Social Science & Medicine 72 (2011) 193–201

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Change in income and change in self-rated health: Systematic review of studies using repeated measures to control for confounding bias

Fiona Imlach Gunasekara*, Kristie Carter, Tony Blakely

Department of Public Health, University of Otago, Wellington School of Medicine and Health Sciences, 3 Mein St, Newtown, Wellington, New Zealand

ARTICLE INFO

Article history:
Available online 20 November 2010

Keywords:
Income
Health change
Longitudinal studies
Systematic review
Self-rated health

ABSTRACT

It is generally assumed that income is strongly and positively associated with health. However, much of the evidence supporting this assumption comes from cross-sectional data or analyses that have not fully accounted for biases from confounding and health selection (the reverse pathway from health to income). This paper reports results of a systematic review of panel and longitudinal studies investigating whether changes in income led to changes in self-rated health (SRH) in adults. A variety of electronic databases were searched, up until January 2010, and thirteen studies were included, using data from five different panel or longitudinal studies. The majority of studies found a small, positive and statistically significant association of income with SRH, which was much reduced after controlling for unmeasured confounders and/or health selection. Residual bias, particularly from measurement error, probably reduced this association to the null. Most studies investigated short-term associations between income and SRH or the effect of temporary (usually one year) income changes or shocks, so did not rule out possibly stronger associations between health and longer-term average income or income lagged over longer time periods. Nevertheless, the true causal short-term relationship between income and health, estimated by longitudinal studies of income change and SRH that control for confounding, may be much smaller than that suggested by previous, mostly cross-sectional, research.

Introduction

Background

The relationship between income and health has been the subject of much research, but most of this research has used cross-sectional data or analysed income and health at only one point in time. Randomised controlled trials of income changes are scarce. A review of those that do exist produced inconclusive evidence of an effect of income supplementation on health outcomes, due to poor quality data (Connor, Rodgers, & Priest, 1999). Another review including all studies, even non-randomised, that evaluated interventions aimed at increasing incomes of those in poverty, also found mixed results and many studies of poor quality (Ludbrook & Porter, 2004). Evidence is suggestive, particularly for the importance of income for the health of those in poverty (Benzeval & Judge, 2001; Kawachi, Adler, & Dow, 2010) but a definite causal relationship between income and health has – perhaps surprisingly – not yet been established.

The most thorough review of income and health using longitudinal data was in 2001, which identified sixteen English-language studies that investigated the relationship between income and health over time (Benzeval & Judge, 2001). However, this review was non-systematic, the last included study was published in 1997 and studies of income and health changes were not the focus, although these types of studies are best placed to control for confounding bias. The majority of studies that included a measure of income change found a significant relationship with health, but the health outcome was usually mortality, so was potentially affected by confounding bias, as unmeasured differences between individuals (associated with income) could account for at least some of the mortality variation.

Systematic reviews on the income–health relationship are otherwise lacking, and reviewing the entire literature on this topic would be a massive undertaking.

However, a review of only repeated measures studies that minimise confounding is feasible, and arguably the most relevant. The major source of confounding is from unmeasured time-invariant confounders, or those fixed characteristics of individuals that are unknown or unquantified, such as personality traits, ability, genetics and beliefs (also known as unobserved heterogeneity) that are associated with both income and health (Imlach Gunasekara,

* Corresponding author. Tel.: +64 43855541.
E-mail address: fiona.imlach-gunasekara@otago.ac.nz (F.I. Gunasekara).

0277-9536/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2010.10.029

Carter, & Blakely, 2008). Methodologically, in observational research, the two best ways to minimise this confounding are: (1) fixed effects models using longitudinal data and repeated measures of both the exposure and the outcome, so that only within-individual changes over time can be analysed, eliminating between-individual differences (Wooldridge, 2002), and (2) instrumental variable (IV) analyses, which may arise in the context of natural experiments (Glymour, 2006), where people are unexpectedly 'randomly assigned' an additional source of money. The IV is an exogenous or external variable, unrelated to the outcome except through the exposure variable, that is used when the exposure and outcome relationship is affected by bias from health selection (the reverse causal pathway from health to income) or unobserved heterogeneity (Angrist & Krueger, 2001). However, strong and valid IVs are notoriously difficult to identify.

This paper presents a systematic literature review of longitudinal studies investigating whether a causal relationship between income and health exists, for the question: 'Do changes in income lead to changes in self-rated health in adults?' Only papers using longitudinal data and repeated measures data analysis are included. This review differs from Benzeval and Judge (2001) because both *change* in income and *change* in health are inclusion criteria, allowing more thorough control of confounding and other biases. Self-rated health (SRH) is the outcome of interest, as it is one of the commonest repeated measures of general health asked in longitudinal surveys. This paper does not attempt to review the effect of income on mental health, the income–health association in children or the effect of wealth on health, all of which are separate but important research questions.

Methodology

Inclusion criteria

The criteria for inclusion were that studies must:

1. Include at least two fiscal income measurements, or an exogenous instrumental variable for income change.
2. Include repeated measures of SRH as the outcome variable, with SRH being some variant of the question: 'In general would you say your health is excellent, very good, good, fair or poor?' Including initial health status as a covariate in a model with health at a later time period as the outcome did not qualify as an analysis of health change.
3. Include adult participants (fifteen years and older).

Search strategy

Searches were performed on the following electronic databases in April 2007: Medline, CINAHL, Econlit, ABI/INFORM Global, Academic Research Library, Dissertations and Theses, Proquest Research Library, Proquest Science Journals, Proquest Social Science Journals, Academic Search Premier, SSRN elibrary, Embase and PsycINFO. Ongoing alerts, using the original search strategies, were reviewed monthly up until January 2010. No language restriction was employed. Reference lists of relevant articles were also examined for papers.

