Published by Oxford University Press on behalf of the International Epidemiological Association
 Advance Access publication 15 June 2009

International Journal of Epidemiology 2009;**38**:1310–1322
doi:10.1093/ije/dyp226

Is education causally related to better health? A twin fixed-effect study in the USA

Takeo Fujiwara[1]* **and Ichiro Kawachi**[2]

Accepted 13 May 2009

Background More years of schooling is generally associated with better health. However, this association may be confounded by unobserved common prior causes such as inherited ability, personality such as patience, or early family circumstances. The twin fixed-effect approach can potentially address this problem by cancelling these factors between twin pairs. The purpose of this study is to identify the causal effects of education on health and health behaviours using a twin fixed-effect approach.

Methods We used twin data from the National Survey of Midlife Development in the United States, 1995–1996. The study population included 302 male [55.6% monozygotic (MZ) and 44.4% dizygotic (DZ)] and 387 female twin pairs (47.3% MZ and 52.7% DZ). A range of health outcomes [perceived global, physical and mental health, body mass index (BMI), waist circumference, waist–hip ratio, number of depressive symptoms] and health behaviours (smoking and physical activity) were examined among twin pairs who were discordant on years of schooling.

Results Among MZ twins, more years of education was associated with better perceived global health. For all other health outcomes/behaviours, the point estimates of the effect of education in the fixed-effect analyses suggested a weak protective association. Among DZ male twins, each additional year of schooling lowered the prevalence of smoking by 32% [odds ratio (OR): 0.68, 95% confidence interval (CI): 0.48–0.97] in the fixed-effect analysis.

Conclusion The widely reported associations between schooling and health outcomes/behaviours may not reflect causal relationships in every instance. Although low statistical power may explain some of the null associations, our twin fixed-effect analyses suggest that at least some cases of the education/health relationship reflect confounding by unobserved third variables.

Keywords Education, schooling, health status, health behaviours, twin study

Introduction

Educational attainment has been reported in the literature to be robustly associated with higher levels of both physical and mental health, as well as healthier patterns of behaviour.[1–7] However, causal inference remains challenging due to confounding of the association by unobserved personal characteristics, such as inherited ability, patience or early family circumstances. For example, although the inverse association

[1] Section of Behavioral Science, Department of Health Promotion, National Institute of Public Health, Saitama, Japan.
[2] Department of Society, Human Development and Health, Harvard School of Public Health, Boston, MA, USA.
* Corresponding author. Section of Behavioral Science, Department of Health Promotion, National Institute of Public Health, 2-3-6 Minami, Wako-shi, Saitama 351-0197, Japan. E-mail: tfujiwara@niph.go.jp

between educational attainment and cigarette smoking is well established, the relationship is not necessarily causal. Farrell and Fuchs[8] examined the relationship between schooling and smoking within a community sample of adults who had all completed 12–18 years of education, and found that educational inequalities in smoking observed in adulthood (mid-20s) were already evident at age 17 when all of the subjects were still in the same grade. In other words, educational inequalities in smoking were already evident even before schooling was actually completed. This suggests that a third variable such as intelligence, or patience (i.e. the ability to put off enjoyment of things in the present in exchange for future benefit)—as opposed to schooling *per se*—might be responsible for the observed association between schooling and cigarette smoking. Indeed, IQ and education are highly correlated and a growing body of research suggests a link between IQ and mortality/morbidity.[9–11]

Recent studies in the USA have attempted to identify the causal effect of schooling on health through the use of instrumental variables, which seek to purge the correlation between education and unobserved confounds such as cognitive ability and personality. Compulsory schooling laws in the state of one's birth have been used as an instrument for education, based on the argument that compulsory schooling laws serve as 'natural policy experiments', whereby children born in states that mandated longer schooling, in fact, ended up with more education. These studies using instrumental variable estimates have found strong associations between education and reduced risk of mortality[12] as well as better cognitive functioning at older ages.[13]

A different approach to identify the causal effect of education on health is the use of twin- or sibling-fixed-effect designs. Krieger *et al*.[14] showed that female twins discordant on occupational class differed with respect to systolic and diastolic blood pressure and low-density lipoprotein cholesterol. In the same study, twins who were discordant on education did not vary by the health outcomes examined. Osler *et al*.[15] reported an analysis of twins who were discordant on social class (defined by type of employment and vocational education). Among dizygotic (DZ) twins, they found associations of occupational class with height and cognitive score for males and physical activity score for females. However, the same relationships were not observed among monozygotic (MZ) twins, leading the authors to conclude that the relationship between occupational class and health is due mainly to selection effects (i.e. those who have better health enjoy higher social class) rather than a causal effect. One limitation of these prior twin studies is that they dichotomized socioeconomic position (high vs low) so they could not assess 'dose–response' of health outcomes across years of schooling. In order to examine the impact of additional years of schooling, a fixed-effect model is superior to stratification analysis.[16]

The National Survey of Midlife Development in the United States (MIDUS) in 1995–1996 includes twin data, which provide detailed educational attainment, health outcomes and health behaviour information. The current study hypothesized that differences in educational attainment influence the health and health behaviours of twins sharing the same genetic background and family environment. The MIDUS twin data were used to identify the causal effect of educational attainment on health and health behaviours among twin pairs, MZ and DZ, who shared the same family environment in childhood. If a positive association was observed in both MZ and DZ twin pairs, this provides evidence of a causal effect of schooling on health. If a positive association was observed only in DZ twins, the association between health and education may be still confounded by genetic factors. If both the DZ and MZ twin pairs showed null associations, this suggests that the relation between schooling and health is biased by unobserved confounders such as inherited ability, early family environment or health selection effects.

Methods

Study population

Details of the sampling procedures have been explained elsewhere.[17] Twin pairs were recruited by using a separate two-part sampling design. The first part involved screening a representative national sample of ~50 000 households for the presence of a twin. The screening was conducted by International Communications Research, the market research group of AUS Consultants and Bruskin Associates. Respondents who indicated the presence of twins in the household or being part of a twin pair themselves were asked permission to be contacted by a research team for inclusion in the first national study of twins. The presence of a twin in the family was reported by 14.8% of the respondents, of whom 60.0% gave permission to be contacted for the twin study. The second part of the twin sample design was carried out by interview staff at the Institute for Social Research (the University of Michigan) who contacted the twin households in order to recruit twins to participate in the survey. The cooperating twins were asked to provide contact information for their co-twins. Almost half (49%) of the first contacts identified twin pairs who did not meet the eligibility criteria of the study (aged between 25 and 74 years, non-institutionalized, living in the continental USA and English speaking). The final response percentage for complete twin interviews was 26%. The final response percentage for the twin pairs varied according to whether the first contact was with a relative of the twin (20% response percentage) or the twin himself or herself (60% response percentage). The final twin sample included a total of 1996 twins, resulting in 998 pairs.

