# Linking studies to assess the life expectancy associated with eighth grade school achievement☆

Robert A. Hahn\*, Sajal K. Chattopadhyay

*Community Guide Branch, Division of Public Health Information Dissemination, Center for Surveillance, Epidemiology, and Laboratory Services, Office of Public Health Scientific Services, CDC, Atlanta, Georgia*

ABSTRACT

Extensive evidence indicates the causal association of school outcomes and long-term health. We combined the findings of two studies by Chetty and colleagues to estimate the life expectancy associated with achievement scores in the eighth grade. We linked the dependent variable of the first study and the independent variable of the second study. The first study (of students in Tennessee) found a positive correlation between school achievement scores in eighth grade and income at age 25–27. Controlling for family background, a one percentile increase in eighth grade test score was associated with an increase of $148 (95% CI: $125, $172) in 2009 $U.S. in mean yearly wages at ages 25–27 years. Based on estimated mean annual income growth of 3.35%, $148 would increase 1.59 fold to $235 (CI: $199, $273) in 14 years, at age 40—$251 (CI: $213, $292) in 2012 $U.S. adjusted for inflation. The second study (of the U.S. population) found that a one percentile household income ($1500 in 2012 $U.S.) was associated with one month life expectancy at age 40. We calculate that an increase in income at age 40 attributable to one percentile increase in eighth grade test scores, i.e., $251, would increase life expectancy by 17% (i.e., $251/$1500) (CI: 14%, 19%) of one month per percentile eighth grade test score. Estimates of long-term health outcomes associated with educational outcomes can be made with caution. Applicability of findings from the Tennessee to the U.S. population is discussed.

## 1. Introduction

A wide array of research has established that educational achievement (e.g., test scores and grades) and attainment (e.g., high school or college completion) are fundamental social determinants of long-term health (Feinstein et al., 2006; Hahn and Truman, 2015; Ross and Wu, 1995; Zajacova and Lawrence, 2018; Kubota et al., 2017). Use of multiple data sources and diverse methods, including experiments, strengthen causal inference (Hahn and Truman, 2015). Educational achievement is linked to lifelong health through three interrelated pathways—(1) development of psychological and interpersonal strengths, such as a sense of control and social support, which contribute to healthy social interactions; (2) problem-solving abilities and the ability to maintain productive work, adequate income—a mediator in the linkage with long-term health, and the health benefits these intermediate outcomes provide; and (3) adoption of healthy behaviors (Hahn and Truman, 2015; Ross and Wu, 1995). The greatest proportion of long-term health benefits of education are effected not by health

education, but by the cultivation of problem solving skills and the facilitation of abilities to regulate emotions, interact effectively, and demonstrate self-efficacy and agency (Ross and Wu, 1995; Mirowsky and Ross, 2017). The Guide to Community Preventive Services (https://www.thecommunityguide.org/) has conducted systematic reviews of educational programs to promote health equity demonstrating the benefits of diverse programs, e.g., early childhood education, programs to improve high school completion, and summer-school-based and after-school programs, in improving educational achievement (Hahn et al., 2016; Hahn et al., 2015; Knopf et al., 2015). Establishing the link between educational achievement and long-term health outcomes was a prerequisite in these reviews.

Several previous studies have estimated the amount or proportion of major, long-term health outcomes associated with educational outcomes. Rostron et al. (2010) linked the Current Population Survey with the National Longitudinal Mortality Study to estimate life expectancy associated with different levels of educational attainment. Their data indicate non-concordance of self-reported educational attainment and

