental separation, occurring during childhood, was associated with increased odds of reporting psychological distress in adulthood in two British birth cohorts, supporting results from many previous studies which suggest that parental separation is associated with depression and anxiety in adult life (e.g. Amato, 1991; Kendler et al., 2002). We found this association despite making some methodological improvements on existing work. It has been suggested that parental separation is associated with the increased reporting of psychological distress in adulthood through pathways of increased disadvantage relative to those from 'intact' families (Wadsworth, Maclean, Kuh, & Rodgers, 1990), including reduced educational attainment (e.g. Ross & Mirowsky, 1999), poorer parent-child relationships (Amato & Sobolewski, 2001), and the inter-generational transmission of divorce (Dronkers & Härkönen, 2008).

The results of this study do not, however, support the 'reduced effect' hypothesis, and thus confirm earlier findings based upon complete case analyses in these cohorts (Ely et al., 1999; Sigle-Rushton et al., 2005). The findings suggest that the association of parental separation with the psychological health of offspring has not reduced, but that as societal divorce rates have increased substantially over this period, more people have been affected. It is conceivable that a 'reduced effect' is counteracted by social changes such as increasing relative financial hardship and changing attitudes towards divorce, that act in opposite directions (Ely et al., 1999).

With regards to gender it was thought that the association between parental separation and adult psychological distress would be greater for women than men, as the inconsistent past evidence on this issue appeared to lean in this direction (Hetherington & Stanley-Hagan, 1999; Pirkola et al., 2005). In this study, however, there was no evidence of differential associations for men and women, supporting previous findings of a lack of gender differences (Amato, 1991; Amato & Booth, 1991; Amato & Keith, 1991; Amato & Sobolewski, 2001; Kendler et al., 2002; Kessler, Davis, & Kendler, 1997; Rodgers, Power, & Hope, 1997; Wallerstein, 1991; Zill, Morrison & Coiro, 1993). It has been shown that girls are more likely to emotionally support a parent, but that boys are more likely to witness conflict (Hetherington & Stanley-Hagan, 1999), and it is possible that pathways linking parental separation to adult psychological distress are different for men and women.

Despite the methodological improvements we made on previous work, the findings of this study need to be interpreted in light of a number of limitations. First, it is possible that participants in the earlier cohort were more likely to experience parental remarriage or re-partnership following separation, due to the increased stigma of single parenthood at the time, which may have buffered the effect of parental separation upon the child's psychological health to some extent. Second, it has been shown in some studies that it is the conflict surrounding relationship breakdown that is responsible for poorer psychological health outcomes in children (Amato, Loomis, & Booth, 1995; Cummings & Davies, 2002). Unfortunately in these studies there are insufficient measures of partnership quality and family conflict to investigate this. In addition it is possible that there may have been a substantial lag time between the sharp increase in divorce rates in the early 1970s and any reducing effect upon children, which the 1970 BCS data was not able to capture. It is therefore possible that if a later cohort were studied, a reduced association might be seen.

There are several major strengths of this study. The first is that information on both parental separation and confounders was collected prospectively, minimising the risk of recall bias. The study used data from two large-

scale longitudinal birth cohorts designed to be representative of the British population of similar ages, which are particularly appropriate for testing these hypotheses. The third advantage is that missing values were not ignored, as they are in many studies, but rather dealt with by multiple imputation methods appropriate under the assumption of missing at random (MAR). It is possible that this assumption does not entirely hold, but it is likely to be more plausible than the assumption that data are missing completely at random (MCAR) – the assumption underlying complete case analyses. Imputation adds power to the analyses while reducing the risk of bias due to selective attrition.

The findings of this study add support to the evidence base addressing the relationship between exposure to parental separation in childhood and psychological distress in adult life. The finding that the association between parental separation and psychological distress has not reduced over time, despite parental separation becoming more common, highlights a real need to support families undergoing problems in order to reduce the impact on children's long-term psychological health.

Acknowledgements

This study was supported by the Economic and Social Research Council grant RES-596-28-001 (ESRC International Centre for Life Course Studies in Society and Health). Thanks to all participants of the National Child Development Study and British Birth Cohort and to the study teams. Thanks to the UK Data Archive for providing access to the data.

References

Amato, P. (1991). Parental absence during childhood and depression in later life. *The Sociological Quarterly, 32*, 543-556.

Amato, P., & Booth, A. (1991). Consequences of parental divorce and marital unhappiness for adult well-being. *Social Forces, 69*, 895-914.

Amato, P., & Keith, B. (1991). Parental divorce and adult well-being: a meta-analysis. *Journal of Marriage and the Family, 53*, 43-58.

Amato, P., Loomis, L., & Booth, A. (1995). Parental divorce, marital conflict and offspring well-being during early adulthood. *Social Forces, 73*, 895-915.

Amato, P., & Sobolewski, J. (2001). The effects of divorce and marital discord on adult children's psychological well-being. *American Sociological Review, 66*, 900-921.

Carpenter, J., & Plewis, I. (2011). Analysing longitudinal studies with non-response: issues and statistical methods. In M. Williams & P. Vogt (Eds.), *The SAGE handbook of innovation in social research methods* (pp. 498-523). London: SAGE publications.

Chase-Lansdale, P., Cherlin, A., & Kiernan, K. (1995). The long-term effects of parental divorce on the mental health of young adults: a developmental perspective. *Child Development, 66*, 1614-1634.

Cummings, E., & Davies, P. (2002). Effects of marital conflict on children: recent advances and emerging themes in process-oriented research. *Journal of Child Psychology and Psychiatry, 43*, 31-63.

Dronkers, J., & Härkönen, J. (2008). The intergenerational transmission of divorce in cross-national perspective: results from the fertility and family surveys. *Population Studies, 62*, 273-288.

Elder, G. (1994). Time, human agency, and social change. *Social Psychology Quarterly, 57*, 4-15.

Elliott, B., & Shepherd, P. (2006). Cohort profile: 1970 British birth cohort. *International Journal of Epidemiology, 35*, 836-843.

Ely, M., Richards, M., Wadsworth, M., & Elliott, B. (1999). Secular changes in the association of parental divorce and children's educational attainment – evidence from three British birth cohorts. *Journal of Social Policy, 28*, 437-455.

Festy, P. (1980). On the context of marriage in Western Europe. *Population and Development Review, 6*, 311-315.

Hetherington, E., & Stanley-Hagan, M. (1999). The adjustment of children with divorced parents: a risk and resiliency perspective. *Journal of Child Psychology and Psychiatry, 40*, 129-140.

Hope, S., Rodgers, B. & Power, C. (1999). Does financial hardship account for elevated psychological distress in lone mothers? *Social Science and Medicine, 49*, 1637-1649.

