



Original Scholarship

Causal Assessment of Income Inequality on Self-Rated Health and All-Cause Mortality: A Systematic Review and Meta-Analysis

MICHAL SHIMONOVICH ,* MHAIRI CAMPBELL,*
RACHEL M. THOMSON,* PHILIP BROADBENT,*
VALERIE WELLS,* DANIEL KOPASKER,*
GERRY McCARTNEY,[†] HILARY THOMSON,*
ANNA PEARCE,* and S. VITTAL KATIKIREDDI*

*MRC/CSO Social and Public Health Sciences Unit, School of Health and Wellbeing, University of Glasgow; [†]School of Social and Political Sciences, University of Glasgow

Policy Points:

- Income is thought to impact a broad range of health outcomes. However, whether income inequality (how unequal the distribution of income is in a population) has an additional impact on health is extensively debated.
- Studies that use multilevel data, which have recently increased in popularity, are necessary to separate the contextual effects of income inequality on health from the effects of individual income on health.
- Our systematic review found only small associations between income inequality and poor self-rated health and all-cause mortality. The available evidence does not suggest causality, although it remains methodologically flawed and limited, with very few studies using natural experimental approaches or examining income inequality at the national level.

Context: Whether income inequality has a direct effect on health or is only associated because of the effect of individual income has long been debated. We aimed to understand the association between income inequality and self-rated health (SRH) and all-cause mortality (mortality) and assess if these relationships are likely to be causal.

Methods: We searched Medline, ISI Web of Science, Embase, and EconLit (PROSPERO: CRD42021252791) for studies considering income inequality and SRH or mortality using multilevel data and adjusting for individual-level socioeconomic position. We calculated pooled odds ratios (ORs) for poor SRH and relative risk ratios (RRs) for mortality from random-effects meta-analyses. We critically appraised included studies using the Risk of Bias

The Milbank Quarterly, Vol. 102, No. 1, 2024 (pp. 141-182)

© 2024 The Authors. *The Milbank Quarterly* published by Wiley Periodicals LLC on behalf of The Milbank Memorial Fund.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

in Nonrandomized Studies – of Interventions tool. We assessed certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation framework and causality using Bradford Hill (BH) viewpoints.

Findings: The primary meta-analyses included 2,916,576 participants in 38 cross-sectional studies assessing SRH and 10,727,470 participants in 14 cohort studies of mortality. Per 0.05-unit increase in the Gini coefficient, a measure of income inequality, the ORs and RRs (95% confidence intervals) for SRH and mortality were 1.06 (1.03-1.08) and 1.02 (1.00-1.04), respectively. A total of 63.2% of SRH and 50.0% of mortality studies were at serious risk of bias (RoB), resulting in very low and low certainty ratings, respectively. For SRH and mortality, we did not identify relevant evidence to assess the specificity or, for SRH only, the experiment BH viewpoints; evidence for strength of association and dose–response gradient was inconclusive because of the high RoB; we found evidence in support of temporality and plausibility.

Conclusions: Increased income inequality is only marginally associated with SRH and mortality, but the current evidence base is too methodologically limited to support a causal relationship. To address the gaps we identified, future research should focus on income inequality measured at the national level and addressing confounding with natural experiment approaches.

Keywords: causality, income inequality, systematic review.

THERE IS A GENERAL ACCEPTANCE THAT THE HEALTH OF INDIVIDUALS IS impacted by the characteristics of the society they live in, such as neighborhood safety and welfare policy.^{1,2} Income inequality, a measure of the distribution of income within a population, has been suggested to be a major determinant of population health.^{3,4} Many studies report that countries with high income inequality also report high rates of infant mortality,⁵ cases of COVID-19,⁶ and unhappiness.⁷

Widening income inequality has been suggested to reduce the health of the whole population across the entire socioeconomic gradient, with psychosocial stress and its adverse consequences impacting the health of everyone.^{8–10} In contrast, others argue that economically unequal societies are more likely to have a high proportion of people experiencing low income and poverty and that the relationship between individuals experiencing poverty and health explains the association between income inequality and population health.^{3,11,12} In this sense, individual income is a “cross-level” confounder, as it affects both income inequality, which is measured across an area, and health, which is measured for individuals.^{13,14} Multilevel data contain both area-level and individual-level information and are needed to disentangle the effects of income inequality on health from the effects of individual-level income and other cross-level factors such as socioeconomic position (SEP).^{15–18}

The debate over whether policymakers should prioritize reducing poverty or income inequality (or both) has been ongoing since the first study that suggested that income inequality exerts a causal effect on health was published in 1992.^{8,9,19,20}

More recently, the debate has centered around the Richard Wilkinson and Kate Pickett 2010 book *The Spirit Level*.²¹ Interest in the possible causal link between income inequality and health^{22,23} remains high^{24–26} because it has been suggested that, if causal, it may result in more than a million excess deaths per year across Organisation for Economic Co-operation and Development countries.²⁰ The recent rise of income inequality both in the United States²⁷ and the United Kingdom²⁸ for the first time in over a decade, and the suspected link between relative income and COVID-19 outcomes, has bolstered calls for policymakers to renew their focus on income inequality.²⁹ Although identifying the evidence needed to understand the causal role income inequality has on health has been a challenge,³⁰ the volume of studies on this topic has increased since the last meta-analyses were published more than a decade ago, as more studies use multilevel data to address confounding.^{1,20,31} A review exploring the evidence to understand the possible causal effects of income inequality on health is therefore needed.

Our aim was to carry out a causal review of income inequality and two important and commonly used indicators of health: self-rated health (SRH) or all-cause mortality (mortality).³² Death is an outcome that is meaningful and easily understood by many, with mortality considered more methodologically reliable than specific causes of death.^{33,34} SRH provides a holistic measure of health, which is a strong predictor of more severe outcomes (including early mortality) and is thought to capture aspects of health that cannot be identified by clinical assessments.³⁵

The aims of this systematic review were to (a) quantify the association of income inequality with individual-level mortality and SRH, exploring factors that may explain differences among studies, and (b) assess evidence for causality using Bradford Hill (BH) viewpoints, a causal framework that is widely used in systematic reviews in population health research.

Methods

Search Strategy and Selection Criteria

A protocol for this systematic review was published¹⁸ and registered with PROSPERO (registration: CRD42021252791). This review follows the Preferred Reporting Items for Systematic Review and Meta-Analysis reporting guidelines³⁶ (see Appendix 1 for checklist) and recommendations from the *Cochrane Handbook for Systematic Reviews of Interventions*.³⁷ A search strategy was developed with the support of an information scientist with terms such as “income,” “income inequality,” “Gini,” “mortality,” and “health” (full strategy in Appendices 2 and 3). We searched for cohort or cross-sectional studies investigating the relationship between income inequality and mortality or SRH in Medline, ISI Web of Science, Embase, and

EconLit published January 1, 1992, to January 1, 2021. Our review was updated to include studies published by June 6, 2022. To be included, studies had to adjust for at least one individual-level measure of SEP (e.g., individual or household income, education, occupational status) using multilevel data with income inequality measured on an aggregate level. We did not include studies that considered specific causes of mortality or indicators of general health other than SRH (e.g., subjective well-being). We included studies using any income inequality metric that was measured at the area level; individual-level measures of inequality (e.g., Yitzhaki index) were excluded.^{38–40} We resolved disagreements over screening and critical appraisal by consensus or in consultation with a third author. Screening was guided by the directed acyclic graph in Appendix 4.

In the initial title-only screening using Endnote, M.S. and P.B. excluded duplicates and clearly irrelevant studies. Two investigators (M.S. and M.C., S.V.K., A.P., R.T., or H.T.) independently screened titles and abstracts in Covidence and then independently screened the retrieved full texts, which appeared to evaluate income inequality and either outcome. Studies were excluded if (a) the majority of the sample was younger than 18 years old, (b) the exposure was not a measure of the distribution of income across an area, (c) the outcome was not mortality or SRH (e.g., subjective well-being and mental health were not included), (d) multilevel data were not used, (e) individual-level SEP was not adjusted for, (f) the study was not published in English and a translation could not be obtained, (g) the study was conducted prior to 1992 (based on one of the first studies suggesting a relationship between income inequality and health¹⁹), or (h) the study was a duplicate. We did not limit studies based on the statistical methods used to account for clustering.¹ A full list of clarifications and minor deviations from our protocol can be found in Appendix 5.

Data extraction was conducted by one author (M.S.) and checked by a second author (M.C., S.V.K., A.P., H.T., or R.T.). We extracted information including author, publication year, study setting, income inequality indicator (e.g., Gini coefficient), analytical strategy to address clustering of individuals within areas, effect estimates for each outcome, and the functional form of the exposure (e.g., categorical or continuous) and outcome (e.g., dichotomous or ordinal; Appendix 6).

Critical Appraisal

Each study included in our review was independently appraised by M.S. and a second reviewer (M.C., S.V.K., A.P., and R.T.) using the Risk of Bias in Nonrandomized Studies – of Interventions (ROBINS-I) tool, as per Cochrane guidance.⁴¹ In ROBINS-I, the nonrandomized study being appraised is compared with a hypothetical randomized controlled trial, known as a target trial, in which practical and ethical issues of randomly assigning income inequality are ignored (full critical appraisal guide in Appendix 7).⁴² The key characteristics of an ideal target trial and potential sources

of bias are used to assess each study's overall risk of bias (RoB) according to seven domains: confounding, selection of participants into the study, classification of intervention, deviations from intended exposure, missing data, measurement of outcomes, and selection of the reported results.⁴³ The following levels of risk are assigned to each domain and then to the study overall: low, moderate, serious, or critical. The study's overall risk is rated on the highest (most severe) domain-level judgment.^{43,44}

As with most areas of population health research, we consider confounding to be one of the most important threats to bias.^{45,46} Most nonrandomized studies will, at best, have a moderate RoB because of confounding, which means that confounding is expected but has been adequately addressed.^{47,48} Studies will be deemed at serious risk of confounding if residual confounding is thought to be likely. In the case of the current review, for example, confounding poses a serious risk when only one indicator of individual-level SEP was adjusted for (e.g., educational attainment) or the variables adjusted for were unreliably measured.^{43,49} Additional guidance on the meaning of low, moderate, serious, and critical RoB for the other six domains are detailed in Appendix 7.

Statistical Analysis

Data Standardization. We aimed to standardize estimates to produce a pooled odds ratio (OR) or risk ratio (RR) per 0.05-unit increase in Gini coefficient for worse SRH and mortality, respectively (additional information on our analysis can be found in Appendices 8-10⁵⁰⁻⁵⁹). A 0.05-unit increase in Gini coefficient is commonly used across population health research^{20,60-63} to describe a change in the distribution of income across an area^{20,64-67} that is considered to be meaningful.⁶⁸ For instance, the increase in the Gini coefficient in the United Kingdom from 0.25 in 1979 to 0.34 in 1990⁶⁹ and the 0.05-unit rise in the United States from 0.36 in 1979 to 0.41 in 1990⁶⁰ are considered large and meaningful.⁶⁹ Hazard ratios and ORs were standardized to RRs based, respectively, on an approach by Zhang and Yu⁶¹ and Shor and colleagues.⁶² SRH is measured using a single-item question about health with several responses from best to worst, such as excellent, very good, good, poor, and very poor, and the measure is highly predictive of mortality.⁶³ Measures of SRH were often dichotomized and, if necessary, estimates were reverse coded to reflect worse SRH. Based on previous findings,²⁰ we expected I^2 , a measure of between-study variation, to be high (>75%).⁷⁰ Given our expectations around heterogeneity, we conducted random-effects meta-analyses using a restricted maximum likelihood approach. Analyses were performed using R Statistical Software⁷¹ using the *meta*⁷² and *dosresmeta*⁷³ packages for the primary meta-analyses, meta-regressions, subgroup analyses, and sensitivity analyses. Data and code are available in a public, open-access repository.⁷⁴

Selection of Multiple Effect Estimates. The effect estimates included in a meta-analysis are typically assumed to be independent, as pooling together nonindependent

estimates can overestimate their precision when synthesized.^{75,76} Nonindependence commonly results from including multiple effect estimates with overlapping samples from the same study (i.e., double counting).⁷⁵ Multiple estimates can be found in a single report or across multiple reports of the same study population or sample.^{77,78} To avoid double counting participants, it is recommended that reviewers select the most relevant, independent estimate(s) from each unique study.⁷⁸ The approach to selecting estimates might introduce bias, such as if estimates are selected in a way that confirms reviewers' perceptions or beliefs.^{75,76} We therefore followed the decision rules described in Appendix 5. The estimates from each unique study and corresponding reports selected for our primary meta-analyses are described in Appendices 9 and 10. These tables note which estimates were included in our sensitivity analyses, exploring if our approach to selecting estimates, as well as other statistical analysis decisions, affects our results.