educational attainment reported on death certificates of between 30% and 51%. Adjusted for this misclassification, at age 45, males and females who completed more than high school can expect to live, respectively, 8.9 and 8.4 years more than non-graduates. Jemal et al. (2008) used 2001 mortality data for persons aged 25–64 years, estimating that, in that age range, approximately 48% of deaths among men and 38% of deaths in women would not have occurred if all segments of the population had experienced the death rates of college graduates. Krueger et al. (2015) linked National Health Interview data from 1986 through 2004 with mortality records for persons aged 25–85 years of age to estimate mortality associated with educational levels. Applying these findings to U.S. educational levels in 2010, approximately 145,243 deaths among those born in 1945 were attributable to having less than a high school degree, similar to the mortality attributable to smoking (Krueger et al., 2015). Both Jemal and Krueger compared mortality by educational attainment levels. Galea et al. (2011) used logically equivalent population attributable risk methods and estimated that approximately 245,000 deaths among adults in the year 2000 were attributable to having less than a high school education, compared with having a high school education or more. Other studies have assessed the association of education with other health outcomes—self-rated health, life expectancy, disability, and chronic conditions (Subramanian et al., 2002; Braveman et al., 2010; Campbell et al., 2014).

This study focuses on educational achievement and its impact on longevity. In contrast to "educational attainment," our exposure variable is a single, point-in-time standardized assessment of accumulated knowledge and capacity assessed at an early age. We used what might be called a "two-stage population attributable risk" approach to estimate the life expectancy associated with a student's school achievement score. We are not aware of prior estimates using income as an intermediating variable to link an educational outcome with life expectancy, a major long-term health outcome. This link enables us to address the implications of educational interventions to boost achievement scores for public health policy and action.

The two studies we link derive from distinct populations—the first from Tennessee students (Chetty et al., 2011) with outcomes we project to age 40, the second from all U.S. adults (Chetty et al., 2016). We use the first study to estimate an effect of early school achievement on adult income; subsequently, we use the second study to obtain the effect of that income on life expectancy. Because Tennessee demographics differ from those of the whole U.S., we allow for effect modification of our findings and propose several ways in which our findings might be adjusted to account for these differences.

As noted before, studies of the long-term health outcomes of exposures such as education commonly use information available on death certificates, but education data on death certificates are commonly in error (Rostron et al., 2010). In this study, we use intermediate income data on individuals whose educational outcome is systematically measured 20 years earlier, linked with a second source.

## 2. Methods

Valid estimation of the proportion of an outcome attributable to an exposure rests on two related assumptions. The first assumption is that the link between the exposure and outcome is causal, i.e., that in the presence of the exposure, the outcome occurs, and that, in the absence of the exposure, some portion of the outcome does not occur—assuming that there are other, independent causes of the same outcome. The second assumption is that the relative risk linking exposure and outcome that is used to estimate attributable risk is adjusted for confounders (Rockhill et al., 1998). Even with randomized experiments, it is difficult, if not impossible, to assure that all confounders are equally distributed between subjects exposed and unexposed to potential confounders. Although all possible confounders are rarely known, those that are recognized and observed can sometimes be controlled for. In
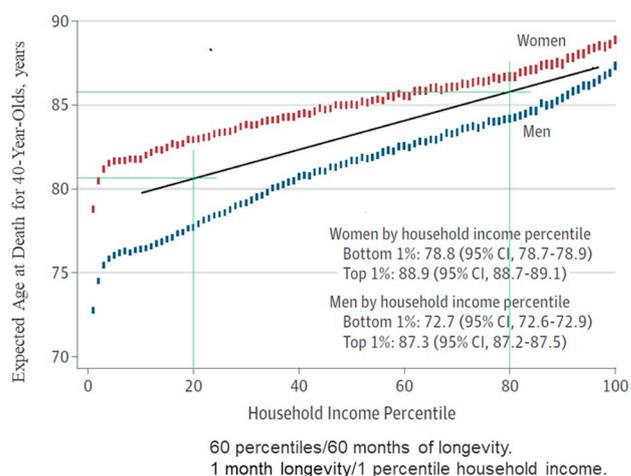
the case of the association between educational achievement and long-term health, parental socioeconomic position is a major confounder (Reynolds and Ross, 1998). In the study we deploy here, the authors use regression analyses to control for parental characteristics including household income and marital status, mother's age at child's birth, indicators for parent's 401(k) savings—an indicator of retirement savings and wealth, and home ownership, as well as student's race, gender, free-lunch status, and age at kindergarten ("KG"), thus improving the likelihood of a valid estimate of attributable risk (Chetty et al., 2011).

