International Journal of Epidemiology, 2019, 1–14
doi: 10.1093/ije/dyz248
Original article

Original article

Reducing socio-economic inequalities in all-cause mortality: a counterfactual mediation approach

Jessica E Laine,[1]* Valéria T Baltar,[2] Silvia Stringini,[3] Martina Gandini,[4] Marc Chadeau-Hyam,[1] Mika Kivimaki,[5] Gianluca Severi,[6] Vittorio Perduca,[6] Allison M Hodge,[7,8] Pierre-Antoine Dugué,[7,8,9] Graham G Giles,[7,8,10] Roger L Milne,[7,8,9] Henrique Barros,[11] Carlotta Sacerdote,[12] Vittorio Krogh,[13] Salvatore Panico,[14] Rosario Tumino,[15] Marcel Goldberg,[16] Marie Zins,[16] Cyrille Delpierre,[17] LIFEPATH Consortium and Paolo Vineis[1]

[1]Department of Epidemiology and Biostatistics, Imperial College London, London, UK, [2]Department of Epidemiology and Biostatistics, Instituto de Saúde Coletiva, Universidade Federal Fluminense, Rio de Janeiro, Brazil, [3]Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland, [4]Epidemiology Unit, ASL TO3 Piedmont Region, Grugliasco, Italy, [5]Department of Epidemiology and Public Health, University College London, London, UK, [6]Centre de Recherche en Épidémiologie et Santé des Populations (CESP, UMR Inserm 1018) Facultés de Medicine, Université Paris-Saclay, Université Paris-Sud, Saint-Aubin, France, [7]Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Australia, [8]Centre for Epidemiology and Biostatistics, School of Population and Global Health, The University of Melbourne, Melbourne, Australia, [9]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Melbourne, Australia, [10]School of Public Health and Preventive Medicine, Monash University Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia, [11]Department of Clinical Epidemiology, Predictive Medicine and Public Health, University of Porto Medical School, Porto, Portugal, [12]Unit of Cancer Epidemiology, AO Citta' della Salute e della Scienza–University of Turin and Center for Cancer Prevention, Turin, Italy, [13]Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy, [14]Department of Clinical Medicine and Surgery, Federico II University, Naples, Italy, [15]Provincial Healthcare Company (ASP) Ragusa, Vittoria, Italy, [16]Population-based Epidemiological Cohorts Unit, INSERM UMS 11, Villejuif, France and [17]LEASP, UMR 1027, Inserm-Université Toulouse III Paul Sabatier, Toulouse, France

*Corresponding author. Department of Epidemiology and Biostatistics, MRC Centre for Environment and Health, Imperial College London, Norfolk Place, London W2 1PG, UK. E-mail: jlaine@imperial.ac.uk

Editorial decision 24 October 2019; Accepted 27 November 2019

Abstract

Background: Socio-economic inequalities in mortality are well established, yet the contribution of intermediate risk factors that may underlie these relationships remains unclear. We evaluated the role of multiple modifiable intermediate risk factors underlying socio-

economic-associated mortality and quantified the potential impact of reducing early all-cause mortality by hypothetically altering socio-economic risk factors.

Methods: Data were from seven cohort studies participating in the LIFEPATH Consortium (total $n = 179\,090$). Using both socio-economic position (SEP) (based on occupation) and education, we estimated the natural direct effect on all-cause mortality and the natural indirect effect via the joint mediating role of smoking, alcohol intake, dietary patterns, physical activity, body mass index, hypertension, diabetes and coronary artery disease. Hazard ratios (HRs) were estimated, using counterfactual natural effect models under different hypothetical actions of either lower or higher SEP or education.

Results: Lower SEP and education were associated with an increase in all-cause mortality within an average follow-up time of 17.5 years. Mortality was reduced via modelled hypothetical actions of increasing SEP or education. Through higher education, the HR was 0.85 [95% confidence interval (CI) 0.84, 0.86] for women and 0.71 (95% CI 0.70, 0.74) for men, compared with lower education. In addition, 34% and 38% of the effect was jointly mediated for women and men, respectively. The benefits from altering SEP were slightly more modest.

Conclusions: These observational findings support policies to reduce mortality both through improving socio-economic circumstances and increasing education, and by altering intermediaries, such as lifestyle behaviours and morbidities.

Key words: Socio-economic inequalities, all-cause mortality, mediation, intervention, health behaviours, causal inference, multiple mediators

Key Messages

- Socio-economic-associated mortality differed based on one's socio-economic position (SEP) or level of education.
- Up to 38% and 34% of the effect of education on mortality was mediated by the joint mediators of smoking, alcohol consumption, adherence to a Western dietary pattern, physical activity, body mass index and morbidities of hypertension, diabetes and coronary artery disease.
- Socio-economic-associated mortality could be reduced both through a hypothetical intervention on SEP or education (natural direct effect) and through the joint mediators considered (natural indirect effect).

Introduction

Those who experience greater socio-economic deprivation have higher risks for poorer health, accelerated biological ageing and early death.[1–5] Several factors underlie these risks, including, but not limited to, differences in access to healthcare and health information,[6] behaviours such as smoking,[7] alcohol intake,[8] dietary habits,[9] physical activity[10] and drug use,[11] greater exposure to environmental and occupational hazards[12] and increased psychosocial stress.[13] Estimating the causal underpinnings of socio-economic-associated morbidity and mortality is complex, as several potential biological mechanisms and health behaviours likely lie on the causal path. Disentangling the specific and integrative causal chains of these factors is further complicated by the disparate methodological approaches that may influence studies' findings, validity and causal conclusions. For example, traditional methodological approaches in mediation, such as the difference and/or product of coefficients, may lead to conclusions that differ between studies (e.g. the magnitude of the direct and indirect effects) and may not be optimal for causal interpretations.[14,15] To reduce the disease burdens from socio-economic inequalities, it is imperative to disentangle these complexities and clarify causal relationships, using counterfactual methods.

Lifestyle factors and health-related behaviours have previously been linked to mortality.[4] Specifically, smoking, alcohol consumption, physical activity and diet have been put forth as putative mediating pathways underlying socio-economic-associated adverse health outcomes and

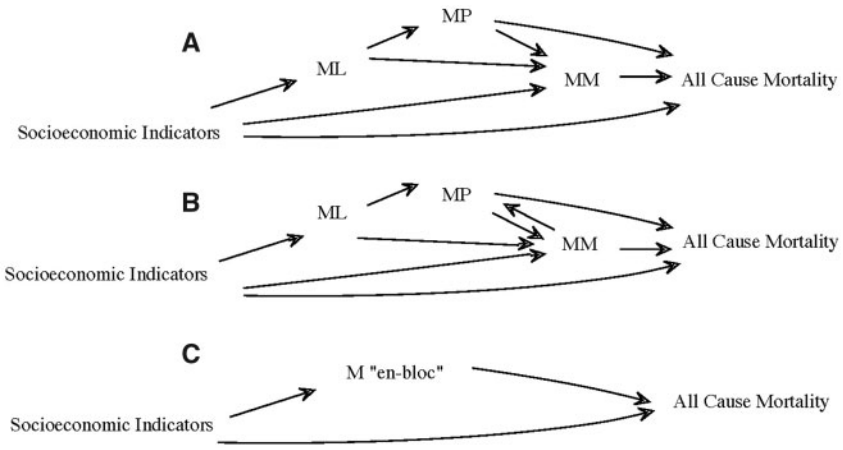

Figure 1. Causal structures of mediators considered: representing two potential causal structures of socio economic position (SEP), mediating lifestyle behaviours (ML), an intermediate phenotype (MP) of body mass index (BMI) and co-morbidities (MM), and the causal structure of all mediators (M) assessed jointly (C) in the present study. (A) displays a Directed Acyclic Graph (DAG) where SEP leads to changes in ML, MP and MM, whereas (B) displays an alternative causal structure where SEP influences ML and ML influences MP, but MM influences MP. Based on these two potential directions (among others not represented) between mediators, we assess all mediators 'en bloc' as displayed in (C).