Since the impact of education on health potentially varies by gender,[18] different-sex DZ twins ($n = 526$; 263 pairs) were excluded. We also excluded twins in whom the zygotic status was unknown ($n = 32$; 16 pairs). Finally, we excluded twin pairs if one of the pair reported that they had lived separately before the age of 14 years ($n = 58$; 29 pairs among MZ and same-sex DZ twins). The final study sample included 1378 individuals or 689 twin pairs (final percentage: 18% of eligible twins)—351 (male: 168, female: 183) pairs of MZ (50.9%) and 338 (male: 134, female: 204) pairs of DZ twins (49.1%).

Measurements

Twin pairs were interviewed by a telephone questionnaire, which inquired about educational attainment, three health outcomes (perceived physical and mental health and number of depressive symptoms) and several types of smoking behaviour (further details below), followed by a self-administered mailed questionnaire, which inquired about four additional health outcomes [perceived global health, body mass index (BMI), waist circumference (WC) and waist–hip ratio (WHR)], as well as physical activities. The telephone interview was completed by 924 twin pairs (92.6%) and the self-administered mail questionnaire was completed by 807 twin pairs (80.9%).

In the telephone survey, educational attainment was assessed by asking the highest grade of school or years of college completed. The responses were categorized into 12 groups: no school/some grade school, eighth grade/junior high school, some high school, general educational development (GED), graduated from high school, 1–2 years college (no degree yet), 3 years of college (no degree yet), graduated 2 years college (vocational or associate degree), graduated 4 or 5 years college (bachelor degree), some graduate school, masters degree, doctor degree (PhD, EdD, MD, DDS, LLB, LLD, JD) or other professional degree. We then assigned the average number of years of schooling to each category as follows: no school/some grade school = 4, eighth grade/junior high school = 8, some high school = 10, GED = 11, graduated from high school = 12, 1–2 years college (no degree yet) = 11.5, $\geqslant$3 years college (no degree yet) = 15.5, graduated 2 years college (vocational or associate degree) = 14, graduated 4 or 5 years college (bachelor degree) = 16.5, some graduate school = 17, masters degree = 18 and doctor degree (PhD, EdD, MD, DDS, LLB, LLD, JD) or other professional degree = 20. Perceived physical health was assessed using a single question: 'In general, would you say your physical health is ....' with Likert-scale responses ranging from poor, fair, good, very good, or excellent.

Perceived mental health was also assessed using a single question: 'What about your mental or emotional health?' with Likert-scale responses ranging from poor, fair, good, very good or excellent. The responses 'poor' and 'fair' were collapsed *a priori* into a single category, since the lowest category (poor) by itself was too small to allow meaningful statistical analysis (0.9%). To assess depressive symptoms, questions in the Composite International Diagnostic Interview Short Form (CIDI-SF) were used. The assessed depressive symptoms were (i) lack of interest in most things, (ii) feeling more tired or having lower energy than usual, (iii) loss of appetite, (iv) sleeping difficulties, (v) difficulty concentrating, (vi) feeling down, bad or worthless and (vii) increased thoughts of death. The total number of depressive symptoms was examined as a mental health outcome.

For smoking, we examined the following behaviours separately: (i) ever smoking, (ii) age at starting to smoke regularly, (iii) current regular smoking, (iv) number of cigarettes smoked per day during the 1 year in life when the individual smoked most heavily, (v) experience of trying to quit smoking and (vi) age at the time when the individual last smoked regularly. The smoking variables were all dichotomous (yes/no), except the number of cigarettes per day (which was asked only for current smokers).

In the mailed questionnaire, perceived global health was assessed using a single question: 'Using a scale from 0 to 10 where 0 means the worst possible health and 10 means the best possible health, how would you rate your health these days?' To calculate BMI, height and current weight were asked. To measure waist and hip circumference, a tape measure was enclosed and respondents were provided with instructions to (i) take the measurements while standing, (ii) avoid measuring over clothing, (iii) record answers to the nearest quarter inch and (iv) follow the figure shown in the questionnaire booklet to guide where to take the measurements. The physical activity was assessed by four questions: 'During the summer, how often do you engage in vigorous physical activity (for example, running or lifting heavy objects) long enough to work up a sweat?'; 'What about during winter—how often do you engage in vigorous physical activity long enough to work up a sweat?'; 'During the summer, how often do you engage in moderate physical activity (for example, bowling or using a vacuum cleaner)?'; and 'What about during the winter—how often do you engage in moderate physical activity?'. The responses were in six categories: several times a week or more, about once a week, several times a month, about once a month, less than once a month and never. These questions have been validated in previous studies,[19–21] taking account of seasonal differences.[22]

Zygosity was determined as identical or fraternal twins using self-report data collected as part of the initial twin-screening questionnaire. In cases where the respondent was unsure, DNA analysis was performed on returned cheek cell samples. Similar methods of diagnosing zygosity have been shown to be >90% accurate in diagnosing twin zygosity.[23]

Analysis

First, all twin pairs were treated as individuals and the association between educational attainment and continuous health outcomes/behaviours were investigated using a generalized estimating equation (GEE) model, which adjusts for the clustering of outcomes within twin pairs. Individual covariates (age, gender, race, working status and marital status) were also adjusted. The number of depressive symptoms was excluded from GEE analysis due to a high degree of right skew. Details of individual covariates are presented in Table 1.

