Social Science & Medicine 334 (2023) 116134

Contents lists available at ScienceDirect

Social Science & Medicine

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

Education and age trajectories of chronic conditions: Are tests of the cumulative advantage and disadvantage hypothesis biased by underreporting?

Liliya Leopold [a,*], Irene G.M. van Valkengoed [b], Henriette Engelhardt [c]

[a] Department of Sociology, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV, Amsterdam, the Netherlands
[b] Department of Public and Occupational Health, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam Public Health Research Institute, the Netherlands
[c] Department of Sociology, Professorship of Demography, University of Bamberg, Germany

ARTICLE INFO

Handling Editor: Social Epidemiology Office

Keywords:
Health inequality
Life course perspective
Hypertension
Diabetes
Cumulative advantage/disadvantage

ABSTRACT

Objective: This study examined the impact of underreporting on tests of the cumulative advantage and disadvantage hypothesis (CAD), which predicts age-related increases in health disparities between individuals with higher and lower education.

Methods: Using the English Longitudinal Study of Ageing (ELSA), we identified underreporting by comparing self-reported hypertension and diabetes with biomedically measured hypertension (systolic blood pressure≥140 mm Hg and/or diastolic blood pressure≥90 mm Hg) and diabetes (fasting glucose level≥7 mmol/l and/or HbA1c≥6.5%). In a sample of 11,859 respondents aged 50 to 85 (54% women, 97% White), we assessed the associations between underreporting and the main analytic constructs in tests of the CAD (education, age, sex, and cohort).

Results: The results showed that self-reported measures underestimated the prevalence of hypertension and diabetes. Underreporting showed weak to moderate associations with the main constructs in tests of the CAD, being more pronounced in individuals with lower education, in older age, in more recent cohorts, and among men. When correcting for underreporting using biomedical measures, the overall prevalence of hypertension and diabetes increased substantially, but education differences in age trajectories of both conditions remained similar.

Conclusions: Underreporting affected conclusions about the prevalence of hypertension and diabetes, but it did not affect conclusions about the CAD hypothesis for either condition.

1. Introduction

During the past decades, many studies have examined how the relationship between education and health changes with age. This research is guided by one of the dominant theoretical frameworks in life-course research – the cumulative advantage and disadvantage (CAD) hypothesis (Dannefer, 2003, 2020; DiPrete and Eirich, 2006; Ferraro et al., 2009). Specifically applied to the association between education and health, the CAD hypothesis asserts that disparities in health determinants, including living and working conditions, exposure to stress, social support, and health behaviors, accumulate gradually between individuals with different levels of education. As a result of these processes, the CAD hypothesis predicts an intracohort increase in health differences between people with lower and higher levels of education (Ross and Wu, 1996).

While many studies have supported this population-level prediction of the CAD (Chen et al., 2010; Kim, 2008; Leopold, 2016; Mirowsky and Ross, 2008; Willson et al., 2007; Yang et al., 2021), conflicting findings of stable or decreasing differences are commonly reported in the literature (House et al., 2005; Sieber et al., 2020; Zhao, 2023). This mixed evidence has fueled an intense debate centered around the question whether alternative patterns arise from methodological issues or reflect substantive processes that counteract the processes postulated by the CAD. Research has shown that patterns of stability or decrease may indeed result from methodological issues, such as confounding of age and cohort effects (Lynch, 2003) and selective attrition (Noymer, 2001).

* Corresponding author.
E-mail address: L.Leopold@uva.nl (L. Leopold).

https://doi.org/10.1016/j.socscimed.2023.116134
Received 15 June 2023; Received in revised form 14 July 2023; Accepted 28 July 2023
Available online 31 July 2023

However, even in studies that appropriately account for age and cohort effects, and address selective attrition, inconsistent patterns persistently emerge. For instance, different countries have displayed increasing, decreasing, or stable patterns in health disparities between educational groups (Leopold, 2018; Van Der Linden et al., 2020). Similarly, diverse patterns have been found among women and men, various ethnic groups, and different birth cohorts (Brown et al., 2012; Taylor et al., 2020; Zhao, 2023). Given that certain methodological concerns have been suitably addressed in these studies, the interpretation of mixed patterns has focused on substantive factors. Cumulative advantage and disadvantage processes in the relationship between education and health are increasingly recognized as malleable through policy interventions, institutional factors, or cultural variations that differ across countries, over time, and between social groups (Dannefer, 2020).

Despite this attention to substantive mechanisms, important methodological issues remain in the study of the CAD hypothesis. A key concern in the empirical tests of the CAD hypothesis addressed in the present study is that almost all previous studies are based on self-reported health measures such as self-rated health (SRH), self-reported functional limitations, activities of daily living (ADL), health-related quality of life (HRQL), and self-reports on chronic conditions (Chen et al., 2010; Dupre, 2007; House et al., 2005; Kim and Durden, 2007; Mirowsky and Ross, 2008; Leopold, 2018; van Kippersluis et al., 2009; Willson et al., 2007). Self-reported measures of health offer important benefits: Their collection is easy and cost-saving, they are commonly available in long-running population-based panel surveys, and they capture individuals' subjective perceptions of health and well-being – a central component in the definition of population health (Mojola et al., 2022; Zajacova et al., 2017; WHO, 1948). Although these benefits are widely recognized, studies have raised doubts about the validity of self-reported measures of health in tests of the CAD, as heterogeneous reporting bias was found in each of the key analytical constructs of the CAD, including education, age, cohort and gender (Molina, 2016; Ziebarth, 2010). Specifically, people with lower education, men, and those observed at older age tended to underreport their health problems (Dowd & Zajacova, 2007, 2010; Peersman et al., 2012). As a result, the assessment of health differences between people with higher and lower levels of education as well as the age trajectory of these differences might be biased by systematic underreporting, and conclusions about the CAD hypothesis may be sensitive to such biases.

Although these concerns have been raised in the field, empirical assessments remain scarce. To our knowledge, only one study has investigated the role of heterogeneous reporting in tests of the CAD (Leopold, 2019). This study compared self-reported measures of general health with maximum grip strength as an objective indicator and found entirely different life-course patterns, particularly among men. Although these results may point to bias in CAD studies based on self-reported health measures, subjective and objective health measures were not equivalent and it thus remained unclear whether the observed discrepancies were caused by heterogeneous reporting or whether they reflected different dimensions of health covered by the subjective and objective measures (Asada et al., 2020).

