The health effects of education: a meta-analysis

Carina A. Furnée1, Wim Groot1, Henriëtte Maassen van den Brink2

Background: There is an abundance of empirical evidence, mainly from the epidemiological and social science literature, on the relation between education and health. Until now a meta-analysis of the relation between education and health was not available. This article presents a meta-analysis of studies that use self-reported health as an outcome variable to quantify the effect of education on health. Methods: Meta-analysis using a random and a fixed-effects model to quantify the marginal effect of education on self-reported health. Results: The results of the meta-analysis show that the quality adjusted life years weight (QALYweight) of a year of education is ~0.036. Some tentative calculations suggest that the cost-benefit ratio of investments in education on health is highly positive. Conclusion: For public policy this implies that a more integrated approach to education and health policies should be taken.

Keywords: education, meta-analysis, self-reported health

Background

Education and health are important characteristics of human capital. Their economic value lies in the effects they have on productivity: they both make individuals more productive. They affect individual well-being and determine the wealth of nations.

There is an evidence to support the claim that there is a positive relation between education and health. High educational attainment improves health directly but also indirectly through work, economic conditions, social-psychological resources and health life style. Further support for the causal relationship between education and health comes from studies on effects of education on health using mortality data. The positive association between education and health is well documented. According to the 2003 Human Development Report education, health, nutrition, water and sanitation complement each other, with investments in any one contributing to better outcomes in the others. This report reveals large differences in education and health outcomes between countries. It also suggests a positive relation between educational attainment—as measured by the net enrolment ratios on primary and secondary education—and health—measured by life expectancy at birth and the infant mortality rate, per country which can partly be attributed to income differences between countries.

Health and prosperity are positively related. One study shows a strong negative association between the log of purchasing power parity adjusted GDP per worker and the percentage of low birth weight. Low-income countries have fewer resources for publicly financed education and health care. Most individuals in low-income countries are unable to purchase education and health care. On the other hand, investing in education and health provide the way out of poverty and are necessary for increasing standards of living.

There are three potential explanations for the positive relation between education and health: (i) better health enables more investment in education, (ii) common factors such as genetic endowment, social background or time preferences may affect health and education in a similar way and (iii) education leads to better health. An abundance of empirical evidence using disaggregated, individual level data, mainly from the epidemiological and social science literature, is available on the relation between socio-economics status (SES) as measured by educational attainment and health. These micro data are suitable for inclusion in a meta-analysis. The purposes of this study are to gain more insight in the magnitude of the effect of education on self-reported health using a fixed effects (FE) and a random effects (RE) model in a meta-analytic procedure and to overcome difficulties in valuing education inputs and health outcomes by means of quantifying the effect of a year of education on individual health status in monetary terms, as described in the literature. So far, a meta-analysis of the relationship between education and health is not available in the literature.

Methods

Many studies on the effects of education on health use a self-assessment of the individual’s perceived health status as the health outcome variable. We use these studies for a meta-analysis of the QALY-weight effect—a health outcome measure that is derived from responses to a self-report of health (SRH) of education. We combine studies that explicitly look at the effects of education on SRH outcomes—either by using a regression-based approach or by comparing the odds for different educational categories—with studies in which SRH is an outcome measure and health is one of the control variables, but in which the relation between the two is not necessarily the main object of study in the article.

Some studies use an age-related question on subjective health rather than a general question. The age-related SRH questions may introduce a source of age norming bias in the results.

The studies were collected by a search in ECONLIT and MEDLINE. Keywords in ECONLIT include: 'self-reported health', 'self-assessed health' and 'subjective health' and in MEDLINE 'self-assessed health and education', 'subjective health and education', 'self-assessed health and schooling', 'self-reported health and schooling' and 'subjective health and schooling'.

For the RE approach we included only studies that used a population-based sample and a regression-based approach such as Ordinary Least Squares (OLS), ordered probit or

References

1 Faculty of Health, Medicine and Life Sciences, BEOZ, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
2 Faculty of Economics and Business, General Economics, University of Amsterdam, Roeterstraat 11, 1018 WB Amsterdam, The Netherlands
Correspondence: Carina A. Furnée, Faculty of Health, Medicine and Life Sciences, BEOZ, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, tel: +31 43 388 1832, fax: +31 43 367 0960, e-mail: c.furnee@beoz.unimaas.nl
logistic regression for analysis. Studies in which education was only included as a control variable or covariate were excluded. For the FE approach all studies reporting odds and/or descriptive statistics on rates for self-reported poor health (SRPH) per category of educational attainment were included to enable the calculation of the odds.

