QUANTITATIVE RESEARCH

What predictors matter: Risk factors for late adolescent outcomes

Elizabeth Wall-Wieler, MSc,[1] Leslie L. Roos, PhD,[1] Dan G. Chateau, PhD,[1] Laura C. Rosella, PhD[2]

ABSTRACT

OBJECTIVES: A life course approach and linked Manitoba data from birth to age 18 were used to facilitate comparisons of two important outcomes: high school graduation and Attention-Deficit/Hyperactivity Disorder (ADHD). With a common set of variables, we sought to answer the following questions: Do the measures predicting high school graduation differ from those that predict ADHD? Which factors are most important? How well do the models fit each outcome?

METHODS: Administrative data from the Population Health Research Data Repository at the Manitoba Centre for Health Policy were used to conduct one of the strongest observational designs: multilevel modelling of large population ($n = 62{,}739$) and sibling ($n = 29{,}444$) samples. Variables included are neighbourhood characteristics, measures of family stability, and mental and physical health conditions in childhood and adolescence.

RESULTS: The adverse childhood experiences important for each outcome differ. While family instability and economic adversity more strongly affect failing to graduate from high school, adverse health events in childhood and early adolescence have a greater effect on late adolescent ADHD. The variables included in the model provided excellent accuracy and discrimination.

CONCLUSION: These results offer insights on the role of several family and social variables and can serve as the basis for reliable, valid prediction tools that can identify high-risk individuals. Applying such a tool at the population level would provide insight into the future burden of these outcomes in an entire region or nation and further quantify the burden of risk in the population.

KEY WORDS: Attention deficit disorder with hyperactivity; education; risk; longitudinal studies

La traduction du résumé se trouve à la fin de l'article.

Can J Public Health 2016;107(1):e16–e22
doi: 10.17269/CJPH.107.5156

A life course approach and linked Manitoba data from birth to age 18 were used to facilitate comparing two important outcomes: high school graduation (an indicator of language/cognitive development) and Attention-Deficit/Hyperactivity Disorder (ADHD) (an indicator of social/emotional health).[1] A substantial literature exists on the biological, health, family and socio-economic determinants of child development. Our approach combines a large number of cases, longitudinal epidemiological data, advanced statistics, and sibling analysis to better compare a number of important risk factors; each variable can be assessed simultaneously controlling for the other measures.[2–4] Using administrative data from the early life course arguably provides better measurement than the usual reliance on retrospective surveys.[3]

Categorizing variables likely to affect child development is inevitably subjective. In this study, biological measures include sex and birth weight (an indicator of infant health). Such health conditions as ADHD, conduct disorders, bipolar disorder, asthma, major injuries, and other major health conditions between ages 0 to 13 are noted as dichotomous (yes/no) variables. At least one "biological" measure reflects social conditions: mother's age at first birth. Classified as family variables are family size/birth order, family structure history (providing details on parental marital status, divorce, remarriage and death), residential mobility, and the presence of serious family problems (reflected in the involvement of the provincial Child and Family Services). Socio-economic measures include an index of neighbourhood socio-economic status (SES) and receipt of income assistance (welfare). The longitudinal data permit looking at earlier manifestations of the late adolescent outcomes: Grade 9 Achievement and ADHD diagnosed in the 0–13 age range.

The impact of adverse childhood experiences – including poor health, family instability and economic adversity – on education and health has been well documented.[5–7] Survey versions of many of the measures in this paper have been used previously.[6] Almost all of the literature is based on research using a smaller *N* and fewer variables; the number of simultaneous controls has been smaller in previous studies. The research on high school graduation and ADHD typically notes such common issues as family instability

Author Affiliations

1. Manitoba Centre for Health Policy, Department of Community Health Sciences, Faculty of Health Sciences, College of Medicine, University of Manitoba, Winnipeg, MB
2. Dalla Lana School of Public Health, University of Toronto, Toronto, ON