Exploring Heterogeneity. As previously noted, we expected the estimates included in our review to vary considerably. Although it is not necessary for estimates to be comparable to be synthesized, considerable heterogeneity can make it challenging to interpret and apply review findings.^{79,80} Therefore, we followed systematic review guidance from Cochrane that underscores the importance of identifying sources of heterogeneity (known as covariates or potential effect modifiers) through meta-regressions and subgroups analyses.⁷⁰ We explored suspected sources of heterogeneity: mean or median Gini coefficient in the sample under study; study RoB (moderate, serious, or critical); global regions where the study is set as classified by the World Bank⁸¹ (North America, Europe and central Asia, east Asia and Pacific, and Latin America and the Caribbean) and those spanning multiple regions; geographical scale, which refers to both the nature of the sample as well as the level at which income inequality was measured (areas defined as within-country, local [e.g., cities, neighborhoods, districts]; within-country, regional [e.g., states, provinces, regions, municipalities, prefectures]; and among countries, national [e.g., continents, world regions, countries]); studies set in the United States vs. outside the United States; adjusted variables (individual-level variables only vs. individual and contextual/area-level variables, with the latter potentially reflecting residual confounding or overadjustment); time between income inequality measure and outcome (time lag ≤ 6 years vs. time lag > 6 years); mean or median follow-up duration; and average age of population (average age 60 years old or younger vs. older than 60 years). Meta-regressions are used to estimate the amount of between-study variance, which can be explained by the covariate.⁸² The regression coefficient estimated from a meta-regression is the response to a one-unit increase in the level of the covariates.⁷⁰ Subgroup analyses differ from meta-regressions, as stratified analyses were conducted for each level of the covariate.⁸³ We explored publication bias using funnel plots and Egger's test for both outcomes.

Causal Assessment

Certainty of Evidence. We considered causality using two complementary⁸⁴ approaches. The first, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach, is considered the gold standard approach used in systematic reviews.⁸⁵ GRADE provides a structured and transparent approach for rating the level of certainty across a body of evidence based on within- and between-study considerations.⁸⁶ The level of certainty is rated as very low, low, moderate, or high and reflects the confidence users have in effect estimates. The body of evidence automatically begins with high certainty and then may be downgraded according to five domains: RoB, inconsistency, indirectness, imprecision, and publication bias.⁸⁵

Weight of Evidence. Our second approach to assessing causality was to use the highly regarded^{87–89} and widely used⁹⁰ BH viewpoints.⁹¹ Epidemiologist Sir Austin Bradford Hill proposed nine characteristics to help researchers discern causal from noncausal associations: strength of association, consistency, specificity, temporality, dose–response gradient, plausibility, coherence, experiment, and analogy.⁹¹ According to Bradford Hill, strength of association refers to a large, but undefined, association between the exposure and outcome that suggests the relationship cannot be explained by a third variable. Consistency means observing similar associations across different settings or population. Specificity refers to evidence of a one-to-one relationship. Temporality indicates that the exposure occurred prior to the outcome. In a dose–response relationship, incremental increases (or decreases) of the exposure are observed alongside incremental increases (or decreases) of the outcome. Plausibility indicates credible explanations of the relationship. Coherence means that assumptions about the causal relationship fit with existing theoretical and empirical evidence. An association observed in an experiment is considered the strongest type of causal evidence. Finally, analogous evidence refers to associations between exposures and outcome with a similar underlying mechanism to the relationship under study.

There are currently no agreed-on guidelines for how BH viewpoints should be used in causal assessment despite being commonly used in systematic reviews across population health. BH viewpoints have often not been applied systematically and transparently,⁹⁰ which is particularly important for a topic as controversial as the relationship between income inequality and health.^{3,92} To improve the use of BH viewpoints, we drew on the principles of process tracing, which is a causal inference method used for case study research, to evaluate the value of evidence for the causal claim.^{93,94} The relative importance of evidence in process tracing is judged by its uniqueness to the hypothesis (if supportive evidence is found, how unlikely it is to be explained by something other than the hypothesis) and its definitiveness (certainty that if contradictory evidence is found, the hypothesis cannot be true).^{93,95} Process tracing classifies evidence into four types according to their combinations of uniqueness and definitiveness, which we applied to each BH viewpoint. In order of

strength, the BH viewpoints by process tracing type are as follows: “doubly decisive” (high uniqueness and definitiveness), “smoking gun” (high uniqueness, low definitiveness), “hoop” (low uniqueness, high definitiveness), and “straw in the wind” (low uniqueness, low definitiveness).⁹⁶ We consider analogy and coherence to be straw-in-the-wind viewpoints, as they can neither disconfirm (low uniqueness) nor confirm (low definitiveness) the hypothesis. The classification assigned to each of the BH viewpoints is described in Table 1. A figure summarizing the relative importance of BH viewpoints by type of process tracing test is in Appendix 11.

Having classified the BH viewpoints by process tracing type, we then explored the evidence for each viewpoint to evaluate our hypothesis: income inequality has an adverse effect on mortality or SRH (each outcome evaluated separately). We did not do this for analogy or coherence because they were considered to be “straw-in-the-wind” viewpoints and therefore have limited value. Using the same principles of process tracing types, finding evidence in support of a highly unique BH viewpoint can help confirm the hypothesis, as supportive evidence is unlikely to be observed if the hypothesis is untrue.⁹³ Finding evidence against the highly definitive BH viewpoints increases the possibility that the hypothesis is untrue. The possible judgments for the remaining BH viewpoints were as follows: “met” (evidence found in support of a BH viewpoint, suggesting the hypothesis is true), “unmet” (evidence found against a viewpoint, suggesting hypothesis is untrue), “inconclusive” (some evidence “met” and some evidence “unmet”), and “no evidence” (absence of evidence to evaluate the BH viewpoint). Judgments were guided by the meta-analysis for each outcome, explorations of heterogeneity (e.g., meta-regression and subgroup analyses), and results of critical appraisal using ROBINS-I. The evidence in support of and against each of the seven BH viewpoints included in the causal assessment are also explained in Appendix 5.⁹³

Results

Included Studies

From the 61,164 studies identified in our search, 9,848 duplicate reports were excluded prior to screening (flow chart in Appendix 12). Of the 51,316 records screened, 50,855 were excluded as irrelevant and 27,912 were excluded at the title-only stage and the remainder at the abstract stage, with an additional 195 excluded because of duplication. We screened 463 full texts (Appendix 13 gives reasons for exclusion) and found 94 eligible studies using multilevel data to evaluate the association between income inequality and mortality (27/94, 28.1%) or SRH (67/94, 71.9%; list of included studies in Appendix 14). We excluded 28 studies with duplicate data on mortality (6/27, 20.0%)^{101–106} or SRH (22/67, 29.7%).^{107–128} Our primary

Table 1. Evidence in Support of and Against Each BH Viewpoint Grouped by the Perceived Uniqueness and Definitiveness of Each Viewpoint

BH Viewpoints	Evidence in Support of BH Viewpoint ^a	Evidence Against BH Viewpoint
Doubly Decisive Viewpoint (High Uniqueness, High Definitiveness)^b		
Experiment	NEs showing adverse effect of increasing income inequality on health given low or moderate RoB.	NEs showing a protective effect of increasing income inequality on health.
Smoking Gun Viewpoints (High Uniqueness, Low Definitiveness)^c		
Strength of association	Studies showing an adverse effect of income inequality on health in which (a) all measured confounding has been appropriately adjusted (indicated by RoB assessment) or (b) additional tests suggesting residual confounding and/or unmeasured confounding have been adequately addressed (e.g., tests using negative exposures/outcome, measures of the minimum observed association unlikely explained by confounding [e.g., E-value] ⁹⁷).	Evidence that an increase in income inequality is associated with an improvement in health measured on the individual level given (a) or (b); put simply, the inverse of evidence in support of strength of association.
Dose–response gradient	We used studies that calculated category-specific exposure–outcome relationships to evaluate the overall dose–response gradient. Categorical studies calculating the association at three or more levels of income inequality for each health outcome compared with a reference exposure level provide insight into the shape (e.g., linearity, threshold effect) of the dose–response gradient. ⁹⁸ As with the strength of association, the shape of the exposure–outcome relationship is considered in light of evidence of adequate adjustments for confounding (e.g., low/moderate RoB).	Evidence of a beneficial dose–response gradient (health improves as income inequality increase) from studies in which the shape is unlikely to be caused by confounding.

Continued

Table 1. (Continued)

BH Viewpoints	Evidence in Support of BH Viewpoint ^a	Evidence Against BH Viewpoint
Specificity	As with many relationships in population health research, we did not expect to identify evidence of a one-to-one relationship between income inequality and health. Thus, evidence in support of specificity is that which suggests measured confounding is not responsible for nonspecific associations. In other words, evidence that residual confounding does not remain after confounding variables have been adjusted, such as from falsification tests (negative controls). ⁹⁹	Failing falsification tests, suggesting that some of the association between income inequality and health can be explained by measured confounding variables even after they have been adjusted for.
Consistency	Consistency is traditionally defined as similar estimates observed in studies covering different settings and populations (especially if bias has been addressed). As considerable heterogeneity is not uncommon in population health research, evidence in support of consistency also includes prespecified plausible explanations for differences in effect sizes (e.g., contextual differences such as geographical scale) supported by meta-regressions and/or subgroup analyses.	Evidence against the hypothesis for consistency based on meta-regressions or subgroup analyses (i.e., unexplained statistical heterogeneity). We do not consider statistical heterogeneity on its own as sufficient for consistency to be unmet.
Hoop Viewpoints (Low Uniqueness, High Definitiveness)^d		
Plausibility	Theoretical explanatory evidence of the mechanism by which income inequality has an adverse effect on health. Empirical evidence exploring the plausibility of the relationship under study (e.g., mediation analysis to look at mechanisms) may strengthen, but is not necessary, for the overall plausibility of the relationship.	Credible evidence that a relationship is not plausible.

Continued

Table 1. (Continued)

BH Viewpoints	Evidence in Support of BH Viewpoint ^a	Evidence Against BH Viewpoint
Temporality	Evidence that indicates that reverse causation is unlikely would provide support for the temporality viewpoint. This includes evidence that shows that (a) income inequality precedes the health outcome (e.g., quasi-experimental/experimental studies); (b) analytical approaches to reduce the possible effect of health on income inequality (e.g., longitudinal studies that adjust for prior or baseline health); or (c) a latent effect of income inequality on health in which an incubation period is thought to suggest, though to a somewhat lesser degree than (a) or (b), that the outcome is unlikely to cause the exposure.	Reverse causation is somewhat unlikely, as only studies that adjusted for individual-level socioeconomic position were eligible, blocking the effect of health prior to income inequality on income inequality. There is also considerable disagreement over the thresholds for a latent effect. As a result, we do not consider an absence of evidence for (b) or (c) as sufficient for temporality to be unmet. Rather, only the inverse of (a) (that health precedes income inequality) would increase our confidence that reverse causation is occurring and result in temporality being unmet.

BH, Bradford Hill; NES, natural experimental study; RoB, risk of bias.

^aEvidence is highly unique if it does not overlap with other theories and highly definitive if *not* finding it is a strong indication that the hypothesis is untrue (used to confirm and disconfirm hypothesis, respectively).¹⁰⁰ As finding evidence for or against the BH viewpoints analogy or coherence will not impact the hypothesis (income inequality adversely affects health) either way, neither are considered.

^bExperiment is the only highly unique and highly definitive viewpoint, as low RoB experimental/quasi-experimental studies, but often unavailable in population health research, suggesting an adverse/beneficial effect is a strong indication that the hypothesis is true/untrue, respectively.

^cSmoking gun viewpoints are highly unique but not definitive. For example, understanding statistical heterogeneity (i.e., consistency is met) may help confirm the hypothesis, but the income inequality and health relationship may be causal even if such evidence is not found (i.e., consistency is unmet).