Table 1 Demographic characteristics of the sample ($N = 1378$)

Variables	Male twins ($N = 604$)	Female twins ($N = 774$)
Demographics		
Zygocity		
Monozygotic	336 (55.6)	366 (47.3)
Dizygotic	268 (44.4)	408 (52.7)
Education		
Schooling year	14.1 (0.12)	13.5 (0.09)
Less than high school	51 (8.8)	83 (11.1)
High school	170 (29.4)	259 (34.7)
Some college	176 (30.5)	242 (32.4)
Graduated college or more	181 (31.3)	163 (21.8)
Age (year)	44.6 (11.8)	44.7 (12.2)
Race		
White	497 (95.6)	634 (92.2)
Black	17 (3.3)	38 (5.5)
Other	6 (1.2)	16 (2.3)
Working status		
Full-time working	481 (82.9)	489 (66.2)
Retired	68 (11.7)	47 (6.4)
Homemaker	0 (0.0)	145 (19.6)
Unemployed	31 (5.3)	58 (7.9)
Marital status		
Married	453 (78.1)	534 (71.5)
Separated	13 (2.2)	18 (2.4)
Divorced	53 (9.1)	79 (10.6)
Widowed	5 (0.9)	39 (5.2)
Never married	56 (9.7)	77 (10.3)

Data are given as N (%) or mean standard deviation (SD), unless otherwise mentioned.

Secondly, the relationships between education and health outcomes/behaviours were re-analysed by using the fixed-effect model among twin pairs. See Greene[24] and Hsiao[25] for a detailed description of the statistical approach used in the fixed-effect model, and Fujiwara *et al.*[16] as an example. Briefly, in the fixed-effect model, the effect of education on health outcomes/behaviours was calculated by cancelling the effect of unknown shared factors, such as genetic or early family environmental influences, that might affect health outcomes/behaviours. An equation representing the association between health and education for each pair of twins (let the first subscript, i, represent the twin pair, and let the second subscript represent either twin 1 or 2 in the pair) can be written as follows:

$$y_{i1} = \beta_{11}x_{i1} + g_{i1} + f_i + \epsilon_{i1}$$
$$y_{i2} = \beta_{12}x_{i2} + g_{i2} + f_i + \epsilon_{i2},$$

where

y is health outcomes/behaviours,

x is the educational attainment,

g and f, respectively, represent unmeasured factors such as genetic endowment g and early family environment f and ϵ represents a normal error term.

In the fixed-effect model, the effects of these unmeasured factors (i.e. fixed effect) can be cancelled by subtracting the equations as follows:

$$y_{i1} - y_{i2} = \beta'_1(x_{i1} - x_{i2}) + (g_{i1} - g_{i2}) + (f_i - f_i) + (\epsilon_{i1} - \epsilon_{i2}),$$

which can be rewritten as

$y_i^* = \beta'_1 x_i^* + g_i^* + f_i^* + \epsilon_i^*$, where the asterisk indicates the difference of variables within each twin pair. As early family environment is the same in both MZ and DZ twins, f_i^* is equal to zero. In MZ twins, the genetic endowment is the same; hence, g_i^* is equal to zero. Therefore, if β'_1 is different between MZ and DZ pairs, the difference can be interpreted to be due to the term g_i^*.

In fixed-effect analysis, years of schooling and health outcomes (perceived global, physical, mental health, BMI, WC, WHR) were treated as continuous variables. The number of depressive symptoms was treated as an ordinal variable (due to a high degree of right skew), and ordered logistic regression was used. In terms of health behaviours, for smoking (other than number of cigarettes per day), variables were treated as binary in the fixed-effect models. The number of cigarettes was also analysed among current smokers as a continuous variable, to check the dose–response effect of education on smoking. For physical activity, first, responses were dichotomized into those who responded 'never' versus any physical activity. Secondly, among those who responded doing any physical activity, items were re-coded into average monthly physical activity using the category midpoints: '0.5' times per month, '1', '3', '4.5' and '13.5'. Detailed analyses were performed to examine the effects of schooling on different aspects of smoking behaviour. In sensitivity analysis, education was treated as a categorical variable (i.e. less than high school, high school, some college, college or more). All analyses were performed by Stata SE version 9.0.

Results

Demographic characteristics including educational attainment are shown in Table 1. Among males, the mean years of schooling was 14.1 and the percentage who graduated college or more was 31.3%, whereas the percentage of those who did not graduate high school was 8.8%. In females, the mean years of schooling was 13.5, and 21.8% graduated college or more and 11.1% did not graduate high school. The majority of the sample was White, full-time workers and married for both sexes.

Table 2 shows the distribution of health outcomes/behaviours. Average BMI values were 26.9 for males

Table 2 Distribution of health outcomes and health behaviours among the study sample ($N = 1378$)

Variables	Male twins ($N = 604$)		Female twins ($N = 774$)	
	Mean	SD	Mean	SD
Health outcomes				
Perceived global health (range 0–10)	7.7	1.4	7.7	1.7
Perceived physical health (range 1–5)	3.7	0.9	3.6	1.0
BMI	26.9	3.9	25.7	5.6
WC (inches)	37.5	4.3	32.9	5.6
WHR	0.96	0.10	0.83	0.11
Perceived mental health (range 2–5)	4.0	0.9	3.8	0.9
Number of depressive symptoms (range 0–7)	0.6	1.6	0.8	2.0
Health behaviours	***N***	**%**	***N***	**%**
Ever smoke				
Smoked never	111	19.3	210	28.2
Smoked ever	465	80.7	534	71.8
Regular smoke				
Have not smoked regularly	280	48.3	418	56.0
Smoked regularly	300	51.7	328	44.0
Age began to smoke regularly[a]	18.4	0.23	19.2	0.25
Current smoking status				
Non-smoking	448	77.2	572	76.5
Current regular smoking	132	22.8	176	23.5
Number of cigarettes per month[b]	28.2	1.3	23.3	0.98
Quit smoking				
Having experience of trying to quit smoking[b]	102	77.3	135	76.7
Vigorous physical activity in summer				
Never	35	6.6	113	16.1
Less than once a month	29	5.5	78	11.1
Almost once a month	30	5.7	58	8.3
Several times a month	58	10.9	98	14.0
About once a week	106	20.0	101	14.4
Several times a week or more	272	51.3	252	36.0
Vigorous physical activity in winter				
Never	38	7.2	125	17.9
Less than once a month	47	8.9	102	14.6
Almost once a month	54	10.2	75	10.7
Several times a month	88	16.6	106	15.2
About once a week	118	22.3	109	15.6
Several times a week or more	184	34.8	182	26.0

(continued)