The purpose of the present study is to evaluate the impact of heterogeneous reporting on tests of the CAD hypothesis. To achieve this, we focused on health outcomes commonly utilized in CAD tests, which are available in longitudinal population-based surveys as both self-reports and conceptually equivalent objective measures. Specifically, we leveraged measures of chronic conditions. Unlike other self-reported measures of health for which objectively measured equivalents are not available, some chronic conditions are directly assessed in population surveys via biomedical tests. In the present study, we focused on hypertension and diabetes. Both conditions are prominent in tests of the CAD (Dupre, 2007, 2008; Leopold, 2018). Each represent highly prevalent causes of death and disability, and each are strongly associated with the lifelong accumulation of social and economic advantages and disadvantages as well as related behavioral patterns proposed as underlying mechanisms by the CAD (Sturm, 2002).

Similar to research on self-rated health and other self-reported measures of health, heterogeneous reporting – particularly underreporting of hypertension and diabetes – was found to be related to the main analytic categories used in the tests of the CAD. Male sex, older age, and lower socioeconomic status were all associated with underreporting of either condition (Okura et al., 2004; Ning et al., 2016; Singh et al., 2022).

The present study builds upon this research to assess how underreporting in self-reported hypertension and diabetes affected conclusions in tests of the CAD. To achieve this aim, we used the English Longitudinal Study of Ageing (ELSA), which allowed us (1) to assess the extent of underreporting in hypertension and diabetes and the degree to which it was related to the main analytic constructs in tests of the CAD, (2) to correct self-reports for underreporting using biomedical markers, and (3) to compare tests of the CAD with and without correction of potential underreporting bias.

2. Data and methods

The analyses presented in this manuscript are pre-registered at the Open Science Framework (OSF). The pre-registration can be found under the following link:https://osf.io/uazrx.

We used data from the English Longitudinal Study of Ageing (ELSA). ELSA is an ongoing longitudinal survey that collects population-based data on a wide range of social, psychological, and health outcomes from samples aged 50 and older living in private households (Banks et al., 2021). Refreshment samples of people recruited in their early 50s were added at waves 3, 4, 6, 7, and 9.

ELSA data have been widely used in studies on the prevalence of chronic conditions (Hamer et al., 2018) and in tests of the CAD hypothesis (Leopold, 2018; Wetzel and Vanhoutte, 2020). For the purposes of the present study, ELSA data are particularly suitable, as they include longitudinal information not only on self-reported, but also on objectively measured chronic conditions.

2.1. Sample

Data on objectively measured chronic conditions were collected in waves 2, 4, 6, 8, and 9. Data on self-reported chronic conditions were collected during the main face-to-face interviews, while data on biomedical measures were collected during follow-up visits by a nurse. Biomedical measures were available for a subset of the face-to-face sample. Because not all eligible respondents participated in the nurse visit, we followed the recommendation of the ELSA user guide and applied baseline survey weights (wave 2, nurse sample) in our descriptive and multivariate analyses (NatCen Social Research, 2022) to correct for self-selection into the nurse sample. Due to technical issues, samples for the analyses of hypertension and diabetes were not identical. 2714 measurements of blood pressure (approximately 7%) were invalid and 7790 measurements (approximately 22%) of blood sugar level were not available either due to technical issues during laboratory analyses, because no blood samples were taken or due to respondents not fulfilling the fasting requirement for tests of the fasting glucose levels. The overall prevalence (i.e., self-reports and biomedical measurements) of hypertension and diabetes were within the range reported in other studies on similar populations in the UK (Hamer et al., 2018; Pan et al., 2020; Whicher et al., 2020). This suggests that ELSA data on self-reported and objectively measured chronic conditions comply with the standards in the field and are suitable for the purposes of the present study.

We restricted the sample to respondents who participated in a nurse visit at least once, and who were at least 50 years old (younger respondents were outside of the sampling frame) but younger than 86 at the time of interview. Finally, we excluded respondents with missing information on education. A flowchart of sample exclusions is shown in

Table A1 in the Appendix. Descriptive statistics are presented in Table 1.

2.2. Measure of education

Following previous studies of CAD using ELSA data (Leopold, 2018; Wetzel and Vanhoutte, 2020), we measured education categorically, drawing on information about respondents' highest levels of education and respondents' age at finishing education. We distinguished between lower, intermediate, and higher levels of education. The bottom category of lower education included respondents below the upper secondary level and/or having finished education before the age of 16; intermediate education consisted of respondents holding upper secondary education or vocational degrees and/or those who were 16–18 years old when finishing their education; the top category of higher education included respondents with a BA degree or higher levels of tertiary education, or in cases of missing data, those who finished education at age 19 or later. Our results were robust to alternative specifications of education categories.

2.3. Age, cohort, and sex

Age was measured in years, ranging from 50 to 85 and centered at the mean of 62 years in the multivariate models. Birth cohort was measured in years ranging from 1939 to 1954 and centered at the mean of 1944 in the multivariate models. Model fit analyses in which various polynomials of age and cohort terms as well as interactions between these terms were tested are reported below. All analyses were run separately for men and women. Information about gender was collected via self-reports. As a robustness check, we have also investigated alternative age and cohort ranges. Inclusion of the full age (50–90) and cohort (1908–1967) ranges did not affect the main conclusions regarding the role of underreporting in tests of the CAD.

2.4. Measures of chronic conditions

We assessed *self-reported hypertension* using survey questions about whether a doctor or a nurse ever told the respondents they had high blood pressure (HBP) and whether they took medication against hypertension. We assessed *biomedically measured hypertension* using data from the nurse visits. The nurse measured systolic and diastolic blood pressure three times with an interval of at least 1 min between the measurements. The average of valid measurements was then used to identify hypertension, clinically defined as systolic blood pressure≥140 mm Hg and/or diastolic blood pressure≥90 mm Hg. HBP is the main diagnostic criterion for hypertension.

For the purposes of our study, we used the biomedical measures to correct our self-reported measures for underreporting. In our main analyses, we directly compared the measure of hypertension based on self-reports only, which was not corrected for underreporting (*uncorrected HBP*), which took value one if respondents reported having been diagnosed with hypertension and/or taking an anti-hypertension medication and value zero otherwise, with the measure of hypertension based on self-reports as well as biomedical measurements, thus, corrected for underreporting (*corrected HBP*), which took value one if respondents either reported having been diagnosed with hypertension and/or taking an anti-hypertension medication and/or had HBP according to the biomedical measurements and value zero otherwise. The results in Table 1 show that the prevalence of HBP increased from 45% when based on self-reports only to 65% when corrected for underreporting using biomedical measures.

We assessed *self-reported diabetes* using survey questions about whether respondents had ever been diagnosed with diabetes by a healthcare professional and whether they took medication for diabetes (e.g., insulin injections or tablets). We assessed *biomedically measured diabetes* using data on levels of fasting blood glucose and HbA1c levels collected from the blood samples taken during the nurse visits. Fasting blood glucose level is the amount of glucose present in the blood after an overnight fast. HbA1c is a measure of the average level of glucose in the blood over the past 2–3 months. In line with the main clinical criteria for diabetes diagnosis, we identified diabetes as fasting glucose levels≥7 mmol/l and/or HbA1c levels≥6.5%.

