Journal of Public Health | pp. 1–9 | doi:10.1093/pubmed/fdv124

# Educational attainment and health transitions over the life course: testing the potential mechanisms

**Marko Elovainio[1,2], Tom Rosenström[1], Christian Hakulinen[1], Laura Pulkki-Råback[1,3], Sari Mullola[1], Markus Jokela[1], Kim Josefsson[1], Olli T. Raitakari[4,5], Liisa Keltikangas-Järvinen[1]**

[1]Institute of Behavioral Sciences, University of Helsinki, PO Box 9, 00014 Helsinki, Finland
[2]National Institute for Health and Welfare, PO Box 30, 00271 Helsinki, Finland
[3]Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki, Finland
[4]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland
[5]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland
Address correspondence to Marko Elovainio, E-mail: marko.elovainio@thl.fi.

## ABSTRACT

**Background** It has been shown that higher education associates with health outcomes, but the less is known about the specific mechanisms mediating this association. We examined whether higher education would associate with long-term health transitions from childhood to adulthood and whether health behaviors, self-esteem, social support and work-related health hazards could mediate or confound this association.

**Methods** The participants were from a population-based sample of 3596 men and women from the Young Finns study aged 3–18 years at the beginning of the study in 1980, and who responded to repeated surveys of educational attainment and self-rated health in four study phases from 1997 to 2012. The associations were tested using multistate Markov models for the health-state transition intensities.

**Results** Our results suggested that a 1-year difference in education was related to a 16% higher transition probability from mediocre to good self-rated health over the 5-year follow-up. Depressive symptoms and job strain seemed to partly mediate or confound the association, but self-esteem and social support did not.

**Conclusions** These results suggest that educational attainment is associated with good self-rated health transitions from childhood to adulthood, and multiple processes rather than a single underlying mechanism are likely to drive the educational differences in self-rated health.

**Keywords** education, Markov-models, socioeconomic, trajectories, young Finns

## Introduction

The relationship between various indicators of socioeconomic status (SES) such as income, education and occupational status and health is well established.[1] This has raised the question of the mechanisms by which this association operates. Although, in some SES indicators, the mechanisms are more inevitable—for example low income determines the ability to consume goods and services which in turn directly affects health[2]—in other indicators, such as education, the association is more distal.

Although the evidence clearly shows that those with more years of schooling tend to have better health and well-being,[3,4] education may, however, be associated with health also through mechanisms unrelated to income. Low education has been associated with behavioral preferences, i.e. health-related

**Marko Elovainio**, Professor of Psychology
**Tom Rosenström**, Research Fellow
**Christian Hakulinen**, Lecturer of Psychology
**Laura Pulkki-Råback**, Senior Research Fellow
**Sari Mullola**, Research Fellow
**Markus Jokela**, Professor of Psychology
**Kim Josefsson**, Research Fellow
**Olli T. Raitakari**, Professor of Cardiovascular Medicine
**Liisa Keltikangas-Järvinen**, Professor of Psychology

behavior,[5] and with higher exposure to occupational health hazards potentially contributing to health problems.[6] Furthermore, differences in other social contexts, social values, social identity and social support[7–9] have been suggested to represent a meditational pathway linking education to health.[10,11] Positive self-esteem has been associated with better academic performance[12–14] as well as with health outcomes.[15,16] Thus, the evidence on the channels for effects of education on health suggests that instead of one single mechanism, multiple mechanisms may operate at the same time through effects at the individual, work and household level.[17–25]

In this study, we investigated the association between the development of educational level and self-rated health, with the goal of examining several potential mediating or confounding factors to better understand possible explaining mechanisms. We tested health behaviors, self-esteem, social support and work-related health hazards (job strain literature[23,24]) as adjusting factors for the association between educational attainment and transitions in self-rated health measured over three study phases (1986, 1989 and 2001). The dataset was derived from general population, representative of Finnish children and adolescents when the study started, and followed up during 21 years.

## Methods

### Participants

The Cardiovascular Risk in Young Finns Study is an ongoing follow-up study of coronary heart disease risk factors in Finnish children, adolescents and young adults.[26] The first cross-sectional study was conducted in 1980 when age cohorts of 3-, 6-, 9-, 12-, 15- and 18-year olds were randomly sampled on the basis of social-security numbers, resulting in a total of 3596 participants. The participants have given a written consent, the study plan has been accepted by local ethics committees, the study protocol of each study phase (1980, 1986, 1989, 1997, 2001 and 2007) followed the guidelines of the World Health Organization and the study conforms with the Helsinki declaration. Treatment of the participants complies with American Psychological Association ethical standards.

