Research Article

Health inequalities: embodied evidence across biological layers

Paolo Vineis[1,2], Cyrille Delpierre[3], Raphaële Castagné[3], Giovanni Fiorito[2], Cathal McCrory[4], Mika Kivimaki[5], Silvia Stringhini[6], Cristian Carmeli[7], Michelle Kelly-Irving[3]

1. MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, W2 1PG UK.
2. Italian Institute for Genomic Medicine, Torino, Italy.
3. UMR LEASP, Université de Toulouse III, UPS, Inserm, Toulouse, France.
4. Department of Gerontology, Trinity College Dublin, Ireland.
5. Department of Epidemiology, University College London.
6. Unit of Population Epidemiology, Division of Primary Care, Geneva University Hospitals, Geneva, Switzerland
7. Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland.

Correspondence and requests for materials should be addressed to Paolo Vineis, School of Public Health, Imperial College, Norfolk Place, W2 1PG London UK, e-mail p.vineis@imperial.ac.uk

Abstract

Rationale - Socioeconomic disparities have been documented in major non-communicable diseases and in their risk factors, such as obesity, hypertension, diabetes, smoking, physical inactivity, unhealthful diet and heavy drinking. However, a key research question has remained unanswered: is there a separate biological embodiment of socio-economic conditions underlying health disparities, additional and independent of those risk factors? As lifelong socioeconomic circumstances cannot be randomised, one way forward is the examination of different biological layers of evidence, including molecular changes.

Method - In this methodological paper we report the association of socio-economic disadvantage with (a) long-term health outcomes, before and after taking risk factors into account; (b) biological intermediaries that increase susceptibility to disease, such as childhood obesity; (c) intermediate circulating biomarkers and omic measurements (transcriptomics, DNA methylation, inflammatory proteins, allostatic load); and (d) immunity. In our Lifepath consortium, these analyses have been performed in several cohort studies, countries and contexts, and at different stages of the life course in up to 1.7 million subjects. The main goal is to test the assumption that each layer (death, functional outcomes, DNA, RNA, proteins, infections) is characterized by different types of bias and confounding, and that consistency across layers reinforces causality assessment.

Results - The findings show consistent associations of social disparities with unfavourable health outcomes spanning inflammatory biomarkers, DNA or RNA-based markers, infection, indicators of physical functioning and mortality. Although each of these associations has a different set of confounders, a dose-response relationship is nevertheless consistently observed, thus showing the power of our multi-layered approach.

Conclusions - This new evidence supports biological embodiment of social disadvantage, in addition to the impact of known (mainly behavioural) risk factors for disease.

Key words: socio-economic disparities, health, aging, embodiment, biomarkers, omics, triangulation

Abbreviations: RF=risk factor, WHIP=Work and Health Histories Italian Panel, NCD=non-communicable diseases, RNA=ribonucleic acid, DNA=deoxyribonucleic acid, COLAUS= Cohorte Lausannoise, AA=age acceleration;

1. Introduction

The multi-layered social environment within which humans exist and live ultimately affects cells, organs, and biological systems (Blane et al, 2013). This concept, known as *embodiment,* was initially developed by Nancy Krieger at the beginning of the century (Krieger, 2005). Establishing a causal connection between social and biological environments is difficult and adds to the existing complexities of causal reasoning in the biological domain alone (Lawlor et al, 2016; Vineis et al, 2017; Munafo & Davey Smith, 2018).

Here, we apply a biologically multi-layered approach to a specific research question, that has been extensively debated: "is there a separate effect of socio-economic conditions on health that is independent of known behavioural risk factors for disease?". Do disadvantaged socio-economic conditions have a direct effect on health and disease not entirely explained by exposure to known risk factors (Marmot, 2004)? For simplicity, here we refer to the main risk factors for non-communicable diseases (NCD), also known as 25x25 risk factors according to the WHO and United Nations strategy (obesity, hypertension, smoking, diabetes, physical inactivity, alcohol intake) (United Nations, 2011). In our previous research we have already suggested that these risk factors explain only a fraction of mortality gradients by socio-economic position (Stringhini et al, 2017). A first aim of this paper is to strengthen our previous observations by referring to several underlying biological layers, beyond a hard outcome like mortality.

A second aim is to reinforce our understanding of causal relationships between social exposures and health through intermediate biological mechanisms.

Observational biomedical research struggles to identify the genuine causal relationships among correlations. This distinction is even more critical for research conducted at the interface between social and biological determinants of health and disease. One proposal to strengthen causal inference in biomedicine is through "triangulation" (Lawlor et al, 2016; Munafo & Davey Smith, 2018). This is "the strategic use of multiple approaches to address one question. Each approach has its own unrelated assumptions, strengths and weaknesses. Results that agree across different methodologies are less likely to be artefacts" (Lawlor et al, 2016). Ideally, results from more than two approaches - which have different and unrelated key sources of potential biases — are compared (Munafo & Davey Smith, 2018). An additional feature is the mixing of qualitative and quantitative types of evidence that bring breadth and depth to the

same research question. Triangulation in research is by no means a new concept, being a staple approach in qualitative research to ascertain the convergence of results and examine rival explanations (Patton, 1999).