In the empirical modelling of the quality of health three concepts are distinguished. The first is the true quality of health \( H^* \). The true quality of health is a latent variable that cannot be observed directly. What we observe is a subjective measure of the quality of health, \( H^s \). The subjective measure of health, \( H^s \), is measured by the response to the self-evaluation question. The latent quality of health variable of estimation \( ii \) in study \( j \) is further determined (years of education \( S \) and a vector of control variables \( X \) that may include income

\[
H^s_{ij} = \beta_0 + \beta_1 S_{ij} + X_{ij} \beta_2 + \mu_i + \xi_{ij}
\]

where \( \beta \) are vectors of coefficients, \( \mu \) is an error term capturing differences in estimates within studies and \( \xi \) is a standard normal distributed random term capturing unmeasured and immeasurable effects on the true health status.

The ordered probit estimations treat the observed health status \( H^s \) as a categorical ordered response variable. The observed health variable is assumed to be related to the latent variable as follows:

\[
H^s = k \leftrightarrow \alpha_{k+1} < H^s \leq \alpha_k, \quad k = 0, \ldots, n
\]

where \( n \) is the number of response categories and \( \alpha_i \) are the threshold levels that demarcate the different response categories. If we further assume that \( \alpha_0 = -\infty \) and \( \alpha_n = \infty \), we get the specification of the ordered probit model.

To determine the impact of a year of education on health, the QALY weight of a year of education is calculated. The QALY weight measures the size of the effect of a year of education on quality of health on a 0–1 scale. Let \( \beta_1 \) be the coefficient that represents the impact of a year of education on health. As the \( \beta \) coefficients are not scaled and can range from \(-\infty \) to \( \infty \), the \( \beta_1 \) coefficient needs to be normalized to produce a QALY weight. We normalize dividing by the difference between the borderline between excellent health and that of a very poor health assuming that an excellent health corresponds to a near perfect health and a very poor health corresponds to near death. The QALY weights are defined as:

\[
QALY = \frac{\hat{\beta}_1}{\alpha_4 - \alpha_1}
\]

In the OLS estimations the subjective health status is treated as a continuous variable. In the OLS estimates the QALY weight is simply determined by the marginal effect of education on health as represented by the \( \beta_1 \) coefficient.

In the logistic regression estimations the subjective measure of health, \( H^s \), is collapsed into a dichotomous variable (usually, health is (very) good vs. less than good, or health is (very) poor vs. better than poor). In most of the studies that use logistic regression education is laid out in at least two but often more dummy variables for level of education based on years of schooling.

The odds ratios are the odds of self-reported poor health given a certain level of educational attainment compared to the odds of self-reported poor health among those with the highest education level. The log odds ratio is widely used in the analysis of data that have dichotomous outcomes and is readily interpretable and has the same meaning across studies.

For a meta-analysis the relationship between the odds for reporting poor health in relation to educational attainment across studies is of interest and can be described by a FE model as:

\[
\ln(\text{odds}) = \alpha_0 + \alpha_1 \cdot y
\]

Where \( \ln(\alpha_0) \) indicates the parameter value for the intercept and \( \alpha_1 \) denotes the parameter for the slope indicating the marginal propensity of education on the SRH outcome. A log transformation is carried out to correct for eventual skewness in the data. Thus, the QALY weights derived by the FE model are simply determined by the marginal that effect education on health as represented by the \( \alpha_1 \) coefficient.

One way to assess the impact of education on health is to calculate the cost-benefit of investing in education to improve health. At the margin, the costs of a year of education are about Euro 6000. The costs per QALY must be compared with the life-time value of a QALY in order to determine whether investments in education are welfare improving. To calculate the monetary value of the education effect of the quality of health, we use the literature on the value of a statistical life year. In a meta-analysis of 33 studies the value of a statistical life lies between $1.5M and $2.5M. At a 5% discount rate this would make the value of a statistical life year somewhere between $76.500 and $127.500.

It should be noted that these figures represent the effect of education on SRH and not on mortality.