Kendler, K., Shether, K., Garner, C., & Prescott, C. (2002). Childhood parental loss and risk for first-onset of major depression and alcohol dependence: the time-decay of risk and sex differences. *Psychological Medicine, 32*, 1187-1194.
Kessler, R., Davis, C., & Kendler, K. (1997). Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey. *Psychological Medicine, 27*, 1101-1119.
Kuh, D., & Maclean, M. (1990). Women's childhood experience of parental separation and their subsequent health and socioeconomic status in adulthood. *Journal of Biosocial Science, 22*, 121-135.
Medeiros, R. (2008). Likelihood ratio tests for multiply imputed datasets. In *2008 Fall North America STATA users group meeting.*
Office for National Statistics (2006). Divorces: 1858-2003, number of couples divorcing, by party petitioning/granted decree.
Office for National Statistics (2010). Marriage, divorce and adoption statistics, England and Wales (Series FM2), No. 35.
Pirkola, S., Isometsä, E., Aro, H., Kestilä, L., Hämäläinen, J., Veijola, J., Kiviruusu, O., & Lönnqvist, J. (2005). Childhood adversities as risk factors for adult mental disorders. *Social Psychiatry and Psychiatric Epidemiology, 40*, 769-777.
Rodgers, B. (1990). Adult affective disorder and early environment. *British Journal of Psychiatry, 157*, 539-550.
Rodgers, B., Power, C,. & Hope, S. (1997). Parental divorce and adult psychological distress: evidence from a national birth cohort: a research note. *Journal of Child Psychology and Psychiatry, 38*, 867-872.
Rodgers, B. & Pryor, J. (1998) *Divorce and separation: the outcomes for children.* York: Joseph Rowntree Foundation.
Ross, C., & Mirowsky, J. (1999). Parental divorce, life-course disruption and adult depression. *Journal of Marriage and the Family, 61*, 1034-1045.
Rubin, D. (1987). *Multiple imputation for non response in surveys.* New York: J. Wiley & Sons.
Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. London: Longmans.
Sacker, A., & Cable, N. (2005). Do adolescent leisure-time physical activities foster health and well-being in adulthood? Evidence from two British birth cohorts. *European Journal of Public Health, 16*, 331-335.
Shepherd, P. (1995). *The National Child Development Study - an introduction to the origins of the study and the methods of data collection*. London: Social Statistics Unit, City University. Published May 1995 [accessed 23 March 2009]. Retrieved from: http://www.cls.ioe.ac.uk/studies.asp?section=0001000200030004
Sigle-Rushton, W., Hobcraft, J., & Kiernan, K. (2005). Parental divorce and subsequent disadvantage: a cross-cohort comparison. *Demography, 42*, 427-446.
Von Hippel, P. (2007). Regression with missing Ys: an improved strategy for analysing multiply imputed data. *Sociological Methodology, 37*, 83-117.
Wadsworth, M., Maclean, M., Kuh, D., & Rodgers, B (1990). Children of divorced and separated parents: summary and review of findings from a long-term follow-up study in the UK. *Family Practice, 7*, 104-109.
Wallerstein, J. (1991). The long-term effects of divorce on children: a review. *Journal of American Academy of Child and Adolescent Psychiatry, 30*, 349-360.
White, I., Royston, P., & Wood, A. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine, 30*, 377-399.
Zill, N., Morrison, D., & Coiro, M. (1993). Long-term effects of parental divorce on parent-child relationships, adjustment and achievement in young adulthood. *Journal of Family Psychology, 7*, 91-103.

Longitudinal and Life Course Studies 2012 Volume 3 Issue 3 Pp 369 – 381 *ISSN 1757-9597*

The HILDA Survey: a case study in the design and development of a successful household panel study

Nicole Watson
Melbourne Institute of Applied Economic and Social Research, University of Melbourne
Mark Wooden
Melbourne Institute of Applied Economic and Social Research, University of Melbourne
m.wooden@unimelb.edu.au

(Received May 2012 Revised September 2012)

Abstract

The Household, Income and Labour Dynamics in Australia (HILDA) Survey is one of only a small number of well-established, large, nationally-representative household panel studies conducted in the world. With annual data collection commencing in 2001 there are now over 10 years of unit record data available to researchers, with the promise of many more to come. While the design of the HILDA Survey owes much to other older household panel studies conducted elsewhere, it has a number of features which make it relatively unique. This paper provides a brief history of the HILDA Survey's progress to date, focusing first on its origins, design, content and data collection processes, before reviewing its achievements with respect to survey response and usage.

Keywords: Australia, HILDA Survey, household panel surveys, longitudinal survey research

Introduction

With over a decade of data collection behind it, the Household, Income and Labour Dynamics in Australia (or HILDA) Survey is one of only a small number of large, nationally-representative household panel studies that can claim to have made the transition into a successful long-running study. An initiative of the Australian Government, the contract to design and manage the study was awarded to the Melbourne Institute of Applied Economic and Social Research (at the University of Melbourne), with the first wave of data collection taking place in 2001. The contract has been renewed or extended on a number of occasions, including most recently in 2012. Currently the HILDA Survey is assured of continuing until at least 2016, with all indications that the study will continue to be funded well beyond that.

This paper provides a brief history of the HILDA Survey's progress to date, focusing first on its origins, design, content and data collection processes, before reviewing its achievements with respect to survey response and usage.

Origins

Australia, like other industrial nations, has long invested heavily in the collection of statistical data. Indeed, Australia arguably has one of the most effective national statistical agencies in the world, the Australian Bureau of Statistics (ABS). Nevertheless, there has always existed one glaring gap in the ABS collection – the absence of any long-running longitudinal survey data collection. Longitudinal survey data collections in Australia have thus mostly been funded and managed by other agencies, and have been designed with quite specific purposes in mind, reflected in studies that focus on relatively small sub-groups of the population (such as youth or recent immigrants) and / or on quite narrowly-defined subject matter (e.g., post-migration settlement or women's health). In the late 1990s, with the greater

emphasis in government in Australia on the need for evidence-based policy, the significance of this data gap became much more obvious. Nowhere was this perhaps noticed more than in the Department of Families and Community Services, which was responsible for delivery of income support and social services to Australian families, and at the time was coordinating a major review of social welfare services and policy.

It was thus against this background that the Australian Government, and more specifically the Department of Families and Community Services, commissioned (following a public tender process) the Melbourne Institute to design and implement the HILDA Survey. As stated in the Request for Tender, the primary objectives were to support research within three broad inter-related areas: family and household dynamics; income and welfare dynamics; and labour market dynamics. At the same time, the winning tender emphasized the advantages of appealing to as broad a cross-section of users within the social sciences as possible, and hence the expansion of the coverage of the survey to topic areas such as subjective well-being. It was also explicitly designed to inform policy development, especially in the areas of economic and social participation, and stronger families. However, it was also recognised that such a survey would not be well suited to the evaluation of a particular set of policy initiatives. Rather the hope was that the HILDA Survey would collect information on a range of policy-relevant variables that would provide baseline information prior to policy changes and assist in the evaluation of future, and unknown, policy changes.

Design

The design of the HILDA Survey owes much to other household panel studies conducted elsewhere in the world, and especially the German Socio-Economic Panel (GSOEP) and the British Household Panel Survey (BHPS) (Haisken-DeNew, 2001; Frick, Jenkins, Lillard, Lipps & Wooden, 2007). Indeed, the winning tender was built on the assumption that the most cost-effective approach, especially given tight timelines, a limited budget and a risk-averse client, was to mimic practices adopted by existing long-running, and thus successful, household panel studies. Thus, like these studies, the HILDA Survey:

i. commenced with a population sample that was intended to be broadly representative of the national population resident in private households;
ii. conducts survey waves on an annual basis;
iii. follows members of the original sample households and their offspring for an indefinite life;
iv. interviews all 'adult' members of the household; and
v. augments the sample in later waves with persons who join a sample household.

It is also common for these studies to introduce new representative samples over time. The GSOEP, for example, has added new representative samples on three occasions since starting in 1984. In this vein the HILDA Survey now has two population samples – the original sample, which commenced in 2001, and a top-up sample that was introduced in 2011.

Reflecting the parallels with its predecessors, in 2007 the HILDA Survey became the 5th member of the Cross-National Equivalence File, which provides harmonised data from comparable nationally representative household panel surveys (Frick et al., 2007). Of course, no two studies are identical. There are many important differences in design across studies, including, for example, the extent of reliance on different survey modes, the use of incentives, and the rules adopted for determining when to re-issue non-responding cases (Watson & Wooden, 2011).

Reference population

The reference population for the initial sample was, with only minor exceptions, all persons residing in private dwellings in Australia in 2001. The exceptions were diplomatic personnel of overseas governments, overseas residents (who intend staying in Australia for less than one year), members of non-Australian defence forces (and their dependents) stationed in Australia, and people living in the most remote and sparsely populated parts of Australia. An essentially identical reference population was used for the wave 11 top-up, but with the obvious difference that all persons had to be residing in private dwellings in Australia in 2011.

The sampling unit

The sampling unit is the household, defined as 'a group of people who usually reside and eat together'. Persons who usually lived in the household but were temporarily absent for work, school or other purposes were treated as part of the household, meaning that a small proportion of interviews had to be conducted at locations other than the household address. Persons who live in more than one household were only treated as members of the household where they spent most of their time. Visitors to the household were not treated as part of the household.