The general structure of the search strategy was 'income' and terms related to income, including appropriate MeSH headings, and depending on the database, AND health AND longitudinal study, with related terms (see Appendix). The initial search in April 2007 produced 3393 references (Fig. 1). After reviewing the initial references and ongoing alerts, only 13 studies met the inclusion criteria.

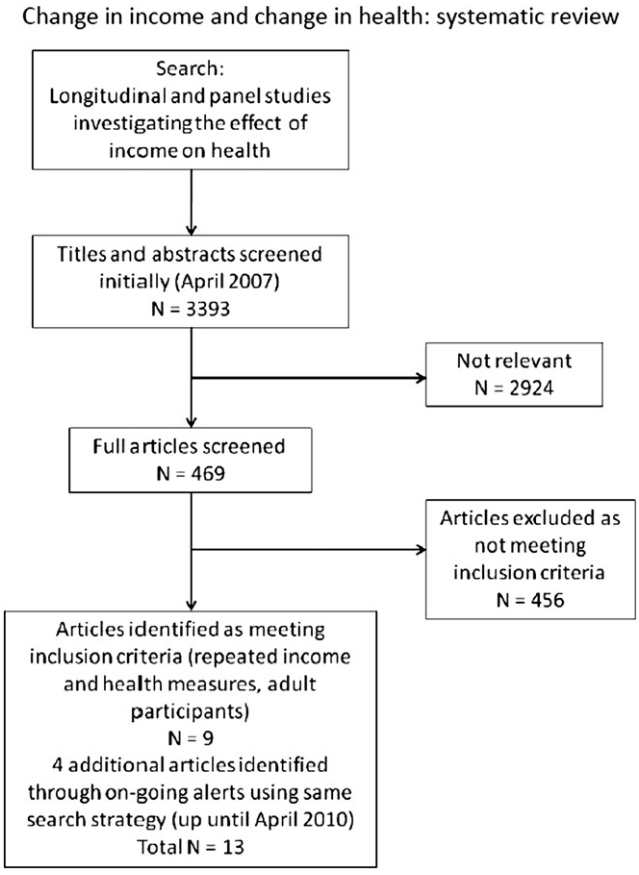

Fig. 1. Process of study selection.

Data extraction and quality assessment

Data were extracted from the included studies onto a form adapted from the Strengthening Reporting of Observational Studies in Epidemiology (STROBE) statement (Vandenbroucke et al., 2007). The 'quality' of the paper was assessed by identifying critical aspects of study design and methods that increase internal validity and minimise bias (preferred to numerical quality scoring with a weighted list of items (Higgins & Green, 2006; Sanderson, Tatt, & Higgins, 2007)), which were:

- Sample size: Larger samples (>10,000) increase precision and power.
- Length of follow up: Extended panels (>20 years) are more able to unravel the complex relationship between income and health.
- Appropriate sampling methods: Ensuring the survey is representative of the target population, to minimise selection bias.
- Appropriate management of attrition: Includes prevention of attrition, to minimise attrition bias.
- Measurement and control of confounders: Includes control for time-invariant and time-varying confounders.
- Testing or control for health selection bias.
- Appropriate measurement of income and health and management of correlation over time: To reduce measurement error/bias.

The quality assessment was applied to all papers that met the inclusion criteria to identify strengths and weaknesses, summarised in Table 1.

Table 1
Summary of papers included in systematic literature review.

Study	Sample	Included variables	Methods and results	Strengths	Limitations[a]
Studies from British Household Panel Survey (BHPS)					
Contoyannis, Jones, and Rice (2004)	BHPS 1991–1998; 48,992 observations from 2780 men and 3340 women ≥ 16 years (balanced panel); 64,053 observations (unbalanced panel)	Outcome: self-rated health (SRH) as five-level ordinal variable; Exposure: logged annual household equivalised income; Confounders: age, sex, ethnicity, education, marital status, household size, number of children in household	Correlated random effects ordered probit models, estimated separately by sex. Income change had small, significant positive association with SRH for men (estimate 0.06, standard error (SE) 0.03 in balanced sample; estimate 0.05, SE 0.02 in unbalanced sample) but for women only in unbalanced sample (estimate 0.04, SE 0.02)	Control of unobserved heterogeneity; investigation into dynamics; reporting of attrition rates and examination of potential attrition bias; sensitivity analyses to assess differential measurement bias in SRH	Health selection bias not controlled for (income exogenous); small-moderate sample size; relatively short length of follow up; labour force status not included
Lorgelly and Lindley (2008)	BHPS 1991–2004; 8645 participants ≥ 16 years (4100 men and 4545 women)	Outcome: SRH as four-level ordinal variable; Exposure: logged annual household equivalised income; Confounders: Gini coefficient, relative income measures, age and quadratic, ethnicity, education, marital status	Correlated random effects ordered probit models,[b] estimated separately by sex. Income change had small, positive, significant association with SRH for men (estimate 0.06, t statistic 3.53) and women (estimate 0.06, t statistic 3.62)	Moderate sample size; good length of follow up; control of unobserved heterogeneity; attrition investigated with variable addition tests	Health selection not explicitly discussed or controlled for (income exogenous); labour force and family status not included; measurement bias
Jones and Wildman (2008)	BHPS 1991–2001; 9902 participants ≥ 16 years; 53,084 observations on men, 62,925 observations on women	Outcome: SRH dichotomised to good/bad health; Exposure: logged annual household equivalised income; Confounders: relative income, age, ethnicity, education, number of children, household size, marital status, region, labour force status	Fixed and random effects linear and semi-parametric models, estimated separately by sex. For men and women, income had a small, positive, significant effect on SRH. One unit (log) income rise increased probability of reporting better SRH by 1%	Moderate initial sample size and length of follow up; control of unobserved heterogeneity; large number of time-varying confounders included	Health selection bias discussed but not controlled for (income exogenous); dichotomising SRH not ideal; measurement bias
Apouey and Clark (2009)	BHPS 1997–2005; 8343 observations from lottery winners ≥ 16 years	Outcome: SRH as five-level ordinal variable; Exposure: logged lottery wins; Confounders: age, sex, ethnicity, education, employment status, marital status, region, number of children, lagged SRH, lagged (log) household income	Random effects ordered probit models using lottery wins as exogenous income 'shock'. For players, winning had a small but insignificant effect on SRH two years later (estimate 0.01, SE 0.01)	Exogenous variable for income change controls for health selection bias; control for large number of measured confounders	Potentially limited external generalisability, as findings may only apply to those who play lotteries; may be residual confounding; small win may lead to null effect; measurement bias
Studies from US Panel Study of Income Dynamics (PSID)					
McDonough and Berglund (2003)	PSID 1967–1996; 7258 household heads ('husbands' in couple households plus all singletons) and partners aged ≥25 years in 1984	Outcome: SRH from 1984 to 1996; Exposure: poverty dynamics based on change in income-to-needs ratio (household income/poverty threshold based on estimated annual needs standards) from 1983 to 1995; Confounders: age, sex, ethnicity, marital status, education	Growth curve models. Change in income-to-needs ratio (increased income) had small, positive, significant effect on SRH (estimate 0.010, SE 0.002). Past experience of poverty (low income) had negative impact on SRH	Long follow up; moderate sample size	Potential bias from unobserved heterogeneity, attrition and health selection (income exogenous); dichotomising income into poverty limits comparability to studies using other income measures; labour force and family status not included; measurement bias

