TITLE PAGE

TITLE

Expected or completed? Comparing two measures of education and their relationship with social inequalities in health among young adults

AUTHOR NAMES

Thierry Gagné, M.Sc[1], Adrian Ghenadenik, M.Sc[1,2], Martine Shareck, Ph.D[3], Katherine L. Frohlich, Ph.D [1]

AFFILIATIONS

1. Département de Médecine Sociale et Préventive, École de Santé Publique de l'Université de Montréal (ESPUM) & Institut de Recherche en Santé Publique de l'Université de Montréal (IRSPUM)

7101, Av. du Parc, Montreal, QC H3N 1X9

2. Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM)

850 Saint Denis Street, Montreal, QC H2X 0A9, Canada

3. London School of Hygiene and Tropical Medicine (LSHTM)

15–17 Tavistock Place, London WC1H 9SH, UK

CORRESPONDING AUTHOR

Thierry Gagné

7101 av du Parc, bureau 3139

Montreal, Quebec (Canada)

H3N 1X9

Telephone: 1.514.343.6411 ext. 4565

E-mail: thierry.gagne@umontreal.ca

TITLE

Expected or completed? Comparing two measures of educational attainment and their relationship with social inequalities in health among young adults

ABSTRACT

Background. Similarly to other age groups, there are significant social inequalities in health among young adults (YA). Education is thought to be the most appropriate indicator of YA socioeconomic status (SES), yet it is often in progress at that age and may not be representative of future achievement. Therefore, scholars have explored YA 'expected' education as a proxy of SES. However, no study has examined how it compares to the more common SES indicator, 'completed' education.

Methods. Using data from 1,457 YA surveyed twice over a two year period, we describe associations between participants' completed and expected education at baseline and completed education at follow-up. We then compare associations between these two measures and three health outcomes – smoking status, self-rated mental health, and participation in physical activity and sports – at baseline and follow-up using regression models.

Results. At baseline, half of the participants were imputed a higher 'expected' level than that 'completed' at that time. In regression models, 'expected' and 'completed' education were strongly associated with all outcomes and performed slightly differently in terms of effect size, statistical significance, and model fit.

Conclusions. 'Expected' education offers a good approximation of future achievement. More importantly, 'expected' and 'completed' education variables can be conceptualized as complementary indicators associated with inequalities in health in YA. Using both may help better understand social inequalities in health in YA.

MANUSCRIPT

Public health research is increasingly focusing on young adults as a distinct population of interest (IOM 2014; Stroud et al. 2015). Several important educational, social, and economic milestones occur during young adulthood (typically defined as the period between ages 18 and 25 years): pursuing higher education; engaging in full-time work; leaving the parental household; entering conjugal relationships; and having children (Cote and Bynner 2008). Increases in the prevalence and incidence of certain health outcomes and behaviours that can become established later in adulthood are also evidenced during this period (IOM 2014). For example, approximately 75% of mental disorders are diagnosed before 24 years of age, hours devoted to physical activity decrease significantly, overweight and obesity rates increase threefold in comparison to those of adolescents, and substance misuse often becomes established in young adulthood (IOM 2014). Moreover, smoking initiation rates are on the rise among young adults who also have the highest smoking prevalence and lowest cessation rates of all age groups (IOM 2014; Bonnie et al. 2007; Freedman et al. 2011).

As is the case in other age groups, there are significant social inequalities in health among young adults for outcomes including smoking (Kestila et al. 2006a; Caban-Martinez et al. 2011; O'Loughlin et al. 2014), physical activity and obesity (Muyle et al. 2009), self-rated health (Kestila et al. 2006b), substance abuse (Redonnet et al. 2012), and sexually-transmitted diseases (Harling et al. 2013). Different indicators of socioeconomic status (SES) have been used to examine these inequalities. For example, studies have documented socioeconomic differences in smoking among young adults based on their employment status, income, school enrolment, educational attainment, and parental education (Caban-Martinez et al. 2011; Lawrence et al. 2007; Dietz et al. 2013; Pampel et al. 2014). Although these indicators can be used to examine social inequalities in health among young adults, several scholars have underlined the need to use SES measures relevant to the age group of interest (Braveman

et al. 2005; Galobardes et al. 2006a; Galobardes et al. 2006b). Indicators used to measure SES in adolescents and adults, such as parental education for the former and income and occupational class for the latter (Galobardes et al. 2006a), may be less relevant to young adults given their growing independence during their transition towards adulthood.

Even though there are few explicit conceptual and empirical guidelines for the measurement of SES in young adults, those in place suggest that educational attainment may be the most appropriate indicator, at least in developed countries (Braveman et al. 2005; Galobardes et al. 2006a; Galobardes et al. 2006b). Education is thought to capture several mechanisms linking SES and health, as it promotes health-related knowledge, values, skills and preferences and provides future occupational opportunities and financial resources as well as psychosocial resources (e.g.: social support, social standing, and sense of control) that allow individuals to avoid unhealthy behaviours and successfully deal with stressors (Braveman et al. 2011; Pampel et al. 2010).