Correspondence: Elizabeth Wall-Wieler, Manitoba Centre for Health Policy, Department of Community Health Sciences, Faculty of Health Sciences, College of Medicine, University of Manitoba, 408-727 McDermot Avenue, Winnipeg, MB R3E 3P5, Tel: 204-330-1776, E-mail: wallwiee@myumanitoba.ca

Acknowledgements: This research has been supported by the Social Sciences and Humanities Research Council, the Canadian Institutes of Health Research, and the Canadian Institute for Advanced Research. Data used in this study are from the Population Health Research Data Repository housed at the Manitoba Centre for Health Policy, University of Manitoba and were derived from data provided by Manitoba Health and Manitoba Education under project #2007/2008-22 and project #2007/2008-20C.

Disclaimer: The results and conclusions are those of the authors and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, or other data providers is intended or should be inferred.

Conflict of Interest: None to declare.

and socio-economic disadvantage without allowing comparison across domains.[7–9] Given the variety of work on high school graduation and ADHD, it makes sense to compare the effects of a uniform set of predictors on these two outcomes. With a common set of variables, we sought to answer the following questions: Do the measures predicting high school graduation differ from those that predict ADHD? Which factors are most important? How well do the models fit each outcome? We use multilevel modelling of large population and sibling samples to answer these questions.[10]

METHODS

Data overview

The Population Health Research Data Repository at the Manitoba Centre for Health Policy provided routinely collected individual information from all provincial residents. A research registry identifying every individual resident in the province, with arrival and departure dates, deaths and births, was created from the provincial health registry and coordinated with Vital Statistics files. Linked data from several ministries provided information on: individual-level health conditions (Manitoba Health, Healthy Living and Seniors), educational achievement (Manitoba Education), economic adversity (Manitoba Jobs and the Economy), and challenged parents (Family Services). Census data on neighbourhood-level household income, education, unemployment and household family structure were also incorporated.

Children were linked to mothers using hospital birth record information, available in essentially all cases.[7] Both the mother's encrypted Personal Health Identification Number and the family registration number were used to define siblings.

Cohort selection

Our population analysis includes those born in Manitoba, Canada between 1984 and 1989 (inclusive), staying in the province up to June of their 18th year, and included in the Repository. Loss to follow-up (from birth to age 18) was roughly 23% from the original cohort, covering over 98% of those born in Manitoba; out-of-province migration is primarily responsible for this attrition.[11] The original cohort totalled 72,634. Children were excluded for intellectual disability ($n = 1,164$); missing data on birth weight ($n = 1,373$); suppressed postal code for children placed with Manitoba Child and Family Services (or Public Trustee) ($n = 218$); missing family structure history ($n = 540$); missing Grade 9 Achievement Index score ($n = 16$); and missing score on the Socio-economic Factor Index ($n = 270$). Children residing in Manitoba's Rural North ($n = 6,920$) were also excluded; many of the schools run by First Nations (Indian) bands in the Rural North do not report graduation figures to the Ministry of Education. Home-schooled children (who do not officially graduate from high school) and those at the few band schools outside the Rural North ($n = 2,449$) were also not included. Supplementary Appendix A (see ARTICLE TOOLS section on journal site) has additional information on exclusions of individuals from the Rural North and band schools. After exclusions, this sample included 62,739 cohort members.

Sibling analyses ($N = 29,444$) compare genetically similar individuals growing up in similar environments, correcting for error from omitted parental variables.[12] With siblings born within six years of each other, the family environment will have had a relatively short time to change (in possibly unmeasured ways). Selecting twins in families with twins and other siblings ensured sufficient sample size.