### Measures

*Self-rated health* was measured by asking a question about perceived general health.[27] The question was 'How would you estimate your current state of health?'.[27] Herein, we studied an item 'Evaluate you general health: 1 = very good, 2 = fairly good, 3 = mediocre, 4 = fairly poor, 5 = very poor' and classified it into good (1 or 2), mediocre (3) and bad (4 and 5). We then assessed temporal state transitions between good, mediocre and bad subjective-health status. Subjective health was assessed in the follow-up years 1986, 1989 and 2001. This single item has been shown to be a powerful predictor of mortality.[28] More than 20 published studies have consistently shown that global self-rated health is an independent predictor of mortality, despite the inclusion of numerous specific health status indicators and other relevant covariates known to predict mortality. Self-perceived health is a particularly valuable measure among younger people, for whom the prevalence of illness is lower than among older adults.[29]

*Body mass index* (BMI) was calculated by dividing participants' weight in kilograms by their squared height in meters (measurements taken in a medical examination in 1986).

*Self-esteem* was assessed in 1986 using a shortened version of the Coopersmith Self-esteem Inventory.[30] The total score consisted of 18 items (e.g. 'I often feel ashamed of myself'; reverse scored), assessed on a 5-point scale; ranging from 'totally disagree' to 'totally agree'. However, self-esteem was not measured from the youngest cohort, i.e. participants who were 9 years old in 1986. Cronbachs' $\alpha$ reliability was good ($\alpha = 0.84$).

*Current smoking* was reported yes/no (in the year 1989). *Alcohol consumption* was assessed also in 1989, by enquiring whether participant had been drunk: 1 = 'never', 2 = 'once', 3 = '2–3 times', 4 = '4–10 times' and 5 = 'more than 10 times'. This variable was recoded into a dichotomous variable computed by setting score 5 of the original item to 1 and scores 1–4 to 0.

*Social support* was assessed in 1997 with the Perceived Social Support Scale-Revised consisting of 12 items ($\alpha = 0.94$) measuring social support received from family and friends. Responses were given on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). Items that indicated social support by family (four items) and items that assessed social peer support were also separately examined.

*Income* in adulthood was classified on an 8-point scale, ranging from annual gross income from 1 (less than €10 000) to 8 (more than €70 000) and measured in 2001.

*Job strain* was measured in 2001 using the 9-item Job Control scale ($\alpha = 0.87$) from the Job Content Questionnaire[23] (the response scale was from 1 = disagree to 5 = agree) and the 3-item Job Demands scale ($\alpha = 0.62$) from the Occupational Stress Questionnaire (response scale from 1 = never to 5 = all the time). These three items correspond to Karasek's Job Content Questionnaire (1985). Job strain was calculated using linear term (job demands − job control). The use of continuous variables has been recommended and when using a linear term of job strain, the contributions of job demands and job control are equally weighed.[31]

*Education* was indicated by two self-reported indicators: (i) years of education and (ii) five educational levels and coded as

(a) comprehensive school, (b) high school or vocational education, (c) bachelor's degree, or studies performed at the university but degree not completed, (d) master's degree and (e) licentiate or doctoral degree. Education was assessed in the year 2001, so that all participants had a reasonable time frame to gather education (age 24–39 years).

*Depression* was assessed using a modified version of Beck's Depression Inventory. Benefits of this measure in the general population have been discussed elsewhere.[32] Because this is not a standard measure, we use the standardized scale with mean of zero and variance of one.

Information about age (years) and gender were also used.

### Statistical analyses

As preliminary analyses, we provided means and standard deviation of all variables and linear regression models predicting self-rated health in 2001 (Table 1). The linear models are used to report crude associations between variables. However, as the variables had very different number of missing values, imputation methodology was required for their meaningful comparison; thus, the analysis was based on multiple imputation, using 'mirt' R package.[33] Imputation model was taken to be a combination of predictive mean matching models as given by the 'quickpred' function; altogether 35 imputation chains were initiated and iterated for 35 iterations, carefully monitoring their convergence and sensibility of distributions of imputed values. Pooled regression coefficients were computed and their significance levels inferred from an $F$ reference distribution.