We use two cohort studies conducted by Chetty and colleagues. The first study (Chetty et al., 2011) extends the experimental STAR (Student/Teacher Achievement Ratio) study designed to determine the effects of class size on student learning. Between 1985 and 1989, 11,571 students in 79 schools in Tennessee were randomly assigned to classrooms of different sizes. Students were tested annually, up to eighth grade, with the Stanford Achievement test in math and reading. Fifty percent of STAR participants were eligible for free or reduced-price lunch (Finn et al., 2001), indicating a high rate of poverty compared with the U.S. population; Chetty et al. noted that the mean wages of STAR participants at ages 25–27 years were lower (by 22.4%) than the mean wages of U.S. residents of the same age. Chetty et al. (2011) used a program developed by the Statistics of Income Division of the Internal Revenue Service (IRS) to link the records of STAR participants with their income (in 2009 dollars) when they reached age 25–27.

Chetty et al.'s (2016) second study, using IRS data from 1999 to 2014 and Social Security mortality records, linked household income at age 40 (in 2012 U.S. dollars) with life expectancy at age 40 in the U.S. population. The researchers plotted this association separately for males and females. The association was linear for the range of data, except for the tails of the plots, which were steeper (Fig. 1). To link this source with income findings from eighth grade achievement test, we approximated a mean effect for males and females combined, ignoring the tails of the distribution.

To estimate the effects of wage growth from age 26 to 40 years, we used an annual wage growth rate of 3.35% (Bureau of Economic Analysis, Current Population Survey, Bureau of Labor Statistics, unpublished data, 1979–2016). To enable direct comparison of incomes between the two studies by Chetty et al., we converted mean wages from the first study expressed in 2009 U.S. dollars to 2012 U.S. dollars using the consumer price index (https://inflationdata.com/Inflation/

## Linking of Two Functions:
## Household Income and Life Expectancy at Age 40 Years



60 percentiles/60 months of longevity.
1 month longevity/1 percentile household income.

Chetty, Raj, et al. "The association between income and life expectancy in the United States, 2001-2014." *JAMA* 315.16 (2016): 1750-1766.

**Fig. 1.** Life expectancy at age 40 associated with household income at age 40 (Chetty et al., 2016). Modified by present authors to indicate overall trends and estimate slope.
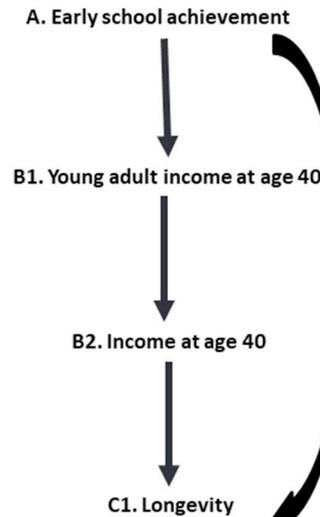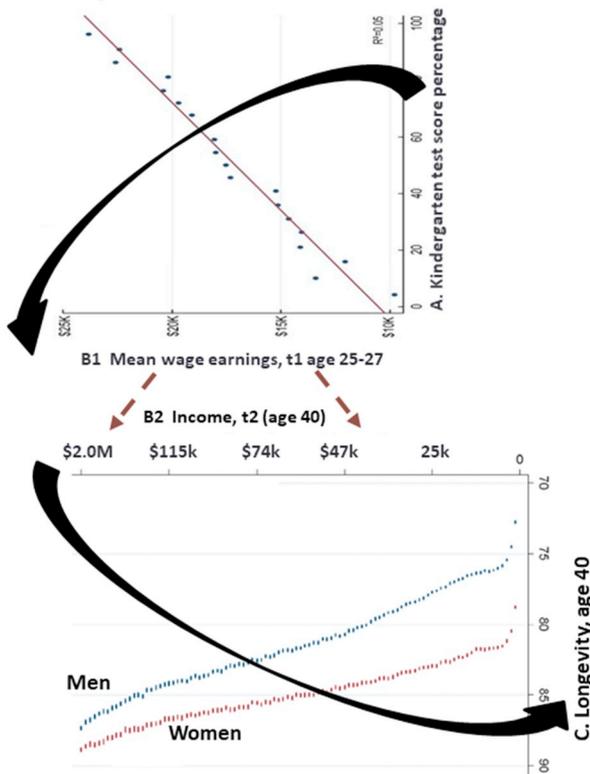
## Plan of Analysis