Variables

The effect of a series of family- and individual-level covariates on failure to graduate from high school (within four years of entering grade 9) and having an ADHD diagnosis in late adolescence (ages 14–18) is examined (this is not necessarily the first diagnosis of ADHD, but a diagnosis occurred in this time period). Variables are displayed in Table 1. The neighbourhood of the oldest child in the cohort at age 17 is used to determine the Socio-economic Factor Index (SEFI) score and rural residency; all other siblings are likely to live in this neighbourhood as well. Neighbourhood SEFI is closely related to family SES as neighbourhoods are based on six-digit postal codes, which rarely include more than 600 individuals. Some variables examining family situations (such as family structure changes, residential mobility and social assistance usage) are classified as individual-level variables to facilitate sibling analysis. We are examining the presence of these predictors before age 14; one of these events could occur when one sibling is younger than 14 and one older. Several variables are defined by The Johns Hopkins University Adjusted Clinical Group software, which groups 14,000 ICD codes into 32 aggregated diagnostic groups to simplify the large number of possible health measures.[13] While the presence of family structure changes, residential mobility and health conditions between birth and age 13 is noted, involvement with Child and Family Services (CFS) and income assistance can only be examined in children ages 9 to 13. Data on these covariates were added to the research repository much later than those of the other covariates. Validity issues surrounding mental illness diagnoses using administrative data are discussed elsewhere.[14]

Procedure

Observational studies are often criticized for an exaggerated emphasis on particular variables.[15] Stronger research designs and statistical innovations can help to better identify risk factors having causal effects.[16] Our population and sibling analyses use multilevel modelling to both a) handle the lack of independence among observations which violates a key assumption of most parametric procedures, and b) clarify "confounding variables at both within- and between-group levels".[17] In the multilevel model framework, neighbourhood and family information were placed on the same (second) level, and individual information on the first level. Neighbourhood information was placed at the second (family) level because neighbourhood is not examined at a specific point in time. The neighbourhood of the oldest child at age 17 depends on his/her birth year, and neighbourhoods change. Uniform access to Canadian health insurance further corrected for the possibility of differential insurance coverage within families;[11] the use of objective measures instead of self-reported or retrospective measures strengthens generalizability. Finally, although siblings share many experiences, these experiences may occur at different times in the developmental process.[18] This variation allows using

Table 1. Characteristics of population sample

	Frequencies/percent of population*	Sibling with different outcomes/predictors†
	n (%)	n
Outcomes		
Failure to graduate high school	12,540 (19.99)	2746
ADHD, 14–18	1999 (3.19)	736
Predictors		
Family level		
Mean standardized SEFI2 score (SD)	0	
Rural	25,172 (40.12)	
Mean mother's age at first birth (SD)	24.19	
Individual level		
Male	31,812 (50.71)	7758
Birth weight, grams		
Birth weight ≤ 2500 g	2809 (4.48)	943
2500 g < birth weight ≤ 3500 g	30,914 (49.27)	5578
3500 g < birth weight‡	29,016 (46.25)	5022
Family size/birth order		
Only child	3964 (6.32)	0
Two children, first born‡	12,272 (19.56)	5305
Two children, second born	12,273 (19.56)	5305
Three children, first born	6285 (10.02)	3257
Three children, second born	6895 (10.99)	5377
Three children, third born	7126 (11.36)	3019
Four children or more, first born	2705 (4.31)	1501
Four children or more, second born	3084 (4.92)	2429
Four children or more, third born	3258 (5.19)	2526
Four children or more, fourth or higher born	4877 (7.77)	1522
Family structure history		
M – parents remain married, ages 0–13‡	41,960 (66.88)	899
M – parental divorce, no further changes, ages 0–13	5973 (9.52)	526
M – parental death, no further changes, ages 0–13	649 (1.03)	81
M – two or more transitions, ages 0–13	1502 (2.39)	184
U – mother remains unmarried, ages 0–13	7588 (12.09)	86
U – mother marries after birth, ages 0–13	3724 (5.94)	729
U – two or more transitions, ages 0–13	1343 (2.14)	246
Moved between ages 0 and 13	39,956 (63.69)	1994
Health conditions		
Had ADHD, ages 0–13	3968 (6.32)	1446
Had conduct disorders and/or ODD, 0–13	4365 (6.96)	1516
Had bipolar disorder, 0–13	292 (0.47)	126
Had asthma, 0–13	20,078 (32)	5025
Had major injuries, 0–13	46,393 (73.95)	5447
Had other major health conditions, 0–13	22,390 (35.69)	5404
Family problems		
Was a child in care, 9–13	959 (1.53)	212
Received services from CFS, 9–13	4524 (7.21)	1040
Economic adversity		
Received income assistance, 9–13	7087 (11.3)	488
Less than average Grade 9 achievement	34,583 (55.12)	5118