The final statistical analysis was based on multistate Markov models for the health-state transition intensities.[34] First, we studied how participants moved between health states, and a model was estimated for the instantaneous transition intensities between states $r$ and $s$,

$$q_{rs}(t, z(t,i)) = \lim_{\delta t \to 0} \frac{P(S(t+\delta t) = s \mid S(t) = r)}{\delta t}$$
$$= q_{rs}^{(0)} \exp\left(\beta_{rs} \cdot z(t,i)\right)$$

Wherein $z(t, i)$ represents value of the covariate $z$ for participant $i$ at the time $t$, $\beta_{rs} \cdot z(t, i)$ is the dot product of estimable

**Table 1** Descriptive statistics and linear models

| *Variable* | n | *Mean* | *SD* | *Model I* B | *Model I* *SE*(B) | *Model I* P | *Model II* B | *Model II* *SE*(B) | *Model II* P |
|---|---|---|---|---|---|---|---|---|---|
| Age in 2001 | 3596 | 31.442 | 4.991 | — | — | — | 0.031 | 0.027 | 0.260 |
| Sex | 3596 | 1.491 | 0.500 | — | — | — | 0.020 | 0.024 | 0.415 |
| Smoke | 2707 | 0.225 | 0.418 | 0.082 | 0.022 | 0.000 | 0.039 | 0.022 | 0.080 |
| Alcohol | 2656 | 0.361 | 0.480 | −0.026 | 0.021 | 0.222 | −0.024 | 0.025 | 0.326 |
| Repeatedly drunk | 2656 | 0.239 | 0.427 | 0.048 | 0.023 | 0.038 | −0.009 | 0.027 | 0.741 |
| Body mass index | 2500 | 20.019 | 3.456 | 0.118 | 0.033 | 0.001 | 0.084 | 0.028 | 0.004 |
| Depression | 2092 | 0.000 | 1.000 | 0.451 | 0.018 | 0.000 | 0.417 | 0.021 | 0.000 |
| Family support | 2691 | 0.000 | 1.000 | −0.095 | 0.021 | 0.000 | 0.003 | 0.021 | 0.903 |
| Friend support | 2687 | 0.000 | 1.000 | −0.110 | 0.023 | 0.000 | −0.019 | 0.024 | 0.426 |
| Education years | 2604 | 14.472 | 3.097 | −0.065 | 0.022 | 0.003 | 0.010 | 0.028 | 0.725 |
| Education level | 2180 | 2.276 | 1.438 | −0.075 | 0.021 | 0.000 | −0.026 | 0.026 | 0.327 |
| Self-esteem | 1898 | 3.682 | 0.492 | −0.163 | 0.026 | 0.000 | −0.020 | 0.027 | 0.471 |
| Job control | 2032 | 3.777 | 0.734 | −0.131 | 0.023 | 0.000 | — | — | — |
| Job demands | 2046 | 2.861 | 0.673 | 0.125 | 0.026 | 0.000 | — | — | — |
| Job strain | 2026 | −0.917 | 0.922 | 0.196 | 0.022 | 0.000 | 0.044 | 0.020 | 0.029 |
| Job passivity | 2026 | −6.639 | 1.063 | 0.004 | 0.024 | 0.878 | — | — | — |
| Income | 2146 | 3.499 | 1.559 | −0.094 | 0.023 | 0.000 | −0.012 | 0.023 | 0.610 |
| Self-rated health 1986 | 1259 | 1.891 | 0.747 | 0.324 | 0.032 | 0.000 | — | — | — |
| Self-rated health 1989 | 1776 | 1.821 | 0.728 | 0.355 | 0.025 | 0.000 | — | — | — |
| Self-rated health 2001 | 2094 | 2.097 | 0.818 | — | — | — | — | — | — |

Column '*n*' provides available sample (nonmissing values) per variable, column 'Mean' proves average of these, 'SD' standard deviation. 'Model I' provides age- and sex-adjusted standardized linear regression coefficients ('*B*'), their standard errors ('SE(*B*)') and Wald-type *P*-values ('*P*'), when predicting subjective health in 2001 ('Health 2001'). Multiple imputation of missing values was used. 'Model II' enters all the variables to a single multivariate linear model. Variables that were linear transformations or repeated measurements of outcome were not included to the model.

regression coefficients $\beta_{rs}$ and the covariate(*s*), and *r* and *s* get values in the set of subjective-health states (good, mediocre and poor). In the initial modeling step, no covariates were included. Second, we tested which of the studied covariates/mediators statistically significantly improved the model. Third, provided that education years and/or level helped to explain the health-state transitions, we tested the possibility of their effect being confounded/mediated by some of the other covariates. This was implemented by testing whether the education variables still improved the model *after* the tested covariate was included *a priori*. Model with education variables and a tested covariate was also compared with the model having only the covariate by computing Akaike's information criteria (AIC) for both models and examining the difference (smaller AIC implies better model).