(*continued on next page*)

Table 1 (*continued*)

Study	Sample	Included variables	Methods and results	Strengths	Limitations[a]
McDonough, Sacker, and Wiggins, (2005)	PSID 1968–1996; 4351 household heads and partners in 1968	Outcome: SRH 1984–1996; Exposure: poverty (defined by estimated annual needs standards) history based on income from 1968 to 1982 and current poverty based on income from 1983 to 1995; Confounders: age, sex, ethnicity, education	Growth curve models. Poverty history did not affect rate of change of SRH over time. 'Current poverty' had a significant negative effect on SRH from 1983 to 1995 for all (estimate −0.08, SE 0.02)	Long follow up which reduces censoring bias	Small sample size; potential bias from unobserved heterogeneity, attrition and health selection (income exogenous); labour force status, marital status or family composition not included; measurement bias
Halliday (2007)	PSID 1978–1997; 6507 men and 7265 women, restricted to 30–60 year olds	Outcome: SRH 1984–1997 dichotomised into good and bad health; Exposure: labour income; Confounders: county level unemployment, age, ethnicity, education, dummy variable indicating zero labour income (whether an individual employed or not)	First-differenced binary logistic models using Arellano–Bond estimator, estimated separately by sex. In men, decreased labour income significantly associated with bad SRH (estimate −0.01, t statistic −2.24) but not for women (estimate 0.00, t statistic 0.06)	Moderate follow up although analyses of SRH only from 1984 to 1997, with income data from previous years used to instrument income; moderate sample size; fixed effects and instrumental variable method control for unobserved heterogeneity and health selection (income endogenous)	Potential for attrition bias; Arellano–Bond estimator may not be ideal as developed for linear outcomes and panel data with many observations but few time periods; marital and family status not included; labour income is only one pathway through which income may affect health; dichotomising SRH not ideal; measurement bias
Study from BHPS and PSID					
Sacker, Wiggins, Bartley, and McDonough (2007)	PSID 1990–2001; 4402 household heads and partners aged 25–55 years with complete data; BHPS 1 991–2002; 4116 individuals aged 25–55 years with complete data; 2-yearly SRH data	Outcome: latent health based on SRH; Exposure: 'low income' (lowest 20% of household income), income divided by square root of household size; Other variables: age, sex, ethnicity, education, work-limiting illness, employment status, occupation	Latent transition analysis (hidden Markov models). Declining health group experienced higher rates of low incomes after 11 years. Proportion of low income in the improved health group fell in US; no change in UK	Moderate length of follow up; latent health states allowed for measurement error in SRH; exploration of health selection and social causation pathways; health trajectories are flexible as are not fixed in one direction over time	Small sample sizes; potential for selection and attrition bias; cannot quantify the effect size of changes in income on health states; 2-yearly data reduces ability to capture changes in SRH over time; direction of income–SRH relationship is not entirely clear
Studies from German Socioeconomic Panel (GSEOP)					
Frijters, Haisken-DeNew, and Shields (2005)	GSEOP 1990–2002 (1984–2002 for West German comparison group); 6198 East German participants aged ≥18 years; 22,461 observations on men, 24,492 observations on women	Outcome: health satisfaction (HS) on 0–10 scale dichotomised with an individual-specific cut-point; Exposure: logged annual household income; Confounders: marital status, number of children, major life events in last 12 months, labour force status, move to West Germany post reunification	Fixed effects ordered logistic regression models, estimated separately by sex. For East Germans, income changes had a small, positive, significant effect on HS for men but not women (1 log unit increase in income led to a 0.08 increase in HS). Same results using different income measures (3 or 5 year average income, continuous income, household equivalised income, pre-tax income, relative poverty measure)	Moderate sample size and length of follow up; control for unobserved heterogeneity; exogenous income increase to account for health selection bias; discussion of measurement bias	Loss of sample over time with replacement of sample (only 26% of East Germans observed in all 13 waves); ordinal fixed effects method makes interpretation of results more complex; income increases in East Germans probably less exogenous over time, introducing bias
Jones and Schurer (2009)	GSEOP 1984–2005; aged ≥ 16 years; 20,950 women, 20,171 men, average 7 years in panel	Outcome: HS on 0–10 scale mapped to five-level SRH; Exposure: logged household equivalised monthly income; Confounders: education, immigrant status, marital status, location in East or West Germany, employment status, household size	Generalised conditional fixed effects logit (GCFEL), estimated separately by sex. No relationship between income and SRH found in any age group. Income increases may be significant for improving health of people in poor health, but not for those already in good health	Large dataset and length of follow up; control for unobserved heterogeneity; many measured confounders included; wide range of methods compared	Use of multinomial models for ordinal outcome; potential attrition bias; health selection bias (income exogenous); measurement bias