Sample selection

The initial sample of households was selected, in 2001, using a multi-stage approach. First, a sample of 488 Census Collection Districts (CDs) was selected from across Australia (each of which consists of approximately 200 to 250 households). To ensure the sample provided adequate coverage of all parts of Australia, the frame of CDs was stratified by State, and within the five most populous States, by metropolitan and non-metropolitan regions. The CDs were sampled with a probability proportional to their size, as measured by the number of dwellings recorded in each CD at the 1996 Census. Second, within each of these CDs, a sample of 22 to 34 dwellings was selected, the precise number depending on the expected response and occupancy rates of the area. The selections were made after all dwellings within each of the CDs were fully enumerated. Finally, within each dwelling, up to three households were selected to be part of the sample. This process resulted in a total of 12,252 addresses being issued to field, though 804 of these addresses were subsequently discovered during the fieldwork to be out of scope (because the dwelling was vacant, not a primary private residence, or all of the occupants did not meet the selection criteria).

The process for selecting the top-up sample for wave 11 was very similar, but with 125 CDs selected (and based on 2006 Census boundaries). A total of 3250 addresses were issued to field. After the removal of addresses determined to be out of scope and adjusting for addresses with multiple households, the initial sample numbered 3117 households.

Following rules

All members of households where at least one person provided an interview in the initial wave form the basis of the panel to be pursued in each subsequent wave. In addition, the sample is gradually extended to include any new household members resulting from changes in the composition of the original households. All children born to, or adopted by, an original sample member become permanent or continuing sample members. Most other persons only remain in the sample for as long as they remain living in a household with a continuing sample member. There are, however, two exceptions to this practice. First, any person who has a child with a continuing sample member is converted to continuing sample member status. Second, from wave 9, new household members who arrived in Australia for the first time after 2001, were also added to the sample on a continuing basis.

Note that while original sample members had to be residents of private households, in subsequent survey waves, sample members are followed to wherever they move, including into institutions. Interviews, however, are not conducted with persons found to be in prison or living overseas.

Survey content

Like other major household panel surveys, the data collected during the HILDA Survey are almost entirely self-reported, and the coverage is both extremely broad and includes topics that are included in every survey wave, and others that appear less frequently. Below, we provide brief summaries of the survey instruments that are used to collect the data. For further information, including paper versions of all the instruments used in every survey wave, readers should consult the HILDA Survey web site at:
http://melbourneinstitute.com/hilda/doc/questionnaires

Annual content: interview components

The main survey instruments were designed to be administered in a personal interview, usually conducted at the home of the respondent. Further, this interview content involved both an individual component, administered to all household members aged 15 years and over, and a shorter

household component administered to just one person in the household.

The household component comprises two separate instruments – the Household Form (HF) and the Household Questionnaire (HQ). The HF is designed to record and verify basic information about the composition of the household immediately after making contact. The HQ collects information about the household rather than about individual household members, and is only administered to one member of the household, though interviewers are given the flexibility to deliver part of this interview to one household member and part to another. In each wave, the HQ contains questions about childcare arrangements, housing, and housing wealth. A very small number of questions about household expenditure (on groceries, food and drink, and meals out) are also included.

The individual interview component involves the administration of either the Continuing Person Questionnaire (CPQ) or the New Person Questionnaire (NPQ). The CPQ is administered to every individual who has responded in a previous wave. The 'core' topics included every year are:

- Education – collects details about any study undertaken since the last interview.
- Employment status – determines employment status using the labour force framework recommended by the International Labour Organization.
- Current employment – collects details about any current employment including, for example, usual hours worked, work schedule, occupation, industry of employer, contractual status, and workplace and firm size. Subjective data about job satisfaction and perceived probability of changing jobs are also collected. For persons that leave a job between interview waves, data are also collected on the reason for ceasing that job.
- Persons not in paid employment – administered to persons not currently employed, this section collects data on recent job search activity, reasons for not looking for work, retirement, and details of any previous job held since the last interview. It is also used to determine unemployment status.
- Other labour market activity – the centerpiece of this section is a calendar which records labour market status and educational attendance for a period extending from 1 July in the year prior to the survey wave, up until the date of interview. Additional information is also collected about government-imposed jobseeker requirements, work-related training (since wave 3) and days of leave from paid work (since wave 5).
- Income – collects detailed information about income by source for the preceding financial year (year ended 30 June) as well as current income information for wages and salaries and government pensions and benefits.
- Family formation – records details about any children. Special emphasis is given to the role of absent parents and the amount of contact with, and support given to, non-resident children.
- Partnering and relationships – records information about changes in marital status, co-residential relationships and changes in those relationships, and marriage expectations.
- Living in Australia – includes short sequences of questions on topics that do not fit easily anywhere else in the survey instruments. Included here are disability, life satisfaction, residential mobility, and since wave 5, caring responsibilities.
- Tracking information – records contact details that could be used to help locate the respondent in the future.
- Interview situation – records interviewer observations about the interview (e.g., the presence of adults during the interview, how cooperative the respondent was, and whether the respondent needed any assistance to complete the interview).

The NPQ has the same format but in addition collects information about respondents' background. Included here are country of birth and language, visa status and migration category (if born overseas), educational attainment, employment history, marital history, family background and parents' characteristics.

Rotating content: interview components

Each wave of the HILDA Survey now includes at least one major topic, typically taking around 10 minutes of interview time per person, which is repeated every four years. These topics are as follows:

- Household wealth – conducted in waves 2, 6 and 10, this module seeks to provide estimates of total household wealth disaggregated by type of asset and liability (Headey, Marks & Wooden, 2005).
- Family formation and fertility – conducted in waves 5, 8 and 11 (but thereafter to be moved on to a 4-year cycle), in this module, additional information about family formation is collected that is not collected on an annual basis. This includes questions on issues such as: recent pregnancies and whether they were intended; use and method of contraception; factors influencing fertility decisions; and return to work following childbirth.
- Retirement from the workforce – conducted in waves 3, 7 and 11, this module is restricted to persons aged 45 years or older and focuses on retirement intentions and decisions, the transition into retirement, expectations about life post-retirement, and the retirement experience.
- Health – introduced for the first time in wave 9, this module includes questions on: expectations about health; difficulties caused by health conditions and disabilities; serious illness conditions; retrospective childhood health; private health insurance; use of healthcare services; diet; and the health status of, and use of health care services by, children in the household.
- Education, skills and abilities – while detailed information about educational attainment is collected every year, relatively little information is collected about cognitive skills and abilities. This module, included for the first time in wave 12, will redress this deficiency. It includes questions on English language and mathematical abilities as well as three short tests of cognitive ability. Also collected is information about the schooling experiences of children in the household, as well as additional information about the education experience of respondents (e.g., field and place of study).

In addition, there are a small number of shorter question sequences included on a rotating basis. Included here are: job-related discrimination (included in waves 8 and 10, but next scheduled to be included in wave 14); intentions and plans regarding mobility, education and work (included in waves 5, 8 and 11 and thereafter to be moved on to a 4-year cycle); and non-coresidential family relationships (included in waves 8 and 12).

Self-completion questionnaire

A feature of the HILDA Survey is that, like the BHPS and Understanding Society panels in the UK, all persons successfully interviewed are also asked to complete a pen-and-paper questionnaire. The instrument in the HILDA Survey, however, is much longer than that administered in the UK studies – it can take anywhere from 20 minutes to 40 minutes to complete – and hence cannot always be completed while the interviewer is in the household.

This self-completion questionnaire (SCQ) consists mainly of questions which are difficult to administer in a time-effective manner in a personal interview, or which respondents may feel slightly uncomfortable answering in a face-to-face interview. The types of topics covered each year include: health status (the SF36 health survey); lifestyle behaviours and outcomes, such as smoking, exercise, alcohol consumption and height and weight; relationship satisfaction; social interaction and support; time use; life events; financial stress; and work-family balance. Other topics appear on a less frequent basis, examples of which include: psychological distress (Kessler 10); religion; neighbourhood characteristics; participation in community activities; and personality traits. In wave 5 the length of the SCQ was expanded from 16 to 20 pages, enabling the inclusion of an inventory of items measuring household expenditure.

