Published in final edited form as:
Epidemiology. 2016 September ; 27(5): 670–676. doi:10.1097/EDE.0000000000000517.

Comparing alternative effect decomposition methods: the role of literacy in mediating educational effects on mortality

Thu T. Nguyen[a], **Eric J. Tchetgen Tchetgen**[b,c], **Ichiro Kawachi**[d], **Stephen E. Gilman**[c,d,e], **Stefan Walter**[a], and **Maria Glymour**[a]

[a]Department of Epidemiology & Biostatistics, University of California, San Francisco

[b]Department of Biostatistics, Harvard T.H. Chan School of Public Health

[c]Department of Epidemiology, Harvard T.H. Chan School of Public Health

[d]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health

[e]*Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Division of Intramural Population Health Research

Abstract

Background—Inverse odds ratio weighting, a newly proposed tool to evaluate mediation in exposure-disease associations, may be valuable for a host of research questions but little is known about its performance in real data. We compare this approach to a more conventional Baron and Kenny decomposition on an additive hazards scale to estimate total, direct, and indirect effects using the example of the role of literacy in mediating the effects of education on mortality.

Methods—Health and Retirement Study participants born in the U.S. between 1900 and 1947 were interviewed biennially for up to 12 years (N=17,054). Literacy was measured with a brief vocabulary assessment. Decomposition estimates were derived based on Aalen additive hazards models.

Results—A one standard deviation difference in educational attainment (3 years) was associated with 6.7 fewer deaths per 1,000 person-years (β=-6.7, 95% CI: -7.9, -5.4). Of this decrease, 1.3 fewer deaths (β=-1.3, 95% CI: -4.0, 1.2) were attributed to the literacy pathway (natural indirect), representing 19% of the total effect. Baron and Kenny estimates were consistent with inverse odds ratio weighting estimates but were more precise (natural indirect effect: -1.2 (95% CI: -1.7, -0.69, representing 18% of total effect).

Conclusion—In a cohort of older Americans, literacy partially mediated the effect of education on mortality.

Despite numerous studies documenting educational inequalities in health, a long-standing gap in knowledge is whether these inequalities are due to differences in cognitive skills acquired through schooling, social norms, health benefits of credentialing, or other

Correspondence: Thu T. Nguyen, Department of Epidemiology & Biostatistics, University of California, San Francisco, 550 16th St, 2nd Floor, San Francisco, CA, 94158-2549. Thu.Nguyen@ucsf.edu.

Conflicts of Interest: None declared.

factors.[1-3] If knowledge- or skills-based mechanisms can be established as important, this will help guide efforts to redress educational inequalities in health. Literacy encompasses a foundational set of skills involved with gathering, processing, synthesizing, and utilizing information that may be critical for health maintenance and promotion. However, very few studies have examined literacy as a possible mediator of the relationship between education and health.[4-6]

Methods to evaluate mediation are hampered by several challenges and have been an important area of development in recent years. One such challenge in conducting mediation analyses is identifying mechanisms in the presence of statistical interaction between the exposures and hypothesized mediators. A new innovation in mediation analyses, the inverse odds ratio weighting approach, does not make parametric assumptions about the joint effect of exposures and mediators, and as a result solves some limitations but may introduce others. In this analysis, we compare a conventional approach to the newly developed technique to quantify the extent to which literacy mediates the relationship between education and mortality.

Inverse odds ratio weighting, developed by Tchetgen Tchetgen,[7] can be used to formally evaluate mediation and has several important strengths. It is a flexible approach that can be implemented with any software that fits standard regression models with weights, such as generalized linear models (GLM), including those with linear or non-linear link functions; models for survival outcomes subject to censoring; and quantile regression. It is agnostic to the presence or absence of exposure–mediator interactions and remains valid in both cases. Inverse odds ratio weighting provides a semiparametric generalization of the well-known Baron and Kenny approach.[8] In contrast to inverse odds ratio weighting, the Baron and Kenny approach is readily applicable only in situations where exposure-mediator interactions are assumed to be absent. Recent work in mediation analyses have adapted a counterfactual approach and extended the work of Baron and Kenny to explicitly account for such interactions[9] by directly modeling and incorporating them in the mediation formula of Pearl.[10] In the context of survival data, mediation analysis has previously been implemented under the Aalen additive hazards model, assuming no interaction between the exposure and mediator.[11] In this paper, we apply an approach we shall refer to as Baron and Kenny, due to its striking similarity to the original approach of Baron and Kenny in the case of standard linear regression analysis, and compare it to inverse odds ratio weighting.