Fischer and Sousa-Poza (2009)	GSEOP 1992–2005; 16–60 years employed only; 95,261 observations (65,332 used for first-differenced model)	Outcome: SRH as five-level continuous variable; Exposure: logged household income; Confounders: job satisfaction (and quadratic), log of age, marital status	Fixed effects linear model, first-differenced linear model. No significant association between income and SRH (estimate 0.01, SE 0.60 in fixed effects; estimate 0.00, SE 0.15 in first-differenced model)	Fixed effects models control for unobserved heterogeneity; moderate sample time frame with large number of observations	Possible selection bias as only employed individuals included; potential attrition bias; treating (ordinal) SRH as linear not ideal; family structure not included; health selection bias (income exogenous); measurement bias
Study from Health and Retirement Study (HRS)					
Frijters and Ulker (2008)	HRS 1992–2002 biannual data (6 waves); 7600 households with a primary participant aged between 51 and 61 years in 1992; 49,666 observations (fixed effects model)	Outcome: SRH dichotomised with an individual-specific cut-point; Exposure: logged household income; Confounders: age, marital status and family, wealth, occupation, number of years worked, smoking, exercise, alcohol consumption	Conditional fixed effects ordered logit model. No significant association between income and SRH (estimate 0.01, SE 0.01). Controlling for unobserved heterogeneity substantially reduced significance and size of the association	Moderate length of follow up; control of unobserved heterogeneity; many measured confounders included; different outcomes and exposures compared	Ordinal fixed effects method makes interpretation of results more complex; possible attrition; health selection bias (income exogenous); measurement bias
Study from Household, Income and Labour Dynamics in Australia (HILDA) Survey					
Buddelmeyer and Cai (2009)	HILDA 2001–2005; balanced panel of 1749 families (8745 observations) headed by persons aged 18–64 years, excluding families headed by full time students and families missing exposure variables	Outcome: SRH of family head, dichotomised into good and ill health. Family head was the oldest male in a family headed by a couple or the oldest person in other family types. Exposure: poverty ('poor' if equivalised family disposable income less than 50% of median income of analysis sample); Confounders: age, sex, education, family type, personality, region, place born	Dynamic simultaneous equations. Family head more likely to report poor health if in poverty the year before (estimate 0.09, SE 0.01); family headed by person in poor health more likely to be in poverty than family headed by person in good health (estimate 0.02, SE 0.01; mean marginal effects); health selection a significant pathway	Simultaneous equation method used to control for health selection bias (poverty endogenous); moderate sample size	Family as unit of analysis limits comparability and inefficient; use of 'family head' can introduce error as family head may not be accurately identified; potential bias from unobserved heterogeneity and attrition; strong assumptions in specifying equations (e.g. SF-36 used as IV for SRH to identify poverty equation); 5 years only; labour force status, marital status, ethnicity not included; dichotomising SRH not ideal

[a] When variables that are commonly included as potential confounders of income–health relationship (e.g. labour force status, marital status) are omitted from the analysis, this is noted as a limitation, due to the possibility of introducing omitted variable bias/confounding.

[b] These models are described as 'fixed effects ordered probit' models in the paper by Lorgelly and Lindley (2008) but approximate a fixed effects model rather than being a true conditional fixed effects model.

Data presentation

The eligible studies applied a range of analytical methods, and transformed income and SRH in different ways, making a meta-analysis of results impossible. Therefore, studies are discussed in a narrative fashion. Several studies presented models that adjusted for correlation from repeated measures as well as pooled models that did not. In such cases, only results from the most appropriate models, using longitudinal data analysis accounting for variability in the data, are discussed.

Results

Thirteen studies were identified as eligible for inclusion in the systematic review (Table 1). Two used natural experiments and measures of exogenous income (lottery wins and the reunification of East and West Germany as proxies for unexpected income windfalls (Apouey & Clark, 2009; Frijters et al., 2005)), although these were not instrumental variable analyses. All of the studies used data from five high quality panel surveys; therefore a brief overview of each survey is given before discussing studies using data from that survey.

British Household Panel Survey

The British Household Panel Survey (BHPS) was used by five studies (Table 1). The BHPS began in 1991, with initially about 10,000 individuals from 5000 households, surveying private households in England, Wales and Scotland. All adult (aged 16 years or older) members of a household containing an original sample member (participants eligible and included in wave one) are interviewed annually. By wave fourteen, 51.5% of all wave one participants were still being interviewed (Uhrig, 2008).

The first study from the BHPS explored the dynamic relationship between income and health using eight waves of data and correlated random effects ordered probit models including lagged SRH (at the previous time period) and initial health (at the first time period) (Contoyannis et al., 2004). The final models also included two measures of income: average income, which was interpreted as long-term or permanent income, and current income, which was a measure of income change. It was found that long-term income was more strongly associated with SRH than income change, with the coefficients for long-term income (which are prone to confounding) about five times larger than those for income change (which are less affected by confounding). Once bias from unmeasured confounders was accounted for, the association between income change and SRH reduced by at least five-fold. Overall, a one unit increase in log income (a 2.7-fold increase) over the population increased the probability of reporting excellent health by approximately 1%.

