American Journal of Epidemiology Advance Access published May 3, 2011

American Journal of Epidemiology

DOI: 10.1093/aje/kwr011

Original Contribution

The Effect of Small Class Sizes on Mortality Through Age 29 Years: Evidence From a Multicenter Randomized Controlled Trial

Peter Muennig*, Gretchen Johnson, and Elizabeth Ty Wilde

* Correspondence to Dr. Peter Muennig, Mailman School of Public Health, Columbia University, 600 West 168th Street, Sixth Floor, New York, NY 10032 (e-mail: pm124@columbia.edu).

Initially submitted July 3, 2010; accepted for publication January 14, 2011.

Limiting the number of students per classroom in the early years has been shown to improve educational outcomes. Improved education is, in turn, hypothesized to improve health. The authors examined whether smaller class sizes affect mortality through age 29 years and whether cognitive factors play a role. They used data from the Project Student Teacher Achievement Ratio, a 4-year multicenter randomized controlled trial of reduced class sizes in Tennessee involving 11,601 students between 1985 and 1989. Children randomized to small classes (13–17 students) experienced improved measures of cognition and academic performance relative to those assigned to regular classes (22–25 students). As expected, these cognitive measures were significantly inversely associated with mortality rates ($P < 0.05$). However, through age 29 years, students randomized to small class size nevertheless experienced higher mortality rates than those randomized to regular size classes (hazard ratio (HR) = 1.58, 95% confidence interval (CI): 1.07, 2.32). The groups at risk included males (HR = 1.73, 95% CI: 1.05, 2.85), whites/Asians (HR = 1.68, 95% CI: 1.04, 2.72), and higher income students (HR = 2.20, 95% CI: 1.06, 4.57). The authors speculate that small classes might produce behavior changes that increase mortality through young adulthood that are stronger than the protective effects of enhanced cognition.

behavior; cognition; educational status; minority health; mortality; poverty; social conditions

Abbreviations: CI, confidence interval; HR, hazard ratio; STAR, Student Teacher Achievement Ratio.

It is widely believed that effective enhancements to our schools would improve not only the economic vitality of the American population but also its health and well-being (1–5). Many schools are overcrowded (6), and many educators believe that this prevents students from reaching their full cognitive potential (7, 8). One route to simultaneously improving educational and public health outcomes may therefore be to reduce the number of students per class in the early grades of elementary school.

A large number of nonexperimental studies have demonstrated that improved cognition and educational attainment are associated with large health benefits in adulthood (5, 9–14). However, the short-term health effects of different schooling policies are largely unknown, and long-term effects have never been evaluated using a randomized trial. The short-term health effects are important to understand because enhancements to schooling produce powerful changes in behaviors, and such behaviors could affect health in unpredictable ways (3, 7). For instance, generally desirable traits, such as extroversion, could increase a youth's exposure to risks inherent in social development, such as experimenting with drinking, substance use, and driving.

We used the Project Student Teacher Achievement Ratio (STAR) randomized controlled trial to determine the causal effect of reduced class size on mortality through age 29 years. Project STAR randomized 11,601 children to receive instruction in small (13–17 students), regular (22–25 students), or regular size classes with a certified teacher's aide (also 22–25 students) within 79 schools in diverse settings. Children assigned to small classes earned higher test scores, achieved higher rates of high school graduation, and were more likely to take college entrance examinations relative to children assigned to regular classes (15–17). However, there was no change in cognitive performance of children