Nonetheless, there are some limitations to its use, chiefly because education is often not completed in a significant proportion of this population. For instance, in the province of Quebec, Canada, approximately 50% of young adults are enrolled in studies in any given year (Lavoie et al. 2010). Moreover, departures from a "standard" duration of studies (e.g.: in Quebec, five years for high school, three years for an undergraduate degree) are increasingly common because of part-time work, health issues, and maternity/paternity leave (Bowen et al. 2011). To overcome these limitations and acknowledge that young adults may not have completed their education, certain scholars have turned to using the highest level 'expected' to be attained as a measure of education (Kestila et al. 2006a; Kestila et al. 2006b; Shareck et al. 2014; Shareck et al. 2015; De Grande et al. 2015; Widome et al. 2013). With this approach, education is measured as the highest level between the level of studies completed and that in which students are enrolled at the time of survey. For example, someone who completed

high school and is currently pursuing a bachelor's degree would be coded as having "some university completed".

The use of this transformation is based on two hypothetical advantages. First, the use of 'expected education' has the potential to reduce SES misclassification that may arise due to the attribution of a lower educational attainment than that actually achieved later on. Second, from a conceptual standpoint, 'expected education' may better represent the meaning of education as an indicator of SES as it taps into young adults' current learning and aspirations as well as the physical and social environments that are being experienced during their ongoing studies. These aspirations and environments are associated with young adults' health behaviours and may mediate the influence of their socioeconomic background beyond their educational attainment at the time of measurement (Tyas and Pederson 1998; Pedersen and von Soest 2014). Nonetheless, even though the use of this indicator has grown over the last decade or so, to our knowledge no studies have empirically tested these hypotheses.

In keeping with the increasing number of studies exploring these two indicators, we propose to address these two knowledge gaps and examine how 'expected education' compares to 'completed education' in the context of social inequalities in health in young adults. We therefore asked the following two questions. The first is methodological: does 'expected education' accurately estimate later 'completed education'? The second question goes back to the above-mentioned public health research needs: does the assessment of social inequalities in health among young adults differ when using 'expected' compared to 'completed' education?

METHODS

Data

We analyzed baseline and two-year follow-up data from the Interdisciplinary Study of Inequalities in Smoking (ISIS), a cohort study established in 2011-2012 with the objective of better understanding the joint contribution of individual and neighborhood factors in shaping social inequalities in smoking among young adults (Frohlich et al. 2015). The target population was non-institutionalized individuals aged 18 to 25 years living in Montreal, Canada, who had resided at their current address for at least one year at the time of first contact. From an initial sample of 6,020 individuals randomly selected from the provincial health insurance program, 2,093 completed the questionnaire (baseline response rate = 38%). Two years later, 1,457 individuals took part in a second wave of data collection (follow-up response rate = 74%). Full details on the cohort sampling and survey procedures are available elsewhere (Frohlich et al. 2015). This study received ethics approval from the provincial information access committee (*Commission d'accès à l'information du Québec*) and the Université de Montréal's ethics board (*Comité d'éthique de la recherche en santé de la Faculté de Médecine*).

Measures

Education variables

'Highest educational level *completed*' was assessed by asking participants "What is the highest level of schooling that you have completed?", with possible answers ranging from 'No school' to 'Earned doctorate'. To measure 'highest educational level *expected* to be completed', we relied on responses to this question, as well as to the question "Are you currently a student (either full-time, part-time, or in an internship program)?" Students were subsequently asked to report the name of the institution they attended. We used the name of the institution reported in this latter question to establish the level taught at the institution, and compared it with their completed level of education. For non-students and students enrolled in an establishment that did not provide an increase in educational level over that

completed, the expected level was the same as the completed level. For students enrolled in an establishment that provided an educational level higher than that already completed, their expected level was coded as the level provided by the establishment in which they were enrolled. For each measure of education, four categories were created: 'Did not finish high school', 'High school completed', 'CEGEP completed' and 'University completed'. CEGEPs (*Collège d'enseignement général et professionel*) are post-secondary educational institutions which provide mandatory pre-university education or vocational training in Quebec, Canada. Details regarding the sources, question labels and value labels for education variables are presented in the Supplementary Material file.

Health outcomes

Based on their relevance to young adults (IOM 2014), three health outcomes were studied: *current smoking status, self-rated mental health* and *participation in physical activity and sports. Current smoking status* was assessed by asking respondents who had smoked at least one entire cigarette in their lifetime whether they currently smoked 'every day', 'occasionally' or 'never'. Those who smoked daily or occasionally were considered to be 'current smokers' while 'non smokers' consisted of never smokers and former smokers. *Self-rated mental health* was measured using the following question: 'Compared to other people your age, would you say that, in general, your mental health is:' with responses on a five-point Likert scale ranging from 'Poor' to 'Excellent'. This variable was dichotomized to compare those in good health ('Excellent', 'Very good' and 'Good') to those in less than good health ('Fair' and 'Poor'). *Participation in physical activity and sports* was assessed by asking respondents whether they regularly engaged in physical activity or sports (yes or no). Details regarding the sources, question labels and value labels for the three outcome variables are presented in the Supplementary Material file.