Mediation Analyses

In this paper, we fit Aalen additive hazards models to estimate hazard differences. The Aalen model is a semiparametric model that estimates the hazard at time t as a linear function of covariates and an unspecified baseline hazard.[12] In comparison to relative measures, absolute effect measures are particularly relevant for policy-making and public health interventions[12] as they more clearly indicate the potential public health impact of intervening on the exposure and/or mediator of interest. One strength of the Aalen model is that it allows for time-varying effects. We implemented the Kolmogorov-Smirnov test of effect heterogeneity to determine whether education has time-dependent effects.

In assessing mediation, we estimated natural direct effects, natural indirect effects, and total effects. The natural direct effect (also known as the pure direct effect) in this paper is the change in the predicted number of deaths per 1,000 person-years per unit change in years of schooling when the mediator, literacy, is set at the value it would take at 12 years of schooling (reference level of the exposure). The natural indirect effect is the change in predicted number of deaths per 1,000 person-years when years of schooling is held at 15 years (one standard deviation above the mean), but literacy is changed from the value it would take at 12 years of schooling to the value it would take at 15 years of schooling. The total effect is the sum of the natural direct and indirect effects.[13] Pure or natural direct and indirect effects are suited to evaluating mechanisms and will be the focus of this paper.[10,14]

Natural direct and indirect effects are not identified without strong assumptions about the absence of unobserved confounding. These include assuming no unmeasured confounding of the relationship between 1) exposure and mediator, 2) mediator and outcome, and 3) exposure and outcome upon conditioning on pre-exposure confounders. We also assume (4) that there are no confounding variables of the mediator outcome relationship, whether observed or unobserved, that are affected by the exposure.[10,15] For a counterfactual formalization of these assumptions, see Pearl.[10]

For assumption 4 to be violated, there must be a variable that is a consequence of education and that goes on to confound the literacy-mortality relationship. We considered the following potential mediator–outcome confounders: mother's and father's education, intelligence, the home environment, family socioeconomic conditions. None of these are post exposure (education), and thus they would not violate assumption 4.

Methods

The sample included 17,054 participants from the Health and Retirement Study (HRS), a longitudinal study of U.S. adults aged 50 and over and their spouses. The first survey wave was collected in 1992, with biennial interviews (or proxy interviews for decedent participants) available through 2010. The sample was restricted to members who were alive and interviewed in 1998. From an initial total sample of 21,384 members alive and interviewed in 1998, we excluded those who were: not born between 1900 and 1947 (n=961), born outside the U.S. (n=2,847), missing place of birth (n=71), missing literacy score (n=1,387), or missing self-reported childhood health (n=25). This brought the final analytic sample to 17,054. HRS was approved by the University of Michigan Health Sciences Human Subjects Committee, and the Harvard School of Public Health Human Subjects Committee determined the current analyses were exempt.

Measures

Exposure and Mediator—The main exposure was educational attainment operationalized as years of schooling, as reported by the respondent at the first HRS interview. We considered alternative modeling strategies, which included adding a quadratic term and modeling education with discontinuities at high school and college completion. However, these different approaches did not result in substantial improvement in model fit with the Akaike Information Criterion remaining virtually unchanged across the different models (eTable 1). Thus, for simplicity and ease of interpretation, we modeled education as a continuous linear term.

Our mediator of interest was literacy. The HRS interview assessed literacy with a set of five vocabulary words of increasing difficulty, which respondents were asked to define. Responses were recorded and subsequently coded according to the degree of accuracy (0=incorrect; 1=partially correct; 2= perfectly correct). Scores had a theoretical range from 0 to 10 with a mean of 5.50 and standard deviation of 2.00. There were two sets of words, one of which was randomly assigned at the participant's first interview. Beginning at Wave 4 (1998), the vocabulary word sets were alternated in successive waves and only asked of re-interviewees who were 65 years of age or older. Our measure of literacy is the average of 1995/1996 (wave 3) and 1998 (wave 4) vocabulary score. This measure was highly correlated with the Wide Range Achievement Test Version 3 (WRAT-3) Reading subtest,[16] a widely used and validated measure of literacy. The WRAT Reading Recognition subtest involves identifying letters and making correct pronunciation of a series of words. Correlation between the average vocabulary score and the WRAT reading subtest total raw score was 0.75 in the subsample of HRS participants with both WRAT and vocabulary score (n=382). We chose to use the average literacy score as it was more highly correlated with the WRAT scores than either the wave 3 or wave 4 measure alone. To facilitate comparison of the natural direct and indirect effects and because our measure of literacy did not have established cut-points, we standardized education and literacy. Effect estimates refer to a one standard deviation difference.