Statistical tests were likelihood ratio (LR) tests between a null model and an alternative model. That is, the test statistic was $-2\log(\mathrm{LR})$, where the LR is the likelihood for null model divided by the likelihood for alternative model. The test statistic grows whenever the likelihood for alternative model does. Missing values in the health-state variable were modeled as censored states, with any underlying state possible.[34] There is no obvious large-sample theoretic method for imputing the covariates in this model, and therefore, missing data in a covariate resulted in list-wise deletion of the observation from the analysis in question. Sum scores of several items may have missing values in just some of the items. In cases with only one missing item value, score was averaged over remaining nonmissing values, whereas for several missing items, the entire score was considered as missing.

Transitions between missing-value state and all other states occurred to both directions, and therefore, we modeled the missing data as censored observations wherein any underlying health state was possible. To obtain a convergence to a feasible transition-time estimates, we needed to allow all instantaneous transitions, that is not only from health state 'good' to 'mediocre' but also from 'good' to 'poor' and vice versa (see Results, Table 3).

## Results

In the univariate age- and sex-adjusted standardized linear regression model, almost all (except alcohol consumption and passive jobs) associated with self-rated health in 2001. In the multivariate model, only BMI, depressive symptoms and job strain were associated with self-rated health in 2001 (Table 1). Table 2 shows the distribution of observed transitions between subjective-health states for the total of 7192 transitions and 10 788 observations from 3596 individuals. Each individual was observed three times, in the year 1986, in 1989 and in 2001. Spearman's rank correlations of the health states across the three follow-ups were $r_{86,89} = 0.361$, $r_{86,01} = 0.249$ and $r_{89,01} = 0.293$.

**Table 2** Observed state transitions and state correlations across the follow-ups

| *From/to* | *Good* | *Mediocre* | *Poor* | *Missing* |
|---|---|---|---|---|
| Good | 1493 | 303 | 27 | 702 |
| Mediocre | 177 | 147 | 20 | 112 |
| Poor | 12 | 25 | 8 | 9 |
| Missing | 1280 | 331 | 47 | 2499 |

Within a year, there was a 90% change to stay in a good health state, 9% change for transition from good to mediocre and 1% change for transition from good to poor subjective-health state. In contrast, the participants were less likely to stay in poor subjective-health state for a year than to transit to a better state within the year. The maximum likelihood estimate of the transition-intensity matrix, and estimated transition probabilities for a 1-year interval (estimated 1 year change), is reported in Table 3.

Having estimated the model without covariates, we proceeded to analyze whether inclusion of some covariates would improve the model. Statistical significance of each covariate was first examined in isolation, as the model complexity exceeded linear regressions and various plausible covariates involved different amounts of missing values (Supplementary Table S1). Only covariates that significantly enhanced model fit were further considered.

Bonferroni-corrected critical significance level for the 15 tests is 0.003 (Table 3), assuming a 0.05 family-wise significance level. Striving for caution and low model complexity, we use this conservative significance level regarding variable choices for multiple-covariate models. Furthermore, the transitions from mediocre to good health were very similarly affected by family support [hazard ratio (HR) = 1.45 (1.01, 2.08)] and friend support [HR = 1.57 (1.15, 2.14)], other transitions not being significantly affected by these covariates. Therefore, family and friend support were combined to a single social support from this point onward [model $-2\log(\mathrm{LR}) = 45.74$; $P = 3.34 \times 10^{-8}$; and $\mathrm{HR}_{\mathrm{mediocre}\rightarrow\mathrm{good}} = 1.49$ (1.08, 2.04)].

Regarding the observed significant effect of education years on health-state transitions, age could clearly be a confounding factor, as both health states and years of education change with age. The association between education years and health-state transitions remained significant despite adjusting for age, however [$-2\log(\mathrm{LR}) = 28.16$, $P = 8.75 \times 10^{-5}$].

**Table 3** Estimated transition intensity matrix and 1-year transition probabilities for subjective-health multistate model, with 95% confidence intervals in parentheses

| *Transition intensities* | | | | *1-Year transition probabilities* | | | |
|---|---|---|---|---|---|---|---|
| *From/to* | *Good* | *Mediocre* | *Poor* | *From/to* | *Good* | *Mediocre* | *Poor* |
| Good | −0.126 (−0.161, −0.098) | 0.115 (0.090, 0.146) | 0.011 (0.003, 0.045) | Good | 0.901 (0.880, 0.917) | 0.089 (0.074, 0.107) | 0.009 (0.005, 0.024) |
| Mediocre | 0.303 (0.204, 0.449) | −0.399 (−0.517, −0.306) | 0.095 (0.049, 0.185) | Mediocre | 0.268 (0.211, 0.348) | 0.686 (0.613, 0.745) | 0.046 (0.025, 0.078) |
| Poor | 1.157 (0.846, 1.583) | 0.009 (0.000, 1.065) | −1.166 (−1.594, −0.854) | Poor | 0.641 (0.483, 0.736) | 0.043 (0.032, 0.363) | 0.316 (0.116, 0.421) |

Then the potentially confounding/mediating effects of other statistically significant covariates on the association between education years and health-state transition intensities were studied by testing whether adding the education-years variable increases the fit of a model that includes age variable and the other covariate in question. Table 4 displays results from these 'confounding' tests.