mortality, as highlighted in a recent systematic review.[16] In addition to their mediating roles, lifestyle behaviours along with morbidities contribute independently to years of life loss. For example, we recently identified that lower socio-economic position (SEP) was associated with a 2.1-year reduction in life expectancy (after adjustment for lifestyle risks), with lifestyle factors and co-morbidities differing in their contribution.[5] These, along with other findings, demonstrate the potential for reducing socio-economic-associated mortality, either directly or indirectly by altering one or more of these modifiable risk factors.[17] Yet, some authors have suggested that socio-economic differentials in health behaviours account for only a modest proportion of social inequalities in overall mortality.[18] Likely inconsistencies in mediated effects from previous studies may have been underestimated by only assessing a single or few mediator(s) and/or by introducing over-adjustment bias by controlling for an intermediate variable (or a descending proxy for an intermediate variable).[19] Controlling for intermediate variables may block some of the effects of socio-economic conditions. The application of empirical methods within the counterfactual mediation framework, along with the assessment of multiple mediators,[20,21] can facilitate the identification and quantification of pathways that underlie socio-economic-related all-cause mortality via potential outcomes that might have been observed had the exposure and/or intermediate risk factor(s) been different and can estimate effects under hypothetical interventions.

In the LIFEPATH project[22]—an international study comprising multiple cohorts—we have previously established an association between socio-economic indicators and all-cause mortality.[5,23] In the present study, we aim to decompose these associations, considering the causal structures posited in Figure 1, to determine the causal processes underlying socio-economically patterned adverse health outcomes that are amenable to intervention. Our approach builds on the emerging body of causal inference for social epidemiology,[24] but aims to explain more fully causal processes by assessing multiple mediators, using multiple large representative cohorts, and by addressing previous methodological limitations. We assess two socio-economic indicators—SEP and education—and apply counterfactual mediation analysis, considering the joint mediating role of known consequences of socio-economic indicators on lifestyle behaviours,[4] such as smoking, alcohol intake, dietary patterns,[25] physical activity[26] and the intermediate phenotype of body mass index (BMI)[27] and morbidities,[28–30] such as hypertension, diabetes and coronary artery disease. We address joint mediation because, as shown in Figure 1, the causal order of mediators to one another can often be of question (Figure 1A and B), thus it can be difficult to obtain path-specific decompositions, irrespective of the causal links between mediators. Therefore, we addressed mediation 'en bloc', where all mediators (M) are considered jointly (Figure 1C). By decomposing these effects, the results of this study could extend into potential social investment and policy recommendations to reduce the burden of socio-economic inequalities on all-cause mortality across populations.

Methods

Cohort selection

Of the 18 cohorts represented in LIFEPATH,[22] only adult cohorts were included for the present study. Of these

cohorts, we further prioritized those with detailed information for each mediator. Based on these criteria, seven cohorts ($n = 179\,090$) are represented in the present study, namely CoLaus ($n = 4605$), E3N ($n = 74\,521$), EPIC-Italy ($n = 33\,151$), EPIPorto ($n = 1780$), Gazel ($n = 15\,585$), MCCS cohort ($n = 41\,412$) and Whitehall II ($n = 8036$). A full description of these cohorts is presented as Supplementary data, available at *IJE* online.

Exposures, mediators and outcome

The proposed causal structure underlying our study is represented by Figure 1. Data from the present study consists of previously harmonized variables described elsewhere[23] and is detailed further as Supplementary data, available at *IJE* online. We considered both individual's SEP and education at baseline as separate socio-economic indicators. This is due to the fact that different health patterns have been observed based on the measure used to represent socio-economic inequalities (e.g. income, occupation, education) and, whereas there are correlations among these measures, each may represent distinct socio-economic information.[31] SEP was based on individuals' last occupational position, dichotomized as lower SEP for manual employees and higher SEP for non-manual employees. This was used as a socio-economic indicator for six out of the seven cohorts, as information was not available for the MCCS cohort. For all cohorts, the individual's education was available and was harmonized and dichotomized as lower education for compulsory education and higher education for non-compulsory education.

A full description of the mediators considered is presented in Supplementary data, available at *IJE* online. ML comprised self-reported smoking (dichotomized as never-smoker and ever-smoker), alcohol intake, measured as units per day (dichotomized as $\leq$ or $>$ 2 drinks per day for men and $\leq$ or $>$ 1 drink per day for women), quartiles of a dietary score that represents a degree of adherence to a Western dietary pattern and physical activity (dichotomized as active or not active). MP (BMI) was categorized as normal ($\leq$25 kg/m^2), overweight (25 to $<$30 kg/m^2) or obese ($\geq$30 kg/m^2). MM comprised hypertensive (defined by individuals having a systolic blood pressure $\geq$140 mm Hg or diastolic blood pressure $\geq$90 mm Hg or self-report use of medication for the treatment of hypertension) or non-hypertensive, diabetic (defined by one of the following states: fasting glucose $\geq$7 mmol/L, 2-hour post-load glucose $\geq$11.1 mmol/L, glycated haemoglobin A1c $>$6.5% or self-reported diabetes) or non-diabetic, and coronary artery disease (CAD) (based on self-report of an event of angina and/or heart attack). All cohorts had longitudinal data available, with data-collection dates ranging from 1989 to 2013 (Supplementary Table 2, available as Supplementary data at *IJE* online). To avoid cross-sectional mediation analyses and the potential for reverse causality, where possible, all mediators followed SEP or education exposure at baseline and preceded mortality; however, this was not true for all cohorts. For example, in CoLaus, EPIC-Italy, EPIPorto and MCCS cohort diet was assessed at baseline. Each set of mediators was considered at only one time point, where the timing of assessment differed across the life course among the cohorts.

Potential confounders were selected based on their previously known association with the exposures, any of the mediators and the outcome, and that do not lie on the causal path, identified by a Directed Acyclic Graph (DAG). Age at baseline (continuous) and marital status at baseline (dichotomized as those who were married or cohabitating or those who were single, separated, divorced or widowed) were considered. The outcome of interest was time-to-death from any cause, with time in years since cohort entry as the underlying timescale. Left truncation was taken into account by adjusting for age at recruitment.

Statistical framework and analysis

All analyses were conducted using R.[32] Adjusted survival curves for all-cause mortality, stratified by sex to investigate the sex-based differences in survival rates, were plotted for both SEP and education using the R package survminer.[33] To estimate the effect of socio-economic indicators and the joint mediating role of all mediators on all-cause mortality, we used natural effect models (NEMs) based on a counterfactual method, established by Lange *et al.*,[34] to estimate the natural direct effect (NDE), natural indirect effect (NIE) and total effect (TE). Formulation of and assumptions underlying natural effects have been discussed in detail elsewhere.[35,36] Briefly, these assumptions are the identification assumption of no unmeasured and uncontrolled confounding of the exposure–mediator, mediator–outcome, mediator–mediator or exposure–outcome relationship, positivity, consistency and sequential ignorability, and are further detailed elsewhere specifically for NEMs.[34] When there is more than one mediator, the sequential ignorability assumption might be violated. To satisfy this assumption, we take a joint mediation approach, where the causal dependence of the joint mediators and the sequential ignorability assumption is now for a vector of mediators instead of a single mediator. Additionally, because the outcome is survival, we also assume that censoring is independent of event time.