^dHoop viewpoints are highly definitive, as no credible explanation of how causality occurs or evidence of reverse causation (i.e., plausibility or temporality are unmet, respectively) disconfirms the hypothesis, but if it is not unique the hypothesis may be untrue even if plausibility or temporality are met.

analyses included 2,916,576 individuals in 38 studies considering SRH^{129–166} and 14 studies with 10,727,470 million individuals considering mortality^{141,167–179} (study characteristics in Appendices 15 and 16, respectively). An additional seven reports considering SRH^{180–186} and eight considering mortality,^{187–194} including one natural experimental study,¹⁹¹ were included in one or more of the subgroup or sensitivity analyses. The natural experimental study exploits a resettlement policy for refugees arriving in Sweden between 1987 and 1994 whose placement in different municipalities and “initial exposure to income inequality is randomly determined conditional on a few key individual characteristics.”¹⁹¹ Despite its favorable study design, estimates from the natural experimental study were excluded from the main analysis, as it used a subset of the Swedish census that overlaps, but is smaller than, another report included in the primary meta-analysis.¹⁷¹

Of the studies considering SRH, most used the Gini coefficient to measure income inequality except for two that used the Theil index^{129,130} and one that used median share.¹⁴¹ All but six (15.8%) were cross-sectional.^{134,137,140,141,145,166} All studies considering mortality were cohort studies, with ten using the Gini coefficient^{167,169,175,177–179} and the other three using the median share.^{141,168,176,187} Categorical measures of income inequality, using various cutoffs, were used by 12 studies considering SRH (31.6%^{137,146,149,151,153,156,157,160,162,164,165}) and five considering mortality (35.7%^{172,175,178}). In the SRH studies, income inequality was predominantly measured within countries at the regional (e.g., states; 14/38, 39.5%^{134,135,140,142,144,149,151,154,156,159,162,166}) or local (e.g., cities, neighborhoods; 16/38, 42.1%^{130,131,136,139,141,146,147,152,153,157,158,160,163}) level, with just eight studies estimating the impact of income inequality across countries and measured at the national level (18.4%^{129,132,133,139,145,148,155,161}). One study¹⁶⁶ measuring income inequality within one country, but on a national level over time, was included in the within-country regional group. One study¹⁴⁵ which included a sample from multiple countries measured income inequality across regional (within-country) areas. Likewise, the mortality studies predominantly considered income inequality on a local (10/14, 71.4%^{141,167,170–177}) or regional (28.6%^{168,169,178,179}) scale. We did not identify any that considered the relationship across multiple countries.

Among studies considering SRH, few (7/38, 18.4%) had moderate RoB,^{139,150,157,159,161} with the majority being rated as serious (24/38, 63.2%^{129,130,132,134,136–141,143–148,151,153–156,162–165}) and a few rated as having critical RoB (7/38, 18.4%).^{131,133,141,152,153,158,161} In contrast, half of the studies considering mortality had moderate RoB (7/14, 50.0%^{141,168,171,173,175,178,179,187}) and half serious RoB (7/14, 50.0%^{167,169,170,172,174,176,177}). Confounding of preexposure variables was the most common source of bias, with bias caused by the selection of participants into the study and bias caused by classification of interventions (e.g., missing or poorly defined information about values of income inequality) also being important issues (see Appendices 17 and 18 for full critical appraisal results). As

noted in Appendix 7, studies that adjusted for variables that we considered to be downstream^{195,196} from income inequality were deemed to have critical RoB caused by confounding (e.g., subjective poverty,¹⁶¹ perceived control,¹³³ and the level of trust in the community¹⁴⁷). Studies that did not sufficiently address confounding, including those evaluating the effects across multiple countries that did not adjust for area-level variables (e.g., GDP¹³⁹) or studies that adjusted for variables that we deemed to insufficiently address individual-level SEP (e.g., only adjusted for age, gender, household income, and family size¹⁴¹), were often considered to have serious RoB.

Meta-Analysis of the Effects of Income Inequality on SRH and Mortality

As shown in Figure 1 and Figure 2, respectively, a 0.05-unit increase in Gini coefficient was associated with poorer SRH (OR 1.06, 95% confidence interval [CI] 1.03-1.08; $I^2 = 96.5\%$; $n = 38$) but less so for mortality (RR 1.02, 95% CI 1.00-1.04; $I^2 = 92.9\%$, $n = 14$). When limited to studies with lower (moderate) RoB, the coefficient reduced very slightly for SRH (OR 1.05, 95% CI 1.03-1.06; $n = 7$) and remained for mortality (RR 1.02, 95% CI 0.99-1.06; $n = 7$), albeit with greater uncertainty.

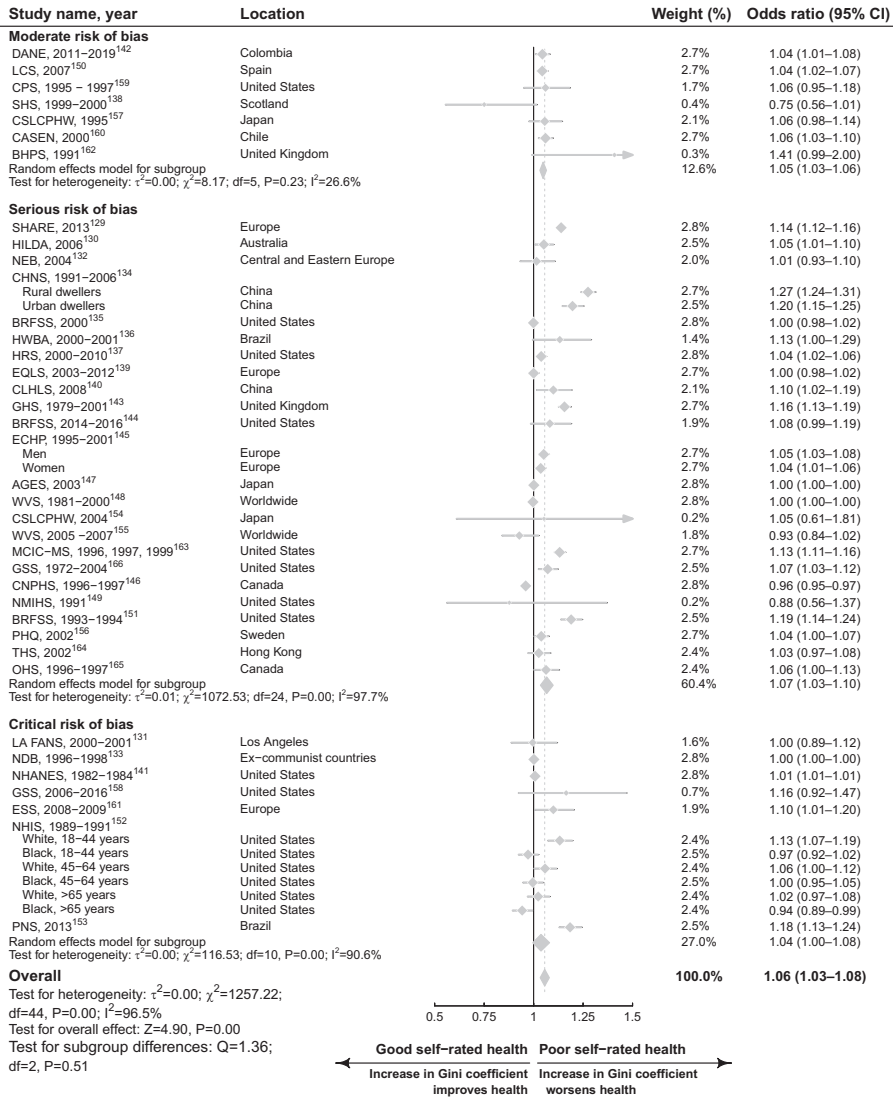
Meta-Regression to Explore Reasons for Heterogeneity

The random-effects meta-regressions, reported in Table 2, indicated that none of the prespecified covariates (geographical scale, world region, age, time lag, length of follow-up, adjusting for contextual variables, or RoB) appear to explain the high heterogeneity found across studies considering SRH ($I^2 = 96.5\%$; $n = 38$) and mortality ($I^2 = 92.9\%$, $n = 14$). Our subgroup analyses similarly did not explain sources of heterogeneity (see Appendices 19 and 20). Moreover, our sensitivity analyses in Appendix 21 indicate that decisions around which data should be in our primary analyses did not impact the overall estimates for either outcome.

Certainty of the Evidence Using the GRADE Approach

We used the GRADE approach to rate our certainty in the evidence that income inequality has a harmful, causal effect on SRH or mortality (evaluated separately; Table 3). We started with a baseline of high certainty in accordance with the ROBINS-I approach.⁴⁷ Across the studies considering SRH, certainty was “downgraded” by a total of four levels caused by RoB, inconsistency and publication bias, which resulted in very low certainty of a causal relationship, which is the lowest

Figure 1. Forest Plot for a Meta-Analysis of Studies Considering Income Inequality and Poor SRH, Stratified by RoB for a 0.05-Unit Increase in the Gini Coefficient. The number of people were 2,916,576, and the number of studies were 38.



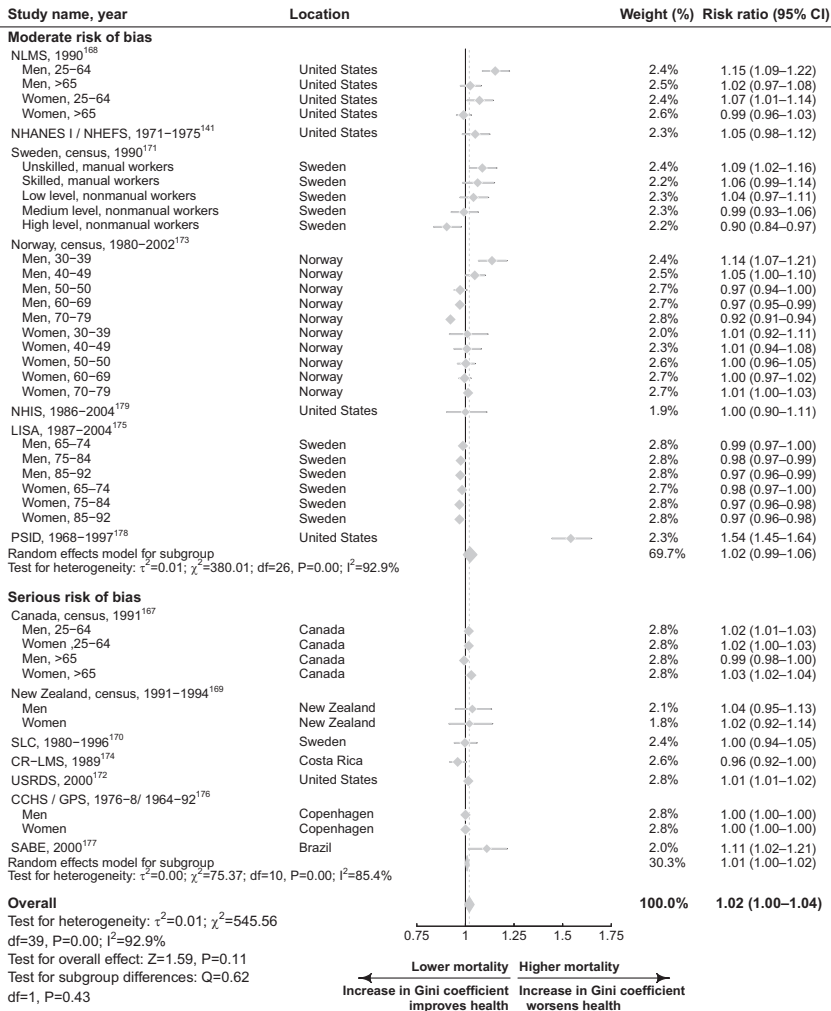
AGES, Aichi Gerontological Evaluation Study Project; BHPS, British Household Panel Survey; BRFSS, Behavioral Risk Factor Surveillance System; CASEN, National

Socioeconomic Characterization Survey; CHNS, China Health and Nutrition Survey; CI, confidence interval; CLHLS, Chinese Longitudinal Healthy Longevity Survey; CNPHS, Canadian National Population Health Survey; CPS, Current Population Survey; CSLCPHW, Comprehensive Survey of Living Conditions of People on Health and Welfare; DANE, Departamento Administrativo Nacional de Estadística; ECHP, European Community Household Panel survey; EQLS, European Quality of Life Surveys; ESS, European Social Survey; GHS, British General Household Survey; GSS, General Social Survey; HILDA, Household Income and Labour Dynamics in Australia; HRS, Health and Retirement Study; HWBA, Health, Well-Being, and Ageing study; LA FANS, Los Angeles Family and Neighborhood Survey; LCS, Life Conditions Survey; MCIC-MS, Metropolitan Chicago Information Center Metro Survey; NDB, New Democracies Barometer; NEB, New European Barometer; NHANES, National Health and Nutrition Examination Survey; NHIS, National Health Interview Survey; NMIHS, National Maternal Infant Health Survey; OHS, Ontario Health Survey; PHQ, Stockholm County Council's Public Health Questionnaire; PNS, National Health Survey; RoB, risk of bias; SHARE, Survey of Health, Ageing and Retirement in Europe; SHS, Scottish household survey; THS, Thematic Household Surveys; WVS, World Values Survey.

possible level of certainty described by GRADE. Most studies considering SRH were deemed to have serious or critical RoB, and as per GRADE guidance, we rated down confidence in the effect estimates caused by RoB by two levels.²⁰⁰ We downgraded certainty because of inconsistency by one level because heterogeneity was high and, based on the results from the meta-regression, remained unexplained for studies considering SRH.²⁰¹ Finally, we rated down certainty by one level because of publication bias after considering the funnel plot and Egger's regression test, which are both used to understand whether the study estimates are related to the sample size (as indicated by some measure of precision [e.g., standard errors]).^{202,203} The Egger's test for funnel plot asymmetry was statistically significant ($p = 0.00041$) and should be considered along visual examination of the distribution of study effect estimates and standard errors in the funnel plot.²⁰⁴ Both suggested that studies with small sample sizes that find no statistically significant association are "missing" (see Appendices 22 and 23 for funnel plots). We rated down our confidence in the effect estimates for mortality because of RoB and inconsistency, resulting in low certainty that the relationship between income inequality and mortality is causal. The evidence for mortality had important methodological limitations with around half of the studies considering deemed to have serious RoB, leading to a downgrade by one level.²⁰⁰ Although studies conducted in the United States that measured the impact of income inequality on mortality reported a larger effect size than studies conducted elsewhere, more studies are needed to be conclusive on this point. We therefore also downgraded confidence in the effect estimates by one level because of inconsistency.²⁰¹ For both SRH and mortality, we did not find any reasons to downgrade certainty because of imprecision

Figure 2. Forest Plot for a Meta-Analysis of Studies Considering Income Inequality and All-Cause Mortality, Stratified by RoB for a 0.05-Unit Increase in the Gini Coefficient.