Analyses

We first used univariate statistics to describe the sample in terms of age, sex, completed and expected education and health outcomes at each time point. Descriptive statistics were then used to assess: (1) at baseline, how many young adults were expected to attain a level of education higher than that already completed; and (2) how many young adults attained their expected level of education two years later. We used a Cohen's kappa as an estimate of overall agreement between measures in cross-tabulations, with results ranging from 0.60 to 0.75 deemed to provide evidence of a satisfactory level of agreement (Landis and Koch 1977; Fleiss 1981).

Associations between each education variable and the likelihood of: (1) being a current smoker (vs. non-smoker); (2) reporting fair or poor mental health (vs. good, very good or excellent) and; (3) not regularly engaging in physical activity (PA) or sports (vs. regularly engaging in such activities) were examined using multivariate Poisson regression models with robust variance estimation. Poisson regression was chosen over logistic regression because it allows for the direct estimation of risk ratios in the form of prevalence ratios (PR) when the outcome is common (usually > 10%) as is the case for the smoking (> 20%) and participation in physical activities and sports (> 40%) outcomes (Barros and Hirakata, 2003; McNutt et al., 2003). Analyses were performed for dependent variables at baseline and at follow-up, in both cases controlling for sex and age at baseline. To compare results we examined point estimates and their 95% confidence intervals in addition to Bayesian information criterion (BIC) values, where lower values are indicative of better model fit. BIC differences between 2 and 6 are considered to provide 'positive' evidence of better model fit, while differences of 6 or higher are deemed to provide 'strong' evidence of better model fit (Raftery 1995). Because there were very few missing cases for each variable, analyses were done using listwise deletion. Descriptive analyses and regression analyses were performed in SPSS and MPlus respectively (IBM 2011; Muthén and Muthén 1998-2013).

RESULTS

1. Sample characteristics

Fifty-eight (3.9%) of the 1,457 participants for whom data was available at both time points reported inconsistencies on education variables (either having attained a lower or an unlikely higher level of education at follow-up than that reported at baseline) and were excluded from the analyses, for a final sample of 1,399 participants.

Descriptive statistics are presented in Table 1. At baseline, participants were on average 21 years old (SD = 2.3) with 58.8% of them being women. Most participants completed post-secondary studies: 41% completed CEGEP and 20% completed some university studies. Twenty percent of participants were smokers at each time point, 9% and 10% declared to be in fair or poor mental health at baseline and follow-up respectively, and 42% and 44% declared not participating in PA and sports at each time point.

Please insert Table 1 here.

2. Completed and expected levels of education

Table 2 presents the correspondence between participants' completed and expected levels of education at baseline. Out of the 1,399 participants, 1,015 (73%) were students. Among these, 713 (70%) were expected to complete a higher level of education than that completed at baseline. The Cohen's kappa value between education variables at baseline (T1) was 0.28.

Please insert Table 2 here.

Table 3 presents the correspondence between the two education variables at baseline and the level of education completed two years later. We compared participants' expected education at baseline to the level completed at follow-up (T2) by examining whether participants were found on the diagonal (indicating that their completed education two years later was 'correctly' estimated from baseline information) or below the diagonal (indicating that their completed education two years later was 'incorrectly' estimated). Analyses revealed that, using the 'expected' codification approach, 397 participants (56%) were attributed an expected level of education at T1 in line with that completed two years later at T2. On the other hand, 310 participants (44%) were attributed an expected level of education at T1 that was not reached two years later at T2. Bivariate tests (not shown) showed that students who did not complete the educational level expected at baseline were more often men ($p <$ 0.001). No significant differences were found with regard to age and the three health variables compared to other students. Cohen's kappa values revealed a small increase in agreement going from 0.53 (between education completed at baseline and completed at follow-up) to 0.61 (between education expected at baseline and completed at follow-up), indicating that the baseline 'expected' education variable provided a moderately better approximation of 'completed' education two years later than the baseline 'completed' education did.

Please insert Table 3 here.

3. Associations between education variables and health outcomes

Table 4 presents prevalence ratios for associations between completed or expected levels of education and each health outcome measured at baseline. Both education variables were significantly associated with all health outcomes. Compared to participants who had completed or were expected to complete some university, those who did not finish high school or were not expected to do so and those who completed only high school or were expected to do so had higher odds of smoking, of having poor self-

rated mental health and of not participating in PA or sports. One difference was found between measures of education when examining point estimates and their statistical significance, but none with regard to their 95% confidence intervals: those who completed CEGEP had significantly higher odds of having poor mental health in comparison to those with some university completed (PR = 2.09, 95%CI (1.15, 3.79)). There was, however, no statistically significant difference among those expected to finish CEGEP and those expected to have some university completed (PR = 1.18, 95%CI (0.77, 1.84)). A comparison of BIC values provided positive evidence of better fit for models predicting current smoking status using 'completed' education as the independent variable.

Please insert Table 4 here.