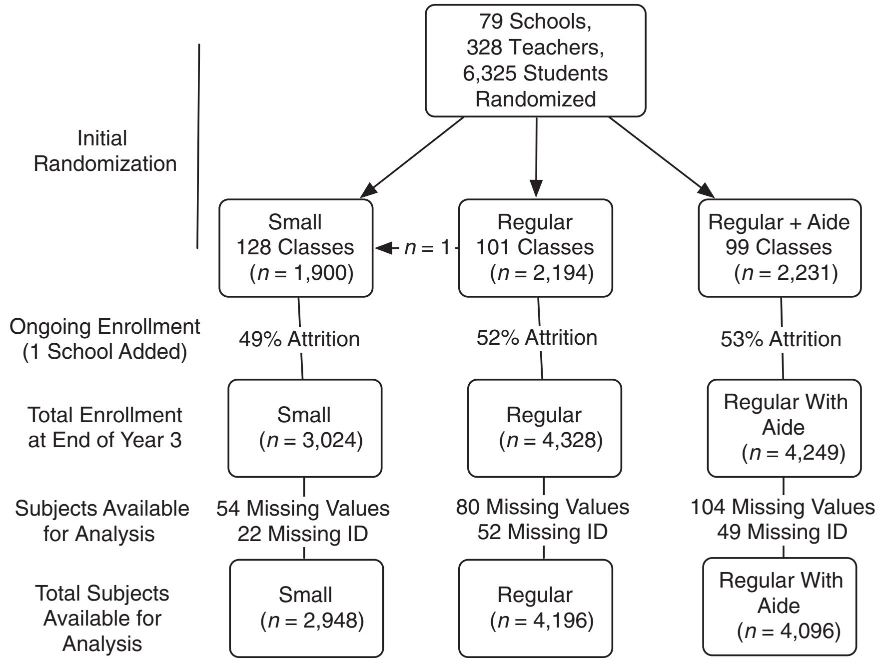

Figure 1. The initial randomization, ongoing enrollment, attrition, and successful electronic data records linkage of subjects in the Project STAR experiment from 1985 to 1989, with follow-up through 2007 for mortality data. ("Missing values" indicates that the Project STAR data set was missing critical information, and "missing ID" indicates that the record was not matched to death certificate data.) STAR, Student Teacher Achievement Ratio.

randomized to regular size classes with an aide compared with regular size classes without an aide (18). Assignment to small class size also appeared to produce changes in many different behavioral traits, such as extroversion and task engagement (15, 19, 20). We linked Project STAR participants to National Death Index records to determine each subject's vital status and cause of death between 1985 and the end of 2007.

MATERIALS AND METHODS

Randomization

The Project STAR schools, located in inner city, urban, suburban, and rural areas, were selected on the basis of their size and their willingness to participate in the study. In the fall of the 1985–1986 school year, both teachers and kindergarten students were randomized to small, regular, or regular with aide classes using random number tables (Figure 1). Randomization occurred within schools. All students were required to remain in their assigned class, and the Project STAR research staff regularly conducted audits to ensure adherence to assignment.

In year 2, as the cohort passed from kindergarten to first grade, some students left Project STAR schools, and additional students who joined Project STAR schools were randomized within schools to each of the 3 experimental groups. This process was repeated in year 3 (second grade) and year 4 (third grade). In addition, all teachers were rerandomized at the beginning of each subsequent school year. All students returned to regular size classes after third grade, when the experiment ended. Initial randomization in kindergarten was successful, with just 1 student switching from a regular class to a small class following assignment (21).

Over the final 3 years of the study, 22.5% of the students changed assignment. Moreover, to alleviate parental concerns regarding the absence of an aide for children assigned to larger classes, researchers rerandomized students between the 2 regular size class types after kindergarten (21). The number and composition of students entering different schools varied considerably over time. Therefore, there were differences in some of the mean sociodemographic characteristics of the students across study arms in the later entry waves (Table 1). Although this variation in treatment allows us to estimate dose-response effects, it also requires that we use an intent-to-treat analysis and precludes a simple comparison between groups.

In the intent-to-treat analysis, some partially treated children are counted among the untreated students, and some partially untreated children are counted among the treated students. The resulting effect sizes in this approach are smaller than they would be had all children remained in their original assignments.