Table 2 Continued

Variables	Male twins (N = 604)		Female twins (N = 774)	
	Mean	SD	Mean	SD
Moderate physical activity in summer				
Never	13	2.5	13	1.9
Less than once a month	10	1.9	4	0.6
Almost once a month	10	1.9	9	1.3
Several times a month	57	10.8	56	8.0
About once a week	91	17.2	173	24.8
Several times a week or more	348	65.8	443	63.5
Moderate physical activity in winter				
Never	15	2.9	20	2.9
Less than once a month	9	1.7	7	1.0
Almost once a month	19	3.6	20	2.9
Several times a month	63	12.0	67	9.5
About once a week	126	23.9	175	24.9
Several times a week or more	295	56.0	413	58.8

[a] $N = 300$ for male and 328 for female.
[b] $N = 132$ for male and 176 for female.

and 25.7 for females, WC was 37.5 inches for males and 32.9 for females, and WHR was 0.96 for males and 0.83 for females. Men and women were fairly comparable on the self-rated measures of global health, physical health and mental health. With regard to smoking habits, the majority of subjects (70–80%) reported having smoked at least some time in the past (ever smoker). The mean age at starting to smoke regularly was 18.4 and 19.2 years, respectively, for males and females. Among current smokers, ~77% reported attempts to quit smoking. More than 50% of male twins reported doing vigorous physical activity in the summer and winter at least once a week. In females, >80% of female twins reported doing moderate physical activity at least once a week in the summer and winter. These percentages were similar to the original MIDUS study, a national representative sample (e.g. 79% of females do moderate physical activity at least once a week in the summer and winter).[26]

The individual-level associations between years of schooling and health outcomes/behaviours adjusted for covariates (race, working status and marital status) as well as the within-twin pair correlation are shown in Table 3. Among males, years of schooling were positively associated with perceived global, physical and mental health, while BMI, WC and WHR were not associated with schooling. Schooling was protective against current smoking: each additional year of schooling decreased the likelihood of being current smoker by 25% [odds ratio (OR): 0.75, 95% confidence interval (CI): 0.68–0.83]. However, schooling was not associated with the intensity of daily smoking among current smokers. Moderate and vigorous physical activity in both summer and winter were positively associated with schooling among males. In females, similar associations were found as with males, but WC and WHR were negatively associated with schooling [i.e. each additional year of schooling reduced WC by 0.29 inches (95% CI: −0.47 to 0.11)].

Table 4 shows the results of the fixed-effect analysis, stratified by gender and zygosity. The point estimates of the associations between schooling and perceived physical and mental health were in a similar direction to the GEE estimates. Similarly, the relationships between schooling and anthropometric values (BMI, WC and WHR) exhibited similar directions of association (higher education → leaner physique). Among male twins, the main finding was an inverse association between schooling and smoking status among DZ twins (OR: 0.68, 95% CI: 0.48–0.97). The point estimate among MZ male twins also showed an inverse association (OR: 0.60, 95% CI: 0.31–1.15). Among those who do moderate physical activity, a dose–response relationship was found between schooling and frequency of moderate physical activity. Among female twins, the associations between education and health outcomes were broadly in a similar direction as the GEE results, except for smoking status where years of schooling and smoking status were not associated among female twins in the fixed-effect model.

To increase the power, sex-combined fixed-effect analysis was performed (results shown in Table 5). It is noteworthy to report that years of schooling positively associated with perceived global health: each additional year of schooling increased the score of perceived global health (range 0–10) by 0.1

Table 3 The association between years of schooling and health outcomes and behaviours among the individual sample adjusted for covariates and within twin pair correlation (continuous variables: GEE model, dichotomous variables: GEE population-average logit model, $N = 1378$)

		Male twins (total $N = 604$)			Female twins (total $N = 774$)	
Outcome variables		β	95% CI		β	95% CI
Health outcomes						
Perceived global health		0.06	0.02 to 0.10		0.07	0.01 to 0.12
Perceived physical health		0.08	0.05 to 0.11		0.08	0.05 to 0.11
BMI		−0.06	−0.18 to 0.06		−0.15	−0.33 to 0.03
WC (inches)		−0.02	−0.15 to 0.10		−0.29	−0.47 to 0.11
WHR		−0.002	−0.005 to 0.001		−0.006	−0.009 to 0.003
Perceived mental health		0.08	0.06 to 0.11		0.06	0.04 to 0.09
Health behaviours						
Dichotomized response (yes/no)		**OR**	**95% CI**		**OR**	**95% CI**
Smoking (yes/no)		0.75	0.68 to 0.83		0.75	0.68 to 0.83
Vigorous physical activity in summer (yes/no)		1.21	1.05 to 1.38		1.12	1.02 to 1.23
Vigorous physical activity in winter (yes/no)		1.21	1.06 to 1.38		1.13	1.02 to 1.24
Moderate physical activity in summer (yes/no)		1.55	1.21 to 2.00		1.30	0.95 to 1.78
Moderate physical activity in winter (yes/no)		1.58	1.24 to 2.01		1.30	1.01 to 1.66
Continuous response among answered 'yes'	***n***	**β**	**95% CI**	***n***	**β**	**95% CI**
Smoking (number of cigarettes/day)	107	−0.69	−1.79 to 0.42	157	−0.82	−1.87 to 0.23
Vigorous physical activity in summer (times/month)	485	−0.003	−0.17 to 0.16	675	0.28	0.09 to 0.47
Vigorous physical activity in winter (times/month)	482	0.06	−0.10 to 0.23	674	0.40	0.22 to 0.59
Moderate physical activity in summer (times/month)	504	0.19	−0.05 to 0.33	673	0.25	0.11 to 0.40
Moderate physical activity in winter (times/month)	500	0.14	−0.01 to 0.30	677	0.29	0.14 to 0.45

Physical activity was values never: 0, less than once a month: 0.5, almost once a month: 1, several times a month: 3, about once a week: 4.5, several times a week or more: 13.5.

(95% CI: 0.0003–0.19). However, in DZ twins, the same association was not found with the point estimate showing an inverse association between years of schooling and perceived global health. Anthropometric outcomes suggested associations between higher years of schooling and lower BMI and WC among MZ pairs. Smoking status was not associated with years of schooling in sex-combined MZ and DZ twins.