In this paper we apply a biologically multi-layered approach to address the consistency of findings across measures of socio-economic conditions and health from populations to molecules, under the assumption that each layer (death, functional outcomes, DNA, RNA, proteins, infections) is characterized by different types of bias and confounding. The paper is based on a large consortium of cohorts, Lifepath, that has brought together data on socio-economic position and health from up to 1.7 million individuals, including extensive data on biomarkers. To show that socio-economic position has an impact at different layers of biology (before reaching the mortality end-point) we use here published results from Lifpeath, that have not been considered as a whole before (under the assumption that the whole is greater than the sum of the parts). To achieve this we have analysed four sets of evidence: (a) the effect of socioeconomic disadvantage measured using different indicators on long-term outcomes, before and after adjustment for behaviourally-related risk factors (b) the effect of socio-economic disadvantage on intermediate biological markers that increase susceptibility to disease, such as childhood obesity; (c) the effect of socioeconomic disadvantage on intermediate blood biomarkers known to be associated with health or ageing and at different omic scales (transcriptomics, DNA methylation, inflammatory proteins, allostatic load); and (d) the impact of socioeconomic disadvantage on risk of infection.

2. Method

2.1. Populations

The Lifepath consortium includes information and biological samples from 8 longitudinal population-based cohorts with extensive phenotyping and repeat biological samples, 9 large longitudinal population-based cohorts with biological samples, and a very large cohort without biological samples (WHIP). Cohorts from the following countries are represented: France, Italy, Portugal, Ireland, UK, Finland, Switzerland and Australia. These cohorts are only a small proportion of all cohorts available in Europe, but we have chosen them for the combination of good measures of socio-economic position, risk factors for NCD and existing biomarkers. The majority of cohorts include "hard" outcomes (diabetes, cancer, cardiovascular disease - CVD, total mortality), and the extensively phenotyped cohorts also include several measurements of

the functional components of healthy ageing, including frailty, impaired vision, cognitive function, renal and brain function, osteoporosis, sleep disturbances and mental health. All age groups are represented with two birth cohorts, one cohort of adolescents and several cohorts encompassing young adults (age 18 and above). Furthermore, there is a strong representation of elderly subjects in 7 cohorts.

For the analyses on mortality we have complemented the Lifepath cohorts with publicly available data from 38 additional cohort studies from the Inter-University Consortium for Political and Social Research and the UK Data Service (Stringhini et al, 2017).

2.2. Socio-economic variables

Our mortality analyses were based on participants whose occupational position was assessed at baseline between 1965 and 2009, depending on the study. All studies included baseline data for socioeconomic position and a mortality follow-up of a minimum of 3 years. We assessed the quality of included studies using the Cochrane Risk of Bias Tool for cohort studies (Stringhini et al, 2017). Our measure of socioeconomic condition is a social class measure based on an individual's last known occupational title at study enrolment, coded into the European Socio-economic Classification (ESEC). This variable was predefined and harmonised across the study cohorts before statistical analyses were performed. Occupational position was aggregated into three categories: higher professionals and managers, higher clerical, services, and sales workers [ESEC class 1, 2, and 3]; small employers and self-employed, farmers, lower supervisors, and technicians [ESEC class 4, 5, and 6]; and lower clerical, services and sales workers, skilled workers, and semi-skilled and unskilled workers [ESEC class 7, 8, and 9]. The use of the last job may entail some degree of misclassification, particularly for people who were not employed in the formal economy at baseline, but then held a position at some point in their lifecourse. Also, we did not make an effort to classify people on the basis of the occupational title of their spouse.

Occupation was used to categorize individuals by social class also for walking speed and all blood-based biomarkers, while in other analyses we used indicators of socio-economic position based on education (education for PRC, Berger et al, 2019; and maternal education for children obesity and height, Table 1).

2.3. Confounding

As we discuss later, adjustment for confounders is not necessarily the best analytical strategy, since SEP is an overarching determinant whose effect on health is likely to me mediated at least in part by risk factors such as smoking or BMI. Proper mediation analysis has been performed elsewhere (Laine et al,2020). Here we aimed to apply an extreme assumption, i.e. to observe the degree of attenuation of the association between SEP and different outcomes after adjustment for several risk factors assumed to be confounders. We used behavioural risk factors for non-communicable diseases to address confounding. Briefly, the 2013–20 World Health Organization (WHO) Global Action Plan for the Prevention and Control of Non-Communicable Diseases (NCDs) targets seven major risk factors, comprising the harmful use of alcohol, insufficient physical activity, current tobacco use, raised blood pressure, intake of salt or sodium, diabetes, and obesity (referred to as the 25×25 risk factors), with the overall aim of reducing premature mortality from noncommunicable diseases by 25% by 2025 (United Nations, 2011; Pearce et al, 2014). For details about categories used for these risk factors we refer to previous papers (Stringhini et al, 2017; Stringhini et al, 2018). Whenever available, we used the same set of potential confounders in all the analyses shown in Table 1.