M = mother married at time of child's birth; U = mother unmarried at time of child's birth.
* Total number of children = 62,739.
† Total number of individuals with siblings in cohort = 31,504.
‡ Reference category for odds ratios.

measures such as residential mobility and family structure changes at the individual level.

Because the same set of predictors and cohort are used to examine two outcomes, adjustments were made for multiple comparisons using the Bonferroni method. A desired confidence level of 95% requires predictors to be significant at $p < 0.025$.[19] SAS version 9.3 was used for all analyses.

RESULTS

In the population cohort, approximately 20% did not graduate from high school in four years, while 3.2% had an ADHD diagnosis between ages 14 and 18. Table 1 presents frequencies of binary variables and means and standard deviations of continuous measures.

Evaluating covariates

Important covariates at birth and during childhood/early adolescence were compared across outcomes. Higher neighbourhood SES decreased the odds of failing to graduate from high school and increased the odds of having an ADHD diagnosis. Rural residency was a protective factor for both outcomes. With increases in mother's age at first birth, the odds of failing to graduate from high school dropped and those of being diagnosed with ADHD grew. Almost all individual measures (26 of 29) were statistically significant predictors of educational achievement; only 11 individual measures were significant for ADHD. Measures of family size/birth order, family structure history, challenged parenting, and economic adversity were particularly important in predicting high school graduation. Mental health conditions in childhood were the strongest

Table 2. Odds ratios of neighbourhood, family and individual covariates (population models)

Model†	Failure to graduate high school		ADHD, 14–18	
	1	2	1	2
Family level				
Standardized SEFI score	1.46**	1.55**	0.80**	0.77**
Rural	0.71**	0.7**	0.84*	0.82**
Mother's age at first birth	0.95**	0.93**	1.02**	1.02*
Individual level				
Male	1.4**	1.79**	1.47**	2.02**
Birth weight, grams				
Birth weight ≤ 2500 g	1.07	1.17*	0.98	0.97
2500 g < birth weight ≤ 3500 g	1.11**	1.14**	0.95	0.96
Family size/birth order				
Only child	1.19*	1.32**	1.06	1.04
Two children, second born	1.19**	1.31**	1.06	1
Three children, first born	1.08	1.08	1.19	1.05
Three children, second born	1.28**	1.33**	1.03	0.88
Three children, third born	1.2**	1.33**	0.95	0.87
Four children or more, first born	1.7**	1.72**	0.85	0.73*
Four children or more, second born	1.73**	1.81**	0.74	0.70*
Four children or more, third born	1.76**	1.92**	0.84	0.69*
Four children or more, fourth or higher born	**2.16****	**2.46****	0.84	0.66**
Family structure history				
M – parental divorce, no further changes, 0–13	1.42**	1.59**	1.32*	1.42**
M – parental death, no further changes, 0–13	1.37*	1.52**	0.6	0.69
M – two or more transitions, 0–13	1.52**	1.71**	1.4*	**1.73****
U – mother remains unmarried	**1.87****	**2.18****	1.09	1.16
U – mother marries after birth, 0–13	1.57**	1.84**	1.38*	1.38**
U – two or more transitions, 0–13	1.82**	**2.07****	1.2	1.39
Moved between ages 0 and 13	1.38**	1.43**	0.99	1.04*
Health conditions				
Had ADHD, 0–13	1.56**	2.03**	**25.23****	NA
Had conduct disorder/ODD, 0–13	1.26**	1.36**	**1.63****	**3.89****
Had bipolar disorder, 0–13	1.78**	2.05**	**1.77***	**2.08****
Had asthma, 0–13	0.94*	0.94*	1.12	1.21**
Had major injuries, 0–13	1.08*	1.15**	1.07	1.18*
Had other major health conditions, 0–13	0.98	1	1.21**	1.34**
Family problems				
Was a child in care, 9–13	**2.11****	**2.44****	**1.61***	**2.35***
Received services from CFS, 9–13	1.17**	1.33**	1.28*	1.51**
Economic adversity				
Received income assistance, 9–13	**2.01****	**2.34****	0.75*	0.82*
Less than average Grade 9 achievement	**9.27****	NA	**3.46****	**4.96****