Prevalence ratios for associations between participants' level of education expected at baseline, completed at follow-up, and health outcomes measured at follow-up are shown in Table 5. In this analysis, education variables were not as systematically associated with the outcomes as in the cross-sectional setting. Examining the educational level completed at follow-up as predictor (the top portion of Table 5), participants who did not complete high school or who only completed high school had higher odds of smoking and not participating in PA in comparison to participants who had some university completed. Using expected educational attainment measured at baseline, participants who were not expected to finish high school or CEGEP also had higher odds of smoking and of not participating in PA and sports at follow-up. We found again one difference in point estimates with regard to statistical significance but no differences in 95% confidence intervals: participants who were expected to have their CEGEP completed had higher odds of not participating in PA or sports compared to participants with some university completed or expected to be completed (PR = 1.16, 95%CI (1.00, 1.34)) whereas there was no significant difference between those who completed their CEGEP and those who completed some university (PR = 1.14, 95%CI (0.97, 1.34)). A comparison of

BIC indices provided strong evidence of a better model fit for models using the 'expected' variable to predict current smoking status, strong evidence of better model fit using the 'completed' variable to predict self-rated mental health and positive evidence of better model fit using the 'expected' variable to predict participation in physical activity and sports.

Please insert Table 5 here.

DISCUSSION

Research on social inequalities in health among young adults faces certain unique challenges. One of these is the usefulness of indicators commonly used to operationalize their socioeconomic status. To advance knowledge in this area, we compared a common measure of education, completed education, with an alternative, expected education, which acknowledges that many young adults may still be pursuing studies. Specifically, we examined whether: (1) 'expected education' accurately estimated later 'completed education' and (2) the assessment of social inequalities in health among young adults differed when using these two measures of education.

'Expected' educational attainment as a proxy of future achievement

A little over half of the sample (i.e. 70% of the 73% who were students) were expected to attain a level of education higher than that completed at baseline. This is likely a reasonable expectation in an urban setting such as Montreal, home to four universities and more students than the provincial average. When compared to the level of education completed two years later, we found that 'expected education' provided a moderately better approximation of future educational achievement than 'completed education' at baseline. This suggests that the 'expected education' measure provides a reasonable estimate of future educational achievement.

Nonetheless, the use of this variable requires the consideration of certain issues. First, many young adults may pursue continuing education or elective classes that do not lead to a diploma, or new diplomas that do not confer a higher level of education. Second, it must be acknowledged that a considerable proportion of young adults will not graduate, even if it was their initial intention (e.g.: in Quebec, Canada, post-secondary graduation rates are 70%) (CRÉPUQ 2006). Finally, given the potential for misclassification, we suggest that researchers exercise care in attributing higher levels of education based on 'expected' education, notably when the establishment in which students were enrolled is not known and attribution of a higher level of education is based solely on student status (i.e. enrolled in studies or not).

'Expected' education as an alternative measure for studying social inequalities in health among young adults

Beyond its accuracy as a proxy for future educational achievement, we suggested that young adults' expected education encompassed beyond their achieved level their learning, educational aspirations and the physical and social environments in which they were studying. In the cross-sectional and prospective scenarios, we found noticeable differences between the two measures, but none allowed us to identify a distinct pattern in their ability to predict the chosen health outcomes. There is reason to believe that a measure of young adults' completed education may not fully capture the sociocultural, financial and psychological resources that it aims to operationalize, and that using the complementary measure of expected educational attainment can contribute to do so. Measures of expected educational attainment may better reflect young adults' social aspirations and current social milieu, and in turn, may allow for a more accurate measurement of social inequalities in health (van Soest & Pedersen, 2014). Scholars have previously advocated for such a perspective by identifying different health-promoting mechanisms based on what is obtained and what is concurrently incorporated with regard to education (Abel 2008, Gagné et al. 2015).

In our study, we found that both education variables were strongly associated with smoking, self-rated mental health, and participation in physical activity and sports. For instance, participants who have not finished high school had more than a three-fold risk of being smokers and reporting poor mental health in comparison to those who continued onwards to university. These results are in line with a rapidly growing literature highlighting young adulthood as an important target group to tackle health inequalities (IOM, 2014; Muyle et al. 2009; Redonnet et al. 2012; Pampel et al. 2014). When comparing results, we found one important difference between educational variables with regard to self-rated mental health in the cross-sectional scenario: those in lower educational categories had a much higher risk of reporting poor mental health when we used those who had completed some university as the reference category instead of those who were expected to in the near future. This suggests that participants who were undertaking undergraduate studies (i.e. who had 'CEGEP' as their completed education but 'some university completed' as their expected education) had a higher risk of reporting poor mental health. Other demographic, socioeconomic and psychosocial characteristics might explain these differences given that college students normally show comparable levels of mental health to most non-college-attending students (Blanco et al., 2008; Kovess-Masfety et al., 2016). However, this is a perfect example to help us understand that educational inequalities in health might be conditional on the transitional stages (in this case, ongoing studies) that young adults experience towards adulthood.