Data linkage

We linked records from Project STAR to death certificates using unique identifiers collected from STAR participants prior to randomization. We determined vital status using a probabilistic matching procedure developed by the National Death Index at the Centers for Disease Control and Prevention. Only 1% of the Project STAR participant records ($n =$

Table 1. Demographic Characteristics, Project STAR, Tennessee, 1985–1989

Group Assignment	Initial No. Randomized	Additional No. Randomized	No. Dead	Actual Class Size	Mean Age at Project STAR Inception, years	% Female	% White/Asian	% Receiving Free Lunch[a]
Students who entered STAR in kindergarten								
Small classes	1,900		23	15.1**	5.4	48.6	68.3	47.1
Regular classes	2,194		16	22.4	5.4	49	67.5	47.7
Regular classes with aide	2,231		21	22.8	5.4	48.3	65.9	50.3
Students who entered STAR in first grade								
Small classes		385	4	15.9**	5.8*	48.7	62.2	59.2
Regular classes		1,206	13	22.7	5.9	43.8	56.2**	62.4
Regular classes with aide		903	16	23.5	5.9	45.6	65.1	60.7
Students who entered STAR in second grade								
Small classes		366	4	15.5**	5.9	42.6	57.4	65.6
Regular classes		654	4	23.7	5.9	45.4	56.7	63.3
Regular classes with aide		659	12	23.6	5.9	46.1	46.5**	65.9
Students who entered STAR in third grade								
Small classes		373	11	15.9**	5.9	43.4	66.9**	59.8*
Regular classes		454	12	22.1	5.9	47.1	57.7	64.5
Regular classes with aide		456	10	22.4	6	45.8	55.8	68.6

Abbreviation: STAR, Student Teacher Achievement Ratio.

* $P < 0.05$; ** $P < 0.01$.

[a] Students who qualify for free lunches are in families earning less than 130% of the federal poverty level.

123) had inadequate identifying information for the mortality analysis, leaving 99% ($n = 11{,}478$) available for matching. (The final analytical sample in Figure 1 is 11,240 because some data on race, gender, and free lunch status were missing.) In total, we identified 146 Project STAR participants as deceased, 25 fewer deaths than would be expected from overall Tennessee mortality data.

Statistical analyses

We conducted our central analyses using Cox proportional hazards models. We used STATA, version 11.0, software (StataCorp LP, College Station, Texas) for all analyses. As noted above, the study design does not permit simple comparisons among the arms of the study. We therefore applied previously validated methods for estimating effects within the Project STAR cohort (16, 21). Specifically, the adjusted analyses assumed the basic (reduced) form:

$$Y_{ics} = \beta 0 + \beta 1 \mathrm{SMALL}_{cs} + \beta 2 \mathrm{REGAIDE}_{cs} + \alpha_s + v_{ics}, \quad (1)$$

where SMALL is a binary variable indicating initial enrollment in small classes, REGAIDE is a binary variable indicating initial enrollment in a regular class with an aide, Y_{ics} represents an outcome for individual i in class c in school s, and α represents a school-by-entry wave fixed effect, so that identification is based on within-school comparisons between students entering Project STAR at the same time. Finally, the error term v contains class-level and individual-level components. We clustered on classroom identifiers to account for these common class-level components.

We included the REGAIDE term a priori for the following reasons: 1) Regular size classes without an aide produced no changes in cognitive performance relative to regular size classes with an aide (15, 18, 19); and 2) we otherwise lack sufficient statistical power for subanalyses. However, we also conducted a secondary analysis that separately compared the small size classroom arm and the regular size classroom arm plus aide arm with the regular size classroom without an aide arm of the study.

In our Cox models, we defined event time as the interval from the start of Project STAR to the time of death, with all surviving subjects censored at year 2008. We conducted the analyses with and without student characteristics (race, gender, and free lunch status) and performed diagnostics to ensure that the data conformed to the assumptions of the proportional hazards model. Separate analyses were conducted by race (with whites and Asians grouped together), age, gender, free lunch status, and school location. As a test of the robustness of our conclusions, we also ran a number of alternative adjusted analyses.