Table 6 shows the association between schooling and detailed smoking behaviours using the fixed-effect model. Schooling was not associated with ever smoking. Among DZ male twins, schooling reduced the likelihood of being a current smoker (OR: 0.66, 95% CI: 0.45–0.96). However, the corresponding association was not observed in MZ male twins, although the point estimate was similar. Among female MZ/DZ twins, schooling was not associated with current smoking status. The point estimate of quitting behaviour among regular smokers was positively associated with schooling among male twins (MZ, OR: 5.76, 95% CI: 0.26–126.4; DZ, OR: 1.49, 95% CI: 0.87–2.53). Finally, education was not associated with the initiation of regular smoking among non-smokers who completed their schooling, although the point estimate among female twins indicated an inverse association between more years of schooling and risk of initiating regular smoking. Trying to quit smoking behaviour was not associated with the schooling year among current smoker either.

Again, to increase the power, the corresponding analysis was performed using the sex-combined sample (Table 7). In the combined DZ sample, schooling reduced the initiation of regular smoking among non-smokers who completed their schooling (OR: 0.74, 95% CI: 0.56–0.98), and the point estimate among MZ twins was in a similar direction (OR: 0.83, 95% CI: 0.53–1.31). For both MZ and DZ twins, there was a suggestion of more years of schooling being associated with lower odds of current smoking and higher odds of quitting.

In sensitivity analysis, education was treated as a categorical variable. The results were essentially similar. Among male twins, there was a suggestion of an inverse association between education and smoking status among DZ twins. Compared with those who

Table 4 Fixed effects of schooling on health among twin pairs (total pairs $N = 689$)

Outcome variables	Male				Female			
	Monozygotic (total pairs $N = 168$)		Same-sex dizygotic (total pairs $N = 134$)		Monozygotic (total pairs $N = 183$)		Same-sex dizygotic (total pairs $N = 204$)	
		β (95% CI)		β (95% CI)		β (95% CI)		β (95% CI)
Health outcomes								
Perceived global health		0.11 (−0.01 to 0.22)		0.17 (−0.17 to 0.50)		0.08 (−0.08 to 0.24)		−0.34 (−0.68 to 0.01)
Perceived physical health		0.07 (−0.01 to 0.15)		0.07 (−0.002 to 0.14)		−0.003 (−0.09 to 0.09)		0.01 (−0.06 to 0.08)
BMI		−0.14 (−0.23 to 0.20)		−0.07 (−0.41 to 0.26)		−0.08 (−0.41 to 0.25)		0.15 (−0.28 to 0.59)
WC (inches)		−0.18 (−0.44 to 0.08)		−0.04 (−0.43 to 0.34)		−0.25 (−0.63 to 0.13)		−0.14 (−0.60 to 0.33)
WHR		−0.01 (−0.02 to 0.003)		0.003 (−0.007 to 0.01)		−0.002 (−0.009 to 0.006)		−0.009 (−0.02 to 0.001)
Perceived mental health		0.04 (−0.03 to 0.11)		0.03 (−0.05 to 0.10)		0.002 (−0.09 to 0.09)		0.05 (−0.02 to 0.12)
Number of depressive symptoms		0.014 (−0.19 to 0.22)		−0.003 (−0.16 to 0.15)		−0.06 (−0.26 to 0.14)		0.01 (−0.13 to 0.15)
Health behaviours								
Dichotomized response (y/n)	*N**	**OR (95% CI)**	*N**	**OR (95% CI)**	*N**	**OR (95% CI)**	*N**	**OR (95% CI)**
Smoking (y/n)	23	0.60 (0.31 to 1.15)	34	0.68 (0.48 to 0.97)	31	1.00 (0.63 to 1.58)	50	1.01 (0.82 to 1.25)
Vigorous physical activity in summer (y/n)	13	1.19 (0.74 to 1.89)	12	1.47 (0.89 to 2.43)	37	0.80 (0.48 to 1.34)	34	1.19 (0.86 to 1.64)
Vigorous physical activity in winter (y/n)	10	1.18 (0.75 to 1.88)	16	1.19 (0.86 to 1.65)	41	0.73 (0.42 to 1.27)	33	1.41 (0.96 to 2.08)
Moderate physical activity in summer (y/n)	6	N/A	4	N/A	3	N/A	7	1.35 (0.38 to 4.83)
Moderate physical activity in winter (y/n)	7	N/A	5	1.38 (0.69 to 2.77)	4	N/A	12	1.06 (0.61 to 1.85)
Continuous response among answered yes	*N**	**β (95% CI)**	*N**	**β (95% CI)**	*N**	**β (95% CI)**	*N**	**β (95% CI)**
Smoking (number of cigarettes/day)	45	−2.12 (−5.63 to 1.38)	52	−0.03 (−0.14 to 0.08)	60	−1.50 (−5.52 to 2.52)	72	−0.01 (−0.20 to 0.18)
Vigorous physical activity in summer (times/month)	157	0.08 (−0.46 to 0.62)	121	−0.24 (−0.81 to 0.33)	172	−0.30 (−0.94 to 0.34)	179	0.51 (−0.07 to 1.09)
Vigorous physical activity in winter (times/month)	155	0.52 (−1.20 to 2.23)	121	−0.19 (−1.76 to 1.38)	171	1.54 (−0.28 to 3.36)	174	1.19 (−0.39 to 2.77)
Moderate physical activity in summer (times/month)	157	0.24 (−0.24 to 0.72)	126	0.51 (0.09 to 0.93)	178	−0.12 (−0.64 to 0.41)	196	−0.009 (−0.43 to 0.41)
Moderate physical activity in winter (times/month)	157	−0.27 (−1.83 to 1.29)	126	1.40 (0.16 to 2.63)	179	0.87 (−0.68 to 2.42)	197	−0.26 (−1.41 to 0.89)

Perceived global, physical and mental health was modelled as linear variable, with higher scores indicating better health. Number of depressive symptoms was modelled as ordinal variable.
*N**: analysed pairs.
N/A: OR was not able to calculate as all the cases responded either yes or no.