We investigate two scenarios of counterfactual exposures, to estimate the counterfactual effects of both socio-

economic indicators, SEP and education (assessed separately). These two scenarios are highlighted to demonstrate differences in effects under two hypothetical actions, the first of no action on socio-economic conditions of lower SEP (SEP^L) or lower education (EDU^L) or an action to increase socio-economic conditions to higher SEP (SEP^H) or higher education (EDU^H). Additionally, we estimate the role of mediators under both actions (of lower and higher socio-economic conditions) and capture the NDE and NIE under joint mediation, which is important as the NDE may depend on the natural level at which the mediator(s) is allowed to vary for NEMs and the NIE may depend on the level to which the exposure is set. Specifically, we estimated the counterfactual survival time (T) that would have been observed had the exposure to SEP or education (A) been set to the contrary a and all mediators (M) to m. For SEP^H/EDU^H, let $A = a$ represent a state in which (if contrary to the fact) a is set to lower SEP or lower education and $A = a^*$ represents a state in which (if contrary to the fact) SEP or education is set to higher, where the reverse is true for the SEP^L and EDU^L. Let M(a) and M(a*) denote the joint mediators with two corresponding potential outcomes. T(a, m) denotes the potential outcome when the exposure is set to a and the mediator is set to m.

Estimands for the NDE, of the socio-economic indicators on all-cause mortality, NIE through the intermediates, TE of both the NDE and NIE, and proportion mediated (PM) were derived based on NEMs, which are conditional mean models for nested counterfactuals.[37–39] These address previous limitations of traditional mediation analyses, based on marginal structural models that directly parameterize the natural direct and indirect effects.[34] This method estimates the parameters of the marginal structural model[37] by using inverse probability of treatment weights, creating a pseudopopulation to remove covariate imbalances between the two exposure groups of lower and higher SEP or education. Specifically, the marginal NEM parameterizes $E(Y(a, M(a^*))) = g(a, a^*; \beta)$, where Y is an outcome and g is a known link function (e.g. logistic for odds ratios), where the data for a are augmented so that $a^* = 1$, and we fit the model to the new data set, with weights (w) for individual i $p(M = mi | A = a^* i, wi)$ $p(M = m\ i | X = xi, wi)$. To obtain estimates for NEM for survival outcomes, we used the approach by Lange *et al.*[34] to (i) fit a parametric Weibull survival model for survival times with either SEP or education (separately), all mediators and confounders, (ii) impute the nested counterfactuals and (iii) fit a Cox model including the observed mortality status, the counterfactual mortality status and all confounders, to calculate hazard ratios (HRs) and 95% confidence intervals (CIs) using a bootstrap procedure with 1000 draws.

Under the different simulated hypothetical actions, we estimated the effect of (SEP^L, EDU^L) for population total effects (TE^L), the NDE (NDE^L) and NIE (NIE^L). Second, to determine the potential effects of a hypothetical intervention (SEP^H, EDU^L), we estimated the effects under an action of increasing SEP or education to a higher status for the TE (TE^H), NDE (NDE^H) and NIE (NIE^H). NEM models were fitted separately for men and women, as sex has been previously demonstrated in LIFEPATH as an effect modifier of socio-economic-associated all-cause mortality.[23] A cohort-specific analysis was performed for each cohort separately, adjusting for potential confounders, age and marital status. In pooled analyses, data across all cohorts were combined, the same covariates of age at baseline and marital status were included in these models, as was a fixed effect for cohort to account for potential cohort effects.

To assess the sensitivity of our results to possible violations of the assumptions of the causal estimands, we addressed two of our causal assumptions in the pooled cohort analyses under the hypothetical action of lower SEP and education. For the first sensitivity analysis, we evaluate the assumption that marital status is a potential confounder of the exposure, mediators and/or outcome. To determine whether marital status acts as a potential mediator, we added marital status as a joint mediator alongside the other mediators. To assess marital status as a potential effect modifier, we ran models separately for men and women, and stratified further by marital status. For the second sensitivity analysis, we addressed the potential relationship between SEP and education. In our primary analysis, we treated education and SEP as separate exposures, but they may influence one another, where education can influence one's SEP. To address this, we ran our models of EDU^L including SEP as a mediator, with and without an interaction term between SEP and education. The causal effects estimated are the NDE of lower education on all-cause mortality and the NIE of joint mediators through the previous intermediates accessed along with SEP.

Results

Characteristics of the population

Summary statistics, including information on mediators, assessed separately for males and females for both SEP and education, and for all individuals presented in Table 1. Summary statistics for each cohort are presented in Supplementary Table 3, available as Supplementary data at *IJE* online. There were more women ($n = 131\,020$) than men ($n = 48\,070$), largely due to the fact that E3N is an all-female cohort. There were more men and women who had lower education (75% and 71%, respectively) than lower SEP (23% and 39%, respectively). At the end of follow-up,

Table 1. Demographic characteristics of participants in LIFEPATH ($n = 179\,089$) included in the present study

	Men ($n = 48\,070$) (26.8%)				Women ($n = 131\,020$) (73.2%)				All ($n = 179\,090$)
	Socio-economic position (SEP) ($n = 31\,089$)		Education ($n = 48\,033$)		Socio-economic position (SEP) ($n = 32\,067$)		Education ($n = 131\,020$)		
	Lower SEP ($n = 7092$) (23%)	Higher SEP ($n = 23\,997$) (77%)	Lower education ($n = 36\,091$) (75%)	Higher education ($n = 11\,942$) (25%)	Lower SEP ($n = 12\,531$) (39%)	Higher SEP ($n = 19\,536$) (61%)	Lower education ($n = 93\,235$) (71%)	Higher education ($n = 37\,785$) (29%)	
All-cause mortality *N* (*N*%)									
Alive	6302 (88.9%)	21 683 (90.4%)	29 646 (82.1%)	10 510 (88.0%)	11 785 (94.0%)	18 586 (95.1%)	83 995 (90.1%)	35 010 (92.7%)	159 162 (88.9%)
Deceased	790 (11.1%)	2314 (9.6%)	6445 (17.9%)	1432 (12.0%)	746 (5.9%)	950 (4.9%)	9220 (9.9%)	2769 (7.3%)	19 866 (11.1%)
Average follow-up time in years (SD)	17.1 (6.1)	20.4 (6.7)	18.8 (6.1)	18.6 (5.8)	15.7 (5.3)	16.6 (5.8)	17.1 (3.9)	17.02 (3.6)	17.5 (4.6)
Baseline characteristics									
Age at baseline, mean (SD)	49 (8.0)	47 (6.8)	51 (8.7)	49 (8.2)	50 (8.8)	48 (8.5)	51 (8.1)	17 (3.6)	50 (8.1)
Marital status, *N* (*N*%)									
Married	5588 (90.3%)	20 764 (90.1%)	29 638 (87.9%)	9687 (82.6%)	10 049 (81.1%)	15 187 (78.5%)	73 157 (81.2%)	27 380 (74.7%)	139 887 (81.2%)
Not married	601 (9.7%)	2283 (9.9%)	4074 (12.1%)	2040 (17.4%)	2337 (18.9%)	4150 (21.5%)	16 936 (18.8%)	9265 (25.3%)	32 324 (18.8%)
Mediating lifestyle behaviours									
Smoking, *N* (*N*%)									
Never-smoker	2114 (29.8%)	8768 (36.5%)	11 951 (33.1%)	5858 (49.0%)	7848 (62.6%)	10 506 (53.8%)	56 955 (61.1%)	18 329 (48.5%)	93 107 (52.0%)
Ever-smoker	4978 (70.2%)	15 229 (63.5%)	24 140 (66.9%)	6084 (50.9%)	4683 (37.4%)	9030 (46.2%)	36 280 (38.9%)	19 456 (51.5%)	85 983 (48.0%)
Alcohol intake, *N* (*N*%)									
≤2 (M)/1 (F) drinks per day	4209 (59.4%)	15 651 (65.2%)	23 238 (64.4%)	8274 (69.3%)	10 112 (80.7%)	15 308 (78.4%)	66 141 (70.9%)	22 609 (59.8%)	120 282 (67.2%)
>2 (M)/1 (F) drinks per day	2883 (40.6%)	8346 (34.8%)	12 853 (35.6%)	3668 (30.7%)	2419 (19.3%)	4228 (21.6%)	27 094 (29.1%)	15 176 (40.2%)	58 807 (32.8%)
Dietary pattern, IQR (SD)									
Western	1.4 (3.4)	0.15 (1.02)	1.3 (8.9)	1.2 (9.8)	1.1 (–5.9)	1.1 (–4.8)	1.2 (–15.5)	2 (0.74)	1.2 (6.1)
Physical activity (PA), *N* (*N*%)									
Not active	4698 (66.3%)	15 095 (62.9%)	22 054 (61.1%)	6120 (51.3%)	9121 (72.8%)	14 232 (72.9%)	40 262 (43.2%)	10 659 (28.2%)	79 114 (44.2%)
Active	2386 (33.7%)	8890 (37.1%)	14 019 (38.9%)	5820 (48.7%)	3401 (27.2%)	5279 (27.1%)	52 946 (56.8%)	27 119 (71.8%)	99 922 (55.8%)
Intermediate phenotype mediator									
BMI, mean (SD)	27.02 (3.7)	26.3 (3.4)	27.1 (3.8)	26.3 (3.7)	26.4 (4.6)	25.5 (4.5)	25.2 (4.7)	23.6 (3.9)	25.3 (4.4)