Limitations

This study has three limitations that should be discussed. First, time between measurements should be considered a potential limitation chiefly due to the time required to obtain certain diplomas. In some cases the 2-year follow-up may have been insufficient to achieve the expected educational level: among the 85 CEGEP and 199 university-level students that were imputed a higher level of education,

those who were in their first year of studies at T1 would have been in their third year of studies at follow-up and therefore would not have been able to yet complete their degree. Because 85% of these 284 participants were still studying two years later, a later time point would have allowed for a better assessment of their actual 'completed' education. Second, whereas results may be representative of urban areas similar to Montreal, Canada (i.e regions in developed countries with similar education systems and outcomes), they may not be generalizable to other regions with substantially different contexts. Third, we adopted a parsimonious approach to modelling and addressed confounding by controlling only for age and sex. It is possible that omitted variables might have influenced our results relative to the association between educational measures and health outcomes.

Conclusion

There is a large scholarship dedicated to understanding and tackling health inequalities in public health, and more work in the operationalization and measurement of socioeconomic characteristics is needed to support it. To our knowledge, this study is the first to inquire into the methodological and conceptual assumptions associated with using 'expected' education to examine social inequalities in health among young adults in comparison to the ubiquitous operationalization of education, 'completed' education. Our findings suggest that the use of 'expected' education as a measure of SES can be a valuable addition to the study of social inequalities in health in young adults, by providing a more reliable appreciation of adult education achievement and tapping into the current aspirations and school environments that young adults continue to experience during their ongoing studies. This is of particular importance in young adults given the many different transitions experienced during this life period and their potential impact on health behaviours and outcomes. We therefore recommend that researchers not only use completed education when the other is available, but that they use these two measures in conjunction whenever possible and report complete results so that readers might compare

them. This can be done in projects explicitly examining the mechanisms linking educational achievement and health but also in many others when performing sensitivity analyses.

REFERENCES

Abel, T. (2008). Cultural capital and social inequality in health. J Epidemiol Community Health, 62(7), e13. DOI: 10.1136/jech.2007.066159

Barros, A. J., Hirakata, V. N. (2003). Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol, 3, 21. DOI: 10.1186/1471-2288-3-21

Blanco, C., Okuda, M., Wright, C., et al. (2008). Mental Health of College Students and Their Non-College-Attending Peers. JAMA Psychiatry: 65(12): 149-1437.

Bonnie, R. J., Stratton, K., Wallace, I. (2007). Ending the Tobacco Problem: A Blueprint for the Nation: Committee on Reducing Tobacco Use: Strategies, Barriers, and Consequences, Institute of Medicine.

Bourdieu, P. (1986). The Forms of Capital (R. Nice, Trans.). In J. E. Richardson (Ed.), Handbook of Theory of Research for the Sociology of Education (pp. 241-258.): Greenword Press.

Bowen, W. G., Chingos, M. M., McPherson, M. S. (2011). Crossing the Finish Line: Completing College at America's Public Universities: Princeton University Press.

Braveman, P., Egerter, S., Williams, D. R. (2011). The social determinants of health: coming of age. Annu Rev Public Health, 32, 381-398. DOI: 10.1146/annurev-publhealth-031210-101218

Braveman, P. A., Cubbin, C., Egerter, S., et al. (2005). Socioeconomic status in health research: one size does not fit all. JAMA, 294(22), 2879-2888. DOI: 10.1001/jama.294.22.2879

Caban-Martinez, A. J., Lee, D. J., Goodman, E., et al. (2011). Health indicators among unemployed and employed young adults. J Occup Environ Med, 53(2), 196-203. DOI: 10.1097/JOM.0b013e318209915e

Cote, J., Bynner, J. M. (2008). Changes in the Transition to Adulthood in the UK and Canada: The Role of Structure and Agency in Emerging Adulthood. J Youth Stud, 11(3), 251-268.

CRÉPUQ. (2006). Le système universitaire québécois : données et indicateurs. (French) Accessed online Jan 25th 2016. URL: http://www.crepuq.qc.ca/IMG/pdf/indicateurs-2.pdf

De Grande, H., Vandenheede, H., Deboosere, P. (2015). Educational Inequalities in the Transition to Adulthood in Belgium: The Impact of Intergenerational Mobility on Young-Adult Mortality in 2001-2009. PLOS One. DOI: 10.1371/journal.pone.0142104

Dietz, N. A., Sly, D. F., Lee, D. J., et al. (2013). Correlates of smoking among young adults: the role of lifestyle, attitudes/beliefs, demographics, and exposure to anti-tobacco media messaging. Drug Alcohol Depend, 130(1-3), 115-121. DOI: 10.1016/j.drugalcdep.2012.10.019

Fleiss, J. L. (1981). Statistical methods for rates and proportions (2nd ed.). New York: John Wiley.