Outcomes—Mortality status and date and month of death were obtained through exit interviews with surviving relatives and linkages to the National Death Index (NDI). For this study, the follow-up period was 1998-2010. Mortality status and date of death was verified through the NDI for deaths occurring between 1998-2008. Linkage between HRS and the NDI had not yet been completed for 2008-2010 at the time of these analyses, and death dates during this period were obtained through exit interviews with surviving relatives. We assumed censoring to be independent of time to death given the exposure and covariates. We additionally assumed the mediator to be independent of censoring given the exposure and covariates. Participants contributed 159,663 person-years, and there were 6,382 deaths.

Covariates—We attempted to control for variables likely to influence education, literacy, and mortality. Covariates included age measured in 1998, sex, race/ethnicity (Non-Hispanic White, Non-Hispanic Black, Hispanic, Other), self-rated childhood health status (dichotomized as excellent, very good, or good versus fair or poor), and five indicators of early life socioeconomic status combined into a single scale: mother's and father's educational attainment, father's occupational status, birth in the southern US, and rural residence during childhood, as in previous research in the Health and Retirement Study.[17] Confirmatory factor analysis was conducted using Mplus Version 7 to reconstruct the early life socioeconomic status (SES) scale, following previous research.[17]

Mediation Analyses—The inverse odds ratio weighting approach of Tchetgen Tchetgen[7] was utilized to estimate natural direct, natural indirect, and total effects. Please see Nguyen et al. for an application of this approach with a binary exposure.[18]

Inverse odds ratio weighting condenses the relationship between exposure and mediators using the odds ratio function as a measure of association into a weight, removing the necessity to specify the regression model for the outcome on exposure and mediator, including any exposure–mediator interactions. The weight (the inverse exposure–mediator odds ratio) given covariates is used to estimate the natural direct effect via regression analysis. Applying the weight renders the exposure and mediator independent, deactivating the indirect pathways involving the mediator.

A key advantage derives from the invariance property of the odds ratio (i.e. the odds ratio for the relationship between two variables is the same regardless of which variable is specified as dependent or independent), which permits estimation of the odds ratio relating an exposure and mediator via multiple logistic regression for binary exposures or via linear regression for continuous exposures.[19]

Compared to the familiar odds ratio formula relating two binary variables, the formula for an odds ratio with two continuous variables simply replaces the probability that the binary variables take the value of 1 with the density at the observed value of each variable, with the reference category still at zero. The odds ratio for the relationship between exposure (X) and mediator (M), conditional on covariates (C) is given by

$$\begin{aligned}\mathrm{OR(X,M|C)} &= \mathrm{f(M|X,C)f(M{=}0|X{=}0,C)/(f(M|X{=}0,C)f(M{=}0|X,C))}\\ &= \mathrm{f(X|M,C)f(X{=}0|M{=}0,C)/(f(X|M{=}0,C)f(X{=}0|M,C))}\end{aligned}$$

where f(A|B) is the density of A given B, for any random variables A and B.

Here we apply inverse odds ratio weighting with a continuous exposure. We use the representation of the odds ratio (above), which we evaluate assuming X is normally distributed. If we first estimate a model for X condition on M and C:

$$\mathrm{X|M,C}=\beta_0+\beta_1 M+\beta'_2 C+e \text{ with } e\sim \mathrm{N}(0,\sigma^2) \quad (1)$$

The conditional inverse odds ratio function 1/OR(X,M|C) relating years of schooling (X) and mediator, literacy (M) within levels of pre-exposure confounders (C) can then be shown to be equal to $\exp(-\beta_1\times X\times M/\sigma^2)$.[7]