The covariates that may mediate, or have overlapping explanatory variance with, education-years' effect on health-state transitions were depressive symptoms, social support, job control and its derivative, job strain. Clear evidence for mediating effects of any single factor were not obtained, however, if 'clear' is taken to mean that $|\Delta - 2\log(\mathrm{LR})|$ should exceed the quantile of $\chi^2$ distribution that provides a statistically significant test-statistic value [i.e. $-2\log(\mathrm{LR})$] with a 0.05 significance level. The 0.95 quantile with 6 degrees of freedom is 12.59. Only job demands covariate seemed to have an effect of this magnitude, but to the direction wherein inclusion of education years enhanced model fit *more* when job demands and age were in null model than when only age was in the null model. In a direct model comparison, education years generally improved the model fit of an examined covariate and age when the covariate was smoking, social support, self-esteem, job demands or gross income. Hence, the effect of education years on health transition was unlikely to be due to these covariates. In contrast, if the null model included depression or job control or strain with age, addition of education-years covariate did not improve the model according to the AIC. Hence, some evidence for mediation or confounding by depression and job control was obtained.

Table 5 shows the multiplicative hazard coefficients for each of the transition intensities for the covariates. Hazard coefficients whose confidence intervals do not overlap with 1 indicate the individual transitions that were clearly affected by the covariate in question (highlighted in bold font). On the one hand, older age was associated with less health-state transitions overall, that is, with a more temporally stable subjective-health status. Although, there seemed to be some trade-off in parameter estimates for transition between mediocre and bad health states with respect to age. On the other hand, greater number of education years was associated with more transitions from mediocre to good health state.

## Discussion

### Main findings of this study

Attained years of education seemed to associate with the preceding 16-year self-rated health-state transitions. On average, an increase of 1 year in education implied 16% higher rate of transition from mediocre to good subjective-health status.

**Table 4** Likelihood-ratio tests for incremental value of education years in explaining self-rated health-state transitions over and above the age and a tested covariate

| *Variable* | *After the covariate* | | *In the subsample* | | *Covar effect* | *Subsample size* | |
|---|---|---|---|---|---|---|---|
| | $-2\log(LR)$ | P-*value* | $-2\log(LR)$ | P-*value* | $\Delta-2\log(LR)$ | *ΔAIC* | *Individuals* |
| Smoking | 28.01 | $9.34\times10^{-5}$ | 27.61 | $1.11\times10^{-4}$ | 0.40 | −16.01 | 2156 |
| Depression | 8.64 | 0.195 | 13.11 | 0.041 | −4.47 | 3.36 | 1989 |
| Social support | 21.09 | 0.002 | 26.16 | $2.07\times10^{-4}$ | −5.07 | −9.09 | 2124 |
| Self-esteem | 13.67 | 0.034 | 16.85 | 0.010 | −3.18 | −1.67 | 1509 |
| Job control | 6.75 | 0.345 | 12.23 | 0.057 | −5.48 | 5.25 | 1932 |
| Job demands | 29.59 | $4.71\times10^{-5}$ | 13.57 | 0.035 | 16.02 | −17.59 | 1946 |
| Job strain | 9.24 | 0.160 | 13.42 | 0.037 | −4.18 | 2.75 | 1926 |
| Income | 24.85 | $3.64\times10^{-4}$ | 16.98 | 0.009 | 7.87 | −12.85 | 1910 |

Because sample size does not stay constant due to different amounts of missing values in different covariates, the statistical significance of education years is tested both against the null model including age and the covariate (after the covariate) and against a null model including only age (alone in the subsample). This way, one can assure that sample attrition does not cause apparent adjustment effects. LR, likelihood ratio; Covar effect, effect of covariate on the likelihood ratio of a null model without education years and an alternative model that includes education years. The test statistic $-2\log(LR)$ is increased whenever modeling of education years improves the model fit. ΔAIC is a difference between AIC for a model with education years, indicating covariate and age with the indicated covariate and age. Negative ΔAIC implies that education years covariate improved the model over just indicated and age covariates.