Note: Statistically significant predictors with odds ratios differing most from zero are bolded.
M = mother married at time of child's birth; U = mother unmarried at time of child's birth.
* $N = 62{,}739$; * $p < 0.0125$; ** $p < 0.001$.
† Model 1 includes "Same but Previous" predictor; Model 2 does not include "Same but Previous" predictor.

predictors of late adolescent ADHD. Tables 2 and 4 present in bold text the odds ratios of the five most important predictors of each outcome.

Presence of the same condition at an earlier age impacted age 14–18 outcomes. For the education outcome, the "same but previous" predictor was "Less than average Grade 9 achievement"; for the ADHD outcome, it was "Had ADHD, Ages 0–13". Outcomes were modelled without the "same but previous" variable in the models (Model 2 in Tables 2 and 4); in the population models, exclusion of "same but previous" variables increased the importance of family structure changes for both outcomes.

Model fit

We used two different statistics to assess model performance. First, a Brier score determined model accuracy; a score of zero indicates complete accuracy and a score of 0.5 indicates complete inaccuracy.[20] Second, a c-statistic (ranging from 0.5 to 1) measured model ability to discriminate between those with and those without the outcome. Both models performed very well when all predictors were included (Brier score < 0.15 and c-statistic well above 0.8); model fit decreased with removal of the "same but previous" variable, but fit remained reasonable (c-statistic over 0.7) (Table 3).[21]

Populations and siblings

Comparisons between population (Table 2) and sibling (Table 4) samples help establish the generalizability of our results through providing somewhat differing perspectives on the life events used. The sibling analysis added the "difference in days between births of siblings" and removed "only child" in the "family size/birth order" measures. With the sample size for the population analyses more than twice as large as that for the sibling analyses, fewer measures in the sibling models were significant (but odds ratios were generally similar). Non-twin siblings were generally less likely to graduate from high school than twins, suggesting higher levels of parental involvement and/or less negative sibling interaction among the relatively small number ($n = 1{,}122$) of twins. Siblings born close together (<842 days) and those born far apart (>1,113 days) had the lowest odds of high school graduation.[22]

Observational studies comparing siblings within families have been considered to have more internal validity than population studies by eliminating "genetic confounds due to passive

Table 3. Fit statistics

Sample	Brier score		C-statistic	
	Population	Siblings	Population	Siblings
Failure to graduate high school				
All predictors	0.0959	0.0856	0.8529	0.8551
Without "Same but Previous" predictor	0.1000	0.0877	0.7953	0.7951
ADHD, 14–18				
All predictors	0.0217	0.0187	0.8943	0.8915
Without "Same but Previous" predictor	0.0267	0.0229	0.7956	0.7900

Table 4. Odds ratios of neighbourhood, family and individual covariates (sibling models)