2.4. Outcomes

Participants' data were linked to national mortality registries that provided information about vital status with the exception of the COLAUS study in which vital status was ascertained through active follow-up. Mean follow-up for mortality ranged between 3.2 years in the National Health Interview Survey 2009, and 27.0 years in men and 29.5 years in women of the Alameda County Study 1965, with a mean across cohorts of 13.3 years [SD 6.4 years]. We refer to previous papers for the measurement of walking speed (Stringhini et al, 2018), obesity in children (McCrory et al, 2017), transcriptomics (Castagne et al, 2016), allostatic load (Barboza Solis et al, 2016), methylation (Fiorito et al, 2017) and EBV infection (Gares et al, 2017).

2.5. Statistical analyses

To address the question of the independent role of socio-economic conditions on health outcomes after adjustment for behavioural risk factors for NCDs, we have analysed four sets of evidence from the Lifepath consortium: (a) the effect of disadvantaged socio-economic conditions on long-term outcomes, before and after adjustment for risk factors; (b) the effect of disadvantaged socio-economic conditions on intermediate biological markers that increase susceptibility to disease, such as childhood obesity; (c) the effect of disadvantaged socio-economic conditions on intermediate blood biomarkers known to be associated with health or ageing and at different omic scales (transcriptomics, DNA methylation, inflammatory proteins,

allostatic load); and (d) the impact of disadvantaged socio-economic conditions on risk of infection.

We present a composite set of results together, which have each been previously published separately.

For papers based on multiple cohorts (Stringhini et al, 2018; Stringhini et al, 2017), analyses were first performed separately in each study, and estimates were subsequently combined in a meta-analysis. In study-specific analyses, we considered the maximum number of participants without missing values for each exposure. To estimate the association between risk factors and mortality, hazard ratios (HR) and years of life lost (YLL) were generated using flexible parametric survival models on the cumulative hazards scale (details in Stringhini et al, 2017). The mutually adjusted models included age, sex, race or ethnicity, marital status, socioeconomic status, and all risk factors as independent variables with total mortality and deaths from cardiovascular disease, cancer, and other causes as outcomes. To enable balanced comparisons between socioeconomic status and risk factors as predictors of cause-specific mortality, these analyses were restricted to a subgroup of participants with complete data for socioeconomic status and the NCD risk factors.

To examine whether the association between socioeconomic status and mortality is attributable to the higher prevalence of the risk factors among disadvantaged socioeconomic position individuals, we repeated the analyses in a subgroup of participants without any risk factors. Analyses were also repeated specifically focusing on premature mortality (<70 years) and by restricting the population to cohorts in which height and weight as well as blood pressure were measured objectively using standard procedures.

The same procedures were used for other outcomes and particularly for walking speed (Stringhini et al, 2018). In the latter, we computed Years of Functioning Lost (YFL) as the difference between the chronological age of the unexposed group equivalent to the walking speed at age 60 (or 85) of the exposed group and age 60 (or 85).

Specific methods used to analyse data on obesity in children (McCrory et al, 2017), transcriptomics (Castagne et al, 2016), allostatic load (Barboza Solis et al, 2016), methylation (Fiorito et al, 2017) and EBV infection (Gares et al, 2017) are reported in the original papers. The present paper is methodological and addresses the specific issue of consistency of findings across biological layers (characterized by different confounders and bias), with the general aim of testing whether SEP is associated to health indicators and biological markers after

adjustment by confounders. More details on methods and specific results are given in the original papers.

3. Results

3.1. Main findings

Results reported in Table 1 show considerable consistency across all findings in our consortium of up to 1.7 million people. Disadvantaged socio-economic conditions were associated with: increased mortality (in 48 cohorts in 3 continents) (Stringhini et al, 2017), increased functional impairment (decrease of walking speed, in 37 cohorts) (Stringhini et al, 2018), increased obesity in children in three cohorts, lower height in five cohorts, increased risk of Epstein Barr Virus (EBV) infection, and consistent changes in a number of intermediate markers (transcriptomics, inflammatory markers, allostatic load and DNA methylation-based age acceleration).