The second study from the BHPS investigated the impact of absolute income and relative deprivation on SRH, using ten waves of data (Jones & Wildman, 2008). The results from fixed effects linear models (using within-individual changes over time) found that for both men and women, (log) annual household income had a small, positive and significant effect on SRH. An increase in one unit (log) income increased the probability of reporting good health by 1% (Jones & Wildman, 2005). The overall mean SRH in this study was 0.7. A one unit increase in log income (using the mean log income of 9.6 (£14,765) this would equate to an increase of £25,370, up to £40,135) would raise this to 0.71 (a 1% increase). This effect was observed either with or without relative deprivation as an additional variable in the model.

The third study also investigated the comparative importance of absolute income, income inequality (measured by the Gini coefficient) and relative deprivation for SRH, using 14 waves of the BHPS (Lorgelly & Lindley, 2008). Only change in the absolute income measure was found to be positively and significantly associated with SRH. No significant relationship with SRH was found for changes in relative deprivation or income inequality. Therefore, this study found a similar result and strength of association to Contoyannis et al. (2004) using a similar method but six more waves of data.

The latter two studies did not include lagged health or initial health, unlike Contoyannis et al. (2004). Including initial health as an attempt to control for health selection has been shown to diminish the effect of income on health (Benzeval & Judge, 2001), but also potentially introduces endogeneity bias as initial SRH is correlated with the outcome of SRH over time. Despite these differences, all studies found similar results of a small effect of income change on SRH, and concluded that unobserved heterogeneity was a major source of bias in the income–health relationship.

The fourth study from the BHPS used the 'random assignment' of lottery winners to investigate the impact of an exogenous or unexpected income rise (income 'shock' of lottery winnings) on a variety of health outcomes, including SRH, over time (Apouey & Clark, 2009). Only those who reported wins were included in the analyses, as it is likely that those who gamble and those who do not differ in unmeasured ways. Ordered probit random effects regression models were used with the outcome being SRH at two time periods subsequent to the win. The average win was £170 and 5% of the winners won more than £500. The best estimate for (log) winnings was of a similar magnitude to the above studies (0.01), but small numbers resulted in a 95% confidence interval of −0.01 to 0.03, that included the null. Beyond SRH, lottery wins had a significant positive effect on mental health, no effect on a number of health problems, but significant negative effects on health behaviours such as smoking and alcohol consumption. The small size of the average win could explain the null effect on SRH.

The fifth study using BHPS data also included data from another panel study (Sacker, Wiggins, Bartley, & McDonough, 2007) and is discussed in the next section.

Panel Study of Income Dynamics

The North American Michigan Panel Study of Income Dynamics (PSID) was used by four studies included in this review (Table 1). The PSID is a long-standing representative longitudinal survey of men, women and children survey begun in 1968 that started with a national sample of nearly 5000 households. Household heads and wives comprise the sample and children are interviewed when they establish families of their own. The survey was designed to track income and employment experiences over time but only collected SRH data from 1984 to 1997. It was an annual survey up until 1997, when it became biannual.

The first study assigned individuals to poverty categories (in poverty or not) at each year, based on estimated standard annual needs and household income level from 1967 to 1983 (McDonough & Berglund, 2003). Many analyses related SRH after 1983 to poverty dynamics up until 1983. Of specific interest to this review, however, was a growth curve model including a dynamic poverty measure, which was annual change in the income-to-needs ratio from 1983 to 1995. This had a small, positive and significant effect on SRH, indicating that SRH improved as income relative to needs increased. The model also included income-to-needs ratio in 1983 as a measure of 'current' poverty or baseline income. This coefficient was more than five times the magnitude of the change in income-to-needs ratio coefficient, consistent with Contoyannis et al. (2004) above which found that average (or baseline) income was more important to SRH than current changes.

The second study also aimed to find out whether different poverty histories affected future health using growth mixture models (McDonough et al. 2005). The definition of poverty was also based on estimated standard annual needs, but participants were classified into four groups depending on their income history from 1968 to 1982: stable non-poor, stable poor, declining and improving. 'Current poverty' (poverty between 1983 and 1995) had a significant negative effect on SRH (also between 1983 and 1995) for all poverty groups. However, the rate of decline of SRH over this time was not significantly different between poverty groups. Poverty history did not exert a differential effect on SRH changes, as was anticipated.

The third study from the PSID took labour income data (income data derived from wages) from 1978 to 1997 and investigated how income 'shocks' affected health, theorising that income shocks primarily arose from unemployment (Halliday, 2007). The models were based on first-differenced equations, where outcome and time-varying exposures were differences from one time period to the next. However, in this study, these differenced outcomes (and labour income exposure variables) were instrumented by time-lagged health and labour income variables, in an application of the Arellano and Bond estimation procedure (Arellano & Bond, 1991). The best specified and most realistic models dichotomised SRH into good and bad health and treated labour income as endogenous, allowing for feedback of income to health (controlling for health selection bias). These models found that a decreased labour income was significantly associated with bad SRH in men, but not for women. The effect size in men was very small.

The last study compared data from the PSID and the BHPS to explore reasons for the differences in health status between the US and the UK (Sacker et al., 2007). Patterns of individual health changes over 11 years were related to measures of socioeconomic position including income, using latent transition analysis to examine transitions to and from latent health good and health states, derived from SRH, over time. At baseline, the group whose health declined over time resembled the group who had stable good health in socioeconomic profile, but 11 years on, the proportion of people with low income increased in the declining health group. The group whose health improved over time was initially similar to those with stable poor health but experienced increased income only in the US, highlighting country-specific factors and context as potentially important. The direction of association between income and health changes was not entirely clear from this analysis.