The number of people were 10,727,470, and the number of studies were 14.



Canada, census, Canadian Census Mortality Follow-up Study; CCHS/GPS, Copenhagen City Heart Study; CI, confidence interval; CR-LMS, Costa Rican Longitudinal Mortality Study; LISA, Linnaeus Database; NHANES I/NHEFS, National Health and Nutrition Examination Survey; NHIS, National Health Interview Survey; NLMS, US National Longitudinal Mortality Study; PSID, Panel Study of Income Dynamics; RoB, risk of bias; SABE, Health, Well-Being, and Aging; SLC, Statistic Sweden's Survey of Living Conditions; USRDS: US Renal Data System.

(e.g., CIs of effect estimates include the null effect)²⁰⁵ or indirectness (e.g., evidence was incomparable with the population of interest).²⁰⁶

Assessment of Causality Using BH Viewpoints

Our assessment using BH viewpoints found no evidence that the effects of income inequality on either SRH or mortality, reported in the meta-analysis, were causal (Table 4). Many of the judgments regarding viewpoints were similar for SRH and mortality. We did not identify any studies that explored whether residual confounding remained after adjusting for hypothesized confounders and, therefore, found no evidence to judge specificity. We considered the RoB alongside the magnitude of the effects and the shape of the relationship to judge strength of association and the dose–response relationship, respectively, and found inconclusive evidence for both viewpoints. Plausibility and temporality were supported because the explanation for how income inequality might harm health was credible and there was no evidence of reverse causation, respectively.

There were two slight differences in our judgments regarding viewpoints for SRH and mortality. First is the experiment viewpoint. We did not identify any quasi-experimental or experimental study considering SRH and therefore could not judge this viewpoint. We identified just one¹⁹¹ natural experimental study, which found that a 0.05-unit increase in the Gini coefficient was associated with a decreased risk of mortality, though the reported estimate was imprecise (RR 0.96, 95% CI 0.87-1.15). We view the findings from this study we identified as potentially important because of its study design as, albeit inconclusive, evidence that income inequality does not have a causal effect on mortality. Second is consistency. Although we deemed the evidence to be unsupportive of the consistency viewpoint for studies measuring SRH because high heterogeneity remained unexplained, we found that some contextual factors may explain the differences across studies considering mortality but that this evidence was inconclusive.

Discussion

Our findings suggest that higher levels of income inequality are associated with small increases in poor SRH and mortality after accounting for individual-level SEP. Most studies considered income inequality measured at the local or regional, rather than the national, level. Based on our assessment using BH viewpoints, we did not find evidence that these effects were causal. A 0.05-unit increase in the Gini coefficient was associated with a 2% (1.02 RR 95% CI 1.00-1.04; Figure 2) increased risk of mortality using estimates from 14 studies with over ten million individuals. Pooled estimates from 38 studies with over two million individuals found that the same

Table 2. Meta-Regressions Investigating Potential Sources of Heterogeneity Between Income Inequality and (a) SRH and (b) Mortality

Covariates	SRH (<i>n</i> = 38)			Mortality (<i>n</i> = 14)			
	OR (95% CI) ^a	<i>p</i> Value ^b	τ^2 ^c	Covariates	RR (95% CI) ^a	<i>p</i> Value ^b	τ^2 ^c
Geographical scale ^d	Ref. group: local level, within-country, <i>n</i> = 16			Ref. group: local level, within-country, <i>n</i> = 10			
Regional level, within-country (<i>n</i> = 14) ^e	1.10 (0.97-1.25)	0.012	0.0038	Regional level, within-country (<i>n</i> = 4)	1.10 (0.96-1.26)	0.0021	0.0043
National level, among-country (<i>n</i> = 8) ^f	1.03 (0.90-1.17)	0.69	–	–	–	–	–
Mean/median income inequality ^g	Ref. group: <0.30 Gini, <i>n</i> = 5			Ref. group: <0.30 Gini coefficient, <i>n</i> = 7			
0.30-0.40 (<i>n</i> = 16)	1.07 (0.93-1.22)	0.32	0.0045	0.30-0.40 (<i>n</i> = 6)	1.03 (0.88-1.20)	0.58	0.0058
>0.40 (<i>n</i> = 17)	1.05 (0.92-1.20)	0.60	–	>0.40 (<i>n</i> = 1)	1.01 (0.82-1.25)	0.99	–
World region ^h	Ref. group: multiple world regions, <i>n</i> = 2			Ref. group: Asia and Pacific, <i>n</i> = 1			
Asia and Pacific (<i>n</i> = 7)	1.10 (0.96-1.26)	0.027	0.0041	Europe and Central Asia (<i>n</i> = 5)	1.00 (0.87-1.15)	0.64	0.0051

Continued

Table 2. (Continued)		Mortality (<i>n</i> = 14)				
SRH (<i>n</i> = 38)		OR (95% CI) ^a	<i>p</i> Value ^b	τ^2 ^c	τ^2 ^c	
Covariates					RR (95% CI) ^a	
					<i>p</i> Value ^b	
					τ^2 ^c	
Europe and Central Asia (<i>n</i> = 11)		1.05 (0.92-1.20)	0.12	-	1.02 (0.86-1.22)	0.95
Latin America and the Caribbean (<i>n</i> = 4)		1.10 (0.95-1.27)	0.042	-		-
North America (<i>n</i> = 14)		1.04 (0.91-1.18)	0.18	-	1.06 (0.92-1.23)	0.61
Study setting		Ref. group: not United States, <i>n</i> = 28			Ref. group: not United States, <i>n</i> = 8	
United States (<i>n</i> = 10)		1.05 (0.92-1.20)	0.73	0.0045	1.08 (0.94-1.24)	0.0042
Average population age ⁱ		Ref. group: ≤60 years old, <i>n</i> = 33			Ref. group: ≤60 years old, <i>n</i> = 12	
>60 years old (<i>n</i> = 6)		1.04 (0.90-1.19)	0.47	0.0044	0.99 (0.86-1.14)	0.046
Covariate adjustments		Ref. group: individual-level variables only, <i>n</i> = 16			Ref. group: individual-level variables only, <i>n</i> = 3	
With area-level variables (<i>n</i> = 22)		1.05 (0.92-1.20)	0.42	0.0045	1.01 (0.89-1.15)	0.0049
Risk of bias		Ref. group: moderate, <i>n</i> = 7			Ref. group: moderate, <i>n</i> = 7	
Serious (<i>n</i> = 24)		1.07 (0.93-1.22)	0.62	0.0044	1.01 (0.87-1.18)	0.74

Continued

Table 2. (Continued)

SRH ($n = 38$)			Mortality ($n = 14$)				
Covariates	OR (95% CI) ^a	p Value ^b	τ^2 ^c	Covariates	RR (95% CI) ^a	p Value ^b	τ^2 ^c
Critical ($n = 7$)	1.04 (0.90-1.19)	0.79	–	–	–	–	–
Time lag	Ref. group: ≤ 6 years, $n = 37$			Ref. group: ≤ 6 years, $n = 13$			
>6 years ($n = 1$)	1.13 (0.94-1.36)	0.30	0.0044	>6 years ($n = 1$)	1.00 (0.80-1.24)	0.78	0.0057
				Follow-Up Time: Ref. Group: <5 Years			
				5-10 years ($n = 5$)	1.02 (0.88-1.19)	0.43	0.0057
				>10 years ($n = 5$)	1.03 (0.88-1.20)	0.33	–

CI, confidence interval; mortality, all-cause mortality; OR, odds ratio; Ref., reference; RR, risk ratio; SRH, self-rated health.

^aOR (95% CI) for each covariate reflects the change in poor SRH per 0.05-unit increase in Gini coefficient. RR (95% CI) for each covariate reflects the change in mortality per 0.05-unit increase in Gini coefficient. The values were calculated using random-effects models with restricted maximum likelihood estimate.

^bFor testing differences, the null hypothesis coefficient has no effect.

^cBetween-study variance explained by the covariates (i.e., residual heterogeneity beyond I^2).

^dGeographical level describing nature of the sample (within-country or among countries) and level at which income inequality was measured (local, regional, or national level) and nature of the sample (within-country or among countries).

^eIncludes one study in which income inequality is measured at the national level over time, but the sample is limited to one country.

^fIncludes one study in which income inequality is measured at a regional level, but the sample spans multiple countries.

^gBased on categories from Lin and colleagues.⁵¹

^hBased on World Bank classifications.¹⁹⁷

ⁱIncludes studies that report multiple levels of average age.^{153,167,168,173}

Table 3. Certainty of Evidence Using GRADE Approach

Outcomes	No. of Participants (No. of Studies)	Anticipated Absolute Effects			Relative Effect (95% CI) ^a	Overall Certainty of Evidence ^b
		No Change in Income Inequality	0.05-Unit Increase in Gini Coefficient			
SRH	2,916,576 (<i>n</i> = 38)	21,590 per 100,000 ^c	22,885 per 100,000	OR 1.06 (1.03-1.08)	Very low ^d	
Mortality (per year)	10,727,470 (<i>n</i> = 14)	925 per 100,000 ^e	944 per 100,000	RR 1.02 (1.00-1.04)	Low ^f	

CI, confidence interval; mortality, all-cause mortality; No., number; OR, odds ratio; RoB, risk of bias; RR, risk ratio; SRH, self-rated health.
^aOR (95% CI) for each covariate reflects the change in poor SRH per 0.05-unit increase in Gini coefficient. RR (95% CI) for each covariate reflects the change in mortality per 0.05-unit increase in Gini coefficient. The values were calculated using random-effects models with restricted maximum likelihood estimate.
^bHigh certainty: we are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect. Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.⁸⁶
^cUsing data from Mutz and Lewis.¹⁹⁸
^dCertainty was rated down by two levels because of the RoB (most studies had serious or critical RoB), one level because of inconsistency (high, unexplained statistical heterogeneity), and one level because of publication bias (Egger's test *p*-value = 0.00041). There was no change for indirectness or imprecision.
^eUsing 2019 data for England from the Office for National Statistics.¹⁹⁹
^fCertainty was rated down by one level because of RoB (half of the studies had serious RoB) and one level because of inconsistency (high, inconclusive explanations of heterogeneity). There was no change for indirectness, imprecision, or publication bias.

Table 4. Causal Assessment for SRH and Mortality Using BH Viewpoints

	SRH Assessment ^a	Mortality Assessment ^a
Doubly Decisive Viewpoint(s) (High Uniqueness, High Definitiveness)^b		
Experiment	No evidence was found.	Inconclusive: we identified one natural experimental study ¹⁹¹ that found that income inequality reduces the likelihood of death, although more studies are needed to draw conclusive judgments.
Smoking Gun Viewpoint(s) (High Uniqueness, Low Definitiveness)^b		
Strength of association	Inconclusive: although there are estimates showing both adverse and protective effects of income inequality, we do not believe the RoB has been properly addressed because estimates from studies with moderate RoB were smaller than from more biased studies.	Inconclusive: we observed a similar and somewhat mild adverse effect of income inequality on mortality regardless of RoB. As there does not appear to be evidence of a protective effect (as illustrated in forest plot in Figure 2), evidence for a strong (adverse) association between income inequality and mortality remains inconclusive.
Consistency	Unmet: none of the meta-regressions, subgroup analyses, or sensitivity analyses convincingly elucidated reasons for the high between-study heterogeneity.	Inconclusive: very high between-study variation ($I^2 = 92.9\%$) may be explained by contextual differences (such as United States vs. non-United-States studies), but more studies are needed to draw conclusive judgments.
Specificity	No evidence was found. We did not identify any studies that used falsification outcomes/exposures to evaluate confounding.	