Results

Table 1 provides the descriptive information about the 40 studies included in the meta-analysis; 35 studies for the RE and 10 studies for the FE approach. These studies provide 64 and 24 estimates of the QALY weight of education, respectively.

From the regression-based studies we derive information on the year of data collection, the country of origin, the econometric method (OLS, ordered probit or logistic regression), sample selection characteristics like gender (male, female, or both genders) and age group (all ages, 60+, younger than 60), the inclusion of control variables for income or parental education, the measurement of the education variable (dummy variables or years of education), the outcome measure (SRH with or without comparison to their own age group), whether the education effect is statistically significant and the size of the QALY weight of a year of education when available. If the education variable was measured by dummy variables, they were converted to years of education (for details see Appendix: http://fdgwbeozrsrv0301.unimaas.nl/people/Publicaties_1.asp?ID=200&Naam1=Furn%28E%294+A%28E%292E).

Some descriptive statistics of the average QALY-weight effect of the studies in the meta-analyses are found in table 1. The un-weighted sample average QALY-weight effects of a year of education are 0.038 and 0.061 for the RE and FE, respectively. The RE estimate of the QALY weight for the entire sample is 0.036 with a standard error of 0.005, while the FE estimate of a QALY weight is 0.061 with a standard error of 0.009. It appears that the sample average of the RE studies is much smaller than the average of the FE studies. However, if we only take the studies that use logistic regression—i.e. studies that are most comparable with the studies using odds ratios—the average QALY weight is 0.058 (with a standard error of 0.007).

The RE estimate of the QALY-weight effect of a year of education is about two-thirds lower in studies using data for the United States (RE coefficient of QALY weight is 0.025) than in studies using data on non-US (predominantly European) countries (0.040), while the data collected for the
Table 1 Means and regression estimates QALY-weights meta-analysis (heteroscedasticity consistent standard errors in brackets)

<table>
<thead>
<tr>
<th>No. of observations</th>
<th>Average QALY weight</th>
<th>Regression estimates QALY weight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RE</td>
<td>FE</td>
</tr>
<tr>
<td>Total sample</td>
<td>64</td>
<td>101</td>
</tr>
<tr>
<td>Country</td>
<td></td>
<td></td>
</tr>
<tr>
<td>European or other non-US country</td>
<td>53</td>
<td>88</td>
</tr>
<tr>
<td>United States</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Econometric method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLS</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Ordered probit</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Odds-Ratio</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>Female</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>Both</td>
<td>22</td>
<td>49</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Aged 60+</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Age &lt; 60</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Controls for income or parental education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Specification of education variable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of education</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Dummy variable(s)</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Health outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otherwise</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Compared to others</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

RE = random effects model; FE = fixed effects model
*significant at 5% level
**significant at 1% level

FE approach show coefficients of QALY weights of 0.098 and 0.054 for the United States and non-US countries, respectively.

Studies using logistic regression yield higher estimates of the education effect on health than studies using OLS or ordered probit. The logistic regression estimates (RE coefficient for studies using logistic regression is 0.061) are more than twice as high as the OLS estimates (0.024) and more than four times as high as the effects in studies using ordered probit regression (0.012).

For the demographic variables, gender and age, we find that the QALY-weight effect is smallest in studies that only use observations on both men and women (0.029). Studies using data on males (0.041) or females (0.045) on average provide higher estimates of the effect of education on self-reported health. The data for the FE approach that reports on an effect of education on self-reported poor health show the opposite effect—for studies using data on both men and women the QALY weights are highest (0.075), while studies using data on men and women separately show QALY-weight effects of 0.051 and 0.052, respectively.

The RE coefficients do not differ much by age group. For studies using samples of individuals older than 60, studies with data of people younger than 60 years and studies using all age groups yield similar effect sizes, on average. For the FE approach we do not have details about the effect of age.

One would expect that the inclusion of income or parental education would lower the QALY-weight effect of education, as without these controls the education effect may pick up some of the income and social-background effects on health as well. This is not what the results of the meta-analysis show, however. The average QALY-weight effect of a year of education is similar in studies that include controls for income or parental education (0.037) and in studies where these variables are not included (0.036).

The specification of the education variable in the equation does appear to matter. In studies that use dummy variables for education the average QALY weight is 0.049, while in studies that use years of education it is 0.020. The data for the FE approach does not have details about the effect of controlling for education.