Model†	Failure to graduate high school		ADHD, 14–18	
	1	2	1	2
Family level	OR	OR	OR	OR
Standardized SEFI score	1.44**	1.53**	0.87**	0.83**
Rural	0.74**	0.72**	0.86	0.85
Mother's age at first birth	0.95**	0.92**	1.03*	1.03*
Difference between siblings' birth dates				
Between 262 and 635 days	1.56**	1.6**	1.17	1.15
Between 636 and 842 days	1.38*	1.36*	1.29	1.32
Between 843 and 1112 days	1.23	1.22	1.13	1.16
More than 1113 days	1.41*	1.42*	1.37	1.44
Individual level				
Male	1.41**	1.84**	1.17	1.59**
Birth weight ≤ 2500 g	1.16	1.27*	0.93	1.02
2500 g < birth weight ≤ 3500 g	1.14*	1.17**	0.86	0.9
Family size/birth order				
Two children, second born	0.97	1.08	1.27	1.33*
Three children, first born	0.98	0.98	1.2	1.21
Three children, second born	1.17	1.22*	1.12	1.08
Three children, third born	1.01	1.09	1.13	1.19
Four children or more, first born	1.69**	1.72**	1.07	0.94
Four children or more, second born	1.71**	1.78**	0.82	0.9
Four children or more, third born	1.6**	1.77**	1.13	1.03
Four children or more, fourth or higher born	**2.21****	**2.5****	0.9	0.82
Family structure history				
M – parental divorce, no further changes, 0–13	1.29**	1.47**	1.27	1.32*
M – parental death, no further changes, 0–13	1.16	1.32	0.64	0.59
M – divorce or death and remarriage, 0–13	1.51**	1.7**	1.3	1.58
U – mother remains unmarried	**1.74****	2**	1.08	1.12
U – mother marries after birth, 0–13	1.69**	**1.92****	1.08	1.06
U – two or more transitions, 0–13	**1.99****	**2.18****	0.78	0.89
Moved between ages 0 and 13	1.41**	1.44**	1.11	1.14
Health conditions				
Had ADHD, 0–13	1.55**	**2.04****	**16.01****	NA
Had conduct disorder/ODD, 0–13	1.35**	1.46**	**1.89****	**3.82****
Had bipolar disorder, 0–13	2*	**2.35****	**2.53***	**2.95****
Had asthma, 0–13	0.95	0.94	1.22*	**1.33****
Had major injuries, 0–13	1.09	1.18**	1.01	1.12
Had other major health conditions, 0–13	0.97	1	**1.39****	**1.45****
Family problems				
Was a child in care of CFS, 9–13	1.7**	1.89**	1.06	1.56
Received services from CFS, 9–13	1.14	1.26**	1.19	1.4*
Economic adversity				
Received income assistance, 9–13	**2.11****	**2.42****	1.07	1.06
Less than average Grade 9 achievement	**10.29****	NA	**3.37****	**4.47****

Note: Statistically significant predictors with odds ratios differing most from zero are bolded.
M = mother married at time of child's birth; U = mother unmarried at time of child's birth.
* $N = 29{,}444$; * $p < 0.025$; ** $p < 0.001$.
† Model 1 includes "Same but Previous" predictor; Model 2 does not include "Same but Previous" predictor.

gene-environment correlation"[16(p. 275)] and by controlling for reciprocal causation.[10] Because stronger research designs tend to show smaller relationships, bootstrap methods were used to determine whether significant differences existed between effects in the population and sibling models (see Supplementary Appendix B).[16] When examining the "failure to graduate from high school" models, a significant decrease in effect was seen in only one predictor – "two children, second born" was significant in the population model but not in the sibling model. The "ADHD, 14–18" model showed no differences. Since only one effect size differed in the population and sibling models, results from using the population sample appear robust and not hindered by familial confounding, social selection, or misclassification for the education and ADHD variables.

Because fit statistics and odds ratios change relatively little between the two types of samples, multilevel population analyses (often drawn from administrative data and consecutive birth cohorts) might be able to replace more traditional sibling samples (typically based on survey data).[23] The greater number of cases provided by administrative data is particularly helpful in assessing importance (and statistical significance) when using many predictors.

Model validation and sensitivity testing

To determine that our model is not over-fit, we used 10-fold cross validation.[21] Differences in c-statistics of the cross-validated model and the original model were small and not statistically significant, indicating that over-fitting was not an issue. Bootstrapping was used to determine the standard errors associated with our estimates; confidence intervals from the original models were very similar to those generated from the bootstrapped models (see Supplementary Appendix C).