Continued

Table 4. (Continued)

	SRH Assessment ^a	Mortality Assessment ^a
Dose-response gradient	Inconclusive: less than one-third of studies examined a dose-response relationship between income inequality categories and SRH. ^{139,146,149,151-153,156,157,160,162,164,165} Moreover, only half reported a nonlinear relationship such that direction of the income inequality and SRH relationship varied across income inequality levels.	Inconclusive: fewer than one-quarter of mortality studies could be used to evaluate a dose-response relationship, and of those four studies that did, the evidence was mixed. Most did not find a dose-response relationship, and one study reported different directions of effect for men and women.
Hoop Viewpoints (Low Uniqueness, High Definitiveness)^b		
Plausibility	Met: there are several credible (though debated) explanations of how income inequality may adversely affect health. Although most studies narratively described explanatory mechanistic evidence, and some included graphical representations of the putative relationship, to some extent, plausibility was established a priori.	
Temporality	Met: most studies (22/38, 57.9%) used different data sources for income inequality and SRH, reducing, though not eliminating, the likelihood of reverse causality. There was no clear explanation as to why some studies found poor SRH increased ^{140,150,159,181} and some that it decreased ^{143,183} when measured more than 6 years after income inequality.	Met: all studies considering mortality were longitudinal, which can be used to evaluate for reverse causation by adjusting for prior health.

BH, Bradford Hill; I^2 , measure of between-study variance (heterogeneity); mortality, all-cause mortality; RoB, risk of bias; SRH, self-rated health.

^aSee Table 1 for evidence in support of each BH viewpoint, meaning that viewpoint has been met, and evidence against each viewpoint, meaning it is unmet. Absence of supportive evidence or evidence against a viewpoint means there is “no evidence”, whereas some evidence against and some evidence in support of it means it is inconclusive.

^bEvidence is highly unique if it does not overlap with other theories (used to confirm hypothesis). Evidence is highly definitive if *not* finding it is a strong indication that the hypothesis is untrue (disconfirms hypothesis).¹⁰¹

change in Gini coefficient increased the odds of poor SRH by 6% (OR 1.06, 95% CI 1.03-1.08; Figure 1). Associations were similar when analyses were repeated within lower RoB studies. The conclusiveness of our findings is constrained by some key evidence gaps, including there being only one natural experimental study for mortality, no experimental or quasi-experimental evidence for SRH, and no studies assessing the impact of income inequality on mortality at the national level among countries.

This review is the first to quantitatively summarize the relationship between income inequality and SRH and mortality in multilevel studies for many years.²⁰ It is also the first to assess causality, meeting the current standards for high-quality systematic reviews, which require methodological rigor and transparent reporting.^{207,208} We used gold standard methods to conduct the review, including using the ROBINS-I tool to critically appraise the included studies and the GRADE approach to evaluate certainty. We were transparent in our causal assessment, describing our methods in a prepublished protocol and carefully documenting minor amendments when necessary.

Our systematic review has several important limitations. First, we did not include non-English language studies, though only six potentially relevant studies were excluded for that reason (listed in Appendix 13). Additionally, most of the studies were set in North America or Europe, which are predominantly made up of high-income countries. Thus, our findings may not be applicable to other world regions, particularly those with many low- and middle-income countries. Moreover, we could not standardize outcomes using measures of income inequality other than the Gini coefficient, meaning that our initial primary meta-analyses assumed equivalent changes for Gini coefficient and non-Gini coefficient measures.⁵⁰ Studies have found that the effects of income inequality on health are similar when income inequality is measured using the Gini coefficient, Theil index, or the median share.^{142,143,146,150,162} The measure of income inequality did not explain differences across studies included in a previous meta-analysis,²⁰ whereas our sensitivity analyses similarly found that our results were not altered by including non-Gini coefficient measures in our primary analyses. The comparability of estimates is a concern even when income inequality is measured using the same indicator. For instance, multiple versions of the Gini coefficient will often arise^{209,210} because of variations in how the income and income-sharing unit are defined,³⁸ which often depend on the availability of data. Quantifying the extent to which these concerns impact our findings is challenging, as it would require reanalysis of error-free data sets, which are unlikely to be identified.²¹¹ Finally, our approach to causal assessment using modified BH viewpoints is novel, and so, other researchers may have applied them differently.

There are several limitations related to the Gini coefficient that are worth noting, as it was the most common measure of income inequality. The Gini coefficient ranges from 0 (reflecting a society in which income is divided equally) to 1 (in which all income is allotted to just one individual),^{212,213} with changes in inequalities occurring

at different ends of the income distribution scale, potentially having different effects on health.²¹² The Gini coefficient, however, cannot distinguish between, for instance, income transferred from the richest to the poorest from changes within the middle of the income distribution scale, which both have the same value.²¹⁴ Additionally, the small association between income inequality and health might be explained by, for instance, income inequality having an adverse effect on the health of people in the bottom half of the income distribution and a beneficial effect on those with the top share of income.²¹⁵ More information is needed to explore this possible explanation, as the Gini coefficient cannot on its own help elucidate whether the effects of income inequality across an entire population differ across subgroups within the population. We compared the effects of income inequality on SRH and mortality between studies using the Gini coefficient and a small number of studies using non-Gini coefficient measures, some of which can account for changes in the concentration of income (see Appendices 19 and 20). Although we found no differences among the subgroups, these findings should be interpreted cautiously given the imbalance in the number of studies using these alternative measures.²¹⁶

There have been several previous reviews evaluating the association between income inequality and SRH or mortality drawing on multilevel studies,^{17,20,31,217,218} although they were all conducted over a decade ago and have important methodological limitations and just one study quantitatively summarized the findings using in meta-analyses.^{219,220} We compare our findings with reviews that predominantly relied on studies using multilevel data that adjusted for individual-level SEP. Our pooled estimates differ slightly from previous meta-analyses estimates, published by Kondo and colleagues in 2008.²⁰ This is perhaps unsurprising given that we included almost double the number of studies. We found a lower association between a 0.05-unit increase in income inequality and mortality (RR 1.02, 95% CI 1.00-1.04 [$I^2 = 92.9\%$]), compared with RR 1.08, 95% CI 1.06-1.10 [$I^2 = 96\%$]). In contrast, we found that the odds of poor SRH given a 0.05-unit increase in Gini coefficient were slightly higher than, but remaining similar to, the estimates reported by Kondo and colleagues (OR 1.06, 95% CI 1.03-1.08 [$I^2 = 96.5\%$]), compared with OR 1.04, 95% CI 1.02-1.06 [$I^2 = 88.8\%$]). Kondo and colleagues did not critically appraise the included studies or explicitly assess causality, but they concluded that income inequality may have a “modest” effect on health.²⁰

In the review by Lynch and colleagues that had characteristics of a systematic review, they found that, for most countries, based on both ecological and multilevel studies, there was little evidence of an effect of income inequality on individual health.¹⁷ In their narrative synthesis, they concluded that low health insurance coverage in the United States may explain why studies set in the United States were more likely to find an adverse effect of income inequality on health than studies set in other wealthy countries (most of which found no effect). We similarly found that studies in the United States were more likely to report stronger

associations for mortality than those set elsewhere, but the measures for SRH were comparable.

Our approach to causal assessment based on BH viewpoints considers the magnitude of the associations between income inequality and SRH and mortality, alongside RoB and possible explanations for heterogeneity. Estimates from studies with moderate RoB were comparable with those at serious RoB, suggesting a level of robustness to important biases such as confounding, a key consideration in nonrandomized studies. Additional evidence from studies using experimental data or addressing unmeasured confounding, an important concern in population health research, is needed to be conclusive. Future research should focus on natural experimental studies, as, though difficult to design or identify, they have the potential to decisively settle the debate over whether income inequality adversely affects health above and beyond the effect of individual income. Additional examples are limited but have included exogenous changes in income inequality during postsocialist transitions in several central and eastern European countries^{221,222} and Swedish resettlement policies placing refugees in municipalities with differing income inequality levels.¹⁹¹

We could find no explanation for the high degree of heterogeneity across studies despite considering numerous possibilities. However, we did find that income inequality has a higher association with mortality for studies set in the United States when compared with those outside. This heterogeneity may be better understood through additional research estimating the impact of income inequality on mortality across multiple countries using multilevel data that would allow greater focus on among-country comparisons. Future research should focus on improving the measurement of political and economic determinants of income inequality, including social security policies²⁰ and macroeconomic development.³¹ Though difficult to measure, contextual factors such as globalization,²⁰⁹ strength of democratic institutes,²²³ or conditions in the labor market^{224,225} may help determine whether findings about the relationship between income inequality and health can be applied to different settings. More research is also needed in low- and middle-income countries to understand if the effects of income inequality on health in these countries differ, as is suspected, from effects in high-income countries.²²⁶

It is important to underscore that our findings are limited to the effect of income inequality on SRH and mortality and that future research may find that income inequality has an effect on health inequalities²²⁷ or on other health outcomes than the ones we chose to focus on (e.g., subjective well-being²²⁸). Our findings are also limited to the effects of income inequality on average health across whole populations, whereas future reviews that are specifically focused on the effect of income inequality across subgroups within populations might find, for instance, that income inequality has a worse effect on the health of those at the lower end of the income distribution

compared with those at the higher end. Some evidence suggests that income inequality affects the health outcomes that are most likely to be impacted by an erosion in social cohesion,²²⁹ including suicide,²³⁰ drug overdoses,²³¹ and mental health,^{8,232,233} although this evidence base might also benefit from a causally informed systematic review.

Conclusion

The relationship between income inequality and health outcomes is complex and has been debated for many years.²³² Here, we present a rigorous and thorough assessment of the potential causal link between income inequality and SRH and mortality after accounting for individual-level SEP. Although we conclude that reducing income inequality per se may not improve health, the evidence base remains uncertain because of a lack of experimental and quasi-experimental evidence and a limited number of multilevel studies examining income inequality and mortality at the national level. In the meantime, policymakers can focus their attention on evidence-based solutions such as reducing poverty²³⁴ and unemployment,²³⁵ which have been shown to improve health across a society.

References

1. Leyland AH, Groenewegen PP. What is multilevel modelling? In: *Multi-level Modelling for Public Health and Health Services Research: Health in Context*. Springer International Publishing; 2020:29-48.
2. Baranyi G, Di Marco MH, Russ TC, Dibben C, Pearce J. The impact of neighbourhood crime on mental health: a systematic review and meta-analysis. *Soc Sci Med*. 2021;282:114106.
3. Lynch JW, Smith GD, Kaplan GA, House JS. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. *BMJ*. 2000;320(7243):1200.
4. Wilkinson RG, Pickett KE. Income inequality and population health: a review and explanation of the evidence. *Soc Sci Med*. 2006;62(7):1768-1784.
5. Curran M, Mahutga MC. Income inequality and population health: a global gradient? *J Health Soc Behav*. 2018;59(4):536-553.
6. Oronce CIA, Scannell CA, Kawachi I, Tsugawa Y. Association between state-level income inequality and COVID-19 cases and mortality in the USA. *J Gen Intern Med*. 2020;35(9):2791-2793.
7. Lous B, Graafland J. Who becomes unhappy when income inequality increases? *Appl Res Qual Life*. 2022;17(1):299-316.
8. Polacko M. Causes and consequences of income inequality – an overview. *Stat Politics Policy*. 2021;12(2):341-357.