We used data from these biomedical measures in order to correct our self-reported measure of diabetes for underreporting. In our main analyses, we directly compared the measure of diabetes based on self-reports, thus not corrected for underreporting (*uncorrected diabetes*), which took a value one if respondents reported having been diagnosed with diabetes and/or taking an anti-diabetes medication and a value zero otherwise, with the measure of diabetes based on self-reports as well as biomedical measurements, thus corrected for underreporting (*corrected diabetes*), which took a value one if respondents either reported having been diagnosed with diabetes and/or taking an anti-diabetes medication and/or had diabetes according to the biomedical measurements and a value zero otherwise. The results in Table 1 show that the prevalence of diabetes was 14% when based on self-reports only and 16% when based on self-reports and biomedical measures.

The associations of underreporting in hypertension and diabetes with the key constructs in tests of the CAD –education, age, cohort, and sex – were assessed in descriptive and multivariate analyses summarized in Figures A1 and A2, and in Tables A2 and A3. Underreporting in HBP and in diabetes was more common at older ages, in more recent cohorts, among men, and among the lower-educated. Education differences in underreporting were small and age patterns of underreporting did not differ between education groups. The only exception was hypertension among women, where the age-related increase in underreporting was slightly more pronounced among the higher educated (Table A3).

Although these results are largely in line with previous research, education was less strongly related to underreporting than found in other studies (Wu et al., 2019). In our sample education differences in underreporting of both chronic conditions were small and not statistically significant for HBP among men (Table A3). Additional analyses showed that these results were robust to sample selection procedures, similar across panel waves, and did not vary depending on the categorization of education.

Table 1
Descriptive statistics.

	Mean	Std. Dev.	Min	Max
Age	69.83	8.25	52	85
Sex				
Male	.47			
Female	.53			
Race				
White	.97			
Non-white	.03			
Year of birth	1940	7.919	1920	1952
Education				
Higher	.14			
Intermediate	.28			
Lower	.57			
Chronic conditions				
Uncorrected HBP	.49			
Corrected HBP	.69			
Uncorrected diabetes	.15			
Corrected diabetes	.17			
Year of the interview	2010	4.82	2004	2019
Duration of participation in years	11.82	4.56	1	16
Dropped out until wave 8	.63			
Died	.02			

Source: ELSA, release 2022, waves 2, 4, 6, 8 & 9, own calculations with weighted data. Note: Descriptive statistics were calculated over the number of observations in the weighted sample used for hypertension analyses. Descriptive statistics were similar when calculated over the number of observations in the weighted sample for diabetes. Descriptive statistics on unweighted date are shown in Table A4 in the appendix.

2.5. Attrition

If those who are more likely to underreport chronic conditions are also more likely to drop out of the panel, the degree to which underreporting is related to age might be underestimated. To assess this potential source of bias, we examined whether dropping out, due to either attrition or death, before the last wave was related to the probability of underreporting while controlling for other variables included in tests of the CAD hypothesis (education, age, and cohort). These analyses showed no associations between underreporting and subsequent dropout. The only exception was diabetes among women, where those who underreported diabetes were slightly more likely to drop out (Table A3).

2.6. Analytic strategy

The impact of underreporting on tests of the CAD hypothesis was assessed by comparing the results in tests of the CAD based on self-reports only, thus, not corrected for underreporting, with those *based on self-reports as well as biomedical measures*, thus, corrected for underreporting. In our investigation of the CAD hypothesis, we employed cohort-sequential hierarchical growth curve models, which have been widely used in previous studies exploring the CAD hypothesis (Yang et al., 2021; Willson et al., 2007). These models offer a robust approach to estimating age trajectories by leveraging within-person variation while capturing cohort and education effects through between-person variation in health levels and age trajectories. Additionally, the inclusion of individual random slopes in these models allowed us to account for panel attrition, as highlighted by Lynch (2003).

Data used in this type of research must meet specific criteria, with the length of observation periods and the range of birth cohorts being particularly important factors (Lynch, 2003). ELSA fulfills these requirements, as the data cover a wide range of birth cohorts observed across long periods. Based on this data structure, we used hierarchical linear probability regression models (HLPM), simultaneously estimating within-person change with age, between-person differences across cohorts and education levels, and interactions between these parameters. All analyses were performed separately for men and women.

The parametrizations of age and cohort effects as well as their interactions with education on each of the outcomes were based on three criteria, (a) similarity between observed and fitted data examined by diagnostic plots, (b) BIC, and (c) model parsimony if models were similar on criterion (a) and did not differ by more than 10 BIC points (Raftery, 1995). We tested for different functional forms of age and cohort, including up to cubic terms for each as well as categorical specifications. We also tested for various possible interactions between age, cohort, and education. All model fit analyses were run separately for women and men. Different functional forms of age and cohort and interactions between these terms and education provided the best model fit for both hypertension and diabetes (see Tables 2 and 3 for details on the specification). Differences in estimates between these models and alternative specifications were minimal and all conclusions regarding substantive tests of the CAD hypothesis and the role of underreporting bias were robust across model specifications.

3. Results

The results of multivariate models testing the CAD hypothesis are presented in Table 2 (for hypertension) and Table 3 (for diabetes) and visualized in Fig. 1 (for hypertension) and Fig. 2 (for diabetes). Figs. 1 and 2 compare the general patterns of results between uncorrected and corrected measures. We mainly focus on these general patterns in our presentation and interpretation of results.

In both figures, the y-axes show predicted probabilities for the prevalence of each outcome. The gray horizontal reference lines ease comparison of results across panels. The curves show age trajectories of hypertension or diabetes respectively for different birth cohorts and education levels. The cohort term was fixed at four values pertaining to those born in 1950 (observed from age 53 to 68), 1940 (observed from age 63 to 77), and 1930 (observed from age 74 to 85). Dashed curves represent age trajectories for lower levels of education, while solid curves represent age trajectories for higher levels of education. Red curves pertain to *uncorrected* measures, black curves to *corrected*

Table 2
Hierarchical Linear Regression Models for Change in High Blood Pressure (HBP) for men and women.