Then, it can be shown that under our assumptions, the estimated natural direct effect on a given scale (e.g. generalized linear models, Cox proportional hazards model or Aalen additive hazards model) is the regression coefficient for the exposure in the weighted regression model for the outcome on the exposure and covariates for example as in Equation 2, with inverse odds ratio weight 1/OR(X,M|C).[7]

$$\gamma(t;X,C)=\lambda_o(t)+\lambda_1 X+\lambda_2{}' C \quad (2)$$

where γ is the hazard function at time t given years of schooling X and covariates C. The total effect is the coefficient for the exposure in the analogous unweighted regression analysis. The natural indirect effect is estimated by taking the difference between the total effect and the natural direct effect on the scale used to obtain direct and total effects. We assume exposure and mediator have time-constant effects. Ninety-five percent confidence intervals for the total, natural direct effects and natural indirect effects are computed via the nonparametric bootstrap.

For comparison, we also implement the Baron and Kenny approach for mediation analyses by fitting two sets of Aalen additive hazards models for the outcome with (3) and without (4) the mediator.

$$\gamma(t;X,M,C)=\sigma_o(t)+\sigma_1 X+\sigma_2 M+\sigma_3{}' C \quad (3)$$

$$\gamma(t;X,M,C)=\lambda_o(t)+\lambda_1 X+\lambda_3{}' C \quad (4)$$

The direct effect is estimated by the coefficient for the exposure in the model with the mediator (Equation 3). The total effect is estimated by the coefficient for the exposure in the model without the mediator (Equation 4), and the indirect effect is given by the taking the difference in the coefficients for the exposures from the two models $(\lambda_1-\sigma_1)(t)$. To investigate whether the total effect of education and the effect of education through literacy varied by educational level, we repeated the analyses stratified by education level (≤12 years vs ≥13).

Age is a potential confounder in many epidemiologic cohort studies including this one. In such instances, age may be added as a covariate to the models, but this approach depends on specifying the correct functional form. One recommended alternative is to treat age as the primary time-scale when modeling time to event data to provide stronger control of confounding by age.[20] To determine whether the results were robust to this modeling choice, both models were fit, including age as a continuous variable as well as treating age as the primary time-scale.

Descriptive analyses were conducted using SAS 9.3. Aalen models were fit using R 2.15.2 (see eAppendix for sample code for mediation analyses).

Results

Descriptive statistics for the study population are presented in Table 1. Mean literacy scores rose with increasing years of education. Thirty-eight percent of the participants died by 2010. The Kolmogorov-Smirnov test of no effect heterogeneity indicated years of schooling did not have time-dependent effects, so we modeled education and covariates as having time constant effects (see online eAppendix for further details).

The estimated total effect of one standard deviation of education (3 years) on mortality was 6.7 fewer deaths per 1,000 person-years (β=-6.7, 95% CI: -7.9, -5.4) (Table 2). When allowing for interaction between years of schooling and literacy, there was evidence of positive interaction, motivating the need to utilize a method that accommodates exposure-mediator interactions. The interaction coefficient was estimated to be 1.4 deaths per 1,000 person-years (95% CI: 0.41, 2.5). The interaction term was positive, while the main effects were negative. This result suggested that the protective effect of literacy on time to death diminished by 1.4 deaths per 1,000 person-years for every unit (one standard deviation) increase in education and vice versa.

Utilizing the Baron and Kenny approach for mediation, 1.2 fewer deaths per 1,000 person-years was attributed to the literacy pathway (β=-1.2, 95% CI: -1.7, -0.69) (natural indirect effect), representing 18% of the total effect of education (Table 2). The inverse odds ratio weighting point estimates were very similar to those obtained through the Baron and Kenny approach but were less precisely estimated. Inverse odds ratio weighting estimates for the natural direct and indirect effects were -5.5 (95% CI: -8.2, -2.6) and -1.3 (-4.0, 1.2) compared to the Baron and Kenny estimates of -5.5 (95% CI: -6.9, -4.1) and -1.2 (95% CI: -1.7, -0.69), respectively. The natural indirect estimates suggest a protective effect of literacy. Although we detected modest exposure–mediator interaction, the consistency of the results between the two approaches indicated that the interaction did not make a substantial difference to the results.

In the current analyses, age was treated as a centered, continuous variable. In sensitivity analyses, we used chronological age as the primary time scale to treat age non-parametrically. The results are qualitatively similar to using follow-up time as the primary time scale, and none of our conclusions changed (see online eAppendix eTable 2).