Mediation

We tested the mediating effects of cognitive and noncognitive variables on mortality using instrumental variables analyses and ordinary least squares (22). In our primary analysis, we tested whether assignment to small classes improved cognition using the following measures: 1) math, verbal, and language test scores from kindergarten to grade 8; 2) a measure of multiple domains of cognitive function developed by Finn and Achilles (23); 3) high school grade

Table 2. Number of Subjects, Number of Deaths, Hazard Ratio, and 95% Confidence Interval for Students Randomized to Small Classes (13–17 Students) Relative to Regular Size Classes (Both With and Without Aide), Project STAR, Tennessee, 1985–1989, With Follow-up Through 2007 for Mortality Data

Characteristic	No. of Observations	No. of Deaths	Hazard Ratio	Confidence Interval
All subjects	11,240	141	1.58*	1.07, 2.32
Gender				
Male	5,941	106	1.73*	1.05, 2.85
Female	5,299	35	0.99	0.43, 2.30
Race				
White/Asian[a]	7,082	89	1.68*	1.04, 2.72
Male	3,751	69	1.64	0.90, 2.99
Female	3,331	20	1.86	0.62, 5.58
Black[a]	4,106	50	1.35	0.64, 2.83
Male	2,159	35	1.93	0.66, 5.63
Female	1,947	15	0.46	0.09, 2.35
Free-lunch status				
Free lunch	6,799	91	1.56	0.95, 2.57
Never free lunch	4,441	50	2.20*	1.06, 4.57
Urbanicity				
Urban[b]	956	8	2.57	0.58, 11.4
Inner city[b]	2,660	34	1.40	0.60, 3.25
Suburban	2,934	32	2.02	0.75, 5.48
Rural	4,690	67	1.46	0.88, 2.42
Exposure, year				
1	3,973	31	1.69	0.78, 3.69
2	2,440	29	1.12	0.29, 4.27
3	1,744	37	1.31	0.47, 3.67
4	3,083	44	1.84	0.95, 3.58
Each added student[c]	11,240	141	0.89	0.83, 0.96

Abbreviation: STAR, Student Teacher Achievement Ratio.
* $P < 0.05$ (significant).
[a] The white/Asian and black subgroups do not sum to 11,240 because they do not include those of other races.
[b] "Urban" and "inner city" are distinct categories.
[c] Basis of the reduced form hazard ratio for the relation between average class sizes and mortality measured by the end of 2007. Refer to the Web Appendix, which is posted on the *Journal*'s Web site (http://aje.oxfordjournals.org/), for the 2-stage least-squares estimate of the impact of each additional child on mortality.

point average; and 4) whether the student graduated from high school. We then examined the extent to which these measures, in turn, predicted all-cause mortality. Finally, we tested the impact of each of these measures on the magnitude of the coefficient in each regression equation. We did test noncognitive measures, but they were administered only to a 20% subset of students and did not have adequate statistical power to test their effects. A complete list of the available measures can be found in the Project STAR User's Guide (24).

Power

Analyses using the entire eligible sample had an 80% power to detect a 0.1 standard deviation in effect size. The subanalyses were adequately powered to detect only large effect sizes. As an approximate rule of thumb, a subsample with 6,000 observations would require a hazard ratio of >1.7 (a 70% increase in mortality) to achieve statistical significance at the $P \leq 0.05$ level.

RESULTS

Among the 6,325 children randomized in kindergarten, there were no significant differences with respect to age, race, and free lunch status among the 3 experimental groups (Table 1). Between 1985 and 2007, there were 42 deaths among the 3,024 Project STAR participants who attended small classes, 45 deaths among the 4,508 participants who attended regular classes, and 59 deaths among the 4,249 participants who attended regular classes with an aide. Sixty percent of the overall deaths were due to shootings and automobile accidents (refer to the Web Appendix, which is posted on the *Journal*'s Web site (http://aje.oxfordjournals.org/)). Although randomization was successful within schools