Various sensitivity analyses determined how different analytic decisions impacted the results. Sensitivity testing examined models that: 1) remove variables only available between ages 9 and 13; 2) use neighbourhood of the oldest child at birth; 3) measure neighbourhood SES using Material and Social Deprivation indices; 4) analyze a population subsample for whom household income was available (details in Supplementary Appendix D). Overall, our findings were minimally affected and very robust. Additional comparisons examined whether predictors differed for males and females (see Supplementary Appendix E). With insufficient numbers of male-only siblings and female-only siblings to permit multilevel analyses, logistic regressions were run. Some differences in predictors were seen between males and females; given the weaker study design, these results are not as robust.

DISCUSSION

The adverse childhood experiences important for each outcome differ. While family instability and economic adversity more strongly affect failing to graduate from high school, adverse health events in childhood and early adolescence have a greater impact on late adolescent ADHD.

Almost every study measure important in predicting high school graduation accords with the perspective put forward by Link and his collaborators.[24] They postulate that individuals of relative wealth and without family problems are disproportionately able to take advantage of government programs – such as (high school) education. Earlier success (as indicated by Grade 9 achievement) predicts later success. Odds ratios show some general similarities across the different measures for predicting graduation, with younger children in large families and children in poor, unstable families at particular risk.

The predictors of ADHD in late adolescence are more difficult to categorize. The family-level analyses indicate the ADHD diagnosis to be somewhat related to access to urban physicians and to having an older mother. Several of the family structure history measures (parental divorce and two or more transitions) remain significant but mother's unmarried status seems less important. This – and the lack of relationship between family receipt of income assistance and the ADHD diagnosis – suggests a smaller role for economic circumstances. Prior health conditions (particularly diagnoses pertaining to externalizing mental health conditions) generate consistently significant odds ratios. The very high odds ratio related to ADHD diagnoses at an earlier age highlight this condition's continuity. Severe family problems significantly contribute in the population analysis; the relatively small number of siblings coded differently on this variable may contribute to the lesser statistical significance of these problems.

CONCLUSION

This study highlights the importance of including both social determinants of health and health conditions in childhood and early adolescence when examining cognitive and emotional outcomes. The large number of cases, which facilitated assessment of the predictors of adolescent intellectual and emotional development, is a major study strength. The longitudinal nature of this research allows making stronger inferences about the impact of childhood adversity on health and education; all of the covariates in the models precede the outcome.[7] Population and sibling samples provide complementary approaches to evaluating the importance of predictors. Our approach should permit both developing valid risk prediction tools (currently used clinically for patient decision making) and identifying those who might benefit from specific interventions.[25] Applying such tools at the population level can give insight into these outcomes for a province, region or nation and further quantify the burden of risk in the population.

Administrative data do have limitations. Symptoms may occur well before an official diagnosis and family problems often precede a formal change (such as divorce or remarriage). Diagnostic identification is less than perfect, and some variables will be lacking. But linkable databases can facilitate "expansion" through the creation of new variables, such as a measure combining residential mobility and school mobility. Over time, additional measures (such as child maltreatment) will become available. Incorporating both surveys and new government datasets (such as housing and justice) into information-rich environments may further increase both the explainable variance and the number of interesting outcomes.

Time-varying measures are critical to better model a life course perspective; however, the timing of such events is an important consideration.[16] Examining several stages of childhood (for instance age groups 0–3, 4–8 and 9–13) would indicate which predictors appear most influential when, and provide important information regarding the appropriate timing of interventions. Maternal data (produced by linking mother's and child's files) should support both a family perspective and new modes of analysis.[23] Unique research possibilities are being created by the availability of longitudinal linkable data.

REFERENCES

1. Hertzman C. Human development, life course, intervention and health. In: Pickles A, Maughan B, Wadsworth M (Eds.), *Epidemiological Methods in Life Course Research*. Oxford: Oxford University Press, 2007; 111–36.
2. Schoeni R, House J, Kaplan G, Pollack H (Eds.). *Making Americans Healthier: Social and Economic Policy as Health Policy*. New York, NY: Russell Sage Foundation, 2008.
3. Currie J, Stabile M. Child mental health and human capital accumulation: The case of ADHD. *J Health Econ* 2006;25(6):1094–118. PMID: 16730082. doi: 10.1016/j.jhealeco.2006.03.001.