Stratified Results

We also conducted the mediation analyses, stratifying by high school completion. The total effect was larger among participants ≤ 12 years of education compared to those with ≥ 13 years of education (≤12: -8.4 (95% CI: -11.5, -6.6); ≥ 13: -3.7 (95% CI: -6.6, -0.80)) (Table 3). Using the Baron and Kenny approach, the estimated natural indirect effect suggested partial mediation among participants with ≤ 12 years of education (-1.7; 95% CI: -2.3, -1.1), representing 20% of the total effect. The inverse odds ratio weighting point estimate for the natural indirect effect was qualitatively similar but had wider confidence intervals (-0.95; 95% CI: -5.5, 4.7). Among those with ≥ 13 years of education, the indirect effects using Baron and Kenny and inverse odds ratio weighting were substantially smaller (Table 3). These results suggested that education was a stronger predictor of mortality among those with 12 or fewer years of education, and the effect of education through the literacy pathway was observed primarily among those with lower levels of education.

In supplemental analyses, we also attempted to investigate measurement error in the mediator. We used the Wide Range Achievement Test Version 3 Reading subtest,[16] a widely used measure of reading ability, as validation data to correct for measurement error in a modified inverse odds ratio weighting approach (see online eAppendix for a description of the approach). Measurement error correction resulted in very wide confidence bounds (see online eAppendix eTables 3-4). We also modified our bootstrapping code to take account of clustering at the household level. The results taking account of clustering were very similar to the original results not correcting for clustering, and none of the conclusions changed (see eTable 5).

Discussion

In a large cohort sample of older Americans, educational attainment was inversely associated with mortality risk, and literacy was a partial mediator of this relationship. In the current analyses, inverse odds ratio weighting results were consistent with Baron and Kenny point estimates but were less precise. Stronger evidence of a total and mediated effect was found among those with 12 or fewer years of education compared to those with 13 or greater years of education. This work adds to a limited body of research examining literacy as mediator of educational attainment and health.

Education is a social determinant of health. More education is associated with longer life expectancy, greater likelihood of engaging in health promoting behaviors, and better health outcomes. Education is also associated with decreased likelihood of smoking, greater likelihood of obtaining health care check-ups and screening, and higher levels of physical activity. Parents' education also has an intergenerational effect on their children's educational attainment and health.[21] Education potentially impacts health and mortality via many pathways. It is associated with increased cognitive functioning[22] and skills such as literacy. It influences occupational opportunities, shaping both material and psychosocial risk factors at work, and earnings.[23] It facilitates access to information as well as resources,[24] including those conveyed via social ties.[25] It can also influence health via psychological processes such as increasing sense of control and social standing.[21]

Literacy can influence health maintenance, health promotion, and ultimately survival in several ways. Literacy increases the capacity to obtain, process, and understand health information.[26,27] Literacy facilitates the comprehension of prescriptions, health care worker's instructions for disease management, printed nutrition information, and publically available health information.[26] Literacy can impact health by influencing access to and utilization of health care, the patient–provider relationship, and self-care.[28] Lower adult literacy is associated with less knowledge of health outcomes such as smoking, hypertension, diabetes, and lower use of screening and preventive services[29] as well as poorer mental and self-rated health and higher hospital admissions.[30]

Education, literacy, intelligence, and cognitive function are associated, but the direction of this association is not clear. In the U.S., the average IQ has risen 20 points over the last 60 years, a rate too rapid to be caused by genetic selection. Increased access to education and changes in public education have been proposed as possible drivers of improvements in IQ.[31,32] Research has also examined the impact of education and literacy on later life cognitive function. Low reading ability was found to predict incidence of dementia[33] and declines in memory,[34] language skills, and executive functioning.[35]

In our analyses, we rely on the assumption of no unmeasured confounding of the relationship between 1) exposure and mediator, 2) mediator and outcome, and 3) exposure and outcome upon conditioning on pre-exposure confounders. One potential confounder relevant for these assumptions is intelligence, which may affect both literacy and subsequent health. However, the Health and Retirement Study does not have an early life measure of intelligence. Although cognitive function later in life is measured, late life cognitive function may be affected by the respondents' educational attainment, literacy level, or age-associated cognitive disorders such as dementia. As a result, late life cognitive function may partially mediate the relationship between education, literacy, and mortality.