	M1a	M1b	M1c	M1d
	Uncorrected HBP Women	Corrected HBP Women	Uncorrected HBP Men	Corrected HBP Men
Age	0.017***	0.021***	0.017***	0.019***
	[0.015,0.019]	[0.019,0.023]	[0.015,0.020]	[0.016,0.021]
Age^2	0.000	−0.000***	−0.000	−0.000***
	[-0.000,0.000]	[-0.000,-0.000]	[-0.000,0.000]	[-0.000,-0.000]
Cohort	0.002	0.005*	0.008**	0.007**
	[-0.002,0.006]	[0.001,0.009]	[0.003,0.012]	[0.002,0.011]
$Cohort^2$	−0.00**	−0.000	−0.000	−0.000
	[-0.0001,-0.000]	[-0.000,0.000]	[-0.000,0.000]	[-0.000,0.000]
Cohort * Age	0.000	0.000	0.000	0.000
	[0.000,0.000]	[0.000,0.000]	[0.000,0.000]	[0.000,0.000]
Higher educ.	−0.070*	−0.106***	−0.077**	−0.059*
	[-0.124,-0.016]	[-0.159,-0.052]	[-0.126,-0.028]	[-0.106,-0.013]
Intermediate educ.	−0.044*	−0.064***	−0.067**	−0.073***
	[-0.082,-0.007]	[-0.100,-0.027]	[-0.111,-0.022]	[-0.115,-0.030]
Higher educ. * Age	−0.003	−0.002	−0.001	−0.000
	[-0.008,0.001]	[-0.006,0.003]	[-0.005,0.003]	[-0.004,0.004]
Intermediate educ. *	−0.003*	−0.001	−0.002	−0.001
Age	[-0.006,-0.000]	[-0.004,0.002]	[-0.006,0.002]	[-0.004,0.003]
Higher educ. * Cohort	−0.004	−0.004	−0.002	−0.003
	[-0.012,0.004]	[-0.012,0.003]	[-0.009,0.005]	[-0.009,0.003]
Intermediate educ. *	−0.002	−0.001	−0.001	−0.002
Cohort	[-0.007,0.003]	[-0.006,0.003]	[-0.007,0.005]	[-0.008,0.003]
Intercept	0.471***	0.660***	0.515***	0.706***
	[0.443,0.480]	[0.626,0.662]	[0.443,0.480]	[0.626,0.662]
N	10671	10671	8796	8796

Note. Data from ELSA, waves 2, 4, 6, 8 & 9. Age, Age2, Cohort and Cohort2 were centered at means, 95% confidence intervals in brackets, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, Source: ELSA, release 2022, waves 2, 4, 6, 8 & 9, own calculations with weighted data.

Table 3
Hierarchical Linear Regression Models for Change in Diabetes for men and women.

	M2a	M2b	M2c	M2d
	Uncorrected Diabetes Women	Corrected Diabetes Women	Uncorrected Diabetes Men	Corrected Diabetes Men
Age	0.010*** [0.008,0.012]	0.012*** [0.010,0.014]	0.014*** [0.011,0.016]	0.016*** [0.014,0.019]
Cohort * Age	0.006*** [0.003,0.008]	0.007*** [0.005,0.010]	0.009*** [0.006,0.012]	0.012*** [0.009,0.016]
Higher educ.	−0.079*** [-0.106,-0.051]	−0.085*** [-0.114,-0.055]	−0.061*** [-0.096,-0.027]	−0.070*** [-0.106,-0.033]
Intermediate educ.	−0.050*** [-0.072,-0.028]	−0.050*** [-0.074,-0.027]	−0.054*** [-0.085,-0.022]	−0.063*** [-0.096,-0.029]
Higher educ. * Age	−0.005** [-0.009,-0.002]	−0.005** [-0.009,-0.002]	−0.007*** [-0.011,-0.004]	−0.008*** [-0.012,-0.004]
Intermediate educ. *	−0.003**	−0.002	−0.005**	−0.005**
Age	[-0.006,-0.001]	[-0.005,0.001]	[-0.009,-0.002]	[-0.009,-0.002]
Higher educ. * Cohort	−0.003 [-0.008,0.002]	−0.003 [-0.008,0.002]	−0.009** [-0.015,-0.003]	−0.010*** [-0.016,-0.004]
Intermediate educ. *	−0.002	−0.000	−0.003	−0.004
Cohort	[-0.006,0.001]	[-0.004,0.003]	[-0.008,0.001]	[-0.010,0.001]
Intercept	0.129*** [0.113,0.144]	0.145*** [0.129,0.162]	0.184*** [0.163,0.205]	0.209*** [0.188,0.231]
N	8229	8229	7023	7023

Note. Data from ELSA, waves 2, 4, 6, 8 & 9. Age, Age2, Cohort and Cohort2 were centered at means, 95% confidence intervals in brackets, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; Source: ELSA, release 2022, waves 2, 4, 6, 8 & 9, own calculations with weighted data.