Data collection

Survey mode and interview length

The principal mode of data collection is face-to-face interviews, usually conducted at the home of the respondent. From wave 1 to wave 8 these were conducted using pen and paper methods. In wave 9 pen and paper was replaced by computer tablet consoles (see Watson 2010 for discussion of the impact of this mode change). Telephone interviews are conducted both as a last resort and to reach sample members who move to locations not covered by the network of face-to-face interviewers. The proportion of interviews

conducted by telephone in wave 1 was negligible, but by wave 10 was around 8%.

Interview length varies from wave to wave, but in general, the aim is to ensure that the average time spent by interviewers in a two-adult household does not exceed 83 minutes.

As already mentioned, all interviewees are also given a self-administered questionnaire to complete. This is either handed directly to the respondent by the interviewer at time of interview or, in the event of telephone interviewees, mailed out after the interview is completed. In the case of personal interviews, if respondents are unable to complete the questionnaire while the interviewer is present, then interviewers are required to make at least one more trip to the household to collect the completed forms. In instances where forms are still not complete, respondents are instructed to return the forms by post in reply-paid envelopes.

Fieldwork

Interviews are conducted on an annual basis, with the fieldwork clustered into three phases spread over the period between August and February the following year.

The data collection task is sub-contracted out to a private research organisation that specialises in survey administration. This organisation is responsible for the recruitment, training and deployment of all interviewers used on the study. In 2009, a change in sub-contractor was required, creating the potential for significant disruption to the continuity of the study. All available evidence, however, suggests this change had minimal, if any, impact on the data collection process (Watson 2010). A likely major reason for this was that a large fraction (72%) of the face-to-face interviewers that worked on the project in wave 8 were subsequently engaged by the new sub-contractor (Roy Morgan Research) to work on the project in wave 9.

Incentives

To encourage response, all households in waves 1 to 4 were paid either $20 or $50 each year they participated, with the higher amount only paid when interviews were completed with all in-scope household members. Payment was made by cheque mailed to households after interviewing was completed. For waves 5 to 8, the incentive was changed to $25 per completed personal interview, with a $25 bonus paid to households where all in-scope household members completed the personal interview. In wave 9, the payment was increased to $30 and, more importantly, where interviews were conducted in person, the payment was made in cash at the time of interview. In wave 13 this amount will rise to $35.

Initial wave response

After adjusting for out-of-scope dwellings and households, and multiple households within dwellings, the number of households identified as in-scope in wave 1 was 11,693. Interviews were completed with all eligible members at 6,872 of these households and with at least one eligible member at a further 810 households. Within the 7,682 households at which interviews were conducted, 13,969 persons were successfully interviewed (out of a total of 15,127 eligible household members). This provides a household response rate of 66% (and an effective individual response rate of 61%).

While non-response was considerable, the characteristics of the initial sample appear to match the broader population quite well. The main exceptions to this are an under-representation of immigrants from a non-English-speaking background, and of residents of Australia's largest city, Sydney (Wooden, Freidin & Watson, 2002).

For the top-up sample introduced in wave 11, interviews were obtained at 2,153 households out of a total of 3,117 selected households identified as in-scope, giving a household response rate of 69.1%. Within these participating households there were 4,280 persons eligible for interview, 93.7% of whom (4,009) were successfully interviewed.

Response rates for the wave 11 top-up sample were thus noticeably higher than obtained in wave 1. Given the accumulated experience of both the survey managers and interviewers, this perhaps should not be surprising. Nevertheless, it contrasts with the international experience from both repeated cross-sections (e.g., de Leeuw & de Heer, 2002) and other household panel studies. The new UK household panel study, Understanding Society, which succeeds and subsumes the BHPS, for example, only obtained a household response rate of 57.2% for its first wave, conducted during 2009 and 2010 (Buck and McFall 2012). By comparison, the equivalent response rate for wave 1 of the BHPS, conducted in 1991, was 74%. A similarly

marked decline in initial wave response is also evident in the successive refreshment samples that have been added to the GSOEP. For the 1998, 2000, and 2006 refreshments these were 54%, 52% and 41% respectively. By comparison, for the original West German sample, and despite the exclusion of all partially responding households, a noticeably higher response rate of 61% was reported for its first survey wave in 1984 (see Frick et al., 2007, Table 1).

Sample attrition and growth

All longitudinal surveys have to confront the problem that with each successive survey wave, some sample members are lost, either because of a failure to locate sample members who have moved, or because sample members withdraw their cooperation. For example, after 10 waves there were 20,287 persons eligible for interview in wave 11, of whom 768 are no longer issued to field because after numerous attempts they could not be located, and a further 3,829 people with whom contact is no longer attempted given previous adamant refusals or long-term irrecoverable illness. With this type of design, however, it is not inevitable that overall sample size will decline. Indeed, since wave 4, sample growth due to changing household composition has more than offset the loss coming from attrition.

A numerical summary of the evolving sample over the first 11 survey waves (but excluding the top-up sample introduced in wave 11) is provided in Table 1. The first row of this table shows that of the 13,969 persons originally interviewed (in wave 1), 8780 were re-interviewed in wave 11. This represents a re-interview rate of 62.9%, rising to 68.8% after excluding deaths and persons known to have moved overseas on a long-term basis. The other rows in the table report comparable figures for persons joining the responding sample at later waves. Note that the gaps between the raw re-interview rates and the adjusted rates reported in the table can be quite large for these new entrants, which reflects the fact that many of these persons only remain in the sample for as long as they co-reside with a continuing sample member.

A better guide to trends in sample attrition is provided by the annual re-interview rate of previous wave respondents. This is graphed in Figure 1, and shows the annual re-interview rate rising from 87% in wave 2 to over 96% by wave 9, and remaining at that level since. While these response rates imply levels of attrition that are non-trivial, they nevertheless compare favourably with the rates achieved over the first 11 waves in other long-running panels, such as the BHPS and the SOEP (Watson and Wooden 2011).

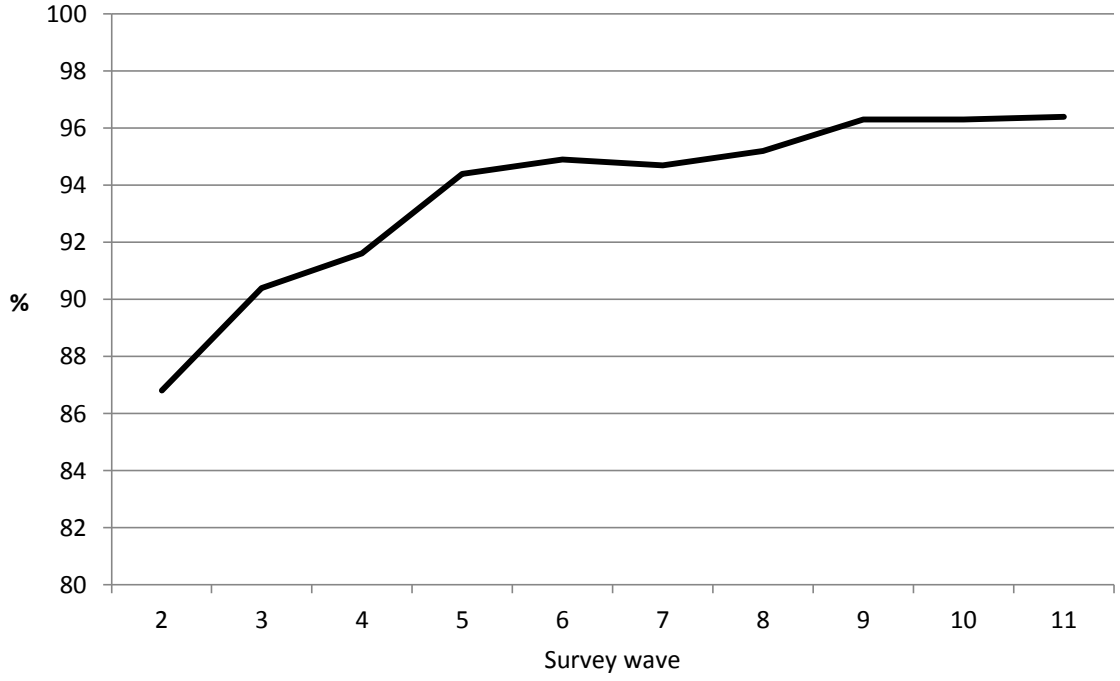
Attrition, however, still has the potential to adversely affect the representativeness of the sample. To help address this problem, considerable effort is put into the construction of population weights that adjust for selective attrition (see below). Despite this, some biases are likely to remain in the sample. Preliminary (and unpublished) research, based on comparisons with Labour Force Survey data from the ABS, suggests that more highly educated, higher income earners are over-represented in the HILDA wave 10 sample. Such differences, however, are both small and uncertain (given differences in the way these variables are defined and measured in the HILDA Survey and by the ABS).