The average QALY-weight effect in studies in which the respondent is asked to compare his/her health with others of the same age is somewhat lower (0.015) than in studies without this age norming (0.039).

The descriptive statistics show differences between studies by country, econometric method and gender. Sample means, however, tell us little about the sources of heterogeneity in the QALY-weight estimates within and between studies. As argued by Thompson and Higgins the appropriate analysis of the heterogeneity in meta-analysis is a RE meta-regression.19

Table 2 contains the regression estimates with the QALY-weight effect as dependent variable. In a first round all of the above mentioned variables are included in a RE meta-regression. As the number of observations is small relative to the number of variables in the meta-regression, we also present results of a more economical model, i.e. a RE model which only includes variables with a t-value > 1.

If we include all variables, four have a statistically significant effect: the country dummy, the variable for the econometric method, the male gender dummy and the dummy variable for controls for income or parental education. If we only include variables with a t-value > 1, only two variables have a statistically significant effect: the country dummy and the dummy variable for the econometric specification. We find
that the QALY-weight effect is 0.008 higher in studies using data for the United States than in studies using data on European countries or Canada. We further find that in studies that used ordered probit regression the QALY-weight effect of a year of education is 0.009 lower than in studies that use OLS.

Controlling for income or parental education does not appear to have a statistically significant effect on the QALY weight, at least in the reduced model. Controlling for country, econometric method and gender yields a QALY-weight effect of a year of education of 0.023 (the intercept term in the regression).

The RE model shows that a year of education improves QALY by 0.036, this is equal to 28 years of education for 1 QALY. In the FE model a year of education improves 0.061 which equals to 16 years of education for 1 QALY. With the costs of a year of education at about Euro 6000 to obtain a QALY worth of health gain, the costs per QALY amount to about 168 000 or 98 000 Euro, respectively.

### Discussion

The available evidence suggests that there is a strong link between education and health. This view is supported by the results of the two meta-analysis carried out in this study. The relationship found in the meta-analysis might actually be an underestimate of the real magnitude of the effect. This is because lower levels of education appear to be associated with underreporting of illness by patients. It appears that educational attainment itself is likely to bias one’s ability to rate their own health status.

SRH is frequently measured by a question that asks respondents to rate their overall health on a 5-point scale from excellent to very poor. This is a rough method for estimating a utility score but there is widespread agreement that this simple global question provides a useful summary of how persons perceive their overall health status. SRH is a powerful predictor of clinical outcome and morbidity.

A limitation of the studies using logistic regression analysis is that they only report odds ratios for different levels of education, i.e. how many times a lower education increases, the probability of being in less than perfect health relative to a higher education. These studies only provide information on the relative effect of education on health and not of the absolute health effect of education. Typically these studies also transform the ordinal variable subjective health (the 5-point self-rating scale from ‘very bad’ to ‘excellent’) into a dichotomous variable (usually a variable indicating whether the health status is ‘excellent’ vs. ‘less than excellent’, or ‘(very) poor’ vs. ‘better than poor’). Dichotomizing an ordinal variable inevitably leads to a loss of information. However, the RE method in this study appears to do well as the outcomes compare well to the outcomes of two studies in the literature on predicting Health-Related Quality of Life (HRQL).

The cost per QALY in this study is compared with the lifetime value of a QALY based on effects of education on quality of life and does not quantify the effects of education on mortality. Therefore, it is likely that the health benefits of education are in fact much larger than the cost per QALY, and that investing in education has an even larger positive cost-benefit ratio. The relation between education and health has important implications for public policy. Public policies tend to be highly compartmentalized: education is the domain of the Ministry of Education, while health care is looked after by the Ministry of Health. What this survey has shown is that there are large spill-over effects between education and health. This implies that education and health policies do not have an effect within their own domain, but that there are large costs and benefits associated with these policies. This entails that these policies should not be looked upon in isolation, but that a more comprehensive or integrated policy approach to education and health is called for.

### Key points

- A meta-analysis on health effects of education has so far not been available.
- Our findings indicate that the cost-benefit ratio of investments in education on health is highly positive.
- For public policy a more integrated approach to education and health policies should be taken.
References


Received 15 November 2007, accepted 27 March 2008