**Table 5** Covariate hazard coefficients for age, education years and education level in a multistate model

| *Subjective-health-state transition* | *Age and education years as covariates for 2180 individuals and 6540 observations* | | *Education years as covariates for 2180 individuals and 6540 observations* |
|---|---|---|---|
| | *1-Year increase in age (95% CI)* | *1-Year increase in education (95% CI)* | *1-Year increase in education (95% CI)* |
| Good → mediocre | **0.85 (0.80, 0.90)** | 1.08 (0.98, 1.18) | 1.07 (0.98, 1.16) |
| Good → poor | 0.38 (0.11, 1.37) | 0.72 (0.50, 1.03) | **0.58 (0.41, 0.81)** |
| Mediocre → good | **0.80 (0.75, 0.86)** | **1.16 (1.06, 1.27)** | **1.15 (1.02, 1.27)** |
| Mediocre → poor | 5.68 (0.84, 38.40) | 0.52 (0.25, 1.09) | 1.01 (0.87, 1.17) |
| Poor → good | 0.37 (0.10, 1.31) | 0.80 (0.57, 1.12) | 0.94 (0.85, 1.03) |
| Poor → mediocre | **7.19 (1.09, 47.66)** | 0.55 (0.29, 1.04) | 0.44 (0.11, 1.72) |

CI, confidence interval.

This main effect was robust to adjusting for gross income, social support, self-esteem, job demands, body mass, alcohol use and smoking status, but was partly explained by or mediated through depressive symptoms and job control aspect of the job strain construct.

Our results are in accordance with the large body of evidence supporting the socioeconomic gradient in health and specifically the predictive effect of SES, indicated here as education, on health outcomes.[4] Furthermore, we found that although less depressive symptoms and control over one's job conditions may act as mechanisms linking higher education to long-term self-rated health transitions, clear and strong evidence for confounding or mediating effects of any single factor was lacking.

## What is already known on this topic

Previously many of the indicators of socioeconomic position, such as education, income and occupational position, have been shown to associate with a large number of various health outcomes.[2] Furthermore, material, biological and psychosocial mediators of the association have been suggested,[35] and furthermore, it has been found that educational attainment or academic performance is associated with these potential mediators.[36] Similar mechanisms have been included in theoretical frameworks, such as the reserve-capacity model.[17,37] Also biological mediators, such as chronic physiological stress reactions, have been suggested.[35]

#### What this study adds

Our results suggest that education really affects self-rated health, which stabilizes during the middle age (cf. Table 4), and that despite multiple mechanisms linking the association, the effect of any single one is relatively small. The effects of mediating variables could be stronger, however, if their accumulation over years and accumulated effects were assessed, whereas here we did not have that data for all the potential confounders or mediators. It has been previously suggested that the association between education and improved health is causal.[38] Although causality was not directly tested here, the current study results also support the notion that exposure to education is tightly linked with *transitions* in health states rather than merely correlated with it. Previously, few prospective studies have adopted similar robust designs to investigate the effects of education on changes in general health status. The association between education and health has been typically investigated using longitudinal design but with traditional statistical methods, i.e. linear regression analysis. Thus, a major strength of this study was the use of a longitudinal data with multiple outcome measures that enabled us to test the health status transitions over a long period of time. Our potential mediators tested were not measured in the same data collection phase, and thus, the design was not perfect, but in principle, the mediators could have been the mechanisms that linked the education to health transitions.

Many of the previous studies have treated self-rated health as a dichotomous variable defining only poor or good health. There is evidence to suggest, however, that self-rated health forms a continuum from poor through average to good health justifying our use of three categories, and further allowing comparability of a five-point and an eight-point measure of self-rated health. Furthermore, it has been shown that different ways of treating the self-rated health variable (dichotomous, ordinal, etc.) offers similar results.[39]

The present modeling framework is able to take nonlinear effects into account, in the sense that different covariates may selectively affect to just some of the subjective-health transitions without affecting others; for example, we observed that age decreased the rate of transiting from mediocre to good health, but increased both mediocre-to-bad and bad-to-mediocre transitions (Table 5). This kind of a finding would not readily arise in a typical linear model, but might well make sense: it is possible that age decreases changes to increase one's subjective health to the highest levels (mediocre → good), but at the same time, increases probability of physical injuries and accidents followed by recovery that cause oscillations between mediocre and bad states. Both in the joint model with age and as a sole covariate, a year more education implied ∼1.15-fold rate of transition from mediocre to good state. Education also seemed to decrease hazard of transitioning directly from good to bad subjective-health state (Table 5). These findings fit with a view that education helps in making moderate improvements to subjective health (mediocre → good) and in preventing severe illness (good → bad), but when faced with severe illness (presumably being in a 'bad' subjective condition), more education does not lead to miraculously fast improvements in comparison with less educated people ('bad → good' nonsignificant).