Table 1. Responding sample size and re-interview rates by survey wave and cohort

	Survey wave										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
2001	13969	11993	11190	10565	10392	10085	9628	9354	9245	9002	8780
%	-	86.8	82.0	78.5	77.9	76.3	73.6	72.0	71.5	70.0	68.8
2002		1048	705	594	572	542	512	483	488	475	461
%		-	83.5	76.2	77.4	75.9	72.9	70.7	72.2	71.0	69.3
2003			833	543	482	429	403	376	383	365	354
%			-	84.8	84.6	81.4	80.0	76.7	78.8	75.9	74.1
2004				706	494	426	408	369	374	362	348
%				-	91.1	86.6	85.9	80.7	82.6	81.7	78.9
2005					819	578	511	462	459	441	407
%					-	89.3	86.3	83.1	86.4	84.3	80.8
2006						845	641	545	525	499	468
%						-	93.7	89.8	89.7	88.6	85.2
2007							686	509	448	427	409
%							-	95.1	91.8	89.9	88.1
2008								687	526	491	444
%								-	94.4	94.2	90.8
2009									853	640	583
%									-	96.7	94.0
2010										824	599
%										-	94.6
2011											749
%											-
Total	13969	13041	12728	12408	12759	12905	12789	12785	13301	13526	13602

***Note**: Re-interview rates are calculated after excluding persons whose status has changed to out-of-scope (due to death, relocation overseas or, for temporary sample members, are no longer living with a continuing sample member).*

Figure 1. Annual re-interview rate (% of all in-scope previous wave respondents)



Panel maintenance

Like other panel studies, the HILDA Survey devotes considerable effort to both tracking sample members over time and encouraging participation.

Contact rates in the HILDA Survey are very high, which reflects a relatively low rate of cases that cannot be found. As already noted, of the total sample eligible for interview in wave 11, just 3.8% are no longer being followed because they cannot be found. The tracking procedures employed include: (i) extensive pre-field office activity, generated by notifications of changes of contact details by sample members, returns to senders from the distribution of annual primary approach letters and mid-wave thank you gifts, and information collected at the previous wave about the likelihood of moving; (ii) seeking information about the whereabouts of the absent sample member during fieldwork from other household members, new residents, neighbours and community resources; (iii) consulting online telephone directories; and (iv) using the extensive contact information collected at the last interview (and especially the contact details of two persons not living with the sample member).

In terms of maximising response once respondents are located, the most significant strategies are:

(i) Use of cash incentives and other small gifts (e.g., a thank you gift sent to respondents around Christmas).

(ii) Maintaining a highly-trained, experienced, motivated and engaged interviewer workforce, who are matched, as far as practicably possible, to responding households over time.

(iii) A three-phase fieldwork structure, which emphasises multiple call attempts and refusal conversion, and includes the use of tailored follow-up letters to sample members who are non-respondents in phase 1, and reallocation of workloads to more experienced or successful interviewers at phases 2 and 3.

(iv) Distribution of a wide variety of materials to sample members to promote interest in and engagement with the study. This includes the use of primary approach letters at each wave, annual newsletters, newspaper articles highlighting the use of HILDA Survey data, a dedicated website for sample members, and even the annual statistical report. A particularly important feature of the annual newsletter has been short contributions from significant Australians (such as the Governor-General and the Governor of the Reserve Bank) testifying to the value of the study.

Data adjustments

Weights

Weights are used to adjust for differential non-response and attrition experienced at both the household- and person-level, as well as to adjust for unequal probabilities of selection. Both cross-sectional and longitudinal weights are provided on the HILDA data files. The cross-sectional weights apply to households, enumerated persons (all individuals in responding households) and responding persons (those individuals providing an interview) in each wave. The longitudinal weights apply to balanced panels of responding persons or enumerated persons from every wave to every other wave, and for the balanced panel of any combination of a pair of waves.

The weights are calculated via a three-step process. For wave 1, the steps are as follows: i) the initial weight is calculated as the inverse of the probability of selection; ii) these initial weights are multiplied by the inverse of the probability each unit had of responding; and iii) the adjusted weights are simultaneously calibrated to known population totals (for household composition, sex and age, labour force status, marital status, and geographic distribution). From wave 2 onwards, the weighting process becomes more complicated, due to both changes in household composition and sample attrition, but essentially the same three-step process is used. The factors used to estimate non-response probabilities, however, are much expanded to include information collected in prior waves.

Users can take into account the complex sample design of the HILDA Survey when calculating standard errors for population estimates in a number of ways (Hayes, 2008). Stratification and cluster variables are provided for those using the Taylor series linearisation method (which can be undertaken in SAS, SPSS and Stata) and replicate weights are provided for those using the Jackknife method (which can be undertaken in Stata or via user-written macros).

Imputation

Even when a respondent provides an interview, they do not necessarily provide answers to all questions asked, resulting in item non-response. The level of item non-response in the HILDA Survey is generally quite low (usually less than 1%), but for questions seeking monetary values (such as income, wealth and expenditure) the rates are much higher. The proportion of individuals, for whom at least some of the information necessary to construct total individual income is missing, has varied between 10% and 16% across the waves. At the household level this rises to between 15% and 21%, since interviews are not always obtained with all adults residing in responding households.

The data files contain both imputed and non-imputed versions of the income, wealth and expenditure derived variables, along with an imputation flag. The imputation uses a method developed by Little and Su (1989) that incorporates (via a multiplicative model) the trend across waves, the individual's departure from that trend, and a residual effect donated from another respondent with complete information. Suitable donors are identified from within imputation classes defined by age. Where a respondent has not provided the monetary component in at least one wave, a nearest neighbour regression method is used to impute starting values. These imputation methods were identified as preserving the best longitudinal and cross-sectional properties of the data in an evaluation study of eight commonly used, longitudinal imputation methods (Watson & Starick, 2011). As this is a single imputation method, the standard errors will be a little understated. The proportion of total individual income that is imputed is between 4% and 7%, as it can often be a small component of income that is missing. At the household level this rises to between 8% and 13%. Users can undertake their own multiple imputation procedures, but we are not aware of any having done so as yet.

Governance structures and scientific stewardship

As a government-owned study, all decisions about the design, content and administration of the HILDA Survey are subject to the approval of the funding agency, the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA). Governance structures, however, were established to ensure the study serves the interests of a much wider group of users, including both other government agencies and the broader academic research community.

The Director of the HILDA Survey has thus been required to report regularly to an advisory group convened by FaHCSIA, but comprising representatives from all of the major government departments and agencies in Australia with an interest in economic and social policy. This group is most important in identifying emerging concerns among policy-makers that the HILDA Survey may be in a position to address. In addition, the Melbourne Institute established two further reference groups, one comprising representatives from the academic research community, to mainly provide advice about survey content, and the other, comprising persons with high level expertise in survey statistics, to provide advice about methodological issues.