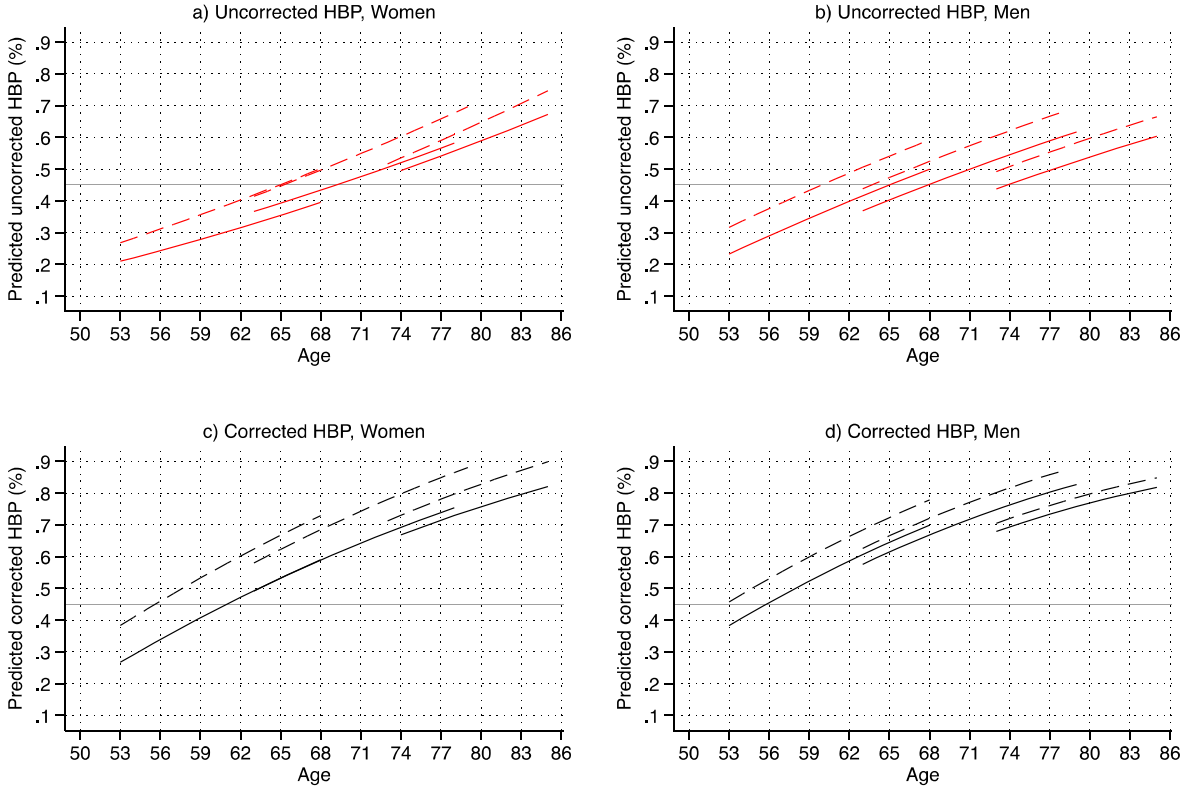

Fig. 1. Model-based prevalence of high blood pressure by education, age, cohort, and sex.
Legend: *Red lines* display the predicted prevalence of uncorrected HBP, thus based on self-reports only (see models M1a, M1c, Table 1). *Black lines* display the predicted prevalence of corrected HPB, thus based on self-reports and biomedical measurements (see models M1b, M1d, Table 1). In each panel, *dashed lines* represent lower levels of education and *solid lines* represent higher levels of education. For visualization, in each panel, the *cohort term* was fixed on four values, representing birth years 1930, 1940, and 1950. Source: ELSA, waves 2, 4, 6, 8 & 9, own calculations with weighted data. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

measures.

3.1. CAD and underreporting bias in hypertension

Fig. 1 shows that the predicted prevalence of hypertension *uncorrected* for underreporting was higher at lower levels of education at the age of initial observation in each cohort among women and men in each of the birth cohorts. Among women, education differences increased slightly with age. This finding is based on a small but statistically significant negative interaction between the age terms and higher education (Table 2, M1a). Among men, education differences in uncorrected HBP remained constant with age in each birth cohort while education differences in uncorrected HBP increased slightly across cohorts.

When *corrected* for underreporting bias using biomedically measured HBP, the prevalence levels of hypertension rose substantially in each of the cohorts and education groups both among women and men, and increasingly so with age. For example, the underestimate of the prevalence of hypertension increased in each comparison group from approximately 10 percentage points at age 50 to approximately 20

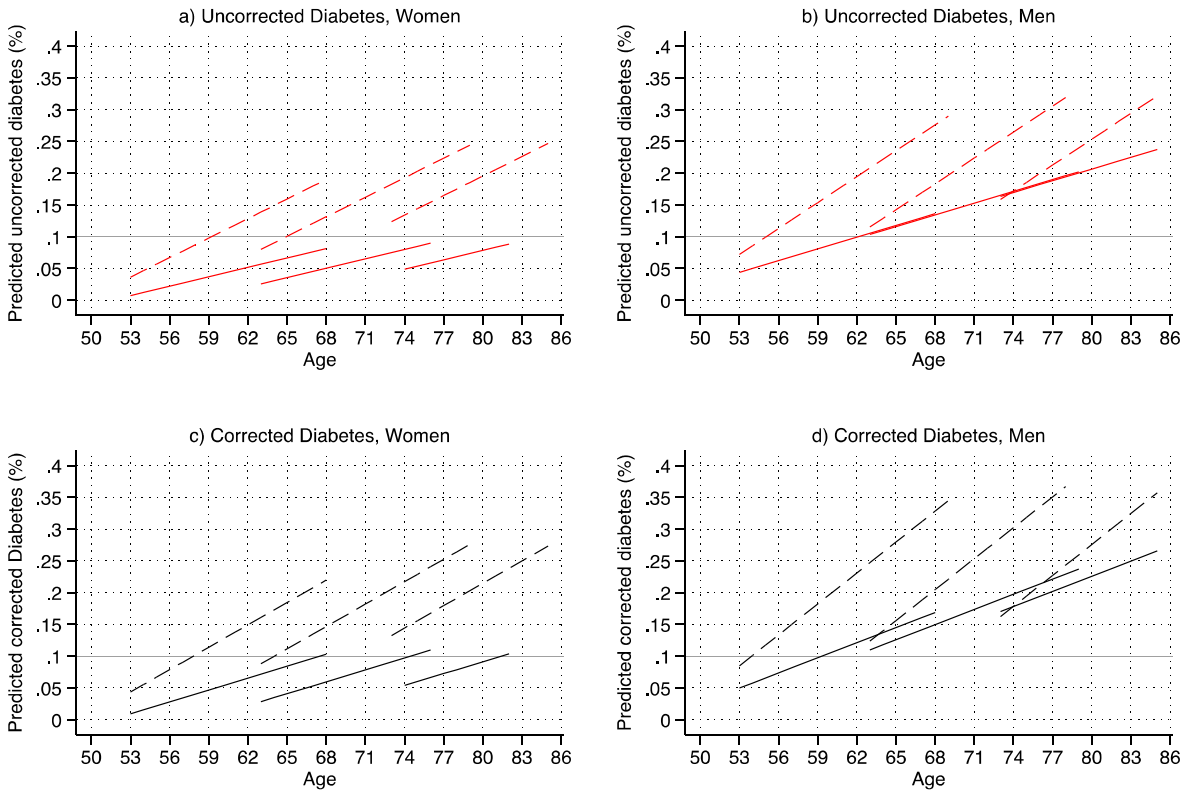

Fig. 2. Model-based prevalence of diabetes by education, age, cohort, and gender.
Legend: *Red lines* display the predicted prevalence of uncorrected diabetes, thus based on self-reports (see models M2a, M2c, Table 2). *Black lines* display the predicted prevalence of corrected diabetes, thus based on self-reports and biomedical measurements (see models M2b, M2d, Table 2). In each panel, *dashed lines* represent lower levels of education and *solid lines* represent higher levels of education. For visualization, in each panel, the *cohort term* was fixed on four values, representing birth years 1930, 1940, and 1950. Source: ELSA, waves 2, 4, 6, 8 & 9, own calculations with weighted data. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

percentage points at age 70.

Conclusions about the CAD hypothesis remained unchanged among men and changed slightly among women when correcting for underreporting. As visible from the comparison between panels a and c in Fig. 1 and models M1a and M1b in Table 2, the slight increase in education differences with age observed among women in uncorrected HPB vanished in models of corrected HBP.