In conclusion, this study presents evidence supporting the association between educational attainment and health transitions from childhood to adulthood. Furthermore, none of the potential mediators derived from the previous empirical evidence and previously presented theoretical frameworks on the pathways from lower SES to poorer health[10] seemed to have a strong mediating effect. Thus, academic attainment predicts positive self-rated health transitions from childhood to adulthood, and multiple processes rather than a single underlying factor are likely to mediate the association.

These findings may have some implications for policy. The general recommendation would be to focus on the accumulation of risk factors and benefits throughout the life course. Such a strategy should ensure that children with health problems are not disadvantaged with respect to educational opportunities. Determined action is also required to remove material and psychosocial mechanisms that may include individual (depressive symptoms, self-esteem) and environmental (social support, positive psychosocial work characteristics) factors likely to underlie the health problems in groups with low educational attainment.

#### Limitations of this study

Our findings should be viewed in the context of some limitations. First of all, most analyses presented here have been conducted on <65% of the original cohort sample. Such levels of attrition may introduce bias in our results. Univariate effect of education years was always also studied in the subsets that included the potential mediating variable, however, thereby ensuring that attrition and covariate adjusting was never confounded with each other. Second, both the initial (no covariates) and final (age and education-years covariates) multi-state models leaved much to hope in terms of model-to-data fit, but nonetheless seemed to capture significant variation in the data and to provide some interpretable and realistic results. Very accurate model-to-data fit would have actually been an unrealistic expectation, given the small number of follow-ups (three) with quite heterogeneous measurement intervals (3 and 12 years). Furthermore, all differences between observed data and model estimates may not represent shortcomings but a legitimate missing-data modeling with respect to self-rated health states (see Methods).

Moreover, we lacked the data for careful separation of baseline developmental phase, or age. Therefore, the results must be interpreted as informing about general temporal transitions during the number of different (relatively young) age periods, rather than, say, specific to transitions from teenhood to young adulthood. Some associations could become clearer if the different phases of life could be studied in finer grain. Bad subjective health is rare in young people, however, thus requiring large samples to evaluate in specific, young age groups.

In the transition analyses, only education years were used due to statistical criteria. Education years is also well defined and clear compared with education level that is dependent on national education system and it may be difficult to determine the actual distances between various level definitions.

## Supplementary data

Supplementary data are available at *PUBMED* online.

## Funding

The research is funded by Academy of Finland (projects 258711, 265869, 128002, 132944, 265977), Signe and Ane Gyllenberg Foundation, the Finnish Work Environment Fund, the EU New OSH ERA research programme and the Juho Vainio Foundation (L.P.-R.).

## References

1 Elovainio M, Ferrie JE, Singh-Manoux A *et al.* Socioeconomic differences in cardiometabolic factors: social causation or health-related selection? Evidence from the Whitehall II Cohort Study, 1991–2004. *Am J Epidemiol* 2011;**174**:779–89.

2 Lynch J, Smith GD, Hillemeier M *et al.* Income inequality, the psychosocial environment, and health: comparisons of wealthy nations. *Lancet* 2001;**358**:194–200.

3 Lleras-Muney A. The relationship between education and adult mortality in the United States. *Rev Econ Stud* 2005;**72**:189–221.

4 Spasojevic J. Effects of education on adult health in Sweden: results from a natural experiment. *Contrib Econ Anal Pol* 2010;**290**:179–99.

5 Johnson W, Kyvik KO, Mortensen EL *et al.* Does education confer a culture of healthy behavior? Smoking and drinking patterns in Danish twins. *Am J Epidemiol* 2011;**173**:55–63.

6 Kivimaki M, Leino-Arjas P, Luukkonen R *et al.* Work stress and risk of cardiovascular mortality: prospective cohort study of industrial employees. *BMJ* 2002;**325**:857.

7 Cassel J. The contribution of the social environment to host resistance: the Fourth Wade Hampton Frost Lecture. *Am J Epidemiol* 1976;**104**:107–23.

8 Cassel J. The contribution of the social environment to host resistance. The Fourth Wade Hampton Frost Lecture. 1976. *Am J Epidemiol* 1995;**141**:798–814; discussion 797.

9 Cobb S. Presidential Address-1976. Social support as a moderator of life stress. *Psychosom Med* 1976;**38**:300–14.

10 Gallo LC, de Los Monteros KE, Shivpuri S. Socioeconomic status and health: what is the role of reserve capacity? *Curr Dir Psychol Sci* 2009;**18**:269–74.

11 Uchino BN. *Social Support and Physical Health. Understanding the Health Consequences of Relationships.* New Haven: Yale University Press, 2004.