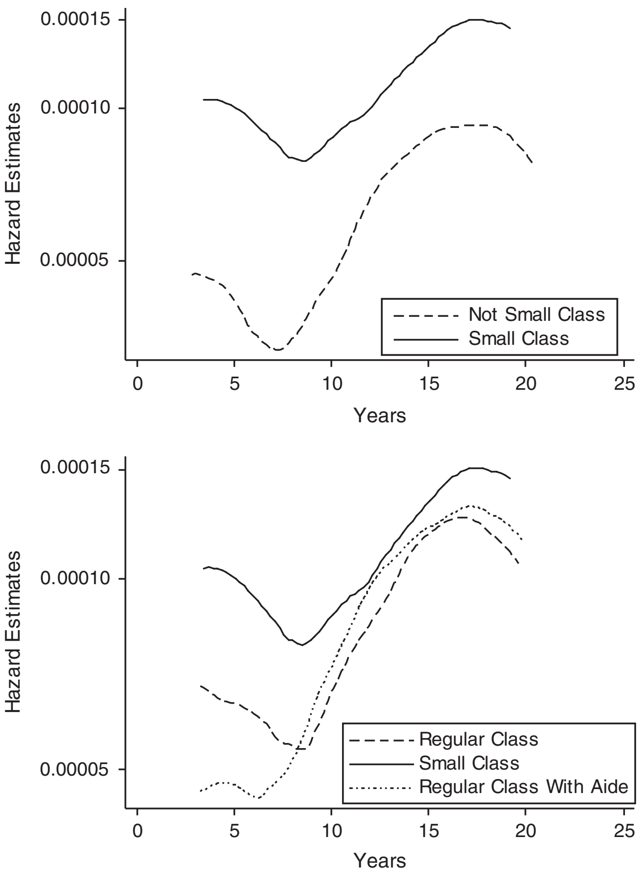

Figure 2. Hazard estimates adjusted for covariates and boundary bias by using the Cox baseline method (regular size classes aggregated (top), separated (bottom), and adjusted by a factor of 10^{10} (0.0000000001) for readability), in the Project STAR experiment from 1985 to 1989, with follow-up through 2007 for mortality data. STAR, Student Teacher Achievement Ratio.

for each additional entry wave, statistically significant sociodemographic differences emerged. These arose because students were not randomized across the entire cohort and because there was significant student mobility between treatment arms over time.

Children assigned to small classes had significantly higher mortality (hazard ratio (HR) = 1.58, 95% confidence interval (CI): 1.07, 2.32) than those assigned to regular classes over 22 years of follow-up (Table 2). Our reduced-form analysis of the relation between mortality and average class size (Equation 1) showed a significant decrease in mortality for every additional student in class (HR = 0.92, 95% CI: 0.87, 0.98). Additional details on dose-response analyses can be found in the Web Appendix.

In the subgroup analysis, males (HR = 1.73, 95% CI: 1.05, 2.85) and whites/Asians (HR = 1.68, 95% CI: 1.04, 2.72) had a significantly elevated risk of premature mortality. Overall, females did not experience a significantly increased risk associated with assignment to small class size. However, this is attributable to a reduced (but not statistically significant) hazard of mortality among black females assigned to small classes. Finally, students assigned to small classes who had never received free lunch also had significantly higher hazards for premature mortality relative to those assigned to regular size classes (HR = 2.20, 95% CI: 1.06, 4.57). Only households with incomes below 130% of the federal poverty level are eligible for free lunch, so students who were never eligible tend to come from economically better-off households.

Children assigned to the small class group in inner-city, urban, suburban, and rural settings all experienced comparable but nonsignificant increases in mortality. Likewise, a nonsignificant increase in mortality for students in small classes was observed across all major groupings of deaths, but it was highest among deaths due to accidents/other causes (a category including motor vehicle accidents, poisoning, drug abuse, drowning, and deaths from fires).

Figure 2 shows survival curves constructed by using the baseline estimates from Cox proportional hazards models correcting for boundary bias (models which do not impose proportionality). The *y*-axis shows point mortality hazards, and the *x*-axis shows follow-up time. Students assigned to regular classes with a certified teaching assistant also experienced higher mortality than students assigned to regular classes, and these differences were statistically significant when compared with regular classes without a teacher's aide (Figure 2, bottom).

In addition, we compared the regular class with no aide separately with the small class and with the regular class with aide in 2 subanalyses. We found that the effect of class assignment to regular with aide (HR = 1.6, 95% CI: 1.1, 2.3) is roughly of the same magnitude as assignment to small class size (HR = 1.6, 95% CI: 1.1, 2.3) when compared with students in the regular class without aide (refer to Web Appendix).