Table 5 Fixed effects of schooling on health among twin pairs (sex-combined)

Outcome variables		Monozygotic (total pairs $N = 351$)			Same-sex dizygotic (total pairs $N = 338$)	
		β	95% CI		β	95% CI
Health outcomes						
Perceived global health		0.10	0.0003 to 0.19		−0.10	−0.34 to 0.14
Perceived physical health		0.04	−0.02 to 0.10		0.04	−0.01 to 0.09
BMI		−0.04	−0.23 to 0.15		0.04	−0.24 to 0.32
WC (inches)		−0.21	−0.43 to 0.01		−0.09	−0.40 to 0.21
WHR		−0.004	−0.01 to 0.002		−0.003	−0.01 to 0.005
Perceived mental health		0.03	−0.03 to 0.08		0.04	−0.01 to 0.09
Number of depressive symptoms		−0.03	−0.17 to 0.11		0.005	−0.10 to 0.11
Health behaviours						
Dichotomized response (yes or no)	**N^***	**OR**	**95% CI**	**N^***	**OR**	**95% CI**
Smoking (yes/no)	54	0.80	0.58 to 1.11	84	0.88	0.74 to 1.03
Vigorous physical activity in summer (yes/no)	50	1.00	0.73 to 1.37	46	1.29	1.00 to 1.65
Vigorous physical activity in winter (yes/no)	51	0.97	0.71 to 1.33	49	1.29	0.99 to 1.67
Moderate physical activity in summer (yes/no)	9	N/A		11	1.68	0.66 to 4.27
Moderate physical activity in winter (yes/no)	11	N/A		17	1.21	0.83 to 1.77
Continuous response among answered yes	**N^***	**β**	**95% CI**	**N^***	**β**	**95% CI**
Smoking (number of cigarettes/day)	105	−1.91	−4.39 to 0.57	124	−0.43	−2.79 to 1.94
Vigorous physical activity in summer (times/month)	329	−0.08	−0.49 to 0.33	300	0.15	−0.26 to 0.56
Vigorous physical activity in winter (times/month)	326	0.97	−0.27 to 2.21	295	0.54	−0.57 to 1.65
Moderate physical activity in summer (times/month)	335	0.09	−0.26 to 0.44	322	0.24	−0.06 to 0.54
Moderate physical activity in winter (times/month)	336	0.28	−0.81 to 1.37	323	0.47	−0.38 to 1.32

N^*: analysed pairs N. Perceived global, physical and mental health was modelled as linear variable, with higher scores indicating better health. Number of depressive symptoms was modelled as ordinal variable.
N/A: OR was not able to calculate as all the cases responded either yes or no.

did not graduate high school, ORs of smoking among high school graduates was 0.45 (95% CI: 0.08–2.57), 0.37 (95% CI: 0.05–2.71) among those with some college and 0.08 (95% CI: 0.01–0.91) among college graduates. The point estimates of OR for overall categorized education on smoking among MZ male twins similarly indicated an inverse association (OR: 0.30, 95% CI: 0.07–1.33). Due to lack of statistical power, we were unable to estimate ORs for high school, some college and college or more graduates in comparison with those who did not graduate high school among MZ male twins. Among DZ male twins, a dose–response relationship was suggested between schooling and frequency of moderate physical activity (in summer, β: 1.47, 95% CI: 0.25–2.69; in winter, β: 1.40, 95% CI: 0.16–2.63). Among female twins, the associations between categorized education health outcomes were in a similar direction to education modelled as a linear variable (years of schooling). In sex-combined analysis, education was associated with better perceived global health: in comparison with less than high school graduates, coefficient of high school graduates was 1.03 (95% CI: 0.20–1.87), some college was 1.27 (95% CI: 0.40–2.14) and college or more was 1.50 (95% CI: 0.48–2.51).

Discussion

At the individual level, we replicated the previously reported positive associations between schooling and health outcomes/behaviours. Thus, higher educational attainment was associated with better perceived global, physical and mental health, more favourable anthropometric measures, less smoking behaviour and frequent physical activity. However, our twin fixed-effect analyses yielded generally a weak association between schooling and health across the range of outcomes examined, though the point estimates generally tended to suggest protective effects. Only perceived global health was associated with years of schooling among sex-combined MZ twins, although the effect size is small (Cohen's $d = 0.11$) and the corresponding association was not found among DZ twins. Regarding schooling and current smoking, we found a protective association among male DZ twins,

Table 6 Fixed effects of schooling on smoking behaviours among twin pairs (total pairs $N = 689$)

	Male						Female					
	Monozygotic			Same-sex dizygotic			Monozygotic			Same-sex dizygotic		
Outcome variables	N^*	OR	95% CI	N^*	OR	95% CI	N^*	OR	95% CI	N^*	OR	95% CI
Ever smoked	26	0.88	0.64 to 1.22	24	1.27	0.83 to 1.97	32	1.24	0.78 to 1.96	49	0.92	0.74 to 1.14
Current smokers, among ever smokers	21	0.60	0.31 to 1.15	33	0.66	0.45 to 0.96	26	1.09	0.68 to 1.76	35	1.08	0.82 to 1.42
Quit regular smoking, among regular smokers	11	5.76	0.26 to 126.4	17	1.49	0.87 to 2.53	10	0.73	0.32 to 1.64	11	0.59	0.29 to 1.20
Smoking initiation among non-smokers who completed schooling	14	1.00	0.55 to 1.81	5	N/A	N/A	20	0.65	0.31 to 1.39	32	0.78	0.58 to 1.06
Try to quit smoking, among current smokers	6	0.70	0.23 to 2.13	5	1.21	0.44 to 3.35	11	N/A	N/A	5	N/A	N/A

N^*: analysed pairs N.
N/A: OR was not able to calculate as all the cases responded either yes or no.

Table 7 Fixed effects of schooling on smoking behaviours among twin pairs (total pairs $N = 689$) (sex-combined)

	Monozygotic			**Same-sex dizygotic**		
Outcome variables	N^*	OR	95% CI	N^*	OR	95% CI
Ever smoked	58	0.99	0.77 to 1.27	73	0.99	0.82 to 1.19
Current smokers, among ever smokers	47	0.83	0.60 to 1.14	68	0.86	0.72 to 1.04
Quit regular smoking, among regular smokers	21	1.31	0.84 to 2.02	28	1.04	0.74 to 1.44
Smoking initiation among non-smokers who completed schooling	34	0.83	0.53 to 1.31	37	0.74	0.56 to 0.98
Try to quit smoking, among current smokers	17	0.41	0.13 to 1.31	10	1.68	0.59 to 4.81

N^*: analysed pairs N.

though not MZ twins. This suggests that the association between schooling and smoking status may be confounded by unobserved genetic factors.