12 Twenge JM, Campbell WK. Self-esteem and socioeconomic status: a meta-analytic review. *Pers Soc Psychol Rev* 2002;**6**:59–71.

13 Orth U, Robins RW, Widaman KF. Life-span development of self-esteem and its effects on important life outcomes. *J Pers Soc Psychol* 2012;**102**:1271–88.

14 Lawrence D. *Enhancing Self-Esteem in the Classroom*. New York: Sage, 2006.

15 Emler N. *Self-Esteem. The Causes and Cost of Low Self-Esteem*. York, UK: York Publishing Services, 2001.

16 Orth U, Robins RW, Trzesniewski KH *et al.* Low self-esteem is a risk factor for depressive symptoms from young adulthood to old age. *J Abnorm Psychol* 2009;**118**:472–8.

17 Matthews KA, Raikkonen K, Gallo L *et al.* Association between socioeconomic status and metabolic syndrome in women: testing the reserve capacity model. *Health Psychol* 2008;**27**:576–83.

18 Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. *PLoS Med* 2010;**7**:e1000316.

19 Croezen S, Picavet HS, Haveman-Nies A *et al.* Do positive or negative experiences of social support relate to current and future health? Results from the Doetinchem Cohort Study. *BMC Public Health* 2012;**12**:65.

20 Pulkki-Raback L, Kivimaki M, Ahola K *et al.* Living alone and antidepressant medication use: a prospective study in a working-age population. *BMC Public Health* 2012;**12**:236.

21 Stringhini S, Berkman L, Dugravot A *et al.* Socioeconomic status, structural and functional measures of social support, and mortality: the British Whitehall II Cohort Study, 1985–2009. *Am J Epidemiol* 2012;**175**:1275–83.

22 Schlossberg NK. *Counseling Adults in Transition: Linking Practice With Theory.* New York: Springer Pub., 2005.

23 Karasek RA, Theorell T. *Healthy Work: Stress, Productivity and the Reconstruction of Working Life*. New York: Basic Books, 1990.

24 Kivimaki M, Nyberg ST, Batty GD *et al.* Job strain as a risk factor for coronary heart disease: a collaborative meta-analysis of individual participant data. *Lancet* 2012;**380**:1491–7.

25 Taris TW, Feij JA. Learning and strain among newcomers: a three-wave study on the effects of job demands and job control. *J Psychol* 2004;**138**:543–63.

26 Raitakari OT, Juonala M, Rönnemaa T *et al*. Cohort profile: the cardiovascular risk in young Finns study. *Int J Epidemiol* 2008;**37**:1–7.

27 Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav* 1997;**38**:21–37.

28 Idler EL, Russell LB, Davis D. Survival, functional limitations, and self-rated health in the NHANES I Epidemiologic Follow-up Study, 1992. First National Health and Nutrition Examination Survey. *Am J Epidemiol* 2000;**152**:874–83.

29 Rahkonen O, Arber S, Lahelma E. Health inequalities in early adulthood: a comparison of young men and women in Britain and Finland. *Soc Sci Med* 1995;**41**:163–71.

30 Coopersmith S. *The Antecedents of Self-Esteem*. New York: Consulting Psychologist Press, 1990.

31 Landsbergis PA, Schnall PL, Warren K *et al*. Association between ambulatory blood pressure and alternative formulations of job strain. *Scand J Work Environ Health* 1994;**20**:349–63.

32 Rosenstrom T, Jylha P, Robert Cloninger C *et al*. Temperament and character traits predict future burden of depression. *J Affect Disord* 2014;**158**:139–47.

33 van Buuren S, Groothuis-Oudshoornvan K. Mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011;**45**:1–67.

34 Jackson CH. Multi-state models for panel data: the msm package for R. *J Stat Softw* 2011;**38**:1–28.

35 Pickett KE, Wilkinson RG. Income inequality and health: a causal review. *Soc Sci Med* 2015;**128**:316–26.

36 Messias E, Eaton WW, Grooms AN. Economic grand rounds: income inequality and depression prevalence across the United States: an ecological study. *Psychiatr Serv* 2011;**62**:710–2.

37 Gallo LC, Bogart LM, Vranceanu A-M *et al*. Socioeconomic status, resources, psychological experiences, and emotional responses: a test of the reserve capacity model. *J Pers Soc Psychol* 2005;**88**: 386–99.

38 Kawachi I, Adler NE, Dow WH. Money, schooling, and health: mechanisms and causal evidence. *Ann N Y Acad Sci* 2010;**1186**: 56–68.

39 Manor O, Matthews S, Power C. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *Int J Epidemiol* 2000;**29**:149–57.