German Socioeconomic Panel

The German Socioeconomic Panel (GSOEP) is a survey using a nationally representative sample of adults (aged 17 years and older) in private residences followed annually from West Germany since 1984 and from East Germany since June 1990, just prior to the official reunification in July 1990 (Wagner, Frick, & Schupp, 2007). The original West German sample included about 12,000 individuals from 6000 households and the East German sample added another 4400 individuals (with additional samples added at later dates). The GSOEP was used by three studies included in this review (Table 1).

The first study used the reunification of East and West Germany in 1990 as a 'natural experiment' because the fall of communism in East Germany was unexpected and resulted in significant rises in income and savings levels for many East Germans over a very short time period (Frijters et al., 2005). The authors argued that no other significant events or structural changes overly influenced health during the study period (e.g. the health system remained the same). Therefore, income increases resulting from the reunification of Germany were considered exogenous for East (but not West) Germans so that models of the income and health relationship for East Germans should be less affected by health selection bias. This assumption could be disputed, especially with as time from reunification advances – the exogeneity of income increases is likely to diminish, introducing bias.

The study used twelve years of data on approximately 6000 East German participants (Frijters et al., 2005). The health outcome was health satisfaction (HS), based on the question: "How satisfied are you at present with your health situation?" with responses from 0 (very unsatisfied) to 10 (very satisfied). Health satisfaction was the only health question asked annually in all waves of the GSOEP. However, the HS question was found to be highly correlated with SRH asked in later waves and sensitivity analyses using the more limited SRH data found very similar results (Frijters et al., 2005). Therefore, HS was considered to be a valid proxy for SRH.

A conditional fixed effects ordinal logistic regression model was used (Ferrer-i-Carbonell & Frijters, 2004) which allowed the cut-point for the outcome to vary for each participant instead of setting a fixed cut-point for the entire sample. This meant that most observations of change could be included in the analysis. Disadvantages included its complexity and computational intensity (Jones & Schurer, 2009), and that the size of the coefficients had no easy interpretation, as they did not directly relate back to changes in aggregate HS (Frijters et al., 2005). The authors addressed this with a causal decomposition analysis, which estimated what changes occurring in the sample led to changes in aggregate HS. Income increases had a positive and significant effect on HS for men, but not for women. However, the size of the effect was small (one log unit increase in income leading to a 0.083 increase in HS, measured on a ten point scale) (Frijters et al., 2005). In a working paper on the same topic, but using a shorter dataset, the authors found insignificant effects in men and women (Frijters, Haisken-DeNew, & Shields, 2003).

The second study using GSEOP data was primarily concerned with comparing a number of different econometric models to investigate the role of unobserved heterogeneity in the relationship between income and health, which was expected to vary over the life course (Jones & Schurer, 2009). Only results of the preferred model, the generalised conditional fixed effects logit, which accounted for unobserved heterogeneity, are discussed here.

The first results looked at the marginal effects of income on the probability of reporting very good health for each age group and by sex. No relationship between income and SRH was found at any age, and controlling for bias from unmeasured confounders was particularly important for middle to older age groups. Additional models found no association between increased income and changes into good or very good health, but a more specific association of increased income with exiting poor or bad health, for men and women, and over all age groups except the youngest (16–30 years).

The last study using GSOEP was primarily focused on the relationship between job satisfaction and various health outcomes, including SRH, but (logged) household income was included in the models (Fischer & Sousa-Poza, 2009). The dataset was restricted to employed participants between the ages of 16 and 60 years in an attempt to reduce impact of dissatisfied participants leaving jobs. However, this may have introduced selection bias into the analysis, as the effect of income on health may be more important for those who are not in the labour force. Two linear fixed effects models were used and both found null effects.

Health and Retirement Study

The Health and Retirement Study (HRS) is a biennial panel study from the US that began in 1992 with a sample of over 22,000 adults

over the age of 50 years, to study health and employment transitions in older people.

The HRS was used by one study (Frijters & Ulker, 2008), which presents, among others, one model relevant to this review, using conditional fixed effects ordinal logistic regression (Ferrer-i-Carbonell & Frijters, 2004). No substantial or significant association between (log) household income and SRH was found.

Household, Income and Labour Dynamics in Australia (HILDA) Survey

The Household, Income and Labour Dynamics in Australia (HILDA) Survey is a national survey that began in 2001 with 7683 households and a household initial response rate of 66%, with nearly 14,000 adults participating in the first wave. The attrition rate from waves one to five was 28% (Melbourne Institute of Applied Economic and Social Research, 2009).

A study using HILDA data investigated the causal relationship between poverty and health, focussing on disentangling the social causation and health selection pathways (Buddelmeyer & Cai, 2009). Two equation simultaneous probit models were used in the analysis, based on the assumption that poverty from the past one year would affect current health but current health status would affect current poverty status. Mean marginal effects from these models included that a family head had a 9% increased probability of poor health after a year in poverty compared to a family head who had not been in poverty. Families headed by a person in ill health had a 2% increased probability of being in poverty compared to families headed by a person in good health. When compared to separate health and poverty models, which did not account for health selection, different estimates were found. The conclusion was that causal effects occurred between poverty and health in both directions.

Discussion

This systematic review addressed the question of whether change in income predicted change in SRH in adults. Only a small number of studies met the inclusion criteria, as much research has not used repeated measures analyses. Two studies that are often cited as providing evidence for a causal relationship between income and health did not meet the criteria for inclusion in this review as they are based on cross-sectional surveys and/or used instrumental variables that are not robust exogenous substitutes for income change (Case, 2001; Ettner, 1996).