Consumer_Price_Index/HistoricalCPI.aspx?reloaded = true). We assessed the life expectancy at age 40 associated with a percentile increase in household income and then calculated the proportion of STAR participants' incremental income by age 40 as a fraction of the percentile household income in the U.S. at age 40 (Chetty et al., 2011) to ascertain the life expectancy at age 40 associated with amount of school achievement in eighth grade. Fig. 2 illustrates our linking of the two sources; the dependent variable in the first study is linked with the independent variable of the second study, thus providing a bridge between eighth grade test score and life expectancy at age 40.

### 3. Results

In Chetty et al.'s extension of the STAR study, KG entry grade test percentile was strongly and positively correlated with mean wage earnings at ages 25–27 years (Fig. 3). When adjusted for fixed class size effects and parent and student background, a one percentile increase in KG achievement test score was associated with an increased wage at ages 25–27 of $93.80 (Table IV in Chetty et al., 2011) in 2009 U.S. dollars. Chetty et al. report that the effect of the KG test score was completely captured by the eighth grade test score, implying in essence that there is a continuity over time linking KG to eighth grade test scores. Thus, one percentile increase in eighth grade achievement test score, also adjusted for parent and student backgrounds, was associated with an increased wage at ages 25–27 years of $148.20 (95% CI: $125, $172).

From this baseline, we calculated what would be the increase in a STAR participant's mean wage at age 40. With a wage growth of 3.35% per year, the mean wage would increase by 1.59 fold ($1.0335^{14}$) over 14 years, or to $235 (CI: $199, $273) in 2009 US dollars. In 2012 U.S. dollars, as used in the second Chetty et al. study, this is equivalent to $251 (CI: $213, $292).

In Chetty et al.'s (2016) study of the association of income and life expectancy at age 40 in the U.S. population, the difference in life expectancy between men aged 40 years at the highest and lowest income

## Correlation of Kindergarten Achievement Score Percentiles and Wage Income at Ages 25–27
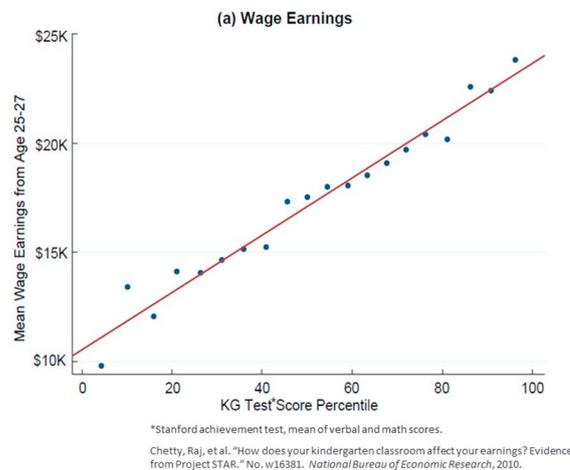


**Fig. 3.** Wage earnings at ages 25–27 years of STAR cohort associated with kindergarten achievement score percentiles (reading and math combined) (Chetty et al., 2011). (Trend line in original; wage earnings in eighth grade almost totally absorbed by kindergarten scores, but eighth grade graphic not provided in Chetty et al., 2011.)

percentiles was 14.6 years; for women aged 40 years, the comparable gap was 10.1 years. Excluding the tails of the association, a one percentile increase in household income, approximately $1500 (in 2012 U.S. dollars), was associated with approximately one month of added life expectancy (Fig. 1).