(Continued)

Table 1. Continued

	Men ($n = 48\,070$) (26.8%)				Women ($n = 131\,020$) (73.2%)				All ($n = 179\,090$)
	Socio-economic position (SEP) ($n = 31\,089$)		Education ($n = 48\,033$)		Socio-economic position (SEP) ($n = 32\,067$)		Education ($n = 131\,020$)		
	Lower SEP ($n = 7092$) (23%)	Higher SEP ($n = 23\,997$) (77%)	Lower education ($n = 36\,091$) (75%)	Higher education ($n = 11\,942$) (25%)	Lower SEP ($n = 12\,531$) (39%)	Higher SEP ($n = 19\,536$) (61%)	Lower education ($n = 93\,235$) (71%)	Higher education ($n = 37\,785$) (29%)	
Normal	2105 (29.7%)	8943 (37.3%)	10 732 (29.7%)	4504 (37.7%)	5394 (43.0%)	10 265 (52.5%)	51 567 (55.3%)	26 796 (70.9%)	93 609 (52.3%)
Overweight	3729 (52.6%)	12 054 (50.2%)	18 484 (51.2%)	5826 (48.8%)	4755 (38.0%)	6415 (32.8%)	28 217 (30.3%)	8390 (22.2%)	60 934 (34.0%)
Obese	1258 (17.7%)	3000 (12.5%)	6875 (19.1%)	1612 (13.5%)	2382 (19.0%)	2856 (14.6%)	13 451 (14.4%)	2599 (6.9%)	24 547 (13.7%)
Co-morbidities mediators									
Hypertension, *N* (*N*%)									
Non- hypertensive	4198 (59.2%)	16 813 (70.1%)	21 125 (58.6%)	8071 (67.6%)	9195 (74.9%)	13 142 (67.3%)	58 917 (63.2%)	1448 (3.2%)	116 579 (65.1%)
Hypertensive	2894 (40.8%)	7184 (29.9%)	14 943 (41.4%)	3863 (32.4%)	3075 (25.1%)	6394 (32.7%)	34 293 (36.8%)	9323 (24.7%)	62 444 (34.9%)
Diabetes, *N* (*N*%)									
Non- diabetic	6389 (95.2%)	21 286 (95.8%)	32 350 (95.0%)	11 088 (97.3%)	11 827 (96.7%)	18 488 (97.7%)	44 065 (96.8%)	9924 (98.5%)	97 861 (96.4%)
Diabetic	320 (4.8%)	935 (4.2%)	1692 (5.0%)	311 (2.7%)	406 (3.3%)	441 (2.3%)	1448 (3.2%)	156 (1.6%)	3639 (3.6%)
Coronary artery disease (CAD), *N* (*N*%)									
No-CAD	1924 (93.1%)	11 484 (93.3%)	21 285 (92.0%)	7728 (94.5%)	2442 (97.7%)	5018 (98.5%)	70 490 (98.0%)	34 275 (99.0%)	133 808 (97.0%)
CAD	143 (6.9%)	823 (6.7%)	1853 (8.0%)	452 (5.5%)	58 (2.3%)	78 (1.5%)	1439 (2.0%)	345 (1.0%)	4095 (3.0%)

there were 19 866 (11.1%) deaths. Participants were an average age of 50 [8.1 standard deviation (SD)] at baseline. The average follow-up time was 17.5 years (4.6 SD).

Total effect of setting SEP and education to a lower level (TE^L) on all-cause mortality

There was lower survival for women and men with lower SEP or education compared higher SEP or education (Figure 2). In the pooled analyses, for SEP, the TE^L HR was 1.06 (95% CI 1.05, 1.07) and 1.30 (95% CI 1.24, 1.30) for all-cause mortality for women and men, respectively, whereas participants with lower education had a TE^L HR of 1.18 (95% CI 1.16, 1.19) and HR of 1.38 (95% CI 1.35, 1.40) compared with those with higher education for women and men, respectively (Table 2). The same pattern for lower vs higher education was apparent in the cohort-specific analysis for men and women, with the magnitude of effect differing by cohort. However, this was not so for SEP among women in Gazel.

Total effect of setting SEP and education to a higher level (TE^H) on all-cause mortality

A decrease in the hazard rate for mortality was observed for both women and men if (contrary to the facts) everyone had a higher SEP or education compared with if (contrary to the facts) everyone had a lower SEP (Table 2). The mortality gradient would decrease to an estimated TE^H HR of 0.94 (95% CI 0.93, 0.95) for women and to an estimated TE^H HR of 0.80 (95% CI 0.78, 0.82) for men by altering the SEP (Table 2). The same trend was observed in the cohort-specific analysis, except for women in Gazel (Table 2). By hypothetically altering education from a lower to a higher status, the mortality gradient would decrease to a TE^H HR of 0.85 for women (95% CI 0.84, 0.86) and to a TE^H HR of 0.71 (95% CI 0.70, 0.74) for men. The same direction for the TE^H HR estimates was observed in the cohort-specific analysis (Table 2).

Decomposition of the effect of an action to increase SEP or education on all-cause mortality through joint mediation

The effect of increasing SEP or education on all-cause mortality has two components: the natural direct effect of SEP or education not through the mediators assessed (NDE^H) and the natural indirect effect through all mediators examined (NIE^H).

For women, in the pooled analysis, under higher SEP, the direct effect (NDE^H) HR was 0.92 (95% CI 0.91, 0.93). This corresponds to the indirect effect under exposure to higher SEP (NIE^H) HR of 1.03 (95% CI 1.01, 1.04). The PM for women is uninterpretable because the direct and indirect effects operate in opposite directions. For education, the direct effect (NDE^H) HR was 0.90 (95% CI 0.89, 0.91). The