Freedman, K. S., Nelson, N. M., Feldman, L. L. (2011). Smoking Initiation Among Young Adults in the United States and Canada, 1998-2010: A Systematic Review. Prev Chronic Dis. DOI: 10.5888/pcd9.110037

Frohlich, K. L., Shareck, M., Vallee, J., et al. (2015). Cohort Profile: The Interdisciplinary Study of Inequalities in Smoking (ISIS). Int J Epidemiol, 1-14. DOI: 10.1093/ije/dyv036

Gagné, T., Frohlich, K. L., Abel, T. (2015). Cultural capital and smoking in young adults: applying new indicators to explore social inequalities in health behaviour. Eur J Public Health, 25(5), 818-823. DOI: 10.1093/eurpub/ckv069

Galobardes, B., Shaw, M., Lawlor, D. A., et al. (2006a). Indicators of socioeconomic position (part 1). J Epidemiol Community Health, 60(1), 7-12. DOI: 10.1136/jech.2004.023531

Galobardes, B., Shaw, M., Lawlor, D. A., et al. (2006b). Indicators of socioeconomic position (part 2). J Epidemiol Community Health, 60(2), 95-101. DOI: 10.1136/jech.2004.028092

Harling, G., Subramanian, S., Barnighausen, T., et al. (2013). Socioeconomic disparities in sexually transmitted infections among young adults in the United States: examining the interaction between income and race/ethnicity. Sex Transm Dis, 40(7), 575-581. DOI: 10.1097/OLQ.0b013e31829529cf

IBM. (2011). IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM.

IOM (Institute of Medicine) and NRC (National Research Council). (2014). Improving the health, safety, and well-being of young adults: Workshop summary. Washington, DC: The National Academies Press.

Kestila, L., Koskinen, S., Martelin, T., et al. (2006a) Influence of parental education, childhood adversities, and current living conditions on daily smoking in early adulthood. Eur J Public Health, 16(6), 617-626. DOI: 10.1093/eurpub/ckl054

Kestila, L., Koskinen, S., Martelin, T., et al. (2006b) Determinants of health in early adulthood: what is the role of parental education, childhood adversities and own education? Eur J Public Health, 16(3), 306-315. DOI: 10.1093/eurpub/cki164

Kovess-Masfety, V., Leray, E., Denis, L., Husky, M., Pitrou, I., Bodeau-Livinec, F. (2016). Mental health of college students and their non-college-attending peers: results from a large French cross-sectional survey. BMC Psychol, 4, 20. DOI : 10.1186/s40359-016-0124-5.

Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Lavoie, K., Bonnelli, H., Gauthier, J., et al. (2010). Portrait de santé des jeunes québécois âgés de 15 à 24 ans. Fédération des Cégeps. 60 pages.

Lawrence, D., Fagan, P., Backinger, C. L., et al. (2007). Cigarette smoking patterns among young adults aged 18-24 years in the United States. Nicotine Tob Res, 9(6), 687-697. DOI: 10.1080/14622200701365319

McNutt, L.-A., Wu, C., Xue, X., Hafner, J.-P. (2003). Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. Am J Epidemiol, 157(10), 940-943.

Mulye, T. P., Park, M. J., Nelson, C. D., et al. (2009). Trends in adolescent and young adult health in the United States. J Adolesc Health, 45(1), 8-24. DOI: 10.1016/j.jadohealth.2009.03.013

Muthén, L. K., Muthén, B. O. (1998-2013). Mplus User's Guide. Seventh Edition. Los Angeles, CA Muthén & Muthén.

O'Loughlin, J. L., Dugas, E. N., O'Loughlin, E. K., et al. (2014). Incidence and determinants of cigarette smoking initiation in young adults. J Adolesc Health, 54(1), 26-32. DOI: 10.1016/j.jadohealth.2013.07.009

Pampel, F. C., Krueger, P. M., Denney, J. T. (2010). Socioeconomic Disparities in Health Behaviors. Annu Rev Sociol, 36, 349-370. DOI: 10.1146/annurev.soc.012809.102529

Pampel, F. C., Mollborn, S., Lawrence, E. M. (2014). Life course transitions in early adulthood and SES disparities in tobacco use. Soc Sci Res, 43, 45-59. DOI: 10.1016/j.ssresearch.2013.08.005

Pedersen, W., von Soest, T. (2014). Tobacco use among Norwegian adolescents: from cigarettes to snus. Addiction, 109(7), 1154-62. DOI: 10.1111/add.12509

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. Sociol Methodol, 25, 111-163.

Redonnet, B., Chollet, A., Fombonne, E., et al. (2012). Tobacco, alcohol, cannabis and other illegal drug use among young adults: the socioeconomic context. Drug Alcohol Depend, 121(3), 231-239. DOI: 10.1016/j.drugalcdep.2011.09.002

Shareck, M., Kestens, Y., Frohlich, K. L. (2014). Moving beyond the residential neighborhood to explore social inequalities in exposure to area-level disadvantage: Results from the Interdisciplinary Study on Inequalities in Smoking. Soc Sci Med, 108: 106-114. DOI: 10.1016/j.socscimed.2014.02.044

Shareck, M., Kestens, Y., Vallee, J., et al. (2015). The added value of accounting for activity space when examining the association between tobacco retailer availability and smoking among young adults. Tobacco Control. DOI: 10.1136/tobaccocontrol-2014-052194

Stroud, C., Walker, L. R., Davis, M., et al. (2015). Investing in the health and well-being of young adults. J Adolesc Health, 56(2), 127-129. DOI: 10.1016/j.jadohealth.2014.11.012

Tyas, S. L., Pederson, L. L. (1998). Psychosocial factors related to adolescent smoking: a critical review of the literature. Tobacco Control, 7(4), 409-420.