Our findings were consistent with the middle-aged Danish twin study.[15] In their study, self-reported health, BMI, depressive symptoms, smoking habits and physical activity were not associated with education after cancelling for genetic and early family environmental factors. A previous study showed that the association between education and smoking is stronger than the association between occupation and smoking.[2] Gilman *et al*.[27] showed robust associations between lower educational attainment and smoking patterns, although adjustment for family environment (using a fixed-effect approach among siblings) failed to confirm the association.[27] A study of female twins in Virginia also showed that both genetic and environmental factors were associated with smoking initiation and nicotine dependence.[28] It is possible that there are health selection effects on the association between education and health, i.e. poor health during childhood, such as asthma or epilepsy, may be prior common causes of both limited educational achievement as well as poor health in adulthood.[29,30] However, it is premature to conclude that education is not causally associated with smoking. In our study, although the fixed-effect model among MZ male twins failed to show a strong association between education and current smoking, the point estimate was nonetheless 0.60, which is in the expected direction. The weak associations reported in our twin fixed-effect analyses may be due to low statistical power, rather than to the absence of a causal effect of education on health outcomes.

In addition, our study showed an association between education and dose of moderate physical activity among male DZ twins, but not male MZ twins, suggesting unmeasured genetic factors, but not early family environment, may influence the association between education and moderate physical activity among males. With regard to females, no association was found between education and physical activity after cancelling genetic and early family environmental factors, although a strong association was found in the individual analysis. This suggests that the associations between education and physical activity in individual analysis were confounded by genetic and early family environmental factors among females. These findings are consistent with previous findings suggesting stronger genetic influences on physical activity among males than females.[31]

In summary, the current study showed possible causal effects of education on perceived global health and on smoking habits among males, but did not suggest direct associations between schooling and the other health outcomes studied. The repeatedly observed association between education and health may be at least partly confounded by genetic or early family environmental factors.[9] Nonetheless, our study also suggested that education may be effective in prevent smoking habits among men.

Our study has several limitations. First, the analyses are cross-sectional, so that reverse causation cannot be ruled out. However, educational attainment was determined before 25 years of age, the youngest age of our sample, so that poor health status (which most people begin to succumb to at older ages) is less likely to have affected educational attainment. The remaining bias is health status during schooling, i.e. those with major illnesses during childhood might have completed fewer years of schooling, and also ended up with lower levels of health in adulthood. Secondly, the analyses assume that twins living together up to aged 14 years shared the same family environment, but living together in childhood does not necessary mean that they shared the same 'rearing' family environment. Thus, for example, the inverse association between schooling and smoking among male twins might be explained by a difference in the rearing environment between twins. Thirdly, BMI, WC and WHR were reported by the respondents, not measured by the investigators. The null associations between schooling and BMI, WC and WHR may be due the reporting bias, as educational attainment could affect self-reporting of health outcomes, thereby diluting the true association between schooling and health outcomes. Other risk factors for poor health, for example, blood pressure or low-density lipoprotein cholesterol level, were not measured. Fourthly, the limited response percentage could have led to selection bias in this study. The educational level among twins in this study was somewhat higher than that in the original MIDUS study. It is also possible that twins who were discordant on educational attainment were less likely to participate in the twin study in order to avoid comparisons with their opposite twin pair. Such selection bias may have contributed to an underestimation of the association between schooling and health/health behaviours.

In conclusion, education was associated with health outcomes in individual-level analysis, but in fixed-effect analysis (i.e. adjusted for genetic and early family environment), the associations were less robust, even though the direction of the associations was similar.

For specific health behaviours, such as current smoking and frequency of moderate physical activity, schooling appeared to be associated in the hypothesized direction among males, though not in females. Taken together, our twin fixed-effect analyses suggest that at least some cases of the education/health relationship reflect confounding by unobserved third variables.

Conflict of interest: None declared.

KEY MESSAGES

- The association between education and health may be confounded by unobserved common prior causes such as inherited ability, personality such as time preference, or early family circumstances.
- The twin fixed-effects approach can potentially address this problem by cancelling these factors between twin pairs.
- The findings showed that among monozygotic twins, more years of schooling was associated with better perceived global health.
- For health outcomes/behaviours, the point estimates of the effect of education suggested a weak protective association.

References

[1] Marmot M, Ryff CD, Bumpass LL *et al*. Social inequalities in health: next questions and converging evidence. *Soc Sci Med* 1997;**44:**901–10.

[2] Winkleby MA, Jatulis DE, Frank E *et al*. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health* 1992;**82:**816–20.

[3] Doornbos G, Kromhout D. Educational level and mortality in a 32-year follow-up study of 18-year-old men in The Netherlands. *Int J Epidemiol* 1990;**19:**374–9.

[4] Kunst AE, Mackenbach JP. The size of mortality differences associated with educational level in nine industrialized countries. *Am J Public Health* 1994;**84:**932–7.

[5] Matthews KA, Kelsey SF, Meilahn EN *et al*. Educational attainment and behavioral and biologic risk factors for coronary heart disease in middle-aged women. *Am J Epidemiol* 1989;**129:**1132–44.

[6] Krokstad S, Kunst AE, Westin S. Trends in health inequalities by educational level in a Norwegian total population study. *J Epidemiol Community Health* 2002;**56:** 375–80.

[7] Scharoun-Lee M, Adair LS, Kaufman JS *et al*. Obesity, race/ethnicity and the multiple dimensions of socioeconomic status during the transition to adulthood: a factor analysis approach. *Soc Sci Med* 2009;**68:**708–16.

[8] Farrell P, Fuchs VR. Schooling and health: the cigarette connection. *J Health Econ* 1982;**1:**217–30.

[9] Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. *BMJ* 2001;**322:** 819.

[10] Hart CL, Taylor MD, Davey Smith G *et al*. Childhood IQ, social class, deprivation, and their relationships with mortality and morbidity risk in later life: prospective observational study linking the Scottish Mental Survey 1932 and the Midspan studies. *Psychosom Med* 2003;**65:** 877–83.