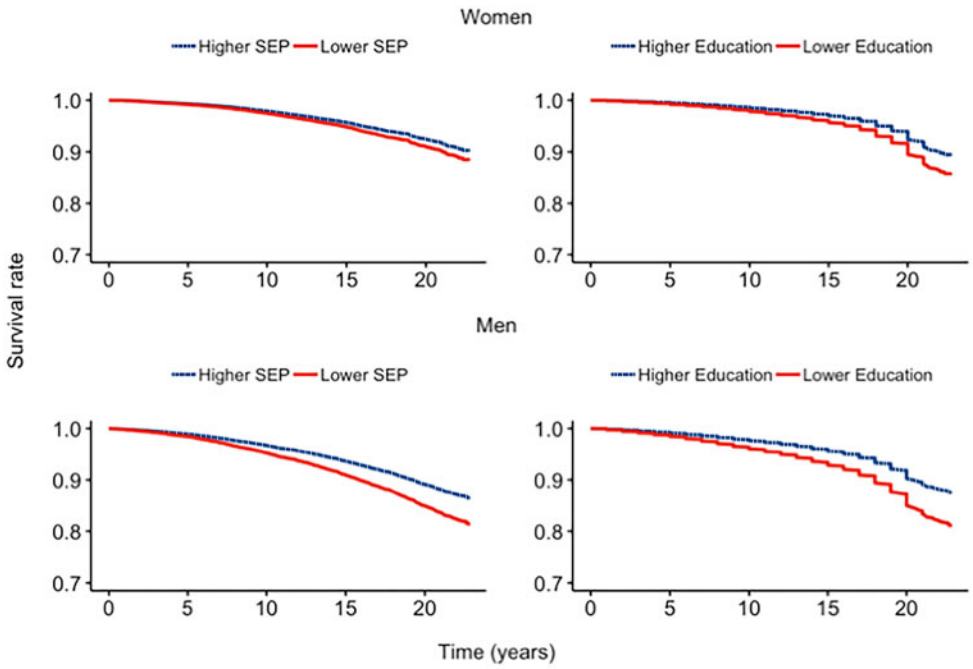

Figure 2. Survival curves for women and men by two socio-economic indicators of socio-economic position (SEP) and education, adjusted for age, marital status and cohort. Higher socioeconomic indicators are represented by the dotted lines and lower socioeconomic indicators are represented by solid lines.

Table 2. Natural effect estimates pooled and by cohort[a] separately for men and women, displaying total, natural direct and indirect effects for socio-economic indicators of socio-economic position and education, and joint mediation by smoking, alcohol consumption, Western dietary pattern, physical activity, body mass index, coronary artery disease, diabetes and hypertension

Socio-economic position[b]							
Cohort	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
	TE^L	TE^H	NDE^H	NDE^L	NIE^H	NIE^L	PM
Pooled[w]	1.06 (1.05, 1.07)	0.94 (0.93, 0.95)	0.92 (0.91, 0.93)	1.09 (1.08, 1.10)	1.03 (1.01, 1.04)	0.97 (0.96, 0.99)	–0.46 (–0.73, –0.23)
EPIC-IT[w]	1.08 (1.07, 1.08)	0.93 (0.92, 0.93)	0.90 (0.89, 0.92)	1.11 (1.09, 1.12)	1.03 (1.01, 1.04)	0.97 (0.96, 0.99)	–0.38 (–0.59, –0.17)
EPIPORTO[w]	2.05 (1.92, 2.23)	0.49 (0.45, 0.53)	0.71 (0.67, 0.75)	1.40 (1.34, 1.48)	0.69 (0.65, 0.72)	1.50 (1.40, 1.54)	0.53 (0.48, 0.58)
GAZEL[w]	0.84 (0.78, 0.91)	1.18 (1.09, 1.28)	1.11 (1.06, 1.17)	0.90 (0.85, 0.95)	1.07 (1.001, 1.13)	0.94 (0.88, 0.99)	0.36 (0.0078, 0.60)
WHITEHALL[w]	1.06 (1.04, 1.08)	0.94 (0.92, 0.96)	0.96 (0.92, 0.99)	1.04 (1.0, 1.08)	0.98 (0.94, 1.03)	1.02 (0.98, 1.06)	0.29 (–0.46, 0.97)
Pooled[m]	1.30 (1.24, 1.30)	0.80 (0.78, 0.82)	0.83 (0.81, 0.84)	1.22 (1.20, 1.24)	0.97 (0.95, 0.99)	1.04 (1.02, 1.10)	0.17 (0.088, 0.24)
CoLaus[m]	1.31 (1.08, 1.54)	0.77 (0.64, 0.92)	0.81 (0.74, 0.89)	1.24 (1.12, 1.36)	0.96 (0.96, 1.008)	1.06 (0.95, 1.17)	0.13 (–0.43, 0.40)
EPIC-IT[m]	1.21 (1.21, 1.23)	0.83 (0.82, 0.83)	0.87 (0.85, 0.89)	1.15 (1.13, 1.17)	0.95 (0.93, 0.97)	1.06 (1.03, 1.08)	0.30 (0.18, 0.40)
EPIPORTO[m]	1.10 (1.01, 1.10)	0.95 (0.91, 1.00)	0.97 (0.91, 1.05)	1.03 (0.95, 1.10)	0.98 (0.90, 1.06)	1.02 (0.94, 1.11)	0.28 (–2.9, 2.80)
GAZEL[m]	1.31 (1.24, 1.38)	0.77 (0.73, 0.81)	0.73 (0.71, 0.75)	1.40 (1.33, 1.42)	1.05 (1.01, 1.10)	0.95 (0.91, 0.99)	1.39 (–0.83, 2.40)
WHITEHALL[m]	1.54 (1.40, 1.70)	0.65 (0.60, 0.71)	0.75 (0.71, 0.79)	1.33 (1.26, 1.39)	0.86 (0.79, 0.93)	1.20 (1.07, 1.30)	0.34 (0.19, 0.46)
Education[c]							
	TE^L	TE^H	NDE^H	NDE^L	NIE^H	NIE^L	PM
Pooled[w]	1.18 (1.16, 1.19)	0.85 (0.84, 0.86)	0.90 (0.89, 0.91)	1.11 (1.10, 1.12)	0.95 (0.94, 0.96)	1.06 (1.04, 1.07)	0.34 (0.29, 0.38)
EPIC-IT[w]	1.10 (0.98, 1.11)	0.96 (0.88, 1.02)	0.98 (0.95, 1.02)	1.02 (0.98, 1.06)	0.97 (0.93, 1.02)	1.03 (0.98, 1.08)	1.39 (–0.83, 2.4)
EPIPORTO[w]	2.75 (2.07, 3.74)	0.36 (0.26, 0.49)	0.58 (0.49, 0.68)	1.71 (1.48, 2.00)	0.62 (0.53, 0.72)	1.61 (1.39, 1.89)	0.47 (0.43, 0.50)
GAZEL[w]	2.67 (2.01, 3.64)	0.51 (0.41, 0.62)	0.55 (0.49, 0.60)	1.65 (1.43, 1.94)	0.93 (0.82, 1.04)	1.62 (1.39, 1.90)	0.10 (–0.8, 0.23
WHITEHALL[w]	1.03 (0.96, 1.13)	0.97 (0.89, 1.1)	1.00 (0.96, 1.06)	0.99 (0.94, 1.05)	0.96 (0.90, 1.03)	1.04 (0.97, 1.10)	0.10 (–0.8, 0.23)
MCCS[w]	1.19 (1.15, 1.22)	0.84 (0.81, 0.87)	0.91 (0.89, 0.93)	1.10 (1.08, 1.12)	0.92 (0.90, 0.95)	1.10 (1.05, 1.11)	0.45 (0.34, 0.51)
E3N[w]	1.16 (1.14, 1.17)	0.87 (0.86, 0.88)	0.92 (0.91, 0.93)	1.10 (1.08, 1.10)	0.94 (0.93, 0.95)	1.06 (1.05, 1.07)	0.42 (0.36, 0.47)
Pooled[m]	1.38 (1.35, 1.40)	0.71 (0.70, 0.74)	0.82 (0.81, 0.83)	1.22 (1.21, 1.24)	0.89 (0.87, 0.90)	1.13 (1.11, 1.15)	0.38 (0.35, 0.42)
CoLaus[m]	1.80 (1.50, 2.23)	0.58 (0.47, 0.70)	0.66 (0.59, 0.73)	1.53 (1.39, 1.71)	0.88 (0.79, 0.98)	1.17 (1.05, 1.32)	0.26 (0.14, 0.36)
EPIC-IT[m]	1.72 (1.44, 2.13)	0.61 (0.55, 0.66)	0.67 (0.64, 0.71)	1.52 (1.38, 1.69)	0.90 (0.85, 0.95)	1.13 (1.02, 1.27)	0.20 (0.11, 0.28)