3.2. Robustness of findings

In each of the analyses we addressed the role of potential confounders. We looked for associations between risk factors and outcomes, which vary by outcome. By definition all "25x25" risk factors are associated with mortality, while determinants of impaired walking speed are partially different, involving musculoskeletal impairment that is not strongly associated with the NCD risk factors (Stringhini et al, 2018). Alcohol was associated with mortality but not with walking speed. Child height is associated with parental height, nutritional status and infectious diseases in infancy (McCrory et al, 2017). None of the NCD risk factors were associated with transcriptomics, inflammation markers or allostatic load in our data.

We also performed a specific search for potential confounders in the case of DNA-based age acceleration, since the latter was associated with SEP and with some NCD risk factors (Fiorito et al, 2018). We evaluated the change in the magnitude of the SEP-age acceleration (AA) association due to the inclusion of NCD risk factors in the models, by comparing the estimated effects (βs) of the basic regression model (adjusted for age and sex) with those of the models including different NCD risk factors (smoking status, BMI, alcohol intake, Mediterranean Diet Score, and physical activity). In the fully adjusted model, the effect size attenuation for the linear trend was 12% (95% CI: −0.02, 0.26; p = 0.09). The most substantial contribution to the reduction of the effect size for the SEP-AA association was due to smoking, which attenuated the estimate for linear trend by 6% (95% CI: −11, 22). Effect size reductions due to the inclusion of other NCD risk factors were small and non-significant.

Finally, we examined the SEP-AA associations after exclusion of individuals exposed to at least one NCD risk factor (i.e. non-obese, non-smokers, physically active, non-drinkers and with a healthy diet). Notably, the association of SEP with AA was still significant in this NCD risk factor free subsample (test for linear trend $\beta = 0.62$; 95% CI: 0.06, 1.18; p = 0.03; N = 1,549).

We also performed extensive sensitivity analyses – in addition to those described above for DNA methylation -, confirming the association of socio-economic conditions with health outcomes in different subsets of the overall sample and using different socio-economic indicators. These are reported in the original papers (Stringhini et al, 2017; Stringhini et al, 2018; McCrory et al, 2017; Castagne et al, 2016; Barboza Solis et al, 2016; Fiorito et al, 2017; Gares et al, 2017).

3.3. Dose-response

The observations above are reinforced by the findings presented in Figure 1, where we considered three levels of socio-economic conditions as defined in Methods. The Figure shows that for all indicators there was a dose-response relationship, hardly compatible with an explanation based on confounding from individual level variables.

4. Discussion

The main finding of this methodological paper is that socioeconomic disadvantage measured using different indicators, is a determinant of poor health through biological embodiment beyond classic behavioural risk factors for non-communicable diseases. We have found consistent results across several biological layers of evidence (deaths, functional outcomes, DNA, RNA, proteins, infection) , in a variety of population contexts, varying cohorts/periods, and using different socioeconomic and health measures. This is what Whewell has named "consilience of induction" (Munafo and Davey-Smith, 2018),

The biologically multi-layered approach we used here involves the ability of the different types of evidence to address different types of bias associated with each layer taken separately. The main criticism to the theory of socio-economic related poor ageing is confounding and residual confounding, the argument being that adjustment for confounders (including the factors we have included here, 25x25 NCD risk factors is imperfect because of measurement error. We believe our approach addresses at least in part these criticisms, since the outcomes we have considered have different sets of confounders and have been studied in different cohorts and

contexts. The main determinants of mortality are the 25x25 risk factors, though some are missing in the UN-WHO definition (United Nations, 2011; Pearce et al, 2014). Determinants of impaired walking speed are partially different, involving musculoskeletal impairment that is not strongly associated with the behavioural risk factors and with mortality itself (Stringhini et al, 2018). In our analysis, alcohol was associated with mortality but not with walking speed. Obesity is an important cause of premature death and disease, starting in childhood, with recent research suggesting it represents a type of low-grade chronic inflammation (McCrory et al, 2017). Again, its determinants are different from those that lead to other outcomes, and obesity is in fact a strong mediator between childhood socio-economic conditions and poor health in adulthood. Finally, we investigated a set of biological markers that may lie on the causal chain from socio-economic disadvantage to poor health. Confounders that we identified were BMI and smoking for methylation markers, but none of the behavioural risk factors for transcriptomics, inflammation markers or allostatic load. The latter markers suggest a central role for inflammation in the pathways leading from disadvantaged socio-economic conditions to poor health. In addition, we performed extensive sensitivity analyses confirming the association of socio-economic conditions with health outcomes in different subsets of the overall sample and using different socio-economic indicators.

In brief, the indicators we studied in association with disadvantaged socio-economic conditions had different sets of confounders. **Figure 1** shows that for all indicators for which information was available there was a dose-response relationship, hardly compatible with an explanation based on confounding. Also, other types of bias might have affected the results of the studies (such as selection or reporting bias), but they could not equally affect this heterogeneous set of investigations, conducted with different designs and methods in different populations.