Data access

A confidentialised, unit-record data file is available, under licence, to bona fide researchers in academia, government and other selected research organisations. To access the data an application must be submitted to FaHCSIA using application forms that can be downloaded from the HILDA Survey website at:
http://melbourneinstitute.com/hilda/data/.
All applicants must agree to terms and conditions of use, which include keeping all unit record data secure, not attempting to identify or disclose the identities of any respondents, and not performing any matching, sharing, merging or linkage with any other individual-level datasets without prior written consent from FaHCSIA. Decisions about who can access the data are solely at the discretion of FaHCSIA.

The key differences between the confidentialised and non-confidentialised versions of the data are the withholding of some variables (notably postcode and precise date of birth), the aggregation of some variables (such as occupation and industry), and the top-coding of some variables (such as income and wealth variables).

Interested persons can apply either through an organisation licence (many Australian universities and government agencies have such a licence) or as an individual. There is a nominal charge to obtain the data (currently AU$121 for overseas-based

users). The data files are supplied on DVD and are provided in SAS, SPSS and STATA formats. Also provided on the DVD is extensive documentation, including coding frameworks, marked-up questionnaires, variable frequencies and the key reference source, the user manual. The data are released on an annual basis in early December each year, but each release includes all unit record data from the previous waves.

Data use and outputs

As shown in Table 2, the number of licensed users of each annual release is now around 500, with the cumulative number of users over the first ten releases (covering the first ten waves of data collection and spanning the period 2001 to 2010) numbering almost 1700.

Not surprisingly, given the size of this community of users, the HILDA Survey data has already spawned a large number of outputs. An attempt at tracking these outputs is maintained on the HILDA Survey website in the form of a bibliography. As at September 2012, this bibliography contained details of some 370 academic journal articles, seven books, 17 chapters in books, and more than 170 reports and other difficult-to-classify publications.

A key feature of this body of research is its breadth, spanning virtually the entirety of the social sciences. Examples of just a select few of the many varied topics covered include: poverty dynamics (Buddelmeyer & Verick, 2008); gender wage inequality (Barón & Cobb-Clark, 2010); transitions in and out of non-standard employment (Buddelmeyer & Wooden, 2011); the changing relationship between cohabitation and divorce (Hewitt & de Vaus, 2009); the association between income inequality and mental health (Bechtel, Lordan & Prasada Rao, 2012); and the adaptation of subjective well-being to major life events (Frijters, Johnston & Shields, 2011). Numerous researchers have also used the data to evaluate specific policy initiatives. Examples include the impact of: Australia's Baby Bonus payments on fertility behaviour (Drago, Sawyer, Seffler, Warren & Wooden, 2011); smoking bans on smoking behavior (Buddelmeyer & Wilkins, 2011); public housing assistance on employment outcomes of labour market program participants (Feeny, Ong, Spong & Wood, 2012); and family benefits payments and child care subsidies on couples' labour supply (Guest & Parr, forthcoming).

The data have also been well used by government, with government employees accounting for 36% of all HILDA Survey data users. Demonstrating a direct impact on policy, however, is more difficult, in part because rarely are policy interventions or changes in policy the result of any single piece of evidence, and in part because much of the use being made of the HILDA Survey data by government happens behind closed doors and so often not available for public scrutiny. We do, however, know that the data have featured prominently in numerous reports published, and inquiries conducted, by the Productivity Commission, Australia's key government agency with a research and advisory role in the area of economic and social policy, covering such issues as paid parental leave, disability care, and the use of labour hire employment. The data on household wealth have also been much used by the Reserve Bank of Australia to assess the amount of risk that households are prepared to accept, and have featured frequently in the Bank's quarterly financial stability reviews. The HILDA Survey data have also figured prominently in submissions to successive Annual Wage Reviews, with the findings from the data analysis highlighted in the decisions handed down by Fair Work Australia in 2011 and 2012. Finally, the data have been much used by the host agency, now the Department of Families, Housing, Community Services and Indigenous Affairs, and have been central to numerous reports, both in-house and commissioned, and in a number of key reviews, and notably the Pension Review (Harmer 2009).

Table 2. Number of licensed users of HILDA Survey data by release

Release	All licensed users	New users	Cumulative total
1	204	204	204
2	265	169	373
3	279	157	530
4	329	176	706
5	387	196	902
6	401	176	1,078
7	455	199	1,277
8	431	125	1,402
9	500	141	1,543
10 (@ 3 Sept 2012)	464	144	1,687

Conclusion

On many, if not most, criteria, the HILDA Survey must be judged a major success. This is most obviously reflected in:

(i) its longevity and ongoing Government support;
(ii) a growing community of data users from within both academia and government;
(iii) the large body of published research evidence that has used the data; and
(iv) rising and relatively high annual re-interview rates.

In part, these outcomes have been a function of the ongoing commitment to both the production of user-friendly data sets and high levels of user support, and a highly successful partnership with the Australian Government. Indeed, the high level of direct support from, and engagement with, government is probably the feature that most distinguishes the HILDA Survey from its international counterparts.

But perhaps most important for the study's success has been the role played by the fieldwork provider and its interviewers. Critical to the success of any survey-based study is the way interviewers interact with sample members, which in turn is a function of how engaged interviewers are with the study (as reflected in interviewer interest in the study and its outcomes, and belief that the objectives of the study are worthwhile). And in a panel study this is even more important given the potential for interviewers to develop ongoing relationships with sample members. Promoting interviewer engagement with the study and nurturing the relationships between interviewers and sample members thus remain the highest priorities of the HILDA Survey management team.

The HILDA Survey has also demonstrated that the often heard claim that populations are becoming increasingly less willing to respond to socio-economic surveys is not entirely justified, with response rates to its wave 11 sample top-up exceeding the rate obtained in wave 1.

The HILDA Survey, however, is not without its weaknesses. Perhaps most significant is the lack of resources devoted to recruiting new immigrants. Like all indefinite life panel designs, the main weakness in the design of the HILDA Survey sample is the lack of any automatic mechanism for adding new immigrants to the sample on an ongoing basis. The study attempted to deal with this by adding a large population replenishment sample in wave 11. Cost considerations, however, are likely to mean that this cannot be replicated every decade, and hence alternative, more cost-effective methods for augmenting the sample with new immigrants will need to be found.

Other weaknesses include a relatively low response rate in the initial survey wave, the absence of any linkages to administrative data sources, and some data items that are prone to high levels of measurement error (most notably the information collected as part of the labour market history calendar and the household expenditure data collected in the SCQ).

Acknowledgements

The HILDA Survey project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs and is managed by the Melbourne Institute of Applied Economic and Social Research.

References

Barón, J.D., & Cobb-Clark, D. (2010). Occupational segregation and the gender wage gap in private- and public-sector employment: A distributional analysis. *The Economic Record*, *86*, 227-246.

Bechtel, L., Lordan, G., & Prasada Rao, D.S. (2012). Income inequality and mental health: empirical evidence from Australia. *Health Economics*, *21*, 4-17.

Buck, N., & McFall, S. (2012). Understanding Society: design overview. *Longitudinal and Life Course Studies*, *3*, 5-17.

Buddelmeyer, H., & Verick, S. (2008). Understanding the drivers of poverty dynamics in Australian households. *The Economic Record*, *84*, 310-321.

Buddelmeyer, H., & Wilkins, R. (2011). Effects of tightening smoking regulations on take-up and cessation of smoking. *Australian Economic Review*, *44*, 167-181.

Buddelmeyer, H., & Wooden, M. (2011). Transitions out of casual employment: the Australian experience. *Industrial Relations*, *50*, 109-130.

de Leeuw, E., & de Heer, W. (2002). Trends in household survey nonresponse: a longitudinal and international comparison. In R.M. Groves, D.A. Dillman, J.L. Eltinge & R.J.A. Little (Eds.), *Survey Nonresponse* (pp. 41-54). New York: John Wiley and Sons.