Finally, we assessed the increase in 40-year-old mean wage of STAR eighth graders as a fraction of the percentile change in household income associated with a month of life expectancy. This fraction is 17%

(i.e., \$251/\$1500) (CI: 14%, 19%), which as a fraction of a month is 5.3 days (CI: 4, 5.9 days). This estimate remained approximately the same when we used the average annual wage growth rates for employees either 25 years and above (3.3%) or between 25 and 35 years (3%).

## 4. Discussion

Conducting primary studies to examine the relationships between educational achievement assessed through a standardized test score and long-term health outcomes is both expensive and time-consuming. We combined secondary information from published literature to link early increases in test scores to long-term increases in life expectancy. A one percentile increase in eighth grade test score was shown to be associated with approximately 5 days of additional life expectancy at age 40, which is 36% of the gain in life expectancy between 2001 and 2014 realized by all women in bottom 5% of the income distribution and 5% for men (Chetty et al., 2016). Our estimate is conservative, as education may have direct influence on both health and duration of life by means of pathways not mediated by income (Ross and Wu, 1995). In addition, while we have used the individual's income at age 40 as a fraction of household income that is the independent variable in Chetty et al.'s longevity study, it is reasonable to suppose that the individual may be married or have others in the household contributing to household income; thus, the individual's income may be the lower bound of household income, and longevity gain per percentile educational score might be greater than that estimated here. On the other hand, a valid estimate may be less than what we find because of confounders we could not control and the possibility of reverse causality, e.g., the contribution of health to opportunities for employment and income. Finally, Chetty et al. (2011) indicated, even with adjustment for potential confounders, KG and eighth grade test scores could explain only 17% of the variation in mean wages at ages 25–27, suggesting the important role of other factors that could not be controlled in their ordinary least-squares regression model. It is also possible that what appears to be an effect of education might be an effect of innate intellectual capacity that leads to educational success; however, there is good evidence that some of the effects of innate intellectual capacity on income are mediated through education (Hahn and Truman, 2015).

For purposes of comparison, smoking of one carton of cigarettes has been estimated to be associated with 1½ days decreased life expectancy at birth (Shaw et al., 2000). In addition, using IMS Health Drug Launches database and longitudinal, disease-level data from 52 countries, an international study found that under conservative assumptions, the average annual increase in life expectancy of the entire population resulting from new drug launches was about 1 week (Lichtenberg, 2005). Note that the life expectancies at birth estimated in these studies would be substantially smaller to be appropriately compared with the life expectancies at age 40 that are the outcome of our analysis.

Our analysis indicates the powerful influence of early education on a key long-term health outcome, substantially prior to the completion of high school. A recent systematic economic review of three types of early childhood education programs (state and district, federal Head Start, and model programs) found future earnings gain was the major benefit among outcomes assessed (Ramon et al., 2018).

We note several limitations to our study. First, we combined percentile test score and mean adult earnings based on STAR (Tennessee) data, but the income-life expectancy relationship in the second study used national level data. We were not able to isolate regional data specific to STAR participants to generate a better representation of the income-life expectancy gradient for STAR participants. Compared with the rest of the nation, the STAR participants were from a relatively low-income population, with 50% eligible for free or reduced-price lunches and 33% of the STAR sample being black. On the other hand, Chetty et al.'s STAR estimates are adjusted for parental socioeconomic position so that the association found between school achievement and income

is independent of this parental status, and likely to be applicable beyond the Tennessee sample.

Chetty's income and life expectancy study includes plots of this association for 4 cities—New York City, San Francisco, Dallas, and Detroit. Of these four, Detroit has the lowest median household income and so might be most similar to the population of Tennessee. Among the four cities, slope generally decreases with city median income. The income–life expectancy plot is steepest for Detroit, and it is plausible that a plot for Tennessee would also be steeper. In this case, our estimate would be conservative.