The absence of an early life measure of intelligence is a data limitation of our study, but we note that there is strong reason to believe that schooling influences cognitive functioning independent of intelligence. Previous research controlling for childhood IQ has found educational attainment to increase cognitive human capital.[22] When comparing schooled and unschooled children in remote rural communities, even small amounts of schooling were associated with higher cognitive functioning.[36] Investigators examining the direct effect of intelligence on health did not find evidence of an effect once education and income were controlled.[37] eFigure presents a direct acyclic graph of our hypothesized relationships between education, literacy, intelligence, later life cognitive function, and mortality (see online eAppendix).

Another potential confounder of the exposure-outcome or mediator-outcome relationship is childhood health status. Self-reported childhood health was included in our models but this measure may be subject to recall bias since it was asked of HRS participants when they were adults.

An additional consideration was survivor bias. Survivor bias in this case would likely underestimate the observed effects of education and literacy. To change the observed percent mediated effect, there would have to be substantial differential survivor bias by education and literacy. We included Health and Retirement Study respondents who were 50 or older in 1998, and it is plausible that respondents' education and literacy level, which were measured at nearly the same time (1996 and 1998 for literacy and 1998 for education), influenced their survival and enrollment into the study. We used a brief vocabulary measure to assess literacy. We did not consider numeracy, health literacy, or other domains that may be relevant for health outcomes.

The choice between various mediation methods rests on evaluating the advantages and limitations of the different approaches, which will vary from study to study. In this paper, we saw that there was a tradeoff between relaxing certain assumptions and optimizing precision in implementing inverse odds ratio weighting or the Baron and Kenny approach. Since inverse odds ratio weighting does not make parametric assumptions about the joint effect of exposures and mediators and can be implemented in a variety of models that accommodate weights such as survival models, quantile regression models, and any generalized linear models with linear or non-linear links, researchers using data with strong exposure-mediator interactions or fitting complex models may find inverse odds ratio weighting more attractive than the Baron and Kenny approach or other related parametric methods.[7] In other situations such as when there is limited sample size, the absence or interactions, or a relatively simple model, the Baron and Kenny approach may be an appropriate option. We implemented the Baron and Kenny approach and inverse odds ratio weighting in R, writing our own analytic code for the analyses. The recently released medflex package in R allows for mediation analyses using imputation strategies for the estimation of natural effects.[38] This study contributes to the growing body of evidence showing the long-term health benefits of years of schooling.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of Funding: Dr. Tchetgen Tchetgen's work was funded by the National Institutes of Health (NIH) grants AI113251 and ES020337.

Dr. Gilman's work was supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development

References

1. Kawachi I, Adler NE, Dow WH. Money, schooling, and health: Mechanisms and causal evidence. Annals of the New York Academy of Sciences. 2010; 1186(1):56–68. [PubMed: 20201868]
2. Chandola T, Clarke P, Morris JN, Blane D. Pathways between education and health: a causal modelling approach. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2006; 169(2):337–359.
3. Ross CE, Mirowsky J. Refining the Assocation between Education and Health: The Effects of Quantity, Credential, and Selectivity. Demography. 1999; 36(4):445–460. [PubMed: 10604074]
4. Schillinger D, Barton LR, Karter AJ, Wang F, Adler N. Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes. Public Health Rep. 2006; 121(3):245–254. [PubMed: 16640146]
5. Pandit AU, Tang JW, Bailey SC, et al. Education, literacy, and health: Mediating effects on hypertension knowledge and control. Patient Education and Counseling. 2009; 75(3):381–385. [PubMed: 19442477]
6. Loucks EB, Gilman SE, Howe CJ, et al. Education and coronary heart disease risk: potential mechanisms such as literacy, perceived constraints, and depressive symptoms. Health Educ Behav. 2015; 42(3):370–379. [PubMed: 25431228]
7. Tchetgen Tchetgen EJ. Inverse odds ratio-weighted estimation for causal mediation analysis. Statistics in Medicine. 2013; 32(26):4567–4580. [PubMed: 23744517]
8. Baron M, Kenny A. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology. 1986; 51:1173–1182. [PubMed: 3806354]
9. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013; 18(2):137–150. [PubMed: 23379553]
10. Pearl, J. Direct and indirect effects. In: Breese, J.; Koller, D., editors. Proceedings of the seventeenth conference on uncertainty in artificial intelligence. San Francisco, CA: Morgan Kaufmann; 2001. p. 411-420.
11. Lange T, Hansen JV. Direct and Indirect Effects in a Survival Context. Epidemiology. 2011; 22(4): 575–581. [PubMed: 21552129]
12. Rod NH, Lange T, Andersen I, Marott JL, Diderichsen F. Additive interaction in survival analysis: use of the additive hazards model. Epidemiology. 2012; 23(5):733–737. [PubMed: 22732385]