Mediation

Assignment to small class size predicted improved cognitive outcomes. Later cognitive measures (after grade 5) were consistently associated with reduced mortality. Including cognitive variables of interest in the main regression specification caused the coefficient on small classes to increase in magnitude, but only by 1%–5%. Thus, the association between enhanced cognition and reduced mortality was modest and significant only after grade 5. Of the noncognitive mediators tested in a secondary analysis, some were significantly associated with assignment to small class size and were also positively correlated with mortality, but none of these correlations was statistically significant. Refer to the Web Appendix for values and additional information.

Model sensitivity

Our results did not change when using probit or ordinary least-squares models. Our full sensitivity analyses (Web Appendix) did not substantively change the conclusions.

DISCUSSION

As noted in previous studies, children randomized to small classes (13–17 students) had improved cognitive and academic outcomes relative to those assigned to larger classes (22–25 students). After grade 5, all measures of cognitive performance were also significantly predictive of lower mortality.

However, assignment to a small class size or a regular size class plus an aide appears to have produced effects that overwhelmed these survival benefits. As a result, overall mortality rates were higher for children randomized to a small class size than for those randomized to a regular class size.

Mortality hazards among students assigned to small classes were high and trended toward statistical significance in all settings (urban, suburban, inner-city, and rural), among all demographic groups except black females, and in different analytical approaches. There is also evidence of a dose-response effect with respect to both total exposure to small class size and the number of students in the classroom. Together, these findings suggest that our results were not chance occurrences.

Likewise, there is no evidence of randomization failure (21, 25). We know that 0.03% of the children switched class type at the outset of the experiment, one of which switched from a regular class to a small class (21), but these small deviations from experimental protocol cannot explain our results.

We speculate that this higher mortality level occurred through noncognitive pathways for the following reasons: 1) Cognitive test performance is consistently inversely associated with mortality after grade 5; 2) some noncognitive traits (flawed as these measures were) were positively correlated with mortality; 3) although the presence of a certified teacher's aide had no impact on educational outcomes (15, 19), the said presence was associated with an increase in mortality that is similar in magnitude to that of small class size assignment; and 4) the groups for which Project STAR was particularly effective at enhancing academic outcomes showed no significant linkage between treatment and mortality.

According to resource substitution theory, whites, males, and the nonpoor student are the students who would be expected to receive fewer cognitive benefits from reduced class size (26). These groups, in fact, did receive fewer cognitive benefits in Project STAR, and they showed less improvement in high school graduation rates or college test taking than less privileged children (15, 27). Therefore, it is conceivable that the cognitive effects of Project STAR were at least partially protective. If so, they may have buffered the impact of assignment to small class sizes on the groups that realized the greatest cognitive benefit (e.g., low-income and minority groups).

It is tempting to speculate that the additional attention children received in their smaller classes—and possibly in the classes with a teacher's aide—helped them to become more outgoing and affirmed their intellectual curiosity (19). It is undoubtedly a good thing that children are affirmed for socializing and for questioning and exploring their environment. However, this will occasionally have negative outcomes. Poisonings, drugs, drinking and driving, and firearms account for the vast majority of deaths in the Project STAR sample. They were also higher in the small and regular plus aide groups. The additional attention received in these settings could plausibly account for a greater degree of exploration (e.g., poisonings in childhood) and extroversion (e.g., social drug use in adolescence). However, this hypothesis remains highly speculative.

Existing prospective data on the correlation between educational attainment and mortality show that the benefits of educational attainment on mortality are modest through the mid-20's and then begin to accrue rapidly thereafter (28). If so, future analyses should show a net benefit associated with assignment to small class size as the subjects age.

Limitations

The major limitation of our work is that we capture only those deaths occurring by young adulthood. Most of the mortality experiences of Project STAR participants lie in the future. In nonexperimental studies, 80% of the excess mortality experienced by high school dropouts has been attributed to heart disease, cancer, infectious disease, lung disease, and injury (29). Most of these diseases manifest during or after midlife, a time period that we cannot yet capture.