Table 2. continued

Education[c]							
	TE^L	TE^H	NDE^H	NDE^L	NIE^H	NIE^L	PM
EPIPORTO[m]	2.20 (1.79, 2.77)	0.46 (0.37, 0.56)	0.52 (0.43, 0.58)	1.93 (1.72, 2.19)	0.88 (0.77, 1.01)	1.14 (0.98, 1.33)	0.15 (–0.09, 0.29)
GAZEL[m]	2.24 (1.84, 2.80)	0.53 (0.50, 0.57)	0.66 (0.63, 0.68)	1.95 (1.75, 2.20)	0.81 (0.78, 0.85)	1.15 (0.99, 1.34)	0.16 –0.012, 0.30)
WHITEHALL[m]	1.10 (1.08, 1.12)	0.91 (0.88, 0.93)	0.94 (0.92, 0.97)	1.06 (1.03, 1.09)	0.96 (0.93, 0.99)	1.04 (1.01, 1.07)	0.40 (0.13, 0.67)
MCCS[m]	1.34 (1.31, 1.37)	0.77 (0.75, 0.79)	0.83 (0.82, 0.85)	1.17 (1.15, 1.19)	0.93 (0.91, 0.95)	1.15 (1.12, 1.17)	0.47 (0.40, 0.52)

[a]Pooled models were adjusted for age, marital status and cohort status, and cohort-specific models were adjusted for age and marital status.
[b]$n = 32\,067$ for women and $n = 31\,089$ for men.
[c]$n = 131\,020$ for women and $n = 48\,033$ for men.
HR, hazard ratio; TE^L, total effect where socio-economic indicators are set to lower status; TE^H, total effect where socio-economic indicators are set to a higher status; NDE^L, natural direct effect where socio-economic indicators are set to a lower status; NDE^H, natural direct effect where socio-economic indicators are set to a higher status; NIE^L, natural indirect effect where socio-economic indicators are set to lower status through all mediators; NIE^H, natural indirect effect where socio-economic indicators are set to higher status through all mediators; PM, the proportion of the effect where socio-economic indicators are set to higher that is mediated through all mediators; [m], men; [w], women.

effect via the mediators when education was higher was that the (NIE^H) HR was 0.95 (95% CI 0.94, 0.96). The PM by all joint mediators was 34%. The direction of the NIE/NDE varied by cohort (Table 2).

For men, in the combined cohort analysis, under higher SEP, the direct effect (NDE^H) HR was 0.83 (95% CI 0.81, 0.84) and the effect via the mediators when exposed to higher SEP was that the (NIE^H) HR was 0.97 (95% CI 0.95, 0.99) (Table 2). The PM by all joint mediators was 17%. Similar results were observed in the cohort-specific analysis, except for the NIE^H for Gazel, which was below the null value (Table 2). For education, the direct effect (NDE^H) HR was 0.82 (95% CI 0.81, 0.83). The effect via the mediators with higher education was that the (NIE^H) HR was 0.89 (95% CI 0.87, 0.89). The PM by all joint mediators was 38%.

Sensitivity analyses

When assessing marital status as a potential mediator (along with the other mediators), there was no difference in the effect estimates compared with adjusting for marital status as a potential confounder for education and for men with SEP; however, there was a slight increase in the NDE, NIE and TE for women with SEP (Supplementary Table 4, available as Supplementary data at *IJE* online). However, when assessing marital status as a potential effect modifier of the relationship between SEP and education, intermediates and all-cause mortality, we found that the HR differed depending on the socio-economic indicator used and by sex (Supplementary Table 5, available as Supplementary data at *IJE* online). To assess the relationship between education and SEP together on all-cause mortality, we ran sensitivity analyses including SEP as a joint mediator between education and all-cause mortality. The direct and TE of lower education on all-cause mortality did not change for men or women; however, the NIE and TE increased slightly by including SEP compared with education modelled alone (Supplementary Table 6, available as Supplementary data at *IJE* online).

Discussion

We investigated the joint-mediatory role of lifestyle factors and co-morbidities underlying the effects of SEP (based on last occupation) and education and all-cause mortality. Overall, we found there was a higher hazard rate for mortality with lower SEP and education compared with higher SEP and education, effects were partially mediated by lifestyle factors and co-morbidities, and effects could be reduced through a hypothetical intervention to increase SEP or education.

The higher hazard rate for mortality observed in the present study with lower SEP and education relative to higher SEP and education is not surprising, as this has been demonstrated within LIFEPATH[5,23] and several other studies (as reviewed by[40]). Similarly to previous studies,[40] we observed differences in the effects of lower SEP and education on mortality by sex, where the hazard rate was higher for men than for women. Additionally, differences in mortality depended on the socio-economic indicator evaluated, where effects were higher when assessing education compared with SEP. This finding supports the argument to assess SEP and education as distinct socio-economic indicators and not interchangeably.[31] Differences

observed between countries, specifically the UK (Whitehall II), compared with others could be partly explained by differences in health behaviours in northern and southern European regions.[41,42]

The effect of education or SEP on all-cause mortality was partially mediated by smoking, alcohol consumption, adherence to a Western dietary pattern, physical activity, BMI and co-morbidities of hypertension, diabetes and CAD, as jointly modelled. This finding is supported by previous studies that estimated potential mediated paths between socio-economic indicators and all-cause mortality, where, in a recent systematic review, smoking, alcohol consumption, physical activity and dietary patterns (assessed separately) contributed independently to socio-economic-associated mortality.[16] The indirect effects were stronger when assessing education compared with SEP, where up to 38% and 34% of the effect of education on mortality was mediated by the joint mediators for both men and women, respectively. A previous study, using a similar counterfactual framework, found that educational attainment had mediating and interacting effects with health behaviours (assessed separately) of smoking, alcohol intake, physical activity and BMI on cause-specific mortality.[17] Sex-based differences in the direct and indirect effects could be attributed to geographical differences, as variations in socio-economic-related behaviours by sex have been previously demonstrated.[43] Differences could also be attributed to sex-related patterns of intermediate lifestyle behaviours, as it has been previously demonstrated that attitudes and behaviours in lifestyle factors differ between women and men.[44]

Notably, we demonstrate the hazard gradient of mortality would substantially decrease if there were hypothetical interventions to raise SEP or education from a lower to higher status, where effects were larger with increased education. Through a scenario of higher education, total effects on mortality could be reduced by 15% for women and by 29% for men. This could be reduced by 5% via the indirect path of the mediators considered and by 10% via all other pathways by for women and by 11% via the indirect path of the mediators considered and by 18% via all other pathways for men by increasing population-level education. Direct effect estimates may differ from previous findings in the same and/or similar populations,[5,23] likely due, at least in part, to our joint mediation approach. In studies where only one mediator is accounted for at a time, the assumption is that each mediator represents independent and non-intertwined causal pathways and, thus, the direct effect associated with one mediator could partially constitute the indirect effect of another mediator.[45]