One limitation of some analyses is that they were cross-sectional (except those on mortality and child height and weight). However, in the consortium we also provided longitudinal evidence on the impact of socio-economic position in early life on diabetes in later life (Kivimaki et al, 2018). We examined 2048 individuals whose risk factor profile was assessed in childhood and in adulthood, and had a clinical examination at follow-up at age 33-48 years. High neighbourhood socioeconomic disadvantage was characterised by decreased homoeostasis model assessment insulin sensitivity, and increased fasting glucose and insulin concentration from early adulthood (27 years; all p<0·03). Individuals consistently exposed to high neighbourhood socioeconomic disadvantage were more likely to be obese (odds ratio [OR] 1·44, 95% CI 1·01-2·06), hypertensive (1·83, 1·14-2·93), have a fatty liver (1·73, 1·11-

2·71), and diabetes (3·71, 1·77-7·75), compared with those who were consistently exposed to low neighbourhood socioeconomic disadvantage (Kivimaki et al, 2018).

Finally, longitudinal evidence was also provided in the Consortium for the relationship between socio-economic disadvantage, allostatic load and mortality (Castagne et al, 2018). We used data from the 1958 British birth cohort including mortality in 8,113 adults after 11 years of follow-up. We analysed 14 biomarkers assayed in blood from a biosample collected at 44 years of age, that overall constituted a measure of allostatic load. Higher allostatic load at 44 years old was a significant predictor of mortality 11 years later [HR = 3.56 (2.3 to 5.53)]. We found that this relationship was related to early-life socio-economic position, adverse childhood experiences and young adulthood health status [HR = 2.57 (1.59 to 4.15)].

4.1. Study limitations

There are limitations in the current analysis. First, we did not consider here contextual (area-based) variables that are associated with socio-economic condition, whether pertaining to air pollution, residential segregation, income inequality, urban vs rural, etc (Beckfield, 2018). Second, , socio-economic position is overarching, i.e. it is not on the same level as other determinants, including environmental and behavioural risk factors. The latter are supposed to be intermediate between socio-economic conditions and outcomes. This causal problem is at the heart of the social-to-biological interface, and cannot be reduced to statistical adjustment (often implying over-adjustment)(Krieger, 2008). Disadvantaged socio-economic conditions simultaneously affect biological systems, exposure to risk factors from conception, and influence social mobility. Health outcomes in adulthood are influenced by exposures to risk factors across the lifecourse. The resources conferred by advantaged socio-economic environments may lead to the development of social and psychosocial strategies to buffer the physiological and psychological challenges over everyday life.

We did not include genetic data in our study. In fact, most epidemiological observational studies lack genetic data, which means that any confounding or bias from genetics is part of the unmeasured confounding problem. By triangulating, and therefore using a variety of social measures (including parental education, social class, own education, area deprivation), combined with different health outcomes in different populations (cohorts, context, time), and observing the consistency of the findings across studies, we can tease out the common effect observed. This helps us to understand the extent of unmeasured confounding, which includes any genetic component. Selection bias and confounding are less likely if studies are replicated

from different countries or continents like in our case (e.g. Europe and Australia): in these countries the genetic background is not too dissimilar, but social and educational contexts are completely different, and educational attainment has considerably improved in the last decades, which goes against the idea that it is genetically driven.

5. Conclusions

In conclusion, evidence from the Lifepath consortium represents an example of triangulation, where different biological layers of evidence (from mortality to molecular markers) converge to suggest that disadvantaged socio-economic conditions are a determinant of poor health originating from biological embodiment beyond behavioural risk factors for non-communicable diseases.

What could be the effect of SEP beyond classic NCD risk factors (mainly behavioural) is a matter of speculation. A certain, probably small, proportion is likely to be related to occupational exposures, another proportion can be attributable to "wear-and-tear" and psychosocial stress (Marmot, 2004), as suggested by our findings on the allostatic load.

Conflicts of interest

The authors declare they have no conflicting interests.

References

Barboza Solís C, Fantin R, Castagné R, Lang T, Delpierre C, Kelly-Irving M., 2016. Mediating pathways between parental socio-economic position and allostatic load in mid-life: Findings from the 1958 British birth cohort. Soc Sci Med 165, 19-27, doi:10.1016/j.socscimed.2016.07.031.

Beckfield J, 2018. *Political sociology & the people's health*. New York: Oxford University Press.

Berger E, Castagné R, Chadeau-Hyam M, Bochud M, d'Errico A, Gandini M, Karimi M, Kivimäki M, Krogh V, Marmot M, Panico S, Preisig M, Ricceri F, Sacerdote C, Steptoe A, Stringhini S, Tumino R, Vineis P, Delpierre C, Kelly-Irving M, 2019. Multi-cohort study identifies social determinants of systemic inflammation over the life course. Nat Commun. Feb 15;10(1):773. doi: 10.1038/s41467-019-08732-x.