[11] Batty GD, Shipley MJ, Mortensen LH *et al*. IQ in late adolescence/early adulthood, risk factors in middle age and later all-cause mortality in men: the Vietnam Experience Study. *J Epidemiol Community Health* 2008;**62:** 522–31.

[12] Lleras-Muney A. Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939. *J Law Econ* 2002;**45:**401–35.

[13] Glymour MM, Kawachi I, Jencks CS *et al*. Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *J Epidemiol Community Health* 2008;**62:**532–7.

[14] Krieger N, Chen JT, Coull BA *et al*. Lifetime socioeconomic position and twins' health: an analysis of 308 pairs of United States women twins. *PLoS Med* 2005;**2:** e162.

[15] Osler M, McGue M, Christensen K. Socioeconomic position and twins' health: a life-course analysis of 1266 pairs of middle-aged Danish twins. *Int J Epidemiol* 2007;**36:** 77–83.

[16] Fujiwara T, Kawachi I. Social capital and health a study of adult twins in the U.S. *Am J Prev Med* 2008;**35:** 139–44.

[17] Brim OG, Baltes PB, Bumpass LL *et al*. *National Survey of Midlife Development in the United States (MIDUS), 1995–1996 Codebook for Part 3, Twins Data*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research, 2003.

[18] Matthews S, Manor O, Power C. Social inequalities in health: are there gender differences? *Soc Sci Med* 1999; **48:**49–60.

[19] Blair SN. How to assess exercise habits and physical fitness. In: Maatarazzo JD, Weiss SM, Herd JA, Miller NE, Weiss SM (eds). *Behavioural Health: A Handbook of Health Enhancement and Disease Prevention*. New York: Wiley, 1984, pp. 424–47.

[20] Kohl HW, Blair SN, Paffenbarger RS, Jr. *et al*. A mail survey of physical activity habits as related to measured physical fitness. *Am J Epidemiol* 1988;**127:**1228–39.

[21] Washburn RA, Goldfield SR, Smith KW *et al*. The validity of self-reported exercise-induced sweating as a measure of physical activity. *Am J Epidemiol* 1990;**132:**107–13.

[22] Marmot MG, Smith GD, Stansfeld S *et al*. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991;**337:**1387–93.

[23] Nichols RC, Bilbro WC, Jr. The diagnosis of twin zygosity. *Acta Genet Stat Med* 1966;**16:**265–75.

[24] Greene WH. *Economic Analysis*. New York: Macmillan, 1993.

[25] Hsiao C. *Analysis of Panel Data*. 2nd edn. Cambridge: Cambridge University Press, 2003.

[26] Grzywacz JG, Marks NF. Social inequalities and exercise during adulthood: toward an ecological perspective. *J Health Soc Behav* 2001;**42:**202–20.

[27] Gilman SE, Martin LT, Abrams DB *et al*. Educational attainment and cigarette smoking: a causal association? *Int J Epidemiol* 2008;**37:**615–24.

28 Kendler KS, Neale MC, Sullivan P *et al*. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychol Med* 1999;**29:**299–308.

29 Austin JK, Huberty TJ, Huster GA *et al*. Academic achievement in children with epilepsy or asthma. *Dev Med Child Neurol* 1998;**40:**248–55.

30 Shapiro AD, Donfield SM, Lynn HS *et al*. Defining the impact of hemophilia: the Academic Achievement in Children with Hemophilia Study. *Pediatrics* 2001;**108:**E105.

31 Maia JA, Thomis M, Beunen G. Genetic factors in physical activity levels: a twin study. *Am J Prev Med* 2002;**23 (2 Suppl):**87–91.

Published by Oxford University Press on behalf of the International Epidemiological Association

Advance Access publication 29 July 2009

International Journal of Epidemiology 2009;**38**:1322–1323
doi:10.1093/ije/dyp264

Commentary: Strengths and limitations of the discordant twin-pair design in social epidemiology. Where do we go from here?

Mia Madsen[1] and Merete Osler[1,2]*

Accepted 23 June 2009

It has become increasingly recognized that many of the associations between various exposures and health outcomes measured in adulthood are confounded by social and behavioural factors acting early in life,[1] and recent research has questioned whether socio-economic position (SEP) in adulthood in fact has an additional impact on adult health beyond the influence of one's childhood environment and genetic make-up.[2]

The traditional way to examine the independent influence of adult SEP on differences in health outcomes is to use multivariable models, which mutually adjust for early and late social factors. However, such analyses may not provide sufficient adjustment for all the factors that constitute childhood social environment, and close correlation between early and late social factors may present serious statistical challenges. Alternative approaches for dealing with confounding are designs that approximate the conditions of a true experiment. One such approach is the discordant twin-pair design employed by Fujiwara and Kawachi[3] in their well-designed and innovative study. Here, the authors compare health status in twin pairs who are discordant on educational status in adulthood, but are matched fully or partly on genetic make-up and rearing environment. Such a study on twin-pair discordant on exposure provides a useful analogue to the idealized counterfactual design, and failure to observe an association within discordant twin pairs would imply that a previously observed association between an exposure and an outcome is attributable to common genetic or shared environmental factors. In addition, differences in the magnitude of association between monozygotic and dizygotic twins may serve to further disentangle the respective effects of genetic and familial environmental effects.

However, as enchanting as this approach may seem, as with everything else, it has a price: in this case, the flip side of the coin is the fact that the statistical power in the fixed-effects analyses is dramatically limited, resulting in highly imprecise estimates. The implication of this is that these 'unconfounded' estimates may in fact be further away from the 'true value', which we are trying to estimate, due to the large imprecision of estimates.[4] A consequence of this is that authors are left with considerable room for interpretation of results and, in some cases, interpretation tends to border on arbitrariness. To illustrate this point, we would like to call attention to a discussion paper by Lundborg,[5] which essentially is based on the same data (MIDUS survey, first data collection wave) as the paper by Fujiwara and Kawachi, although dizygotic same-sex twin pairs were not included. Despite the fact that the results in the two papers appear to be fairly similar, the authors

1 Danish Ageing Research Centre, Institute of Public Health, University of Southern Denmark, Odense, Denmark.

2 Research Center for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark.

* Corresponding author. Research Center for Prevention and Health, Glostrup, University Hospital, 2600 Glostrup, Denmark. E-mail: m.osler@pubhealth.ku.dk