This study is not without limitations. Our hypothesis focused on two socio-economic indicators. This, along with the binary modelling of SEP and education, may limit capturing the full social stratification process and socio-economic-gradients, and an underestimate of the effects on all-cause mortality. Additionally, a single measurement of occupation for SEP may not capture the dynamics of multiple socio-economic factors, such as interactions with income. There is the potential for unmeasured confounders and/or other casual intermediates not assessed, such as early-life socio-economic and health factors, social networks, access to healthcare, psychosocial factors and other morbidities such as cancer. This may lead to an underestimation of mediating factors and an overestimation of the NDE of SEP or education on all-cause mortality. We combined cohort data from different time periods and geographical locations, and assessed the mediators at one time point; such an approach has limitations.[46] We may not have captured historical differences in environments (e.g. dietary patterns), there might be varying degrees in the precision of the measurements and we did not capture events occurring between observations, which collectively may affect the results of the study. However, most of the data were collected using similar protocols, limiting differences in exposure, mediators and covariates across cohorts, and we did account for potential cohort effects in the models. Our study did not account for certain dynamics within longitudinal settings where the exposures, mediators and confounders (and subsequent relationships between them) may be time-varying. This is particularly important to consider for life-course epidemiology and for the DAGs developed to inform analyses, where we tend to focus on the time for which observations are available in data.[47] Methods to address multi-time-point data using causal inference within a survival context have proven difficult. At least one method[48] has recently been put forth to estimate the direct and indirect effects integrated over time, producing cumulative effect estimates using an additive hazards model. However, application of this method to the present study is limited, as it has not been extended to cover multiple mediators. Additionally, as is typical with many longitudinal studies,[49] we had limited time-varying information in our data, underscoring an anticipated area for future research. There is the potential for measurement error of the exposure and/or mediators and there are limitations in dichotomizing mediators, possibly leading to residual confounding and underestimation of the indirect effects, and overestimation of the direct effects,[50] though this is not always so. Furthermore, there are other multi-mediator methods within causal inference to estimate interventions, including estimating the controlled direct effects,[51] and (in)direct interventional effects.[52] Currently, these have not been applied to settings with more than two mediators and these methods may provide less information regarding the causal mechanisms, making the estimates

potentially less generalizable. Future studies will need to consider such limitations, including the potential for exposure–mediator, exposure–covariate and mediator–mediator interactions. It is possible that intermediate behaviours and morbidities assessed in the present study may act as both a confounder and also a mediator, particularly if they are measured at the same time; e.g. one's BMI could influence their SEP.[21] Given that all mediators (except for diet in some cohorts) were assessed post baseline when the exposures were assessed, it is unlikely that they are confounders of the relationship between socio-economic indicators and mortality; however, to fully address this, we would need information on the mediators collected before the exposure.[53] Future studies will benefit from having multiple waves with repeated exposure and potential mediator/confounding data. Lastly, we postulate a hypothetical intervention on SEP or education, but do not define or operationalize how such an intervention could be formulated for policies. Raising SEP or education requires interventions with likely imperfect compliance, and education and SEP are themselves intermediate roles between policy measures (such as incentives) and mortality. Therefore, the quantitative results of this study enhance the understanding of inequalities in health, but should not be taken literally for policy considerations. Despite these limitations, the effect estimates in the present study are robust and important for understanding etiological pathways of socio-economic-related all-cause mortality.

A major strength of the study includes the use of counterfactual NEM to address several issues that may ensue from using traditional methods,[54] such as the difference and product methods for mediation. NEM models allow for the decomposition of effects in the presence of exposure/mediator interaction, overcoming a limitation of difference and product methods, and structural equation modelling.[55] Notably, we attempt to eliminate sources of bias introduced from adjusting for mediators[15] and issues that arise when ignoring more than one mediating pathway (i.e. interest in only one mediator) by evaluating joint mediation 'en bloc'.[37] This approach is particularly useful when the directionality between multiple mediators is in question, is robust to unmeasured common causes of mediators and is identifiable even if we relax assumptions to allow for unmeasured confounding.[56] Additionally, joint mediation can ease computational challenges for path-specific decomposition within the constraints of current mediation methods. When earlier mediators may affect subsequent ones, there are $(2n)!$ (where n is the number of mediators considered) ways of decomposing a total causal effect into a sum of path-specific effects.[57] Thus, assessing multiple mediators can become cumbersome, exemplified by the eight mediators in the present study resulting in 256 path-specific effects. However, a con to this approach is that we do not know the entire decomposition of the mediated effects. We assessed mediation as a whole where the indirect effect is through at least one (or all) of the mediators and the effect is not further disentangled into the contribution of all specific pathways through which a single mediator transmits the exposure effect on the outcome. Even in light of assessing path-specific multiple mediators, another consideration still in question is the ability to meet the identification criteria for conditional path-specific effects and is an area for future research.[56] An additional strength is the assessment of both education and occupational positions as socio-economic indicators. Education and occupation indicate different underlying concepts that may lead to distinct causal pathways. Additionally, based on a sensitivity analysis, we find that marital status may act as an effect modifier of the relationship between socio-economic indicators and all-cause mortality. Future studies, with more power, will need to further investigate this.

In conclusion, our findings validate the role of lower SEP and education on all-cause mortality. Our results also suggest that the lifestyle intermediates, BMI and co-morbidities contribute to an indirect effect and an increased HR for all-cause mortality. This work, alongside other models of social causality, can assist in inductive causal reasoning of socially patterned diseases and mortality risks.

Acknowledgements

LIFEPATH Consortium (in alphabetical order): Harri Alenius, Mauricio Avendano, Valéria Baltar, Mel Bartley, Henrique Barros, Murielle Bochud, Cristian Carmeli, Luca Carra, Raphaele Castagné, Marc Chadeau-Hyam, Franç Oise Clavel-Chapelon, Giuseppe Costa, Emilie Courtin, Cyrille Delpierre, Angela Donkin, Angelo D'Errico, Pierre-Antoine Dugué, Paul Elliott, Giovanni Fiorito, Silvia Fraga, Valérie Garès, Martina Gandini, Graham Giles, Marcel Goldberg, Dario Greco, Allison Hodge, Maryam Karimi, Michelle Kelly-Irving, Piia Karisola, Mika Kivimaki, Vittorio Krogh, Jessica Laine, Thierry Lang, Richard Layte, Benoit Lepage, Johan Mackenbach, Michael Marmot, Carlos de Mestral, Cathal McCrory, Roger Milne, Peter Muennig, Wilma Nusselder, Salvatore Panico, Dusan Petrovic, Silvia Polidoro, Martin Preisig, Olli Raitakari, Ana Isabel Ribeiro, Fulvio Ricceri, Erica Reinhard, Oliver Robinson, Jose Rubio Valverde, Carlotta Sacerdote, Roberto Satolli, Gianluca Severi, Martin J Shipley, Silvia Stringhini, Rosario Tumino, Joannie Tieulent, Salvatore Vaccarella, Anne-Claire Vergnaud, Paolo Vineis, Peter Vollenweider, and Marie Zins.

Supplementary data

Supplementary data are available at *IJE* online.

Funding

This work was supported by the European Commission under Horizon 2020 (grant number 633666) and J.E. Laine is supported

by a Medical Research Council Early Career Fellowship (MR/M501669/1).

Conflict of interest: None declared.