3.2. CAD and underreporting bias in diabetes

As shown in Fig. 2, the predicted probability of diabetes *uncorrected* for underreporting was higher among men as compared to women and increased with age among higher-educated as well as lower-educated women and men. The age-related increase in the predicted probability of diabetes was much steeper among lower-educated women and men than among their higher-educated counterparts. This pronounced age increase in education differences was consistent with the CAD hypothesis. It did not differ substantially across cohorts among women, but it intensified across cohorts among men.

When *corrected* for underreporting bias, the predicted probabilities of diabetes increased even more steeply with age, although this difference in age slopes was modest in size. The patterns of education differences remained largely unchanged (see also the nearly identical coefficients of the education and age terms and the interaction terms in Table 2, M2 and M4).

4. Discussion

The CAD hypothesis predicts an increase in health disparities between individuals with higher and lower levels of education. Empirical tests of this hypothesis have yielded mixed results, with patterns of stability or decrease also being commonly observed. The nature of these mixed findings has sparked a debate. Methodological factors have played a significant role in explaining the inconsistencies. Well-known methodological issues associated with testing the CAD hypothesis are confounding of age and cohort effects and selective attrition. Recent research has examined the potential impact of heterogeneous reporting as a source of the mixed evidence (Leopold, 2019). The present study contributes to this recent line of investigation by examining the influence of underreporting on tests of the CAD hypothesis in the context of self-reported chronic conditions, specifically hypertension and diabetes.

The findings showed that the prevalence of both conditions was substantially underestimated in self-reported measures, increasingly so with age, and to a larger extent among men. Underreporting by lower-educated respondents was only slightly more pronounced than among those with higher levels of education. Overall, conclusions about the CAD hypothesis regarding education differences in hypertension and diabetes were not substantially biased by underreporting. Even after correcting for underreporting bias using biomedical measurements, we found large educational differences in hypertension and diabetes. For diabetes, these differences increased strongly with age, supporting the CAD hypothesis. For hypertension, these differences remained largely stable with age, contradicting the CAD hypothesis.

For the context of the present study, we conclude that underreporting does not constitute an important source of bias or a potential reason for mixed findings in tests of the CAD hypothesis. As the main source of underreporting of chronic conditions is a lack of awareness of having a disease (Singh et al., 2022), the relatively small education differences in underreporting found in the present study may be due to the favorable health care context of the UK, where access to high-quality

health care is universal (Zhou et al., 2019) and health care utilization differs little by education (Thomson et al., 2018). In other health care contexts, underreporting in chronic conditions might be a more serious source of bias. For instance, studies based on Chinese CHARLS data reported a strong association between underreporting and socioeconomic status (Ning et al., 2016), possibly reflecting more stratified access to and utilization of health care. Future studies should investigate the extent to which underreporting influences conclusions about the CAD hypothesis in institutional and cultural contexts other than the UK.

Next to highlighting the potential role of underreporting, our results have implications for interpretating the nature of mixed evidence in tests of the CAD hypothesis. Previous explanations mainly targeted variations between population groups (cohorts, genders, ethnic groups, countries) and focused on institutional and cultural factors. The present study found mixed patterns within the same population but in different domains of health. These results suggest that differences in conclusions regarding the CAD hypothesis between the two chronic conditions are due to differences in substantive mechanisms underlying diabetes and hypertension. When comparing diabetes and hypertension, there are similarities but also important differences in known risk factors. Both chronic conditions require a genetic predisposition and react to diet and physical activity. Yet, the risk of developing diabetes is more strongly related to diet, physical activity, and obesity whereas the risk of developing hypertension is more strongly related to alcohol consumption, smoking, and stress (Cuffee et al., 2014; Messerli et al., 2007). It is possible that in the UK context, alcohol consumption, smoking or other risk factors that are more strongly linked to hypertension are less stratified by education than diet, physical activity, and obesity. In addition, risk factors common to both chronic conditions may lead more often and at an earlier age to hypertension than to diabetes. The increase in education differences in the prevalence of hypertension takes place much earlier in life as compared to diabetes. In older age, ceiling effects may prevent education differences in hypertension from increasing further.

Future research on age change in education differences in diabetes and hypertension should address differences in such underlying mechanisms to understand why the pattern predicted by the CAD is present in diabetes but not in hypertension. At a more general level, our findings suggest that predictions of the CAD with regard to the underlying processes of accumulation of health-related advantages and disadvantages would benefit from being tailored to specific conditions or domains of health.

Although ELSA provided the highest-quality data for the purposes of the present study, our correction for underreporting bias has limitations. First, our biomedical tests for hypertension were taken on only one occasion. For a clinical diagnosis, high blood pressure should be established on several occasions. It is unclear whether measuring hypertension on one occasion only led to an over- or underestimation of overall prevalence. Second, it is important to note that data on blood tests were not available for the entire sample of individuals who participated in nurse visits. Although the proportion of missing information resulting from absent blood tests was relatively low, missing data were more prevalent among respondents with lower levels of education. Consequently, our analysis may slightly underestimate disparities in underreporting of diabetes related to education. Moreover, respondents for whom valid data on blood tests were available were somewhat selective. In particular, self-reported diabetes was less prevalent among those who fulfilled the fasting requirements for the tests of fasting blood glucose levels. It is thus possible that underreporting of diabetes based on fasting blood glucose levels was underestimated in the ELSA data.

While the results of the present study suggest that underreporting bias in not a major reason behind the mixed evidence offered by tests of the CAD hypothesis for chronic conditions, more research is needed to address the role of heterogeneous reporting biases in other health measures. Our findings may not extend to more general and subjective measures of population health, such as self-rated health, health-related quality of life (HRQL), or measures of limitations in physical or mental functioning (e.g., ADL). As self-reported chronic conditions capture recall of very specific health issues as diagnosed by a medical practitioner, little room is left for reporting biases related to the interpretation of the survey questions or response categories (Zajacova et al., 2017). Several quantitative and qualitative studies have shown that such reporting issues are related to education, age, cohort, and sex (Altman et al., 2016; Krause and Jay, 1994; Schnittker and Bacak, 2014; Zajacova et al., 2017). Longitudinal surveys similar to ELSA include a wide range of subjective reports and objective tests of physical impairment (e.g., self-reports and objective tests of grip strength, chair raise, leg raise, walkability, etc.), which can be used in order to understand the extent to which the results in tests of the CAD are driven by differences in interpretations and response styles – as opposed to substantive mechanisms of accumulation of health-related advantages and disadvantages.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data availability

ELSA data are available from the provider upon registration. The code used for the present study will be made available on the OSF and on the personal Webpage for the main authror.

Acknowledgements

This research was funded by the German Research Foundation, Grant Nr. EN 424/12–1. Health Measures and Health Inequality Over the Life Course.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2023.116134.