Blane D, Kelly-Irving M, d'Errico A, Bartley M, Montgomery S, 2013. *Social-biological transitions: how does the social become biological?* Longitudinal and Life Course Studies. Volume 4 Issue 2 Pp 136 – 146. doi: http://dx.doi.org/10.14301/llcs.v4i2.236;

Castagné R, Kelly-Irving M, Campanella G, Guida F, Krogh V, Palli D, Panico S, Sacerdote C, Tumino R, Kleinjans J, de Kok T, Kyrtopoulos SA, Lang T, Stringhini S, Vermeulen R, Vineis P, Delpierre C, Chadeau-Hyam M., 2016. Biological marks of early-life socioeconomic experience is detected in the adult inflammatory transcriptome. Sci Rep 6, 38705, doi:10.1038/srep38705.

Castagné R, Garès V, Karimi M, Chadeau-Hyam M, Vineis P, Delpierre C, Kelly-Irving M, Lifepath Consortium, 2018. Allostatic load and subsequent all-cause mortality: which biological markers drive the relationship? Findings from a UK birth cohort. Eur J Epidemiol. May;33(5):441-458.

Fiorito G, Polidoro S, Dugué PA, Kivimaki M, Ponzi E, Matullo G, Guarrera S, Assumma MB, Georgiadis P, Kyrtopoulos SA, Krogh V, Palli D, Panico S, Sacerdote C, Tumino R, Chadeau-Hyam M, Stringhini S, Severi G, Hodge AM, Giles GG, Marioni R, Karlsson Linnér R, O'Halloran AM, Kenny RA, Layte R, Baglietto L, Robinson O, McCrory C, Milne RL, Vineis P., 2017. Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. Sci Rep 7, 16266, doi:10.1038/s41598-017-16391-5.

Gares, V., Panico, L., Castagne, R., Delpierre, C. & Kelly-Irving, M, 2017. The role of the early social environment on Epstein Barr virus infection: a prospective observational design using the Millennium Cohort Study. Epidemiol Infect 145, 3405-3412, doi:10.1017/S0950268817002515.

Kivimäki M, Vahtera J, Tabák AG, Halonen JI, Vineis P, Pentti J, Pahkala K, Rovio S, Viikari J, Kähönen M, Juonala M, Ferrie JE, Stringhini S, Raitakari OT, 2018. Neighbourhood socioeconomic disadvantage, risk factors, and diabetes from childhood to middle age in the Young Finns Study: a cohort study. Lancet Public Health. Aug;3(8):e365-e373

Krieger, N, 2005. Embodiment: a conceptual glossary for epidemiology. J Epidemiol Community Health 59, 350-355, doi:10.1136/jech.2004.024562.

Krieger, N, 2008. Proximal, distal, and the politics of causation: what's level got to do with it? Am J Public Health 98, 221-230, doi:10.2105/AJPH.2007.111278.

Laine, JE, Baltar VT, Chadeau-Hyam M, Stringini S, Lifepath Consortium, Vineis P, 2020. Reducing Mortality from Socioeconomic Inequalities. Int. J. Epidemiol., in press

Lawlor, D. A., Tilling, K. & Davey Smith, G, 2016. Triangulation in aetiological epidemiology. Int J Epidemiol 45, 1866-1886, doi:10.1093/ije/dyw314.

Marmot, M. G, 2004. The status syndrome: how social standing affects our health and longevity. 1st American edn, Times Books.

McCrory, C., Leahy, S., Ribeiro, A., Fraga, S., Barros, H., Vineis, P. & Layte, R. for the LIFEPATH consortium (2019). Maternal educational inequalities in measured body mass index trajectories in children in three European countries. Paediatr Perinat Epidemiol. 2019;33:226–237. https://doi.org/10.1111/ppe.12552

McCrory C, O'Leary N, Fraga S, Ribeiro AI, Barros H, Kartiosuo N, Raitakari O, Kivimäki M, Vineis P, Layte R; Lifepath Consortium., 2017. Socioeconomic differences in children's growth trajectories from infancy to early adulthood: evidence from four European countries. J Epidemiol Community Health 71, 981-989, doi:10.1136/jech-2016-208556.

Munafo, M. R. & Davey Smith, G, 2018. Robust research needs many lines of evidence. Nature 553, 399-401, doi:10.1038/d41586-018-01023-3.

Patton, M.Q., 1999. Enhancing the quality and credibility of qualitative analysis. Health services research 34 (5 Pt 2):1189-1208.