Overall, the majority of studies (ten of thirteen) found that income change had a small, statistically significant positive association with SRH, although it was not significant for women in three studies (Contoyannis et al., 2004; Frijters et al., 2005; Halliday, 2007). Of the studies that did not find a statistically significant result, one was on a sample of only employed individuals, which introduced selection issues (Fischer & Sousa-Poza, 2009); and two others found a significant association of income with mental health, but not SRH (Apouey & Clark, 2009; Frijters & Ulker, 2008). Average income had much greater impact on SRH than income change (Contoyannis et al., 2004; McDonough & Berglund, 2003), also found in the review by Benzeval (Benzeval & Judge, 2001), suggesting either that average income is more important for health than fluctuations in income (and less affected by measurement error, that may bias estimates of income change to the null), or that the average income–SRH association is subject to substantial residual confounding.

Statistically significant results are not always practically significant, if effect sizes are so small as to have little relevance when applied to real world situations. Most models used log transformed income, and a common finding from the BHPS was that a one unit increase in log income would be required to see a 1% increase in SRH (a small effect). The study of lottery wins (Apouey & Clark, 2009) raised the possibility that income works in different ways on mental health (positively) compared to health behaviours (negatively). One might hypothesize that an association with SRH is actually the net effect of these two opposing processes. Other studies of lottery winnings have investigated the impact of lottery wins on a composite health measure, mortality and mental health (Gardner & Oswald, 2007; Lindahl, 2005). These studies have found significant positive effects of the income windfall on health, with modest effect sizes.

One problem with studying income windfalls from natural experiments (such as lotteries) is that they may not be associated with the same health behaviours occurring upon the expectation of permanent increases in income, even within the same individuals. The separation of income into transitory (as exemplified by the income windfall) and permanent change components has been noted in the earnings dynamics literature (Meghir & Pistaferri, 2004). Natural experiments may be advantageous in terms of exogeneity but limited in terms of generalisability.

Although included studies attempted to minimise bias, none controlled for all types of systematic bias. Unobserved heterogeneity was eliminated with fixed effects methods or parameterised with correlated random effects models in over half of studies. Some attempt to account for health selection was made in about a third of studies. Investigation into attrition was undertaken in several studies but measurement error and how this might affect results was rarely mentioned. Many studies omitted one or more variables that are often considered important confounders of the income–health relationship, such as individual-level labour force status, marital status or family structure.

The major limitation of these studies was the general lack of consideration of measurement error as a probable source of bias, as both income and SRH may be affected by measurement error (Bound, Brown, Mathiowetz, & Leamer, 2001). This may cause either differential (i.e. measurement error of outcome depends on level of exposure, and vice versa) or non-differential measurement bias, leading to unpredictable effects on estimates (Hernán & Cole, 2009). Income estimates from fixed effects models, that deal well with the problem of confounding but less well with measurement error in repeated income measures, are likely to be biased towards the null (Deaton, 1995, chap. 33). None of the included studies account for correlation of income over time, although the literature on earning dynamics suggests this is another potential source of measurement error (Abowd & Card, 1989). The impact of measurement error in longitudinal studies of income and SRH must be minimised to achieve a better causal estimate of the income–health association. Related to this is the issue of power, especially for shorter panel studies using fixed effects models, which may have insufficient within-individual changes to demonstrate an association between income and health, causing a null result.

In any systematic review, publication bias is a potential problem. It is always possible that there are unpublished results on this subject that have not been included. However, given the small number of panel studies that could be used to answer the question, this review is likely to represent much of what is known at this time. It is possible, however, that null findings have not been published. Bias may have also been introduced by how studies were assessed for inclusion into the systematic review. At the initial review stage, when abstracts and titles were being appraised against the inclusion criteria, a study could have been missed because of a poorly worded abstract. However, after searching multiple databases and reference lists of many relevant articles, it is unlikely that a major study was overlooked.

Conclusions

Five longitudinal surveys from four different countries have been analysed in thirteen studies to give an overall conclusion that in most instances increases in income have a small positive association with SRH. Controlling for bias from unobserved heterogeneity and/or health selection reduced the association of income with SRH. However, none of these studies controlled for all potential systematic error, so this small effect may be affected by residual bias, particularly measurement error, which is likely to have biased the results to the null. The relatively small number of studies eligible for inclusion in this review, and the notable superiority of these studies over cross-sectional studies, suggests more research of this nature should be undertaken, using methods accounting for major sources of bias, particularly confounding, health selection and measurement error.

Acknowledgments

The Health Research Council of New Zealand funded the SoFIE-Health project (reference 08/048) and the PhD thesis for which this literature review was performed. Ken Richardson, Ivy Liu and Steve Stillman provided valuable assistance in interpreting methods. Publication was supported by the University of Otago Research Committee, by means of the University of Otago Postgraduate Publishing Bursary.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.socscimed.2010.10.029.

References

Abowd, J. M., & Card, D. (1989). On the covariance structure of earnings and hours changes. *Econometrica, 57*(2), 411–445.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: from supply and demand to natural experiments. *The Journal of Economic Perspectives, 15*(4), 69–85.

Apouey, B., & Clark, A. E. (2009). *Winning big but feeling no better? The effect of lottery prizes on physical and mental health*. PSE working paper 09. Paris: Paris School of Economics (Ecole normale supérieure).

Arellano, M., & Bond, S. (1991). Some test of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277–297.

Benzeval, M., & Judge, K. (2001). Income and health: the time dimension. *Social Science & Medicine, 52*(9), 1371–1390.