Widome, R., Wall, M. M., Laska, M. N., et al. (2013). Adolescence to young adulthood: when socioeconomic disparities in substance use emerge. Subst Use Misuse, 48(14), 1522-1529. DOI: 10.3109/10826084.2013.800885

TABLE 1

Descriptive statistics for 1,399 young adults participating in the *Interdisciplinary Study of Inequalities in Smoking*, Montreal, Canada, 2011-2014

Variable	Baseline (T1) *n* (%)	Two-year follow-up (T2) *n* (%)
Sex		
Woman	822 (58.8)	822 (58.8)
Man	577 (41.2)	577 (41.2)
Age		
mean (SD)	21.4 (2.3)	23.4 (2.3)
Completed education		
Less than high school	83 (5.9)	52 (3.7)
High school	456 (32.6)	221 (15.8)
CEGEP	575 (41.1)	684 (48.9)
Some University	281 (20.1)	437 (31.2)
Missing	4 (0.3)	5 (0.4)
Expected education		
Less than high school	49 (3.5)	---
High school	151 (10.8)	---
CEGEP	549 (39.2)	---
Some University	646 (46.2)	---
Missing	4 (0.3)	---
Currently studying	1,105 (72.6)	777 (55.5)
Missing	9 (0.6)	13 (0.9)
Current smokers	286 (20.4)	279 (19.9)
Missing	5 (0.4)	5 (0.4)
Self-rated mental health (fair or poor)	123 (8.8)	146 (10.4)
Missing	10 (0.7)	9 (0.6)
Not participating in PA or sports	581 (41.5)	681 (44.2)
Missing	21 (1.5)	24 (1.7)

PA = Physical activity, SD = Standard deviation

TABLE 2

Correspondence between education variables at baseline (T1)

Level of education *expected* at baseline (T1)

Cohen's Kappa = 0.28 (n = 1,395)

Level of education *completed* at baseline (T1)	Less than High School *n* (%)	High School *n* (%)	CEGEP *n* (%)	University *n* (%)	Total *n* (%)
Less than High School	49 (3.5)	29 (2.1)	5 (0.4)	0	83 (5.9)
High School	0	122 (8.7)	314 (22.5)	20 (1.4)	456 (32.7)
CEGEP	0	0	230 (16.5)	345 (24.7)	575 (41.2)
University	0	0	0	281 (20.1)	281 (20.1)

The light gray area (n = 682) on the diagonal represents participants who were not imputed a higher level of education because they were not studying or because the institution where their studies took place did not award a higher level of education. The darker grey area (n = 713) above the diagonal represents participants who were imputed a higher level of education based on their student status and the institution where they were studying.

TABLE 3

Comparison of education variables at baseline (T1) regarding their correspondence with completed education at follow-up (T2)

Level of education *completed* at follow-up (T2)
Cohen's Kappa = 0.53 (*n* = 1390)

Level of education *completed* at baseline (T1)	Less than High School *n* (%)	High School *n* (%)	CEGEP *n* (%)	University *n* (%)	Total *n* (%)
Less than High School	50 (3.6)	27 (1.9)	4 (0.3)	0	81 (5.8)
High School	0	194 (14.0)	256 (18.4)	6 (0.4)	456 (32.8)
CEGEP	0	0	422 (30.4)	151 (10.9)	573 (41.2)
University	0	0	0	280 (20.1)	280 (20.1)

Level of education *completed* at follow-up (T2)
Cohen's Kappa = 0.61 (n = 1390)

Level of education *expected* at baseline (T1)	Less than High School *n* (%)	High School *n* (%)	CEGEP *n* (%)	University *n* (%)	Total *n* (%)
Less than High School	36 (2.6)	11 (0.8)	2 (0.1)	0	49 (3.5)
High School	13 (0.9)	113 (8.1)	24 (1.7)	0	150 (10.8)
CEGEP	1 (0.1)	85 (6.1)	457 (32.9)	4 (0.3)	547 (39.4)
University	0	12 (0.9)	199 (14.3)	433 (31.1)	644 (46.3)

The light gray area on the diagonal represents participants who were correctly classified by the expected variable based on their completed education two years later; the darker gray area (n = 310) below the diagonal represents participants who expected a higher level of education that was not completed two years later.