13. Pearl J. The causal mediation formula--a guide to the assessment of pathways and mechanisms. Prev Sci. 2012; 13(4):426–436. [PubMed: 22419385]
14. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology. 1992; 3(2):143–155. [PubMed: 1576220]
15. Avin, C.; Shpitser, I.; Pearl, J. Identifiability of Path-Specific Effects. Edinburgh; Scotland: 2005.
16. Wilkinson, GS. WRAT3: Wide Range Achievement Test Administration Manual. Wilmington, DE: Wide Range, Inc; 1993.
17. Nandi A, Glymour MM, Kawachi I, VanderWeele TJ. Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. Epidemiology. 2012; 23(2):223–232. [PubMed: 22317806]
18. Nguyen QC, Osypuk TL, Schmidt NM, Glymour MM, Tchetgen Tchetgen EJ. Practical guidance for conducting mediation analysis with multiple mediators using inverse odds ratio weighting. Am J Epidemiol. 2015; 181(5):349–356. [PubMed: 25693776]
19. Tchetgen Tchetgen EJ. On the interpretation, robustness, and power of varieties of case-only tests of gene-environment interaction. Am J Epidemiol. 2010; 172(12):1335–1338. [PubMed: 21098632]
20. Cologne J, Hsu WL, Abbott RD, et al. Proportional hazards regression in epidemiologic follow-up studies: an intuitive consideration of primary time scale. Epidemiology. 2012; 23(4):565–573. [PubMed: 22517300]
21. Egerter, S.; Braveman, P.; T, S-N.; Grossman-Kahn, R.; Dekker, M. [March 20, 2012] Education Matters for Health. Robert Wood Johnson Foundation: Commission to Build a Healthier America. 2009. http://www.commissiononhealth.org/PDF/c270deb3-ba42-4fbd-baeb-2cd65956f00e/Issue%20Brief%206%20Sept%2009%20-%20Education%20and%20Health.pdf
22. Herd P. Education and health in late-life among high school graduates: Cognitive versus psychological aspects of human capital. J Health Soc Behav. 2010; 51(4):478–496. [PubMed: 21131622]
23. Lynch, J.; Kaplan, G. Socialeconomic position. In: Berkman, LF.; Kawachi, I., editors. Social Epidemiology. New York New York: Oxford University Press; 2000. p. 13-35.
24. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. Health Aff (Millwood). 2002; 21(2):60–76. [PubMed: 11900187]
25. Ross CE, Wu C. The Links Between Education and Health. American Sociological Review. 1995; 60(5):719–745.
26. Weiss BD, Hart G, Pust RE. The relationship between literacy and health. J Health Care Poor Underserved. 1991; 1(4):351–363. [PubMed: 1932458]
27. Baker D, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. Archives of Internal Medicine. 2007; 167(14):1503–1509. [PubMed: 17646604]
28. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. American Journal of Health Behavior. 2007; 31:S19–26. [PubMed: 17931132]
29. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med. 2004; 19(12):1228–1239. [PubMed: 15610334]
30. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med. 2011; 155(2):97–107. [PubMed: 21768583]
31. Blair C, Gamson D, Thorne S, Baker D. Rising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. Intelligence. 2005; 33(1):93–106.
32. Flynn, JR. Are We Getting Smarter? Rising IQ in the Twenty-First Century. Cambridge University Press; 2012.
33. Schmand B, Smit JH, Geerlings MI, Lindeboom J. The effects of intelligence and education on the development of dementia. A test of the brain reserve hypothesis. Psychol Med. 1997; 27(6):1337–1344. [PubMed: 9403905]

34. Mathews M, Abner E, Caban-Holt A, Dennis BC, Kryscio R, Schmitt F. Quality of education and memory test performance in older men: the New York University Paragraph Recall Test normative data. Curr Alzheimer Res. 2013; 10(7):776–783. [PubMed: 23906000]

35. Manly JJ, Schupf N, Tang MX, Stern Y. Cognitive decline and literacy among ethnically diverse elders. J Geriatr Psychiatry Neurol. 2005; 18(4):213–217. [PubMed: 16306242]

36. Stevenson H, Chen C, Booth J. Influences of schooling and urban-rural residence on gender differences in cognitive abilities and academic achievement. Sex Roles. 1990; 23(9–10):535–551.