9. Walasek L, Brown GDA. Income inequality and social status: the social rank and material rank hypotheses. In: Jetten J, Peters K, eds. *The Social Psychology of Inequality*. Springer International Publishing; 2019:235-248.
10. Wilkinson RG, Pickett KE. The enemy between us: the psychological and social costs of inequality. *Eur J Soc Psychol*. 2017;47(1):11-24.
11. Ifcher J, Zarghamee H, Graham C. Income inequality and well-being in the U.S.: evidence of geographic-scale- and measure-dependence. *J Econ Inequal*. 2019;17(3):415-434.
12. Subramanian SV, Kawachi I. Income inequality and health: what have we learned so far? *Epidemiol Rev*. 2004;26(1):78-91.
13. Blakely TA, Woodward AJ. Ecological effects in multi-level studies. *J Epidemiol Community Health*. 2000;54(5):367-374.
14. Bergh A, Nilsson T, Waldenström D, eds. The ecological fallacy: what conclusions can be drawn from group averages? In: *Sick of Inequality? An Introduction to the Relationship between Inequality and Health*. Edward Elgar Publishing; 2016:72-81.
15. Gravelle H. How much of the relation between population mortality and unequal distribution of income is a statistical artefact? *BMJ*. 1998;316(7128):382-385.
16. Kawachi I, Blakely TA. Commentary: when economists and epidemiologists disagree...: comments on Mellor and Milyo. *J Health Polit Policy Law*. 2001;26(3):533-541.
17. Lynch J, Smith GD, Harper S, et al. Is income inequality a determinant of population health? Part 1. A systematic review. *Milbank Q*. 2004;82(1):5-99.
18. Shimonovich M, Pearce A, Thomson H, McCartney G, Katikireddi SV. Assessing the causal relationship between income inequality and mortality and self-rated health: protocol for systematic review and meta-analysis. *Syst Rev*. 2022;11(1):20.
19. Wilkinson RG. Income distribution and life expectancy. *Br Med J*. 1992;304(6820):165-168.
20. Kondo N, Sembajwe G, Kawachi I, van Dam RM, Subramanian SV, Yamagata Z. Income inequality, mortality, and self rated health: meta-analysis of multi-level studies. *BMJ*. 2009;339:b4471.
21. Wilkinson R, Pickett K. *The Spirit Level: Why Greater Equality Makes Societies Stronger*. Bloomsbury Publishing; 2009.
22. Coburn D. Income inequality, welfare, class and health: a comment on Pickett and Wilkinson, 2015. *Soc Sci Med*. 2015;146:228-232.
23. Eckersley R. Beyond inequality: acknowledging the complexity of social determinants of health. *Soc Sci Med*. 2015;147:121-125.
24. Lynch J. The political economy of health: bringing political science in. *Annu Rev Polit Sci (Palo Alto)*. 2023;26:389-410.
25. Kawachi I, Subramanian SV. Income inequality. In: Berkman LF, Kawachi I, Glymour MM, eds. *Social Epidemiology*. 2nd ed. Oxford University Press; 2014:126-152.

26. Melita D, Rodríguez-Bailón R, Willis GB. Does income inequality increase status anxiety? Not directly, the role of perceived upward and downward mobility. *Br J Soc Psychol.* 2023;62(3):1453-1468.
27. Semega J, Kollar M. *Income in the United States: 2021.* Report P60-276. US Census Bureau; 2022.
28. Office for National Statistics. *Household Income Inequality, UK: Financial Year Ending 2022.* Office for National Statistics; 2023.
29. McFarland MJ, Hill TD, Montez JK. Income inequality and population health: examining the role of social policy. *J Health Soc Behav.* 2022;64(1): 2-20.
30. Macinko JA, Shi L, Starfield B, Wulu Jr. JT Income inequality and health: a critical review of the literature. *Med Care Res Rev.* 2003;60(4):407-452.
31. Zheng H. Do people die from income inequality of a decade ago? *Soc Sci Med.* 2012;75(1):36-45.
32. Wu S, Wang R, Zhao Y, et al. The relationship between self-rated health and objective health status: a population-based study. *BMC Public Health.* 2013;13(1):320.
33. McLeod C, Norman R, Litton E, Saville BR, Webb S, Snelling TL. Choosing primary endpoints for clinical trials of health care interventions. *Contemp Clin Trials Commun.* 2019;16:100486.
34. DeSalvo KB, Fisher WP, Tran K, Bloser N, Merrill W, Peabody J. Assessing measurement properties of two single-item general health measures. *Qual Life Res.* 2006;15(2):191-201.
35. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav.* 1997;38(1):21-37.
36. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev.* 2021;10(1):89.
37. Higgins JP, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions.* Version 6. Cochrane Collaboration; 2019.
38. Trapeznikova I. Measuring income inequality. IZA World of Labor. 2019. Accessed January 4, 2023. <https://wol.iza.org/articles/measuring-income-inequality/long>
39. Chen X. Relative deprivation and individual well-being: low status and a feeling of relative deprivation are detrimental to health and happiness. *IZA World Labor.* 2015;2015:140.
40. Eibner C, Sturn R, Gresenz CR. Does relative deprivation predict the need for mental health services? *J Ment Health Policy Econ.* 2004;7(4):167-175.
41. Sterne JAC, Hernán MA, McAleenan A, Reeves BC, Higgins JPT. Assessing risk of bias in a non-randomized study. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions.* Version 6.3. Cochrane Collaboration; 2022:621-641.
42. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol.* 2016;79:70-75.
43. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919.

44. Igelström E, Campbell M, Craig P, Katikireddi SV. Cochrane's risk of bias tool for non-randomized studies (ROBINS-I) is frequently misapplied: a methodological systematic review. *J Clin Epidemiol*. 2021;140:22-32.
45. Thomson H, Craig P, Hilton-Boon M, Campbell M, Katikireddi SV. Applying the ROBINS-I tool to natural experiments: an example from public health. *Syst Rev*. 2018;7(1):1-12.
46. Hilton Boon M, Thomson H, Shaw B, et al. Challenges in applying the GRADE approach in public health guidelines and systematic reviews: a concept article from the GRADE Public Health Group. *J Clin Epidemiol*. 2021;135:42-53.
47. Schünemann HJ, Cuello C, Akl EA, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol*. 2019;111:105-114.
48. Valentine JC, Thompson SG. Issues relating to confounding and meta-analysis when including non-randomized studies in systematic reviews on the effects of interventions. *Res Synth Methods*. 2013;4(1):26-35.
49. Tony B, Darren H, Alistair W. Confounding by socioeconomic position remains after adjusting for neighbourhood deprivation: an example using smoking and mortality. *J Epidemiol Community Health*. 2004;58(12):1030.
50. Kawachi I, Kennedy BP. The relationship of income inequality to mortality: does the choice of indicator matter? *Soc Sci Med*. 1997;45(7):1121-1127.
51. Lin Y, Zhang Q, Chen W, Ling L. The social income inequality, social integration and health status of internal migrants in China. *Int J Equity Health*. 2017;16(1):139.
52. World Bank Development Research Group. Gini index. 2022. Accessed April 25, 2022. World Bank. [iresearch.worldbank.org/PovcalNet/index.htm](https://research.worldbank.org/PovcalNet/index.htm)
53. Income distribution database. Organisation for Economic Co-operation and Development Statistics. 2022. Accessed January 5, 2023. <https://stats.oecd.org/Index.aspx?DataSetCode=IDD>
54. Breau S. Rising inequality in Canada: a regional perspective. *Appl Geogr*. 2015;61:58-69.
55. Lynch JW, Kaplan GA, Pamuk ER, et al. Income inequality and mortality in metropolitan areas of the United States. *Am J Public Health*. 1998;88(7):1074-1080.
56. Gustafsson B, Johansson M. In search of smoking guns: what makes income inequality vary over time in different countries? *Am Sociol Rev*. 1999;64(4):585-605.
57. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol*. 2014;14(1):135.
58. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol*. 2006;59(1):7-10.

59. Executive Director's Office for Germany. *World Bank Annual Report 2021*. World Bank; 2021.
60. Fritzell J. Income inequality trends in the 1980s: a five-country comparison. *Acta Sociol.* 1993;36(1):47-62.
61. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA.* 1998;280(19):1690-1691.
62. Shor E, Roelfs D, Vang ZM. The "Hispanic mortality paradox" revisited: meta-analysis and meta-regression of life-course differentials in Latin American and Caribbean immigrants' mortality. *Soc Sci Med.* 2017;186:20-33.
63. Ganna A, Ingelsson E. 5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study. *Lancet.* 2015;386(9993):533-540.
64. Gadalla TM, Fuller-Thomson E. Examining the lag time between state-level income inequality and individual disabilities: a multilevel analysis. *Am J Public Health (N Y).* 2008;98(12):2187-2190.
65. Hamilton TG, Kawachi I. Changes in income inequality and the health of immigrants. *Soc Sci Med.* 2013;80:57-66.
66. Steinbeis F, Gotham D, von Philipsborn P, Stratil JM. Quantifying changes in global health inequality: the Gini and Slope Inequality Indices applied to the Global Burden of Disease data, 1990–2017. *BMJ Glob Health.* 2019;4(5):e001500.
67. Tan AX, Hinman JA, Abdel Magid HS, Nelson LM, Odden MC. Association between income inequality and county-level COVID-19 cases and deaths in the US. *JAMA Network Open.* 2021;4(5):e218799-e218799.
68. Furceri D, Loungani P, Ostry JD, Pizzuto P. Will COVID-19 Have long-lasting effects on inequality? Evidence from past pandemics. *J Econ Inequal.* 2022;20(4):811-839.
69. Gottschalk P, Smeeding TM. Cross-national comparisons of earnings and income inequality. *J Econ Lit.* 1997;35(2):633-687.
70. Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.3. Cochrane Collaboration; 2022:241-284.
71. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2022.
72. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health.* 2019;22(4):153-160.
73. Crippa A, Orsini N. Multivariate dose-response meta-analysis: the *dostresmeta* R package. *J Stat Softw.* 2016;72(1):1-15.
74. Income inequality and health review repository. Updated 2024. Accessed January 4, 2024. <https://github.com/mshimonovich/income-inequality-sr>
75. Hussein H, Nevill CR, Meffen A, et al. Double-counting of populations in evidence synthesis in public health: a call for awareness and future methodological development. *BMC Public Health.* 2022;22(1):1827.

76. Senn SJ. Overstating the evidence—double counting in meta-analysis and related problems. *BMC Med Res Methodol.* 2009;9:1-7.
77. Bender R, Bunce C, Clarke M, et al. Attention should be given to multiplicity issues in systematic reviews. *J Clin Epidemiol.* 2008;61(9):857-865.
78. López-López JA, Page MJ, Lipsey MW, Higgins JPT. Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Res Synth Methods.* 2018;9(3):336-351.
79. Thompson SG. Systematic review: why sources of heterogeneity in meta-analysis should be investigated. *BMJ.* 1994;309(6965):1351-1355.
80. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ.* 2011;342:d549.
81. World Bank. *World Bank Annual Report 2021.* World Bank; 2021.
82. Borenstein M, Higgins JPT, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Res Synth Methods.* 2017;8(1):5-18.
83. Borenstein M, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis.* Wiley; 2009.
84. Schünemann H, Hill S, Guyatt G, Akl EA, Ahmed F. The GRADE approach and Bradford Hill's criteria for causation. *J Epidemiol Community Health.* 2011;65(5):392-395.
85. Schünemann HJ, Higgins JPT, Vist GE, et al. Completing 'summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions.* Version 6.3. Cochrane Collaboration; 2022:375-402.
86. Hultcrantz M, Rind D, Akl EA, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol.* 2017;87:4-13.
87. Howick J, Kelly P, Kelly M. Establishing a causal link between social relationships and health using the Bradford Hill Guidelines. *SSM Popul Health.* 2019;8:100402.
88. Olsen J, Jensen UJ. Causal criteria: time has come for a revision. *Eur J Epidemiol.* 2019;34(6):537-541.
89. Rothman KJ. The wrong message from the wrong talk. *Observational Studies.* 2020;6(2):30-32.
90. Shimonovich M, Pearce A, Thomson H, Katikireddi SV. Causal assessment in evidence synthesis: a methodological review of reviews. *Res Synth Methods.* 2022;13(4):405-423.
91. Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965;58(5):295-300.
92. Kawachi I, Kennedy BP. Income inequality and health: pathways and mechanisms. *Health Serv Res.* 1999;34(1 Pt 2):215-227.
93. Collier D. Understanding process tracing. *PS Polit Sci Polit.* 2011;44(4):823-830.
94. Mahoney J. The logic of process tracing tests in the social sciences. *Sociol Methods Res.* 2012;41(4):570-597.