Drago, R., Sawyer, K., Sheffler, K., Warren, D. & Wooden, M. (2011). Did Australia's Baby Bonus increase fertility intentions and births? *Population Research and Policy Review*, *30*, 381-397.

Feeny, S., Ong, R., Spong, H., & Wood, G. (2012). The impact of housing assistance on the employment outcomes of labour market programme participants. *Urban Studies*, *49*, 821-844.

Frick, J.R., Jenkins, S.P., Lillard, D.R., Lipps, O., & Wooden, M. (2007). The Cross-National Equivalent File (CNEF) and its member country household panel studies. *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, *127*, 627-654.

Frijters, P., Johnston, D.W. & Shields, M.A. (2011). Life satisfaction dynamics with quarterly life event data. *Scandinavian Journal of Economics*, *113*, 190-211.

Guest, R., & Parr, N. (forthcoming). Family policy and couples' labour supply: an empirical assessment. *Journal of Population Economics*.

Harmer, J. (2009). *Pension Review Report, 27 February 2009*. Canberra: Department of Families, Housing, Community Services and Indigenous Affairs.

Haisken-DeNew, J.P. (2001). A hitchhiker's guide to the world's household panel data sets. *Australian Economic Review*, *34*, 356-366.

Hayes, C. (2008). HILDA standard errors: user guide. HILDA Project Technical Paper Series no. 2/08. Melbourne: Institute of Applied Economic and Social Research, University of Melbourne.

Headey, B., Marks, G., & Wooden, M. (2005). The structure and distribution of household wealth in Australia. *Australian Economic Review*, *38*, 159-175.

Hewitt, B., & de Vaus, D. (2009). Change in the association between premarital cohabitation and separation, Australia 1945-2000. *Journal of Marriage and Family*, *71*, 353-361.

Little, R.J.A., & Su, H.L. (1989). Item non-response in panel surveys. In D. Kasprzyk, G.J. Duncan, G. Kalton, & M.P. Singh. (Eds.), *Panel Surveys* (pp. 400-425). New York: John Wiley and Sons.

Watson, N. (2010). The impact of the transition to CAPI and a new fieldwork provider on the HILDA Survey. HILDA Survey Discussion Paper no. 2/10. Melbourne: Institute of Applied Economic and Social Research, University of Melbourne.

Watson, N., & Starick, R. (2011). Evaluation of alternative income imputation methods for a longitudinal survey. *Journal of Official Statistics*, *27*, 693-715.

Watson, N., & Wooden, M. (2011). Re-engaging with survey non-respondents: the BHPS, SOEP and HILDA Survey experience. Melbourne Institute Working Paper Series no. 2/11. Melbourne: Institute of Applied Economic and Social Research, University of Melbourne.

Wooden, M., Freidin, S., & Watson, N. (2002). The Household, Income and Labour Dynamics in Australia (HILDA) Survey: wave 1. *The Australian Economic Review*, *35*, 339-348.

Book Review

History of the Avon Longitudinal Study of Parents and Children (ALSPAC) c. 1980-2000.
Overy, C., Reynolds, L.A., & Tansey, E.M. (eds.) 2012. Wellcome Witnesses to Twentieth Century Medicine, vol. 44. London: Queen Mary, University of London. 122pp. ISBN 9780 90223 878 7

John Bynner
Institute of Education, London
jbynner@slls.org.uk

There are many ways of writing the history of a research programme, project or research centre depending on author predilections, research team preferences and the aims of the exercise in terms of the messages that the sponsors want to get across to the target audience. The Wellcome Witnesses to Twentieth Century Medicine history series, compiled by the History of Modern Bioemedicine Group at Queen Mary College, London use, as the source for the history, a recorded conversation between key actors involved in the project, lasting for some hours. The monograph devoted to the life history of the Avon Longitudinal Study of Parents and Children (ALSPAC) – the 44th in the series - comprises the edited version of such a conversation that took place in the Wellcome Trust headquarters in London in May 2011. Nineteen people representing a wide range of ALSPAC interests, expertise and experience took part, including the original director, Jean Golding and the current director, George Davey Smith.

ALSPAC's key role in the development of birth cohort study strategy makes such a conversation particularly worthy of attention, while recognising the limitations of the discussion method in providing a full account. Inevitably, even in the hands of such an excellent chairman as Professor Catherine Peckham, Emeritus Professor of Paediatric Epidemiology at the London Institute of Child Health, a conversation on this scale can be controlled only to a certain extent. Consequently, some topics such as genetics attract a great deal of attention while others such as the important "Children in Focus" sub-study is mentioned only briefly but never fully explained and survey response barely features at all. A conventional history will include within the text, a lot of relevant facts about the subject. In the monograph, such information is supplied in footnotes, along with explanations of key terms and concepts and additional information about the affiliations and biographies of the speakers, and of others referred to in the discussion who played a crucial role. These are backed-up further by seven appendices and a bibliography.

The benefit of the method is the rich seam of personal reminiscences from key personalities in a major scientific venture revealing dogged perseverance, if not genius, in surmounting the obstacles that needed to be overcome. The weakness of the approach is a kind of lumpiness in the narrative, involving the reader in digging into footnotes that may or may not contain the information sought. If you are looking for measures of ALSPAC performance such as attrition rates, you will not find them here. Nor will you find much on the broader context of especially the British series of post-war birth cohort studies starting at 12 year intervals, 1946, 1958, 1970, 2000 and 2012 (listed in an appendix but not discussed). ALSPAC starting in 1991-92 is now seen as playing an important part in the series as the 'missing cohort'. We might assume the 1970 birth study, which grew up in the Department of Child Health in Bristol under the direction of Neville Butler, might have had some influence on ALSPAC planning, yet this survey similarly receives barely a mention.

But then maybe such an approach to understanding ALSPAC is not such a bad idea: technicalities and backcloths are the job of another kind of history e.g. Jean Golding' history of ELSPAC (Golding, 1989). The Wellcome discussion format is better seen as eliciting a collective oral history bearing comparison with such classics as the US 'Looking at Lives' (Phelps, Furstenberg & Colby, 2002) and closer to home the edited video transcript of a discussion in 1982 between the directors of the 1946, 1958 and 1970 birth cohort studies, James Douglas, Mia Kellmer Pringle and Neville Butler (Bynner & Goldstein, 1998). Personality emerges as a key factor in moving these pioneering studies along.

Few people who know about ALSPAC in general terms, or even use its data for their own research, have much idea of the saga involved in getting it going. This is why the period examined embraces only 10 years of the actual longitudinal survey, now in its 20th year, with the rest devoted to the preceding 10 years that it took to get the required funding together. Though talked about in the 1970s as Catherine Peckham recalls, the main driver for the project was actually the rejection in 1980, by the UK Department of Health, of a proposal by Jean Golding for a new national birth cohort study due to begin in 1982. Such a study would have continued the British birth cohort studies series of studies at 12 year intervals. Largely by chance, what emerged instead was an international commission from the World Health Organisation to test the feasibility of a comparative area-based birth cohort study focused on child health in a number of European countries. £5000 was made available for piloting a self-completion instrument in Greece, Russia and the UK. The comparative framework for the whole survey finally took substance in the form of the European Longitudinal Study of Parents and Children (ELSPAC), involving eight countries - and funded separately by each of them. This initiative provided the starting point for ALSPAC, the most successful of the individual studies that survived.

If WHO supplied the seed corn, inspiration for development arose more from a conversation between Jean Golding with geneticist Marcus Pembrey in 1988. Their idea was to break new ground by going beyond the scope of the earlier birth cohort studies in Britain that follow the developing child from birth, in favour of recruiting the individual cohort member from the first notification of pregnancy. In other words, the idea was to move away from the focus on the birth of a child and what followed, to the child's development from conception. And unlike the other ELSPAC studies, there was also to be much more emphasis in ALSPAC on the collection of biological samples early in the baby's life. The starting point was urine and blood samples from the mother during pregnancy and at the time of the birth, and storage of a sample of cord blood and the placenta. This was to be followed by the collection from the child, of blood samples for the extraction of DNA, and the regular collection of a range of genetic and non-genetic biological samples, and physical assessments throughout the child's development.