37. Link BG, Phelan JC, Miech R, Westin EL. The Resources That Matter: Fundamental Social Causes of Health Disparities and the Challenge of Intelligence. Journal of Health and Social Behavior. 2008; 49(1):72–91. [PubMed: 18418986]

38. Steen, J. [November 10, 2015] Medflex. 2015. https://github.com/jmpsteen/medflex

Table 1

Demographic characteristics of the analytic sample

	N	(%)	Mean Literacy Score (SD)
N	17054	100	
Years of schooling			
<6	505	3	2.6 (2.0)
6-8	1588	9	3.8 (1.9)
9-11	2591	15	4.6(1.8)
12	5998	35	5.5 (1.7)
>12	6372	37	6.5 (1.6)
Male	7052	41	
Birth year			
<1914	1117	7	
1914-1921	2460	14	
1922-1930	4044	24	
1931-1941	6800	40	
1942-1947	2633	15	
Non-Hispanic White	13821	81	
Non-Hispanic Black	2401	14	
Hispanic	611	4	
Other	220	1	

Table 2

Total, natural direct and indirect effects for mortality (1,000 person-years) estimated with Baron and Kenny and inverse odds ratio weighting approaches using follow-up years as the timescale

Treatment Effects (1,000 person-years)[a,b] (N=17,054)	Baron and Kenny	95% CI[c]	Inverse Odds Ratio Weighting	95% CI[c]
Total[d]	-6.7	(-7.9, -5.4)	-6.7	(-7.9, -5.4)
Natural Direct	-5.5	(-6.9, -4.1)	-5.5	(-8.2, -2.6)
Natural Indirect	-1.2	(-1.7,-0.69)	-1.3	(-4.0, 1.2)

[a]Exposure: z-scored years of schooling with one standard deviation = 3 years of schooling
Mediator: z-scored literacy score with one standard deviation =2 on a 10-point scale.

[b]Covariates: age, sex, race/ethnicity, self-rated childhood health status, and five indicators of early life SES combined into a single scale including mother's and father's educational attainment, father's occupational status, birth in southern US, and rural residence during childhood.

[c]Obtained using nonparametric bootstrap.

[d]Estimation of total effects is the same for Baron and Kenny and inverse odds ratio weighting.

Table 3

Total, natural direct and indirect effects for mortality (1,000 person-years) using Baron and Kenny and inverse odds ratio weighting among participants with ≤12 and 13+ years of schooling

Treatment Effects (1,000 person-years)[a,b]	Baron and Kenny	95% CI[c]	Inverse Odds Ratio Weighting	95% CI[c]
		≤ 12 years of schooling (N=10,682)		
Total[d]	-8.4	(-11.5, -6.6)	-8.4	(-11.5, -6.6)
Natural Direct	-6.7	(-9.7, -4.8)	-7.4	(-14.3, -2.9)
Natural Indirect	-1.7	(-2.3,-1.1)	-0.95	(-5.5, 4.7)
		13+ years of schooling (N=16,611)		
Total[d]	-3.7	(-6.6, -0.80)	-3.7	(-6.6, -0.80)
Natural Direct	-3.4	(-6.5, -0.60)	-3.1	(-6.6, 0.85)
Natural Indirect	-0.23	(-0.91,0.58)	-0.60	(-3.2, 1.7)

[a]Exposure: z-scored years of schooling with one standard deviation = 3 years of schooling
Mediator: z-scored literacy score with one standard deviation =2 on a 10-point scale.

[b]Covariates: age, sex, race/ethnicity, self-rated childhood health status, and five indicators of early life SES combined into a single scale including mother's and father's educational attainment, father's occupational status, birth in southern US, and rural residence during childhood.

[c]Obtained using nonparametric bootstrap.

[d]Estimation of total effects is the same for Baron and Kenny and inverse odds ratio weighting.