95. Beach D, Pedersen RB. Developing empirical tests of causal mechanisms. In: Beach D, Pedersen RB, eds. *Process-Tracing Methods: Foundations and Guidelines*. University of Michigan Press; 2019.
96. Punton M, Katharina W. *Straws-in-the-Wind, Hoops and Smoking Guns: What Can Process Tracing Offer to Impact Evaluation?* Centre for Development Impact; 2015.
97. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med*. 2017;167(4):268-274.
98. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose–response data. *Stata J*. 2006;6(1):40-57.
99. Weiss NS. Can the “specificity” of an association be rehabilitated as a basis for supporting a causal hypothesis? *Epidemiology*. 2002;13(1):6-8.
100. Beach D, Pedersen RB. Developing empirical tests of causal mechanisms. In: Beach D, Pedersen RB, eds. *Process-Tracing Methods: Foundations and Guidelines*. University of Michigan Press; 2013.
101. Edvinsson S, Lundevaller EH, Malmberg G. Do unequal societies cause death among the elderly? A study of the health effects of inequality in Swedish municipalities in 2006. *Glob Health Action*. 2013;6:1-9.
102. Eibner C, Evans WN. Relative deprivation, poor health habits, and mortality. *J Hum Resour*. 2005;40(3):591-620.
103. Kim D. The associations between US state and local social spending, income inequality, and individual all-cause and cause-specific mortality: the National Longitudinal Mortality Study. *Prev Med*. 2016;84:62-68.
104. Kinge JM, Vallejo-Torres L, Morris S. Income related inequalities in avoidable mortality in Norway: a population-based study using data from 1994–2011. *Health Policy*. 2015;119(7):889-898.
105. Lochner K, Pamuk E, Makuc D, Kennedy BP, Kawachi I. State-level income inequality and individual mortality risk: a prospective, multilevel study. *Am J Public Health*. 2001;91(3):385-391.
106. Waitzman NJ, Smith KR, Stroup A. The direct and indirect effects of metropolitan area inequality on mortality. a hierarchical analysis. *Ann N Y Acad Sci*. 1999;896:347-349.
107. Aida J, Kondo K, Kondo N, Watt RG, Sheiham A, Tsakos G. Income inequality, social capital and self-rated health and dental status in older Japanese. *Soc Sci Med*. 2011;73(10):1561-1568.
108. Blakely TA, Kennedy BP, Glass R, Kawachi I. What is the lag time between income inequality and health status? *J Epidemiol Community Health*. 2000;54(4):318-319.
109. Blakely TA, Kennedy BP, Kawachi I. Socioeconomic inequality in voting participation and self-rated health. *Am J Public Health*. 2001;91(1):99-104.
110. Blakely TA, Lochner K, Kawachi I. Metropolitan area income inequality and self-rated health - a multi-level study. *Soc Sci Med*. 2002;54(1):65-77.
111. Fang H, Rizzo JA. Does inequality in China affect health differently in high-versus low-income households? *Appl Econ*. 2012;44(9):1081-1090.
112. Hou F, Chen J. Neighbourhood low income, income inequality and health in Toronto. *Health Rep*. 2003;14(2):21-34.

113. Li H, Zhu Y. Income, income inequality, and health: Evidence from China. *J Comparative Econ.* 2006;34(4):668-693.
114. Lopez R. Income inequality and self-rated health in US metropolitan areas: a multi-level analysis. *Soc Sci Med.* 2005;61(10):2267-2409.
115. Lorgelly PK, Lindley J. What is the relationship between income inequality and health? Evidence from the BHPS. *Health Econ.* 2008;17(2):249-265.
116. Mansyur C, Amick BC, Harrist RB, Franzini L. Social capital, income inequality, and self-rated health in 45 countries. *Soc Sci Med.* 2008;66(1):43-56.
117. McLeod CB, Lavis JN, Mustard CA, Stoddart GL. Income inequality, household income, and health status in Canada: a prospective cohort study. *Am J Public Health.* 2003;93(8):1287-1293.
118. Mellor JM, Milyo J. Income inequality and health status in the United States - evidence from the current population survey. *J Hum Resour.* 2002;37(3):510-539.
119. Mellor JM, Milyo J. Is exposure to income inequality a public health concern? Lagged effects of income inequality on individual and population health. *Health Serv Res.* 2003;38(1):137-151
120. Oshio T, Kobayashi M. Income inequality, perceived happiness, and self-rated health: evidence from nationwide surveys in Japan. *Soc Sci Med.* 2010;70(9):1358-1366.
121. Pasqualini M, Lanari D, Minelli L, Pieroni L, Salmasi L. Health and income inequalities in Europe: what is the role of circumstances? *Econ Hum Biol.* 2017;26:164-173.
122. Rozer J, Kraaykamp G, Huijts T. National income inequality and self-rated health: the differing impact of individual social trust across 89 countries. *Eur Soc.* 2016;18(3):245-263.
123. Rozer JJ, Volker B. Does income inequality have lasting effects on health and trust? *Soc Sci Med.* 2016;149:37-45.
124. Shi L, Starfield B. Primary care, income inequality, and self-rated health in the United States: a mixed-level analysis. *Int J Health Serv.* 2000;30(3):541-555.
125. Soobader MJ, LeClere FB. Aggregation and the measurement of income inequality: effects on morbidity. *Soc Sci Med.* 1999;48(6):733-744.
126. Subramanian SV, Kawachi I. Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the United States. *Health Place.* 2006;12(2):141-156.
127. Ting S, Zang W, Chen C, Chen D. Income distribution and health: what do we know from Chinese data? *PLoS One.* 2022;17(1):e0263008.
128. Zagorski K, Evans MDR, Kelley J, Piotrowska K. Does national income inequality affect individuals' quality of life in Europe? inequality, happiness, finances, and health. *Soc Indic Res.* 2014;117(3):1089-1110.
129. Adeline A, Delattre E. Some microeconomic evidence on the relationship between health and income. *Health Econ Rev.* 2017;7(1):27.

130. Baum S, Kendall E, Parekh S. Self-assessed health status and neighborhood context. *J Prev Interv Community*. 2016;44(4):283-295.
131. Bjornstrom EES. The neighborhood context of relative position, trust, and self-rated health. *Soc Sci Med*. 2011;73(1):42-49.
132. Bobak M, Murphy M, Rose R, Marmot M. Societal characteristics and health in the former communist countries of Central and Eastern Europe and the former Soviet Union: a multilevel analysis. *J Epidemiol Community Health*. 2007;61(11):990.
133. Bobak M, Pikhart H, Rose R, Hertzman C, Marmot M. Socioeconomic factors, material inequalities, and perceived control in self-rated health: cross-sectional data from seven post-communist countries. *Soc Sci Med*. 2000;51(9):1343-1350.
134. Cai JL, Laporte A, Zhang L, et al. Impacts of absolute and relative income on self-rated health in urban and rural China. *Int J Health Serv*. 2022;52(1):129-140.
135. Chen Z, Gotway Crawford CA. The role of geographic scale in testing the income inequality hypothesis as an explanation of health disparities. *Soc Sci Med*. 2012;75(6):1022-1031.
136. Chiavegatto ADP, Lebrão ML, Kawachi I. Income inequality and elderly self-rated health in Sao Paulo, Brazil. *Ann Epidemiol*. 2012;22(12):863-867.
137. Choi H, Burgard S, Elo IT, Heisler M. Are older adults living in more equal counties healthier than older adults living in more unequal counties? A propensity score matching approach. *Soc Sci Med*. 2015;141:82-90.
138. Craig N. Exploring the generalisability of the association between income inequality and self-assessed health. *Soc Sci Med*. 2005;60(11):2477-2488.
139. Evans MDR, Kelley J, Kelley CGE, Kelley SMC. Income inequality in the great recession did not harm subjective health in Europe, 2003–2012. *Appl Res Qual Life*. 2020;15(5):1451-1473.
140. Feng Z, Wang WW, Jones K, Li Y. An exploratory multilevel analysis of income, income inequality and self-rated health of the elderly in China. *Soc Sci Med*. 2012;75(12):2481-2492.
141. Fiscella K, Franks P. Individual income, income inequality, health, and mortality: what are the relationships? *Health Serv Res*. 2000;35(1):307-318.
142. Góngora-Salazar P, Casabianca MS, Rodríguez-Lesmes P. Income inequality and self-rated health status in Colombia. *Int J Equity Health*. 2022;21(1):69.
143. Gravelle H, Sutton M. Income, relative income, and self-reported health in Britain 1979–2000. *J Health Econ*. 2009;18(2):125-145.
144. Haithcoat TL, Avery EE, Bowers KA, Hammer RD, Shyu CR. Income inequality and health: expanding our understanding of state-level effects by using a geospatial big data approach. *Soc Sci Comput Rev*. 2021;39(4):543-561.
145. Hildebrand V, Van Kerm P. Income inequality and self-rated health status: evidence from the European Community Household Panel. *Demography*. 2009;46(4):805-825.

146. Hou F, Myles J. Neighbourhood inequality, neighbourhood affluence and population health. *Soc Sci Med.* 2005;60(7):1557-1569.
147. Ichida Y, Kondo K, Hirai H, Hanibuchi T, Yoshikawa G, Murata C. Social capital, income inequality and self-rated health in Chita peninsula, Japan: a multilevel analysis of older people in 25 communities. *Soc Sci Med.* 2009;69(4):489-499.
148. Jen MH, Jones K, Johnston R. Global variations in health: evaluating Wilkinson's income inequality hypothesis using the World Values Survey. *Soc Sci Med.* 2009;68(4):643-653.
149. Kahn RS, Wise PH, Kennedy BP, Kawachi I. State income inequality, household income, and maternal mental and physical health: cross sectional national survey. *BMJ.* 2000;321(7272):1311-1315.
150. Karlsdotter K, Martin JJM, Gonzalez M. Multilevel analysis of income, income inequalities and health in Spain. *Soc Sci Med.* 2012;74(7):1099-1106.
151. Kennedy BP, Kawachi I, Glass R, Prothrow-Stith D. Income distribution, socioeconomic status, and self rated health in the United States: multilevel analysis. *BMJ.* 1998;317(7163):917-921.
152. LeClere FB, Soobader MJ. The effect of income inequality on the health of selected US demographic groups. *Am J Public Health (N Y).* 2000;90(12):1892-1897.
153. Massa KHC, Pabayo R, Chiavegatto Filho ADP. Income inequality and self-reported health in a representative sample of 27 017 residents of state capitals of Brazil. *J Public Health (Oxf).* 2018;40(4):e440-e446.
154. Oshio T, Kobayashi M. Income inequality, area-level poverty, perceived aversion to inequality, and self-rated health in Japan. *Soc Sci Med.* 2009;69(3):317-326.
155. Qi YQ. The impact of income inequality on self-rated general health: evidence from a cross-national study. *Res Soc Stratif Mobil.* 2012;30(4):451-471.
156. Rostila M, Kolegard ML, Fritzell J. Income inequality and self-rated health in Stockholm, Sweden: a test of the 'income inequality hypothesis' on two levels of aggregation. *Soc Sci Med.* 2012;74(7):1091-1098.
157. Shibuya K, Hashimoto H, Yano E. Individual income, income distribution, and self rated health in Japan: cross sectional analysis of nationally representative sample. *BMJ.* 2002;324(7328):16-19.
158. Sommet N, Elliot AJ. The effects of U.S. county and state income inequality on self-reported happiness and health are equivalent to zero. *Qual Life Res.* 2022;31(7):1999-2009.
159. Subramanian SV, Blakely T, Kawachi I. Income inequality as a public health concern: where do we stand? Commentary on "Is exposure to income inequality a public health concern?" *Health Serv Res.* 2003;38(1 Pt 1):153-167.
160. Subramanian SV, Delgado I, Jadue L, Vega J, Kawachi I. Income inequality and health: multilevel analysis of Chilean communities. *J Epidemiol Community Health.* 2003;57(11):844-848.