Pearce N, Ebrahim S, McKee M, Lamptey P, Barreto ML, Matheson D, Walls H, Foliaki S, Miranda J, Chimeddamba O, Marcos LG, Haines A, Vineis P., 2014. The road to 25x25: how can the five-target strategy reach its goal? Lancet Glob Health 2, e126-128, doi:10.1016/S2214-109X(14)70015-4

Stringhini, S., Carmeli, Jokela M, Avendaño M, Muennig P, Guida F, Ricceri F, d'Errico A, Barros H, Bochud M, Chadeau-Hyam M, Clavel-Chapelon F, Costa G, Delpierre C, Fraga S, Goldberg M, Giles GG, Krogh V, Kelly-Irving M, Layte R, Lasserre AM, Marmot MG, Preisig M, Shipley MJ, Vollenweider P, Zins M, Kawachi I, Steptoe A, Mackenbach JP, Vineis P, Kivimäki M, for the LIFEPATH consortium‡, 2017. Socioeconomic status and the 25 x 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. Lancet, doi:10.1016/S0140-6736(16)32380-7.
Stringhini S, Carmeli C, Jokela M, Avendaño M, McCrory C, d'Errico A, Bochud M, Barros H, Costa G, Chadeau-Hyam M, Delpierre C, Gandini M, Fraga S, Goldberg M, Giles GG, Lassale C, Kenny RA, Kelly-Irving M, Paccaud F, Layte R, Muennig P, Marmot MG, Ribeiro AI, Severi G, Steptoe A, Shipley MJ, Zins M, Mackenbach JP, Vineis P, Kivimäki M; LIFEPATH Consortium, 2018. Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: multi-cohort population based study. BMJ 360, k1046, doi:10.1136/bmj.k1046.

United Nations, 2011. High level meeting of the General Assembly on the Prevention and Control of Non-communicable Diseases. New York. http://www.who.int/nmh/events/un_ncd_summit2011/speeches/en/index.html

Vineis, P., Illari, P. & Russo, F, 2017. Causality in cancer research: a journey through models in molecular epidemiology and their philosophical interpretation. Emerg Themes Epidemiol 14, 7, doi:10.1186/s12982-017-0061-7.

Table 1. Associations between socio-economic condition and different health and biological outcomes. The reference category is always the less disadvantaged (for definition see text). Minimally adjusted estimates are adjusted by age and sex in most comparisons (race when relevant), and by age, age^2, height and year of birth in papers [9] and [10]. Fully-adjusted estimates are adjusted by 25x25 risk factors for non-communicable diseases (obesity, hypertension, smoking, diabetes, physical inactivity, alcohol intake), where appropriate. For studies in children models were adjusted for mother's age, parity, smoked during pregnancy, alcohol during pregnancy and marital status. All estimates are reported with their 95% confidence intervals.

Outcome	Cohorts – exposure – units	Male	Female
Relationship between disadvantaged socio-economic conditions & long-term outcomes			
Total mortality (Stringhini et al, 2017) (_countries:_ UK, France, Switzerland, Portugal, Italy, USA, Australia)	48 cohorts - most disadvantaged vs most advantaged social position (occupation) – Hazard Ratio (HR) and 95%CI in brackets	Minimally adjusted HR=1.42 (1.38 ; 1.45) Fully-adjusted HR=1.29 (1.23 ; 1.36)	Minimally adjusted HR=1.34 (1.28 ; 1.39) Fully adjusted HR=1.20 (1.13 ; 1.28)
Walking speed (meters/second) (Stringhini et al, 2018) (_countries_ UK, France, USA, Mexico, China, Ghana, India, Russia, South Africa, Costa Rica, Taiwan, Ireland, Austria, Belgium, Denmark, Germany, Greece, Israel, Italy, the Netherlands, Portugal, Spain, Sweden, and Switzerland)	37 cohorts - most disadvantaged vs most advantaged SEP (occupation) - Years of Function Lost (YFL) by age 60 (95%CI) estimated from a mixed model with walking speed as a continuous variable	Minimally adjusted YFL 6.6 (5.0-9.4). Fully adjusted YFL 4.8 (3.7 ; 6.5)	Minimally adjusted YFL 4.6 (3.6 ; 6.2) Fully adjusted YFL 3.3 (2.5 ; 4.4)
Relationship between disadvantaged socio-economic conditions & intermediate risk factors for later health outcomes			
Increase in Body Mass Index (2019)	G21 (Portugal) cohort aged 7 - most disadvantaged vs most advantaged - Increase in Kg/m^2 (95%CI)	BMI increase=0.42 (0.24; 0.60)	BMI increase=0.71 (0.50 ; 0.91)