Second, although high school completion and subsequent college education are important mediators that link percentile change in entry level or eighth grade test scores and change in mean earnings, only 26.4% of students in the STAR sample were enrolled in college at age 20 (year 2000), and 45.5% of students were enrolled in college at some point between 1999 and 2007, compared with 57.1% in the same cohort of the U.S. population. It is not clear if there would be any differential impacts of education on income and life expectancy based on these specific characteristics of the STAR participants.

Third, although STAR data allowed for control of some key confounders in the association of early education achievement and mean wage at ages 25–27 years, the second study linking income and life expectancy did not present results based on regression modeling or adjustment for potential confounders. As Chetty and colleagues note: (Chetty et al., 2011) "Because of such unmeasured confounding factors, the causal effects of income on life expectancy are likely to be smaller than the associations documented in this study." Fourth, our crude estimate based on the two-stage population attributable risk approach may compound the inaccuracies of measurement of each component study. In addition, there are several approximations in our analysis. Fifth, in our use of Chetty et al.'s life expectancy study in the U.S. population, to estimate a mean effect, we have drawn a straight line between male and female plots that are not precisely linear, and, as noted, we have ignored the tails of these plots. The tails are steeper, indicating a greater positive effect of household income on life expectancy. Given the low SES background of STAR participants, is it possible that even with increase in mean incomes, the STAR adults would be placed closer to the tails? In that case again, our estimate would be conservative.

Life expectancy effects of education are more likely to occur for preventable causes, such as tobacco-related causes, rather than causes for which preventive measures are not known (Phelan et al., 2004). Part of the "mechanism" through which education affects health is via choice of a healthy environment, acquisition of preventive and curative services, and adoption of healthy behaviors—all of these are more likely in more educated and higher income individuals (Ross and Wu, 1995). The study of Chetty et al. (2016) does not examine separately cause-specific components of life expectancy changes, and we thus could not address this hypothesis.

The outcomes of educational exposures are associated with long-term health outcomes of magnitudes similar to those of major, established health risk factors, such as cigarette smoking, suggesting that educational interventions promoting educational outcomes may promote long-term health at the same time. As noted earlier, diverse educational interventions have been shown to promote educational achievement and attainment and are recommended by the Task Force on Community Preventive Services. It may be productive to consider education as a public health intervention (Freudenberg and Ruglis, 2007; Allensworth et al., 2011). Thus, for several reasons, public health practitioners might fruitfully collaborate with education policy makers and practitioners to promote long-term health along with educational benefits.

# References

Allensworth, D., Lewallen, T.C., Stevenson, B., Katz, S., 2011. Peer reviewed: addressing the needs of the whole child: what public health can do to answer the education sector's call for a stronger partnership. Prev. Chronic Dis. 8 (2), 1–6.

Braveman, P.A., Cubbin, C., Egerter, S., Williams, D.R., Pamuk, E., 2010. Socioeconomic disparities in health in the United States: what the patterns tell us. Am. J. Public Health 100 (S1), S186–S196.

Campbell, F., Conti, G., Heckman, J.J., et al., 2014. Early childhood investments substantially boost adult health. Science 343 (6178), 1478–1485.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., Yagan, D., 2011. How does your kindergarten classroom affect your earnings? Evidence from Project STAR. Q. J. Econ. 126 (4), 1593–1660.

Chetty, R., Stepner, M., Abraham, S., et al., 2016. The association between income and life expectancy in the United States, 2001-2014. JAMA 315 (16), 1750–1766.

Feinstein L, Sabates, R, Anderson, TM, Sorhaindo, A, and Hammond, C. What are the effects of education on health? 2006; Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium, OECD.

Finn, J.D., Gerber, S.B., Achilles, C.M., Boyd-Zaharias, J., 2001. The enduring effects of small classes. Teach. Coll. Rec. 103 (2), 145–183.