161. Vauclair CM, Marques S, Lima ML, et al. Perceived age discrimination as a mediator of the association between income inequality and older people's self-rated health in the European region. *J Gerontol B Psychol Sci Soc Sci*. 2015;70(6):901-912.
162. Weich S, Lewis G, Jenkins SP. Income inequality and self rated health in Britain. *J Epidemiol Community Health*. 2002;56(6):436-441.
163. Wen M, Browning CR, Cagney KA. Poverty, affluence, and income inequality: neighborhood economic structure and its implications for health. *Soc Sci Med*. 2003;57(5):843-860.
164. Wong IO, Cowling BJ, Lo SV, Leung GM. A multilevel analysis of the effects of neighbourhood income inequality on individual self-rated health in Hong Kong. *Soc Sci Med*. 2009;68(1):124-132.
165. Xi GL, McDowell I, Nair R, Spasoff R. Income inequality and health in Ontario - a multilevel analysis. *Can J Public Health*. 2005;96(3):206-211.
166. Zheng H. Rising U.S. income inequality, gender and individual self-rated health, 1972–2004. *Soc Sci Med*. 2009;69(9):1333-1342.
167. Auger N, Hamel D, Martinez J, Ross NA. Mitigating effect of immigration on the relation between income inequality and mortality: a prospective study of 2 million Canadians. *J Epidemiol Community Health*. 2012;66(6):e5.
168. Backlund E, Rowe G, Lynch J, Wolfson MC, Kaplan GA, Sorlie PD. Income inequality and mortality: a multilevel prospective study of 521 248 individuals in 50 US states. *Int J Epidemiol*. 2007;36(3):590-596.
169. Blakely T, Atkinson J, O'Dea D. No association of income inequality with adult mortality within New Zealand: a multi-level study of 1.4 million 25–64 year olds. *J Epidemiol Community Health*. 2003;57(4):279.
170. Gerdtham U-G, Johannesson M. Absolute income, relative income, income inequality, and mortality. *J Hum Resour*. 2004;39(1):228-247.
171. Henriksson G, Allebeck P, Weitoft GR, Thelle D. Are manual workers at higher risk of death than non-manual employees when living in Swedish municipalities with higher income inequality? *Eur J Public Health*. 2007;17(2):139-144.
172. Kimmel PL, Fwu CW, Eggers PW. Segregation, income disparities, and survival in hemodialysis patients. *J Am Soc Nephrol*. 2013;24(2):293-301.
173. Kravdal O. Does income inequality really influence individual mortality? Results from a 'fixed-effects analysis' where constant unobserved municipality characteristics are controlled. *Demogr Res*. 2008;18:205-232.
174. Modrek S, Dow WH, Rosero-Bixby L. Long-term association of economic inequality and mortality in adult Costa Ricans. *Soc Sci Med*. 2012;74(2):158-166.
175. Ng N, Lundevaller E, Malmberg G, Edvinsson S. Income inequality and old-age mortality in Sweden: do regional development and lagged effect matter? *Health Place*. 2020;64:102384.
176. Osler M, Prescott E, Gronbaek M, Christensen U, Due P, Engholm G. Income inequality, individual income, and mortality in Danish adults: analysis of pooled data from two cohort studies. *BMJ*. 2002;324(7328):13-16.

177. Pabayo R, Chiavegatto Filho AD, Lebrao ML, Kawachi I. Income inequality and mortality: results from a longitudinal study of older residents of Sao Paulo, Brazil. *Am J Public Health (N Y)*. 2013;103(9):e43-e49.
178. Zhao L, Hessel P, Simon Thomas J, Beckfield J. Inequality in place: effects of exposure to neighborhood-level economic inequality on mortality. *Demography*. 2021;58(6):2041-2063.
179. Zheng H. Do people die from income inequality of a decade ago? *Soc Sci Med*. 2012;75(1):36-45.
180. Caicedo-Velasquez B, Restrepo-Mendez MC. The role of individual, household, and area of residence factors on self-rated health in Colombian adults: a multilevel study. *Biomedica*. 2020;40(2):296-308.
181. Latif E. Income inequality and health: panel data evidence from Canada. *BE J Econ Anal Policy*. 2015;15(2):927-959.
182. Massa KHC, Chiavegatto ADP. Income inequality and self-reported health among older adults in Brazil. *J Appl Gerontol*. 2021;40(2):152-161.
183. Pei X, Rodriguez E. Provincial income inequality and self-reported health status in China during 1991-7. *J Epidemiol Community Health*. 2006;60(12):1065-1069.
184. Subramanian SV, Kawachi I. The association between state income inequality and worse health is not confounded by race. *Int J Epidemiol*. 2003;32(6):1022-1028.
185. Subramanian SV, Kawachi I, Kennedy BP. Does the state you live in make a difference? Multilevel analysis of self-rated health in the US. *Soc Sci Med*. 2001;53(1):9-19.
186. Zimmerman FJ, Bell JF. Income inequality and physical and mental health: testing associations consistent with proposed causal pathways. *J Epidemiol Community Health*. 2006;60(6):513-521.
187. Fiscella K, Franks P. Poverty or income inequality as predictor of mortality: longitudinal cohort study. *BMJ*. 1997;314(7096):1724-1727.
188. Dahl E, Elstad JI, Hofoss D, Martin-Mollard M. For whom is income inequality most harmful? A multi-level analysis of income inequality and mortality in Norway. *Soc Sci Med*. 2006;63(10):2562-2574.
189. Daly MC, Duncan GJ, Kaplan GA, Lynch JW. Macro-to-micro links in the relation between income inequality and mortality. *Milbank Q*. 1998;76(3):315-339.
190. Elstad JI. Does the socioeconomic context explain both mortality and income inequality? Prospective register-based study of Norwegian regions. *Int J Equity Health*. 2011;10:7.
191. Grönqvist H, Johansson P, Niknami S. Income inequality and health: lessons from a refugee residential assignment program. *J Health Econ*. 2012;31(4):617-629.
192. Henriksson G, Allebeck P, Weitoft GR, Thelle D. Income distribution and mortality: implications from a comparison of individual-level analysis and multilevel analysis with Swedish data. *Scand J Public Health*. 2006;34(3):287-294.

193. Kahn HS, Patel AV, Jacobs EJ, Calle EE, Kennedy BP, Kawachi I. Pathways between area-level income inequality and increased mortality in U.S. men. *Ann N Y Acad Sci.* 1999;896:332-334.
194. Lau EW, Schooling MC, Tin KY, Leung GM. Income inequality and cause-specific mortality during economic development. *Ann Epidemiol.* 2012;22(4):285-294.
195. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology.* 2009;20(4):488-495.
196. Tennant PWG, Murray EJ, Arnold KF, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *Int J Epidemiol.* 2020;50(2):620-632.
197. The world by income and region. The World Bank. Accessed April 25, 2022. <https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html>
198. Mutz J, Lewis CM. Cross-classification between self-rated health and health status: longitudinal analyses of all-cause mortality and leading causes of death in the UK. *Sci Rep.* 2022;12(1):459.
199. Deaths in the UK from 1990 to 2020. Office for National Statistics. February 5, 2021. Accessed January 4, 2023. <https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/deathsintheukfrom1990to2020>
200. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence - study limitations (risk of bias). *J Clin Epidemiol.* 2011;64(4):407-415.
201. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence - inconsistency. *J Clin Epidemiol.* 2011;64(12):1294-1302.
202. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence - publication bias. *J Clin Epidemiol.* 2011;64(12):1277-1282.
203. van Aert RCM, Wicherts JM, van Assen MALM. Publication bias examined in meta-analyses from psychology and medicine: a meta-meta-analysis. *PLoS One.* 2019;14(4):e0215052.
204. Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics.* 2018;74(3):785-794.
205. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol.* 2011;64(12):1283-1293.
206. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence - indirectness. *J Clin Epidemiol.* 2011;64(12):1303-1310.
207. Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Overviews of reviews. In: Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.4. Cochrane Collaboration; 2023.

208. McCartney G, Hearty W, Arnot J, Popham F, Cumbers A, McMaster R. Impact of political economy on population health: a systematic review of reviews. *Am J Public Health (N Y)*. 2019;109(6):e1-e12.
209. Atkinson AB. A critique of the transatlantic consensus on rising income inequality. *World Econ*. 2001;24(4):433-452.
210. Anand S, Segal P. What do we know about global income inequality? *J Econ Lit*. 2008;46(1):57-94.
211. Brakenhoff TB, Mitroiu M, Keogh RH, Moons KGM, Groenwold RHH, van Smeden M. Measurement error is often neglected in medical literature: a systematic review. *J Clin Epidemiol*. 2018;98:89-97.
212. Fernando GDM. Income inequality measures. *J Epidemiol Community Health*. 2007;61(10):849.
213. Cowell F. Inequality and poverty measures. In: Adler MD, Fleurbaey M, eds. *The Oxford Handbook of Well-Being and Public Policy*. 3rd ed. Oxford University Press; 2016.
214. Hoffmann R, Borassio DC. Sensitivity of inequality measures considering regressive transfers with fixed relative income distance. *Metron*. 2020;78(3): 279-296.
215. Sommet N, Morselli D, Spini D. Income inequality affects the psychological health of only the people facing scarcity. *Psychol Sci*. Published online October 12, 2018. <https://doi.org/10.1177/0956797618798620>.
216. Richardson M, Garner P, Donegan S. Interpretation of subgroup analyses in systematic reviews: a tutorial. *Clin Epidemiol Glob Health*. 2019;7(2):192-198.
217. Wagstaff A, Doorslaer EV. Income inequality and health: what does the literature tell us? *Annu Rev Public Health*. 2000;21(1):543-567.
218. Deaton A. Health, inequality, and economic development. *J Econ Lit*. 2003;41(1):113-158.
219. McCartney G, Hearty W, Arnot J, Popham F, Cumbers A, McMaster R. Impact of political economy on population health: a systematic review of reviews. *Am J Public Health*. 2019;109(6):e1-e12.
220. Pickett KE, Wilkinson RG. Income inequality and health: a causal review. *Soc Sci Med*. 2015;128:316-326.
221. Mahutga MC, Bandelj N. Foreign investment and income inequality: the natural experiment of Central and Eastern Europe. *Int J Comp Sociol*. 2008;49(6):429-454.
222. Mavridis D, Mosberger P. *Income Inequality and Incentives: The Quasi-Natural Experiment of Hungary 1914–2008*. World Wealth & Income Database; 2017. *WID.world Working Paper Series*; no 2017/17.
223. Milanovic B, Gradstein M, Ying Y. *Democracy and Income Inequality: An Empirical Analysis*. The World Bank; 2001. *Policy Research Working Paper Series*; no 2561.
224. Van der Hoeven R. *Labour Markets and Income Inequality What Are the New Insights after the Washington Consensus?* United Nations University, World Institute for Development Economic Research; 2000. *World Institute for Development Economic Research Working Papers*; no 295541.

225. Leibbrandt M, Woolard I. The labour market and household income inequality in South Africa: existing evidence and new panel data. *J Int Dev.* 2001;13(6):671-689.
226. Houweling TAJ, Kunst AE. Socio-economic inequalities in childhood mortality in low- and middle-income countries: a review of the international evidence. *Br Med Bull.* 2010;93(1):7-26.
227. Wildman J. Modelling health, income and income inequality: the impact of income inequality on health and health inequality. *J Health Econ.* 2003;22(4):521-538.
228. Ngamaba KH, Panagioti M, Armitage CJ. Income inequality and subjective well-being: a systematic review and meta-analysis. *Qual Life Res.* 2018;27(3):577-596.
229. Claire B, Brendan TS, Elaine H, Ambikaipakan S, Paul JV, Roman P. Investigating the association between income inequality in youth and deaths of despair in Canada: a population-based cohort study from 2006 to 2019. *J Epidemiol Community Health.* 2023;77(1):26.
230. Miller JR, Piper TM, Ahern J, et al. Income inequality and risk of suicide in New York City neighborhoods: a multilevel case-control study. *Suicide Life Threat Behav.* 2005;35(4):448-459.
231. Nandi A, Galea S, Ahern J, Bucciarelli A, Vlahov D, Tardiff K. What explains the association between neighborhood-level income inequality and the risk of fatal overdose in New York City? *Soc Sci Med.* 2006;63(3):662-674.
232. Rowlingson K. *Does Income Inequality Cause Health and Social Problems?* Joseph Rowntree Foundation; 2011.
233. Tibber MS, Walji F, Kirkbride JB, Huddy V. The association between income inequality and adult mental health at the subnational level—a systematic review. *Soc Psychiatry Psychiatr Epidemiol.* 2022;57(1):1-24.
234. Thomson RM, Igelström E, Purba AK, et al. How do income changes impact on mental health and wellbeing for working-age adults? A systematic review and meta-analysis. *Lancet Public Health.* 2022;7(6):e515-e528.
235. Popham F, Gray L, Bambra C. Employment status and the prevalence of poor self-rated health. Findings from UK individual-level repeated cross-sectional data from 1978 to 2004. *BMJ Open.* 2012;2(6):e001342.

Funding/Support: This work is supported by Wellcome Trust (218105/Z/19/Z and 205412/Z/16/Z), NHS Research Scotland (SCAF/15/02), Medical Research Council (MC_UU_00022/2), Chief Scientist Office (SPHSU17) and the European Research Council (949582). The funders played no active role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication. All authors have completed the ICMJE uniform disclosure form at <http://www.icmje.org/disclosureofinterest/> and declare no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, and no other relationships or activities that could appear to have influenced the submitted work.

Conflict of Interest Disclosures: No disclosures were reported.

Address correspondence to: Michal Shimonovich, MRC/CSO Social and Public Health Sciences Unit, School of Health and Wellbeing, University of Glasgow, Clarice Pears Building, 90 Byres Rd, Glasgow G12 8TB (email: Michal.shimonovich@glasgow.ac.uk).

Supplementary Material

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/journal/10.1111/1468-0009> (ISSN)1468-0009:

Online Appendix