A second potential limitation of our analysis using the Cox model is that it relies on the assumption of proportionality in the hazards of death for the treatment and control groups. Because the treatment was assigned within school and entry year, unadjusted hazard or survival curves, which do not account for these covariates, will be biased. Therefore, we constructed adjusted hazard curves using the Cox baseline method in which we compare those in small classes with those in regular and regular with aide classes (combined and separately). However, although they may be statistically indistinguishable from each other along the entire curve, the adjusted regular and regular with aide curves appear to cross around year 8 (age 12–13 years). This may limit the inferences one can draw from students assigned to the regular plus aide group.

In addition, the study was not blinded. This would not affect our conclusions; even if the Hawthorne effect were in play, the children responded to treatment by outperforming their peers on all available academic measures. Another limitation is that, although our study was a multicenter trial, the findings may not be generalizable outside of Tennessee. Finally, in the mediation analysis, noncognitive measures were available only for a nonrandom sample consisting of fewer than 20% of the total sample. This greatly limited our ability to draw conclusions surrounding noncognitive mediators.

Conclusions

This is the first study to examine the impact of improvements in educational quality on mortality using a multicenter randomized control trial. It is also the first study to examine the impact of improvements in educational quality on the health of young children and adolescents. Several studies using variations in compulsory schooling laws as natural experiments have shown that educational attainment increases adult health (9, 14, 30), although one study found no effect (31).

Our study suggests that small classes lead to higher mortality among younger people, but that this increased mortality probably occurs through noncognitive pathways. Most successful educational interventions produce a wide array of cognitive and noncognitive effects (3). In an age of intensive pressure for success in elite schools and broad experimentation in the education sector, it is important to understand how our schooling policies affect health throughout the life course.

Nevertheless, although the effect sizes that we observe are large, mortality in childhood is rare, so the excess mortality is quite small. If anything, our findings highlight the need for reducing childhood accidental deaths due to inadequate firearm and automobile safety—the 2 leading causes of death among Project STAR participants. Moreover, any excess mortality associated with schooling interventions must be weighed against the lifelong benefits with respect to job security, health insurance, higher income—and probably better health later in life—enjoyed by children attending smaller classes.

ACKNOWLEDGMENTS

Authors affiliation: Mailman School of Public Health, Columbia University, New York, New York (Peter Muennig, Gretchen Johnson, Elizabeth Ty Wilde).

The authors would like to thank the William T. Grant Foundation for their generous financial support.

The William T. Grant Foundation played no role in the design or analysis of this study.

Conflict of interest: none declared.