References

Altman, C.E., Van Hook, J., Hillemeier, M., 2016. What does self-rated health mean? Changes and variations in the association of obesity with objective and subjective components of self-rated health. J. Health Soc. Behav. 57 (1), 39–58. https://doi.org/10.1177/0022146515626218.

Asada, Y., Grignon, M., Hurley, J., Kirkland, S., 2020. Cautionary tails of grip strength in health inequality studies: an analysis from the Canadian longitudinal study on aging. Soc. Sci. Med. 265, 113382 https://doi.org/10.1016/j.socscimed.2020.113382.

Banks, J., Batty, D.G., Breedvelt, J., Coughlin, K., Crawford, R., Marmot, M., Nazroo, J., Oldfield, Z., Steel, N., Steptoe, A., Wood, M., Zaninotto, P., 2021. English Longitudinal Study of Ageing: Waves 0-9, 1998-2019 [data collection], 37th Edition. UK Data Service. https://doi.org/10.5255/UKDA-SN-5050-24. SN: 5050.

Brown, T.H., O'Rand, A.M., Adkins, D.E., 2012. Race-ethnicity and health trajectories: tests of three hypotheses across multiple groups and health outcomes. J. Health Soc. Behav. 53 (3), 359–377. https://doi.org/10.1177/0022146512455333.

Chen, F., Yang, Y., Liu, G., 2010. Social change and socioeconomic disparities in health over the life course in China: a cohort analysis. Am. Socio. Rev. 75 (1), 126–150. https://doi.org/10.1177/0003122409359165.

Cuffee, Y., Ogedegbe, C., Williams, N.J., Ogedegbe, G., Schoenthaler, A., 2014. Psychosocial risk factors for hypertension: an update of the literature. Curr. Hypertens. Rep. 16, 1–11. https://doi.org/10.1007/s11906-014-0483-3.

Dannefer, D., 2003. Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory. J. Gerontol.: Ser. Bibliogr. 58 (6), 327–337. https://doi.org/10.1093/geronb/58.6.S327.

Dannefer, D., 2020. Systemic and reflexive: foundations of cumulative dis/advantage and life-course processes. J. Gerontol.: Ser. Bibliogr. 75 (6), 1249–1263. https://doi.org/10.1093/geronb/gby118.

DiPrete, T.A., Eirich, G.M., 2006. Cumulative advantage as a mechanism for inequality: a review of theoretical and empirical developments. Annu. Rev. Sociol. 32 (1), 271–297. https://doi.org/10.1146/annurev.soc.32.061604.123127.