TABLE 4

Associations between health outcomes at baseline (T1) and education variables

	Current smoking		Fair or poor mental health		Not participating in PA or sports	
	PR	95% CI	PR	95% CI	PR	95% CI
Educational level *completed* at baseline (T1)						
Less than high school	**3.37**	**2.35, 4.84**	**3.53**	**1.60, 7.77**	**1.44**	**1.10, 1.88**
High school	**1.68**	**1.18, 2.39**	**2.73**	**1.35, 5.53**	**1.29**	**1.04, 1.59**
CEGEP	1.12	0.81, 1.56	<u>**2.09**</u>	<u>**1.15, 3.79**</u>	1.09	0.91, 1.32
Some University	ref.		ref.		ref.	
BIC	1480.038*		868.976		2195.183	
Educational level *expected* at baseline (T1)						
Less than high school	**3.33**	**2.45, 4.53**	**2.51**	**1.26, 5.01**	**1.41**	**1.08, 1.86**
High school	**1.76**	**1.29, 2.41**	**2.01**	**1.18, 3.42**	**1.30**	**1.08, 1.58**
CEGEP	1.24	0.96, 1.60	1.18	0.77, 1.84	1.04	0.89, 1.20
Some University	ref.		ref.		ref.	
BIC	1484.757		870.397		2194.745	
BIC Δ	4.72		1.42		0.44	

Multivariate Poisson regression with a robust variance estimation; adjusted for age and sex at baseline; bolded regression coefficients are significant at the α = 0.05 level; models with the lowest BIC value are considered to better fit the data; underlined regression estimates are considered statistically significant for only one of the two education variables.

PA = Physical activity

PR = Prevalence ratio, CI = Confidence interval, BIC = Bayesian Information Criteria

BIC Δ ≥ 2 = positive evidence of better model fit

TABLE 5

Associations between health outcomes at follow-up (T2) and education variables

	Current smoking		Fair or poor mental health		Not participating in PA or sports	
	PR	95% CI	PR	95% CI	PR	95% CI
Educational level *completed* at follow-up (T2)						
Less than high school	**3.33**	**2.37, 4.68**	1.62	0.78, 3.37	**1.58**	**1.22, 2.05**
High school	**1.57**	**1.13, 2.19**	1.40	0.85, 2.32	**1.28**	**1.06, 1.55**
CEGEP	1.21	0.91, 1.61	0.94	0.61, 1.43	1.14	0.97, 1.34
Some University	ref.		ref.		ref.	
BIC	1462.947		979.437**		2248.620	
Educational level *expected* at baseline (T1)						
Less than high school	**2.92**	**2.11, 4.06**	1.61	0.81, 3.21	**1.38**	**1.05, 1.83**
High school	**1.60**	**1.16, 2.19**	1.07	0.63, 1.82	**1.41**	**1.18, 1.68**
CEGEP	1.12	0.86, 1.45	0.93	0.63, 1.38	<u>**1.16**</u>	<u>**1.00, 1.34**</u>
Some University	ref.		ref.		ref.	
BIC	1456.016**		987.382		2245.580*	
BIC Δ	6.93		7.95		3.04	

Multivariate Poisson regression with a robust variance estimation; adjusted for age and sex at baseline; bolded regression coefficients are significant at the α = 0.05 level; models with a lower BIC value are considered to better fit the data; underlined regression estimates are considered statistically significant for only one of the two education variables.
PA = Physical activity
PR = Prevalence ratio, CI = Confidence interval, BIC = Bayesian Information Criteria
* BIC Δ ≥2 = positive evidence of better model fit; ** BIC Δ ≥6 = strong evidence of better model fit

SUPPLEMENTARY MATERIAL

QUESTION AND VALUE LABELS OF VARIABLES USED IN THE STUDY

Variable	Source	Question labels	Value labels (code)
Age	Taken from the provincial universal health insurance program database	---	18-25
Sex	Taken from the provincial universal health insurance program database	---	M (0) F (1)
Education	Questionnaire from a provincial health agency (Santé Québec)	"What is the highest level of education you have completed?"	1. No school, or only kindergarten (1) 2. Elementary school (1) 3. Secondary 4 or less (10th grade or less) (1) 4. Secondary 5 (11th grade) (2) 5. Diploma or certificate of studies in a technical program at a CEGEP, a trade school, a commercial or private college, a technical institute, or a nursing school (3) 6. Diploma or certificate of studies in a general program at a CEGEP (3) 7. University undergraduate certificate (4) 8. Bachelor's degree (4) 9. Degree in medicine, dentistry, veterinary medicine, optometry or chiropracty (4) 10. University graduate certificate (4) 11. Master's degree (4) 12. Earned doctorate (4)
Student status	Developed by the research team	"Are you currently a student (either full-time, part-time or an internship program?"	Yes (1) No (0)
Education establishment	Developed by the research team	"What is the name of the institution you attend for your studies, including the campus and the building (if these apply)?"	Name of the institution: (string) Name of the campus: (string) Name of the building: (string)
	Developed by the research team	"What is the address of this study location? If you are studying at home or doing a distance learning program, please indicate it here."	Number and/or street name: (string) Intersection: (string) Closest landmark: (string) Neighborhood: (string) City: (string)
Current smoking	Questionnaire from	"Currently, do you smoke	1. Every day (1)

	the federal survey agency (Statistics Canada)	cigarettes every day, sometimes or never?"	2. Sometimes (1) 3. Never (0)
Self-rated mental health	Questionnaire from the federal survey agency (Statistics Canada)	"Compared to other people your age, would you say that, in general, your mental health is:"	1. Excellent (0) 2. Very good (0) 3. Good (0) 4. Fair (1) 5. Poor (1)
Participation to physical activity and sports	Developed by the research team	"Do you regularly engage in physical activity or sports?"	Yes (1) No (0)