REFERENCES

1. Muennig P, Woolf SH. Health and economic benefits of reducing the number of students per classroom in US primary schools. *Am J Public Health*. 2007;97(11):2020–2027.
2. Woolf SH, Johnson RE, Phillips RL Jr, et al. Giving everyone the health of the educated: an examination of whether social change would save more lives than medical advances. *Am J Public Health*. 2007;97(4):679–683.
3. Levin H, Belfield C. *The Price We Pay: Economic and Social Consequences of Inadequate Education*. Washington, DC: Brookings Institution Press; 2007.
4. Kolata G. A surprising secret to long life: stay in school. *New York Times.* January 3, 2007(http://www.nytimes.com/2007/01/03/health/03aging.html?pagewanted=all). (Accessed March 20, 2009).
5. Grossman M. Education and nonmarket outcomes. In: Hanushek E, Welch F, eds. *Handbook of the Economics of Education*. Vol 2. Amsterdam, the Netherlands: Elsevier; 2006.
6. Alonso G, Anderso N, Su C, et al. *Our Schools Suck: Students Talk Back to a Segregated Nation on the Failures of Urban Education*. New York, NY: New York University Press; 2009.
7. Currie J. Early childhood education programs. *J Econ Perspect*. 2001;15(2):213–238.
8. Levin HM, Belfield C, Muennig P, et al. The public returns to public educational investments in African-American males. *Econ Educ Rev.* 2007;26:700–709.
9. Lleras-Muney A. The relationship between education and adult mortality in the United States. *Rev Econ Stud*. 2005;72(1):189–221.
10. Muennig P, Schweinhart L, Montie J, et al. Effects of a prekindergarten educational intervention on adult health: 37-year follow-up results of a randomized controlled trial. *Am J Public Health*. 2009;99(8):1431–1437.
11. Kawachi I, Adler NE, Dow WH. Money, schooling, and health: mechanisms and causal evidence. *Ann N Y Acad Sci.* 2010;1186:56–68.
12. Levin H, Belfield C, Muennig P, et al. *The Costs and Benefits of an Excellent Education for America' s Children*. New York, NY: Teachers College; 2006.
13. Muennig P, Robertson D, Johnson G, et al. The effect of an early education program on adult health: the Carolina Abecedarian Project randomized controlled trial. *Am J Public Health*. 2011;101(3):512–516.
14. Muennig P. How education produces health: a hypothetical framework. *Teach Coll Rec*. 2007;14606:1–17.
15. Finn JD, Gerber SB, Achilles CM. The enduring effects of small classes. *Teach Coll Rec*. 2001;103(2):145–183.
16. Krueger AB, Whitmore DM. The effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from Project STAR. *Econ J*. 2001;111(468):1–28.
17. Finn JD, Gerber SB, Boyd-Zaharias J. Small classes in the early grades, academic achievement, and graduating from high school. *J Educ Psychol*. 2005;97(2):214–223.
18. Gerber SB, Finn JD, Achilles CM, et al. Teacher aides and students' academic achievement. *Educ Eval Policy Anal*. 2001;23(2):123–143.
19. Finn JD, Achilles CM. Tennessee's class size study: findings, implications, misconceptions. *Educ Eval Policy Anal*. 1999; 21(2):97–109.
20. Finn JD, Pannozzo GM. Classroom organization and student behavior in kindergarten. *J Educ Res.* 2004;98(2):79–92.
21. Krueger AB. Experimental estimates of education production functions. *Quart J Econ*. 1999;114(2):497–531.
22. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6): 1173–1182.
23. Finn J, Achilles C. Answers and questions about class size: a statewide experiment. *Am Educ Res J*. 1990;27(3):557–577.
24. Finn JD, Boyd-Zaharias J, Fish RM, et al. Project STAR and beyond: database user's guide. Lebanon, TN: HEROS, Inc, 2007. (http://www.heros-inc.org/starUsersGuide.pdf). (Accessed March 29, 2009).
25. Hanushek EA. Evidence, politics, and the class size debate. In: Mishel L, Rothstein R, eds. *The Class Size Debate*. Washington, DC: Economic Policy Institute; 2002:37–66.
26. Ross CE, Mirowsky J. Sex differences in the effect of education on depression: resource multiplication or resource substitution? *Soc Sci Med*. 2006;63(5):1400–1413.
27. Krueger AB, Whitmore DM. Would smaller classes help close the black-white achievement gap? Princeton, NJ: Princeton University Industrial Relations Section, 2001. (http://www.irs.princeton.edu/pubs/pdfs/451.pdf). (Accessed August 11, 2009).
28. Muennig P, Fiscella K, Tancredi D, et al. The relative health burden of selected social and behavioral risk factors in the United States: implications for policy. *Am J Public Health*. 2010;100(9):1758–1764.
29. Wong MD, Shapiro MF, Boscardin WJ, et al. Contribution of major diseases to disparities in mortality. *N Engl J Med*. 2002;347(20):1585–1592.
30. Almond D, Mazumder B. *How Did Schooling Laws Improve Long-Term Health and Lower Mortality?*. Chicago, IL: Federal Reserve Bank of Chicago; 2006. (Working Paper 2006-23).
31. Clark D, Royer H. evidence from the United Kingdom. *The effect of education on adult mortality and health*. San Francisco, CA: Federal Reserve Bank of San Francisco; 2008. (http://www.frbsf.org/economics/conferences/0806/royer.pdf). (Accessed June 4, 2010).