Freudenberg, N., Ruglis, J., 2007. Peer reviewed: reframing school dropout as a public health issue. Prev. Chronic Dis. 4 (4), 1–11.

Galea, S., Tracy, M., Hoggatt, K.J., DiMaggio, C., Karpati, A., 2011. Estimated deaths attributable to social factors in the United States. Am. J. Public Health 101 (8), 1456–1465.

Hahn, R.A., Truman, B.I., 2015. Education improves public health and promotes health equity. Int. J. Health Serv. 45 (4), 657–678.

Hahn, R.A., Knopf, J.A., Wilson, S.J., et al., 2015. Programs to increase high school completion: a Community Guide systematic health equity review. Am. J. Prev. Med. 48 (5), 599–608.

Hahn, R.A., Barnett, W.S., Knopf, J.A., et al., 2016. Early childhood education to promote health equity: a community guide systematic review. J. Pub. Health Manag. Pract.: JPHMP 22 (5), E1–E8.

Jemal, A., Thun, M.J., Ward, E.E., Henley, S.J., Cokkinides, V.E., Murray, T.E., 2008. Mortality from leading causes by education and race in the United States, 2001. Am. J. Prev. Med. 34 (1), 1–8 e7.

Knopf, J.A., Hahn, R.A., Proia, K.K., et al., 2015. Out-of-school-time academic programs to improve school achievement: a community guide health equity systematic review. J. Pub. Health Manag. Pract. 21 (6), 594–608.

Krueger, P.M., Tran, M.K., Hummer, R.A., Chang, V.W., 2015. Mortality attributable to low levels of education in the United States. PLoS One 10 (7), 1–13 e0131809.

Kubota, Y., Heiss, G., MacLehose, R.F., Roetker, N.S., Folsom, A.R., 2017. Association of educational attainment with lifetime risk of cardiovascular disease: the atherosclerosis risk in communities study. JAMA Intern. Med. 177 (8), 1165–1172.

Lichtenberg, F.R., 2005. The impact of new drug launches on longevity: evidence from longitudinal, disease-level data from 52 countries, 1982–2001. Int. J. Health Care Finance Econ. 5 (1), 47–73.

Mirowsky, J., Ross, C.E., 2017. Education, Social Status, and Health. Routledge, New York.

Phelan, J.C., Link, B.G., Diez-Roux, A., Kawachi, I., Levin, B., 2004. "Fundamental causes" of social inequalities in mortality: a test of the theory. J. Health Soc. Behav. 45 (3), 265–285.

Ramon, I., Chattopadhyay, S.K., Barnett, W.S., Hahn, R.A., Community Preventive Services Task Force, 2018. Early childhood education to promote health equity: a community guide economic review. J. Pub. Health Manag. Pract. 24 (1), e8–e15.

Reynolds, J., Ross, C.E., 1998. Social stratification and health: education's benefit beyond economic status and social origins. Soc. Probl. 25, 221–247.

Rockhill, B., Newman, B., Weinberg, C., 1998. Use and misuse of population attributable fractions. Am. J. Public Health 88 (1), 15–19.

Ross, C., Wu, C.L., 1995. The links between education and health. Am. Sociol. Rev. 60 (5), 719–745.

Rostron, B., Boies, J.L., Arias, E., 2010. Education reporting and classification on death certificates in the United States. Natl. Center Health Stat. Vital Health Stat 2 (151), 1–14.

Shaw, M., Mitchell, R., Dorling, D., 2000. Time for a smoke? One cigarette reduces your life by 11 minutes. BMJ: Brit. Med. J. 320 (7226), 53.

Subramanian, S.V., Kim, D.J., Kawachi, I., 2002. Social trust and self-rated health in US communities: a multilevel analysis. J. Urban Health 79 (1), S21–S34.

Zajacova, A., Lawrence, E.M., 2018. The relationship between education and health: reducing disparities through a contextual approach. Annu. Rev. Public Health 1 (39), 273–289.