Dowd, J.B., Zajacova, A., 2007. Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? Int. J. Epidemiol. 36 (6), 1214–1221. https://doi.org/10.1093/ije/dym214.
Dowd, J.B., Zajacova, A., 2010. Does self-rated health mean the same thing across socioeconomic groups? Evidence from biomarker data. Ann. Epidemiol. 20 (10), 743–749. https://doi.org/10.1016/j.annepidem.2010.06.007.
Dupre, M.E., 2007. Education differences in age-related patterns of disease: reconsidering the cumulative disadvantage and age-as-leveler hypotheses. J. Health Soc. Behav. 48 (1), 1–15. https://doi.org/10.1177/002214650704800101.
Dupre, M.E., 2008. Educational differences in health risks and illness over the life course: a test of cumulative disadvantage theory. Soc. Sci. Res. 37 (4), 1253–1266. https://doi.org/10.1016/j.ssresearch.2008.05.007.
Ferraro, K.F., Shippee, P.T., Schafer, M.H., 2009. Cumulative inequality theory for research on aging and the life course. In: Bengtson, V.L., Gans, D., Putney, N.M., Silverstein, M. (Eds.), Handbook of Theories of Aging, second ed. Springer Publishing Company, New York, pp. 413–433.
Hamer, M., Terrera, G.M., Demakakos, P., 2018. Physical activity and trajectories in cognitive function: English Longitudinal Study of Ageing. J. Epidemiol. Community Health 72 (6), 477–483. https://doi.org/10.1136/jech-2017-210228.
House, J.S., Lantz, P.M., Herd, P., 2005. Continuity and change in the social stratification of aging and health over the life course: evidence from a nationally representative longitudinal study from 1986 to 2001/2002 (Americans' changing lives study). J. Gerontol. B Psychol. Sci. Soc. Sci. 60 (Ii), 15–26. https://doi.org/10.1093/geronb/60.special_issue_2.s15.
Kim, J., 2008. Intercohort trends in the relationship between education and health: examining physical impairment and depressive symptomatology. J. Aging Health 20 (6), 671–693. https://doi.org/10.1177/0898264308321004.
Kim, J., Durden, E., 2007. Socioeconomic status and age trajectories of health. Soc. Sci. Med. 65 (12), 2489–2502. https://doi.org/10.1016/j.socscimed.2007.07.022.
Krause, N.M., Jay, G.M., 1994. What do global self-rated health items measure? Med. Care 32 (9), 930–942. https://doi.org/10.1097/00005650-199409000-00004.
Leopold, Liliya., 2016. Cumulative advantage in an egalitarian country? Socioeconomic health disparities over the life course in Sweden. J. Health Soc. Behav. 57 (2), 257–273.
Leopold, L., 2018. Education and physical health trajectories in later life: a comparative study. Demography 55 (3), 901–927. https://doi.org/10.1007/s13524-018-0674-7.
Leopold, L., 2019. Health measurement and health inequality over the life course: a comparison of self-rated health, SF-12, and grip strength. Demography 56 (2), 763–784. https://doi.org/10.1007/s13524-019-00761-x.
Lynch, S.M., 2003. Cohort and life-course patterns in the relationship between education and health: a hierarchical approach. Demography 40 (2), 309–331. https://doi.org/10.1353/dem.2003.0016.
Messerli, F.H., Williams, B., Ritz, E., 2007. Essential hypertension. Lancet 370 (9587), 591–603. https://doi.org/10.1016/S0140-6736(07)61299-9.
Mirowsky, J., Ross, C.E., 2008. Education and self-rated health: cumulative advantage and its rising importance. Res. Aging 30 (1), 93–122. https://doi.org/10.1177/0164027507309649.
Mojola, S.A., Ice, E., Schatz, E., Angotti, N., Houle, B., Gómez-Olivé, F.X., 2022. The meaning of health in rural South Africa: gender, the life course, and the socioepidemiological context. Popul. Dev. Rev. 48, 1061–1095. https://doi.org/10.1111/padr.12494.
Molina, T., 2016. Reporting heterogeneity and health disparities across gender and education levels: evidence from four countries. Demography 53 (2), 295–323. https://doi.org/10.1007/s13524-016-0456-z.
NatCen Social Research, 2022. User Guide to the Nurse Visit Datasets (Waves 2, 4, 6, 8 and 9).
Ning, M., Zhang, Q., Yang, M., 2016. Comparison of self-reported and biomedical data on hypertension and diabetes: findings from the China Health and Retirement Longitudinal Study (CHARLS). BMJ Open 6 (1), e009836. https://doi.org/10.1136/bmjopen-2015-009836.
Noymer, Andrew., 2001. Mortality Selection and Sample Selection: A Comment on Beckett. J. Health Soc. Behav. 42 (3), 326–327.
Okura, Y., Urban, L.H., Mahoney, D.W., Jacobsen, S.J., Rodeheffer, R.J., 2004. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J. Clin. Epidemiol. 57 (10), 1096–1103. https://doi.org/10.1016/j.jclinepi.2004.04.005.
Pan, H., Hibino, M., Kobeissi, E., Aune, D., 2020. Blood pressure, hypertension and the risk of sudden cardiac death: a systematic review and meta-analysis of cohort studies. Eur. J. Epidemiol. 35 (5), 443–454. https://doi.org/10.1007/s10654-019-00593-4.
Peersman, W., Cambier, D., De Maeseneer, J., Willems, S., 2012. Gender, education and age differences in meanings that underlie global self-rated health. Int. J. Publ. Health 57 (3), 513–523. https://doi.org/10.1007/s00038-011-0320-2.
Raftery, A.E., 1995. Bayesian model selection in social research. Socio. Methodol. 25, 111–163. https://doi.org/10.2307/271063.
Ross, C.E., Wu, C.-L., 1996. Education, age, and the cumulative advantage in health. J. Health Soc. Behav. 37 (1), 104–120. https://doi.org/10.2307/2137234.
Sieber, S., Cheval, B., Orsholits, D., van der Linden, B.W.A., Guessous, I., Gabriel, R., Kliegel, M., von Arx, M., Kelly-Irving, M., Aartsen, M.J., Boisgontier, M.P., Courviosier, D., Burton-Jeangros, C., Cullati, S., 2020. Do welfare regimes moderate cumulative dis/advantages over the life course? Cross-national evidence from longitudinal SHARE data. J. Gerontol.: Ser. Bibliogr. 75 (6), 1312–1325. https://doi.org/10.1093/geronb/gbaa036.
Singh, S.K., Sharma, S.K., Mohanty, S.K., Mishra, R., Porwal, A., Gulati, B.K., 2022. Inconsistency in prevalence of hypertension based on self-reports and use of standard tests: implications for large scale surveys. SSM-Population Health 19, 101255. https://doi.org/10.1016/j.ssmph.2022.101255.
Sturm, R., 2002. The effects of obesity, smoking, and drinking on medical problems and costs. Health Aff. 21, 245–253. https://doi.org/10.1377/hlthaff.21.2.245.
Taylor, M.G., Min, S.N., Reid, K.M., 2020. Cumulative inequality at the end of life?: racial disparities in impairment in the time before death. J. Gerontol.: Ser. Bibliogr. 75 (6), 1292–1301. https://doi.org/10.1093/geronb/gby129.
Thomson, K., Hillier-Brown, F., Todd, A., McNamara, C., Huijts, T., Bambra, C., 2018. The effects of public health policies on health inequalities in high-income countries: an umbrella review. BMC Publ. Health 18 (1), 1–21. https://doi.org/10.1186/s12889-018-5677-1.
Van Der Linden, B.W.A., Sieber, S., Cheval, B., Orsholits, D., Guessous, I., Gabriel, R., et al., 2020. Life-course circumstances and frailty in old age within different European welfare regimes: a longitudinal study with SHARE. J. Gerontol.: Ser. Bibliogr. 75 (6), 1326–1335. https://doi.org/10.1093/geronb/gbz140.
van Kippersluis, H., Van Ourti, T., O'Donnell, O., van Doorslaer, E., 2009. Health and income across the life cycle and generations in europe. J. Health Econ. 28 (4), 818–830. https://doi.org/10.1016/j.jhealeco.2009.04.001.
Wetzel, M., Vanhoutte, B., 2020. Putting cumulative (dis)advantages in context: comparing the role of educational inequality in later-life functional health trajectories in England and Germany. PLoS One 15 (12), e0244371. https://doi.org/10.1371/journal.pone.0244371.
Whicher, C.A., O'Neill, S., Holt, R.I.G., 2020. Diabetes in the UK: 2019. Diabet. Med. 37 (2), 242–247. https://doi.org/10.1111/dme.14225.
Willson, A.E., Shuey, K.M., Elder, G.H., 2007. Cumulative advantage processes as mechanisms of inequality in life course health. Am. J. Sociol. 112 (6), 1886–1924. https://doi.org/10.1086/512712.
World Health Organization, 1948. Constitution of the World Health Organization, vols. 1–18. WHO, Geneva. https://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf?ua=1.
Wu, A.S., Dodhia, H., Whitney, D., Ashworth, M., 2019. Is the rule of halves still relevant today? A cross-sectional analysis of hypertension detection, treatment and control in an urban community. J. Hypertens. 37 (12), 2470–2480. https://doi.org/10.1097/HJH.0000000000002192.
Yang, Y.C., Walsh, C.E., Johnson, M.P., Belsky, D.W., Reason, M., Curran, P., Aiello, A.E., Chanti-Ketterl, M., Harris, K.M., 2021. Life-course trajectories of body mass index from adolescence to old age: racial and educational disparities. Proc. Natl. Acad. Sci. USA 118 (17), e2020167118. https://doi.org/10.1073/pnas.2020167118.
Zajacova, A., Huzurbazar, S., Todd, M., 2017. Gender and the structure of self-rated health across the adult life span. Soc. Sci. Med. 187, 58–66. https://doi.org/10.1016/j.socscimed.2017.06.019.
Zhao, Y., 2023. Socioeconomic positions and midlife health trajectories in a changing social context: evidence from China, 1991–2006. J. Health Soc. Behav. 64 (1), 39–61. https://doi.org/10.1177/00221465221150381.
Zhou, B., Danaei, G., Stevens, G.A., Bixby, H., Taddei, C., Carrillo-Larco, R.M., 2019. Long-term and recent trends in hypertension awareness, treatment, and control in 12 high-income countries: an analysis of 123 nationally representative surveys. Lancet 394 (10199), 639–651. https://doi.org/10.1016/S0140-6736(19)31145-6.
Ziebarth, N.R., 2010. Measurement of health, health inequality, and reporting heterogeneity. Soc. Sci. Med. 71 (1), 116–124. https://doi.org/10.1016/j.socscimed.2010.03.016.