The idea attracted support across a wide range of interests leading to a bank of expert advice to draw upon. But, as David Gordon in the preface to the monograph makes clear, attracting funding for the study was a different matter. It required "a long memory, a conscience and willingness to work round mindless rules and regulations, and a disdain for action that is not backed by evidence". On the basis of what is reported I would add to that: "the need for: a clear vision of where the future lay, long before it was realisable in terms of scientific products, serendipity in taking advantage of everything that could possibly progress the work, and doggedness of a high order in resisting objections to what was being proposed". The record of 258 failed applications for grants and 176 that were successful is just one indication of the battles that had to be won.

The key problem in funding terms was a mismatch between the funding model customary in medicine, of a scientific project with hypotheses to be tested and data collected within a clearly defined analytic framework, and what amounts to setting up a small business. The birth cohort study needs to conserve resources not only to do the immediate research at hand, but to sustain the project on a long-term basis, so that the widest range of research rewards, some serendipitous, can be fully gained from it. Notably, it is not until relatively recently in the UK, that, as George Davey Smith observes, the recognition of the need to build the necessary infrastructure for a sustainable 'longitudinal resource' (as it is now called), has been endorsed by the Research Councils and Government, and has become the basis of the substantial joint funding now committed to it.

It would be difficult to overestimate the importance of the 'incurable optimism' of Jean Golding in sticking to her guns over the first 20 years of development, not least through building the right alliances. Without the support of the University of Bristol's Vice-Chancellor and Director of Finance, through effectively underwriting the study through dark times, when one month fixed-term contracts for staff were not uncommon, ALSPAC is unlikely to have begun, let alone survived. A degree of stability was finally achieved when first the Wellcome Trust then the Medical Research Council (MRC) started supplying core funding for genetic data collection.

The conversation brings these key challenges to centre stage and provides a vivid picture of the achievement in overcoming them. But as it emerges at the end, the key protection came ultimately from introducing charges as part of grant applications for access to ALSPAC data. First the MRC, then the Economic and Social Research Council (ESRC), agreed to accept such charges of up to £40,000 for the projects they agreed to fund.

Once the project was up and running with the first data collection scheduled for the period 1991-1992, the sample design and operational strategies for recruitment and retention of families to the study came into play. In the earlier birth cohort studies the sample was defined by all births in Great Britain in a single week. ALSPAC was based in the now defunct administrative area of Avon, comprising Bristol, Bath and surrounding areas in Gloucestershire and Somerset (population of 1 million individuals in the 1991 census) and extended the period of recruitment to a whole year, 1991-1992.

At this stage the advantage of an area study becomes obvious in the sense that the whole network of facilities and services that families are connected to in relation to the birth of a child can be tapped into as the means of making contact with expectant mothers. Such a network - extending to subsequent child care and schooling - supplies the foundations for continuing contact with the family, for the medical and other assessments of that the study demands. In the case of expectant mothers, general practitioners followed by midwives were typically the main point of first contact, with follow-up and filling of gaps by members of the team recruited especially for their inter-personal skills. Promotional literature in multiple translations, distributed through libraries, GP surgeries and other media was also used to recruit mothers who might have fallen through the net.

Retention techniques inherited from other birth cohort studies, such as birthday cards and regular feedback to parents, were also built into the study's approach. Much attention was also paid to child-friendly clinical procedures, to make the experience of assessment as enjoyable as possible. Another feature common to effective birth cohort studies is boosting the feeling of 'specialness', extending, in the case of an area study, to local 'ownership' of a project of much importance to medical and health service improvement.

The management of feedback also confronted the problem of what medical information to supply to families in the event of serious problems being identified in the child's development. It is at this stage that another initiative, widely copied elsewhere, comes into the picture - an ethics committee established specifically for the project, but also, unusually, covering legal advice and scientific advice alongside ethical rulings. In this respect, the committee served as a kind of interface with the officially established Local Research Ethics Committee (LREC) operated in the UK by the National Health Service to govern research involving human subjects.

LRECs can often be seen more as obstacles to rather than facilitators of research. The ALSPAC experience shows how an internally appointed committee can actually work most effectively in collaboration with a research team in solving problems. Members of the committee clearly became as committed to the success of the project as the scientific team running it, and this contributed in a major way to the success of the project that continues to the present day.

The central principle that emerged early on in the committee's work was that every consideration of what was scientifically desirable, and therefore might be permissible in the study, was secondary to the interests of the child. Thus it was mandatory to ask the child's permission, from as early an age possible, for the application of any medical assessment procedure in the course of data collection. Thus the procedure for collecting blood, for example, went through a number of iterations until the means was found of making the use of a syringe for this purpose virtually unnoticed. The collection of first (milk) teeth by parents, who posted them to the team for the reward of a badge for the child, was also aided by one of the team acting the part of 'tooth fairy' for the study.

Such sensitivity and consideration for the child's needs – of vital importance to long term participation in the survey - was carried over in the cohort's teenage years to collective consultation through the Teenage Advisory Panel (TAP) covering all aspects of the survey.

The committee's only notable failing was to decide against recruitment of fathers at the same time as the mothers in the study. As the team acknowledge, genetic data for the whole family would have offered much added value over that of

the mother and baby alone and would now be central to such a study's design.Apart from the acclaim attached to the work of the ethical committee, tribute is also paid to the massive technological effort that went into data entry, processing and labelling of especially the biological data. This included the use of robotics from industry at the storage and data preparation stages. It is hard to remember now that in the beginning stages of the project in the early 1980s, cohort studies were still a long way from the massive high-speed computing facilities that enable data processing to be carried out anywhere with relative ease. At the beginning of ALSPAC's development, survey research was barely past the era of punched cards, card sorting machines and batch processing through mainframe computers, all of which added time and cost to the data collection and analysis enterprise.

This account rightly ends with recognition of perhaps ALSPAC'S biggest achievement (as the last part of the discussion makes clear), namely the collection and use of the biological samples collected in pregnancy and after. This resource now underpins a major programme of leading edge genetic investigations. One of the key strengths of ALSPAC is that the individual genotype is embedded in a series of exposures to environmental influences in every domain of life over a period of what is now 20 years – an unsurpassed resource for epidemiological understanding.

To summarise twenty years' history in a few paragraphs does insufficient justice to a project that helped move birth cohort studies to a new level of scientific significance, paving the way for the large-scale cohort studies in the USA and UK targeting up to 100,000 pregnancies. The collective oral history that the monograph reflects is in no sense a complete picture of what a full appraisal of ALSPAC would encompass. That we may hope is still to come.

What such a history does is illuminate facets of experience that no other method could match. We learn what it was like to work on the project and how, and why, key decisions were taken. The insights gained provide guidance about the pitfalls, as well as how to avoid them, in a 'real life' context that text book descriptions often fail to get across. They also contribute to strategy for understanding the human condition in challenging established paradigms of scientific value and legitimacy in favour of a more holistic, life course-orientated approach to developmental science. Those who contributed to the ALSPAC endeavour and provided the where-with-all for its achievement deserve the gratitude of those who have reaped the benefit.

References

Bynner, J. and Goldstein, H. (1998) Three generations of children – an edited transcript of a video recording made in March 1982 of the then directors of the three major British cohort studies in conversation. *Paediatric and Perinatal Epidemiology*, *12, Supplement*, 15-30.

Overy, C., Reynolds, L.A., & Tansey, E.M. (eds.) 2012. *History of the Avon Longitudinal Study of Parents and Children (ALSPAC) c. 1980-2000*. Wellcome Witnesses to Twentieth Century Medicine, vol. 44. London: Queen Mary, University of London.

Golding, J. (1989) European longitudinal study of pregnancy and childhood. *Paediatric and Perinatal Epidemiology, 3*, 460-9.

Phelps, E., Furstenberg, F.F. Jnr., & Colby, A. (2002). *Looking at lives: American Longitudinal studies of the twentieth century.* New York: Russell Sage Foundation.