	GUI (Ireland) cohort aged 13 - most disadvantaged vs most advantaged - Increase in Kg/m^2 (95%CI)	BMI increase=0.90 (0.60 ; 1.19)	BMI increase=1.31 (1.00 ; 1.62)
	MCS (UK) aged 14 - most disadvantaged vs most advantaged- Increase in Kg/m^2 (95%CI)	BMI increase=0.75 (0.52 ; 0.97)	BMI increase=0.76 (0.53 ; 1.00)
Decrease in height (McCrory et al, 2017)	G21 (Portugal) cohort aged 7 - most disadvantaged vs most advantaged - decrease in cm (95%CI)	Decrease in cm=0.49 (0.04 ; 0.94)	Decrease in cm= 0.05 (-0.50 ; 0.39)
	GUI (Ireland) cohort aged 13 - most disadvantaged vs most advantaged - decrease in cm (95%CI)	Decrease in cm=1.64 (0.83 ; 2.46)	Decrease in cm=1.58 (1.01 ; 2.14)
	MCS (UK) aged 11 - most disadvantaged vs most advantaged- decrease in cm (95%CI)	Decrease in cm=1.38 (1.04 ; 1.72)	Decrease in cm=0.94 (0.57 ; 1.31)
	Epiteen (Portugal) aged 13 - most disadvantaged vs most advantaged - decrease in cm (95%CI)	Decrease in cm=1.95 (0.48 ; 3.42)	Decrease in cm=2.46 (1.47 ; 3.44)
	Young Finns (Finland) aged 12 - most disadvantaged vs most advantaged - decrease in cm (95%CI)	Decrease in cm=0.63 (0.27 ; 1.53)*	Decrease in cm=1.62 (0.86 ; 2.38)
Relationship between socio-economic conditions & blood biomarkers			
Increase in allostatic load score (Barboza Solis et al, 2016)	NCDS (UK) social class IV & V semi-unskilled vs I & II professional/managerial – regression coefficient from multivariable regression β (95%CI) per unit increase in allostatic load score	β = 0.32 (0.09 ; 0.54)	β = 0.30 (0.06 ; 0.53)
	NCDS (UK) social class III skilled manual vs I & II professional/managerial - regression coefficient from multivariable regression β (95%CI) per unit increase in allostatic load score	β = 0.27 (0.09 ; 0.45)	β = 0.29 (0.09 ; 0.49)
	NCDS (UK) social class III skilled non-manual vs I & II professional/managerial - regression coefficient from multivariable	β = 0.05 (-0.19 ; 0.29)	β = 0.12 (-0.14 ; 0.39)

	regression β (95%CI) per unit increase in allostatic load score		
Increase in C reactive protein (Berger et al, 2019) (countries: Italy, Switzerland, UK)	Across 6 cohorts - Low vs high education level - regression coefficient from multivariable regression β (95%CI) per unit increase in CRP	Minimally adjusted: β = 0.28 (0.18; 0.37). Fully adjusted: β = 0.13 (0.05; 0.21)	Minimally adjusted: β = 0.33 (0.21; 0.45) Fully adjusted: β = 0.14 (0.04; 0.25)
	Across 6 cohorts - Medium vs high education level - regression coefficient from multivariable regression β (95%CI) per unit increase in CRP	Minimally adjusted: β = 0.15 (0.08; 0.21) Fully adjusted: β = 0.08 (0.03; 0.14)	Minimally adjusted: β = 0.15 (0.02; 0.27) Fully adjusted: β = 0.07 (-0.04; 0.18)
		Males & Females combined	
Risk of Infection with Epstein Barr Virus (Gares et al, 2017)	MCS (UK) aged 3 -most disadvantaged social class vs most advantaged (occupation) – Odds Ratio (OR) and 95%CI in brackets Fully adjusted model includes household socio-economic position, area environment home environment (overcrowding, temperature in the baby room), housing tenure	Minimally adjusted OR=1.24 (1.02 ; 1.50) Fully adjusted OR=1.15. (0.94 ; 1.41)	
Epigenetic age acceleration (DNA methylation) (Fiorito et al, 2017) (countries: Italy, Ireland, Australia)	3 cohorts - intermediate social class vs advantaged (occupation): increase in number of years (biological age acceleration)	Minimally adjusted Years increase=0.75 (0.17 ; 1.39) Fully adjusted Years increase= 0.78 (0.19–1.37)	
	3 cohorts - disadvantaged social class vs advantaged (occupation): increase in number of years	Minimally adjusted Years increase=0.99 (0.39 ; 1.59) Fully adjusted Years increase=0.93 (0.32-1.54)	
Inflammatory transcriptome (RNA)	EPIC-Italy - disadvantaged father's occupation vs advantaged - regression	β =0.35 (0.04 ; 0.66)	

score (Castagne et al, 2016) (countries Italy and Canada)	coefficient from multivariable regression β (95%CI) per unit increase in transcriptomic score	
	GSE15180 dataset - disadvantaged father's occupation vs advantaged - regression coefficient from multivariable regression β (95%CI) per unit increase in transcriptomic score	β =0.60 (0.11 ; 1.09)
* up to 2.35 at 21 years		

Figure 1 - legend

Dose-response relationship between three levels of socio-economic conditions (as defined in the text) and outcome variables: A. mortality; B. walking speed; C. height in children; D. blood levels of C-reactive protein. Outcome variables and cohorts of origin are described in Table 1.