Bound, J., Brown, C., Mathiowetz, N., & Leamer, J. J. H. A. E. (2001). Measurement error in survey data. In J. Heckman, & E. E. Leamer (Eds.), *Handbook of econometrics* (pp. 3705–3843). Amsterdam: Elsevier.

Buddelmeyer, H., & Cai, L. (2009). *Interrelated dynamics of health and poverty in Australia*. Discussion paper series. IZA DP no. 4602. Bonn: Institute for the Study of Labor.

Case, A. (2001). *Does money protect health? Evidence from South African pensions*. NBER working papers no. 8495. Cambridge, Massachusetts: National Bureau of Economic Research, Inc (NBER).

Connor, J., Rodgers, A., & Priest, P. (1999). Randomised studies of income supplementation: a lost opportunity to assess health outcomes. *Journal of Epidemiology & Community Health, 53*, 725–730.

Contoyannis, P., Jones, A. M., & Rice, N. (2004). The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics, 19*(4), 473–503.

Deaton, A. (1995). Data and econometric tools for development analysis. In B. Jere, & T. N. Srinivasan (Eds.), *Handbook of development economics* (pp. 1785–1882). Elsevier.

Ettner, S. L. (1996). New evidence on the relationship between income and health. *Journal of Health Economics, 15*(1), 67–85.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal, 114*(497), 641–659.

Fischer, J. A. V., & Sousa-Poza, A. (2009). Does job satisfaction improve the health of workers? New evidence using panel data and objective measures of health. *Health Economics, 18*(1), 71–89.

Frijters, P., Haisken-DeNew, J. P., & Shields, M. (2003). *Estimating the causal effect of income on health: Evidence from post reunification East Germany* (pp. 1–36). Centre for Economic Policy Research, Research School of Social Sciences, Australian National University. Discussion papers 465.

Frijters, P., Haisken-DeNew, J. P., & Shields, M. A. (2005). The causal effect of income on health: evidence from German reunification. *Journal of Health Economics, 24*(5), 997–1017.

Frijters, P., & Ulker, A. (2008). Robustness in health research: do differences in health measures, techniques, and time frame matter? *Journal of Health Economics, 27*(6), 1626–1644.

Gardner, J., & Oswald, A. J. (2007). Money and mental wellbeing: a longitudinal study of medium-sized lottery wins. *Journal of Health Economics, 26*(1), 49–60.

Glymour, M. M. (2006). Natural experiments and instrumental variable analyses in social epidemiology. In J. M. Oakes, & J. S. Kaufman (Eds.), *Methods in social epidemiology* (pp. 429–460). San Francisco: Jossey-Bass.

Halliday, T. J. (2007). *Income volatility and health*. IZA discussion paper no. 3234. Bonn: Institute of the Study of Labor.

Hernán, M. A., & Cole, S. R. (2009). Invited commentary: causal diagrams and measurement bias. *American Journal of Epidemiology, 170*(8), 959–962.

Higgins, J., & Green, S. (2006). *The Cochrane handbook for systematic reviews of interventions 4.2.6* [Updated September 2006]. Available on-line at. Chichester: The Cochrane Collaboration. www.cochrane.org/resources/handbook/.

Imlach Gunasekara, F., Carter, K., & Blakely, T. (2008). Glossary for econometrics and epidemiology. *Journal of Epidemiology and Community Health, 62*, 858–861.

Jones, A. M., & Schurer, S. (2009). How does heterogeneity shape the socioeconomic gradient in health satisfaction? *Journal of Applied Econometrics*.

Jones, A. M., & Wildman, J. (2005). *Disentangling the relationship between health and income*. HEDG working papers, 05/07. York: Health, Econometrics and Data Group, University of York.

Jones, A. M., & Wildman, J. (2008). Health, income and relative deprivation: evidence from the BHPS. *Journal of Health Economics, 27*(2), 308–324.

Kawachi, I., Adler, N. E., & Dow, W. H. (2010). Money, schooling, and health: mechanisms and causal evidence. *Annals of the New York Academy of Sciences, 1186*, 56–68.

Lindahl, M. (2005). Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income. *Journal of Human Resources, 40*(1), 144–168.

Lorgelly, P. K., & Lindley, J. (2008). What is the relationship between income inequality and health? Evidence from the BHPS. *Health Economics, 17*(2), 249–265.

Ludbrook, A., & Porter, K. (2004). Do interventions to increase income improve the health of the poor in developed economies and are such policies cost effective? *Applied Health Economics and Health Policy, 3*(2), 115–120.

McDonough, P., & Berglund, P. (2003). Histories of poverty and self-rated health trajectories. *Journal of Health & Social Behavior, 44*(2), 198–214.

McDonough, P., Sacker, A., & Wiggins, R. D. (2005). Time on my side? Life course trajectories of poverty and health. *Social Science & Medicine, 61*(8), 1795–1808.

Meghir, C., & Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica, 72*(1), 1–32.

Melbourne Institute of Applied Economic and Social Research. (2009). *HILDA survey: Annual report 2008*. Victoria: Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Sacker, A., Wiggins, R., Bartley, M., & McDonough, P. (2007). Self-rated health trajectories in the United States and the United Kingdom: a comparative study. *American Journal of Public Health, 97*(5), 812–818.

Sanderson, S., Tatt, I. D., & Higgins, J. P. T. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *International Journal of Epidemiology, 36*(3), 666–676.

Uhrig, N. (2008). *The nature and causes of attrition in the British Household Panel Survey*. No. 2008-05. Essex: Institute for Social and Economic Research, University of Essex.

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gotzsche, P. C., Mulrow, C., Pocock, S., et al. (2007). Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *PLOS Medicine, 4*(10), e297.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – scope, evolution and enhancements. *Schmollers Jahrbuch, 127*(1), 139–169.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: MIT Press.