References

1. Chetty R, Stepner M, Abraham S *et al*. The association between income and life expectancy in the United States, 2001-2014. *JAMA* 2016;**315**:1750–66.
2. Gallo V, Mackenbach JP, Ezzati M *et al*. Social inequalities and mortality in Europe: results from a large multi-national cohort. *PLoS One* 2012;**7**:e39013.
3. Mackenbach JP, Stirbu I, Roskam AJ *et al*. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008;**358**:2468–81.
4. Stringhini S, Sabia S, Shipley M *et al*. Association of socioeconomic position with health behaviors and mortality. *JAMA* 2010;**303**:1159–66.
5. Stringhini S, Carmeli C, Jokela M *et al*. Socioeconomic status and the 25×25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *Lancet* 2017;**389**:1229–37.
6. Nunes BP, Thumé E, Tomasi E, Duro SM, Facchini LA. Socioeconomic inequalities in the access to and quality of health care services. *Rev Saúde Pública* 2014;**48**:968–76.
7. Gilman SE, Abrams DB, Buka SL. Socioeconomic status over the life course and stages of cigarette use: initiation, regular use, and cessation. *J Epidemiol Community Health* 2003;**57**:802–08.
8. Bellis MA, Hughes K, Nicholls J, Sheron N, Gilmore I, Jones L. The alcohol harm paradox: using a national survey to explore how alcohol may disproportionately impact health in deprived individuals. *BMC Public Health* 2016;**16**:111.
9. Darmon N, Drewnowski A. Does social class predict diet quality? *Am J Clin Nutr* 2008;**87**:1107–17.
10. Chinn DJ, White M, Harland J, Drinkwater C, Raybould S. Barriers to physical activity and socioeconomic position: implications for health promotion. *J Epidemiol Community Health* 1999;**53**:191–92.
11. Redonnet B, Chollet A, Fombonne E, Bowes L, Melchior M. Tobacco, alcohol, cannabis and other illegal drug use among young adults: the socioeconomic context. *Drug Alcohol Depend* 2012;**121**:231–39.
12. Evans GW, Kantrowitz E. Socioeconomic status and health: the potential role of environmental risk exposure. *Annu Rev Public Health* 2002;**23**:303–31.
13. Muntaner C, Eaton WW, Miech R, O'Campo P. Socioeconomic position and major mental disorders. *Epidemiol Rev* 2004;**26**:53–62.
14. Kaufman JS, Maclehose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiol Perspect Innov* 2004;**1**:4.
15. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002;**31**:163–65.
16. Petrovic D, de Mestral C, Bochud M *et al*. The contribution of health behaviors to socioeconomic inequalities in health: a systematic review. *Prev Med* 2018;**113**:15–31.
17. Nordahl H, Lange T, Osler M *et al*. Education and cause-specific mortality: the mediating role of differential exposure and vulnerability to behavioral risk factors. *Epidemiology* 2014;**25**:389–96.
18. Lantz PM, House JS, Lepkowski JM, Williams DR, Mero RP, Chen J. Socioeconomic factors, health behaviors, and mortality: results from a nationally representative prospective study of US adults. *JAMA* 1998;**279**:1703–08.
19. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009;**20**:488–95.
20. Pearl J. An introduction to causal inference. *Int J Biostat* 2010;**6**:7.
21. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction*, 1st edn. Oxford: Oxford University Press, 2015.
22. Vineis P, Avendano-Pabon M, Barros H *et al*. The biology of inequalities in health: the LIFEPATH project. *Longit Life Course Stud* 2017; **8**:33.
23. d'Errico A, Ricceri F, Stringhini S *et al*. Socioeconomic indicators in epidemiologic research: a practical example from the LIFEPATH study. *PLoS One* 2017;**12**:e0178071.
24. Kaufman JS. Commentary: causal inference for social exposures. *Annu Rev Public Health* 2019;**40**:7–21.
25. Mayén AL, Marques-Vidal P, Paccaud F, Bovet P, Stringhini S. Socioeconomic determinants of dietary patterns in low- and middle-income countries: a systematic review. *Am J Clin Nutr* 2014;**100**:1520–31.
26. Gidlow C, Johnston LH, Crone D, Ellis N, James D. A systematic review of the relationship between socio-economic position and physical activity. *Health Educ J* 2006;**65**:338–67.
27. Devaux M, Sassi F. Social inequalities in obesity and overweight in 11 OECD countries. *Eur J Public Health* 2013;**23**:464–69.
28. Williams J, Allen L, Wickramasinghe K, Mikkelsen B, Roberts N, Townsend N. A systematic review of associations between non-communicable diseases and socioeconomic status within low- and lower-middle-income countries. *J Glob Health* 2018;**8**:020409.
29. Sommer I, Griebler U, Mahlknecht P *et al*. Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. *BMC Public Health* 2015;**15**:914.
30. Di Cesare M, Khang YH, Asaria P *et al*. Inequalities in non-communicable diseases and effective responses. *Lancet* 2013;**381**:585–97.
31. Geyer S, Hemström O, Peter R, Vågerö D. Education, income, and occupational class cannot be used interchangeably in social epidemiology: empirical evidence against a common practice. *J Epidemiol Community Health* 2006;**60**:804–10.
32. Team RC. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2017. https://www.R-project.org/.
33. Alboukadel K, Kosinski M., Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. In R package version 0.4.3, 2019. https://CRAN.R-project.org/package=survminer.
34. Lange T, Hansen KW, Sørensen R, Galatius S. Applied mediation analyses: a review and tutorial. *Epidemiol Health* 2017;**39**:e2017035.
35. Pearl, J. Direct and indirect effects. In: Breese J, Koller D (eds). *Proceedings of the Seventeenth Conference on Uncertainty in*

Artificial Intelligence (UAI-01). San Francisco, CA: Morgan Kaufmann, 2001, pp. 411–20.
36. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology* 1992;**3**:143–55.
37. Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol* 2012;**176**:190–95.
38. Lange T, Rasmussen M, Thygesen LC. Assessing natural direct and indirect effects through multiple pathways. *Am J Epidemiol* 2014;**179**:513–18.
39. Steen J, Loeys T, Moerkerke B, Vansteelandt S. Flexible mediation analysis with multiple mediators. *Am J Epidemiol* 2017; **186**:184–93.
40. Mustard CA, Etches J. Gender differences in socioeconomic inequality in mortality. *J Epidemiol Community Health* 2003;**57**: 974–80.
41. Stringhini S, Dugravot A, Shipley M *et al*. Health behaviours, socioeconomic status, and mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. *PLoS Med* 2011;**8**:e1000419.
42. Mackenbach JP, Valverde JR, Artnik B *et al*. Trends in health inequalities in 27 European countries. *Proc Natl Acad Sci U S A* 2018;**115**:6440–45.
43. Read S, Grundy E, Foverskov E. Socio-economic position and subjective health and well-being among older people in Europe: a systematic narrative review. *Aging Ment Health* 2016;**20**:529–42.
44. Varì R, Scazzocchio B, D'Amore A, Giovannini C, Gessani S, Masella R. Gender-related differences in lifestyle may affect health status. *Ann Ist Super Sanita* 2016;**52**:158–66.
45. Imai K, Yamamoto T. Identification and sensitivity analysis for multiple causal mechanisms: revisiting evidence from framing experiments. *Polit Anal* 2013;**21**:141–71.
46. Allen J, Inder KJ, Lewin TJ *et al*. Integrating and extending cohort studies: lessons from the eXtending Treatments, Education and Networks in Depression (xTEND) study. *BMC Med Res Methodol* 2013;**13**:122.
47. Aalen OO, Røysland K, Gran JM, Kouyos R, Lange T. Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms. *Stat Methods Med Res* 2016;**25**: 2294–314.
48. Aalen OO, Stensrud MJ, Didelez V, Daniel R, Røysland K, Strohmaier S. Time-dependent mediators in survival analysis: modeling direct and indirect effects with the additive hazards model. *Biom J* 2019.
49. Lange T, Rod NH. Commentary: causal models adjusting for time-varying confounding—please send more data. *Int J Epidemiol* 2019;**48**:254–67.
50. Ogburn EL, VanderWeele TJ. On the nondifferential misclassification of a binary confounder. *Epidemiology* 2012;**23**:433–39.
51. VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. *Epidemiol Methods* 2014;**2**:95–115.
52. Vansteelandt S, Daniel RM. Interventional effects for mediation analysis with multiple mediators. *Epidemiology* 2017;**28**: 258–65.
53. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019;**34**:211–19.
54. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**: 1173–82.
55. VanderWeele TJ. Invited commentary: structural equation models and epidemiologic analysis. *Am J Epidemiol* 2012;**176**: 608–12.
56. Steen J, Vansteelandt S. Graphical models for mediation analysis. arXiv:2018.
57. Daniel RM, De Stavola BL, Cousens SN, Vansteelandt S. Causal mediation analysis with multiple mediators. *Biometrics* 2015; **71**:1–14.