4. Adler N, Stewart J. Health disparities across the lifespan: Meaning, methods, and mechanisms. *Ann N Y Acad Sci* 2010;1186:5–23. PMID: 20201865. doi: 10.1111/j.1749-6632.2009.05337.x.
5. Currie J. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *J Econ Lit* 2000;41(1):87–122. doi: 10.1257/jel.47.1.87.
6. Koenen K, Rudenstine S, Susser E, Galea S (Eds.). *A Life Course Approach to Mental Disorders*. New York, NY: Oxford University Press, 2014.
7. Duncan G, Murnane R (Eds.). *Whither Opportunity?* New York, NY: Russell Sage Foundation, 2011.
8. Jutte D, Roos N, Brownell M, Briggs G, MacWilliam L, Roos L. The ripples of adolescent motherhood: Social, educational and medical outcomes for children of teen and prior teen moms. *Acad Pediatr* 2010;10(5):293–301. PMID: 20674531. doi: 10.1016/j.acap.2010.06.008.
9. Levine J, Pollack H, Comfort M. Academic and behavioral outcomes among the children of young mothers. *J Marriage Fam* 2001;63(2):355–69. doi: 10.1111/j.1741-3737.2001.00355.x.
10. Lahey B, D'Onofrio B. All in the family: Comparing siblings to test causal hypotheses regarding environmental influences on behavior. *Curr Dir Psychol Sci* 2010;19(5):319–23. PMID: 23645975. doi: 10.1177/0963721410383977.
11. Oreopoulos P, Stabile M, Walld R, Roos L. Short, medium, and long term consequences of poor infant health: An analysis using siblings and twins. *J Hum Resour* 2008;43(1):88–138. doi: 10.3368/jhr.43.1.88.
12. Moffitt R. Remarks on the analysis of causal relationships in population research. *Demography* 2005;42(1):91–108. PMID: 15782897. doi: 10.1353/dem.2005.0006.
13. *The Johns Hopkins ACG Case-Mix System (Version 6.0 Release Notes)*. Baltimore, MD: Johns Hopkins University, 2003.
14. Martens PJ, Fransoo R, McKeen N, The Need to Know Team, Burland E, Jebamani L, et al. *Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study*. Winnipeg, MB: Manitoba Centre for Health Policy, 2004.
15. Mann C. Observational research methods. Research design II: Cohort, cross sectional, and case-control studies. *Emerg Med J* 2003;20(1):54–60. PMID: 12533370. doi: 10.1136/emj.20.1.54.
16. Jaffee S, Strait L, Odgers C. From correlates to causes: Can quasi-experimental studies and statistical innovations bring us closer to identifying the causes of antisocial behavior?*Psychol Bull* 2012;138(2):272–95. PMID: 22023141. doi: 10.1037/a0026020.
17. Agnew-Blais J, Seidman L. Attention deficit/hyperactivity disorder over the life course. In: Koenen K, Rudenstine S, Susser E, Galea S (Eds.), *A Life Course Approach to Mental Disorders*. New York, NY: Oxford University Press, 2014; 156–64.
18. Turkheimer E, Waldron M. Nonshared environment: A theoretical, methodological, and quantitative review. *Psychol Bull* 2000;126(1):78–108. PMID: 10668351. doi: 10.1037/0033-2909.126.1.78.
19. Smith M. Multiple Comparisons. Available at: http://www.ma.utexas.edu/users/mks/384E07/multcomp.pdf (Accessed March 12, 2015).
20. Green K. *Assessing Probabilistic Forecasts About Particular Situations*. Melbourne, Australia: Business and Economics Forecasting Unit, 2004.
21. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. New York, NY: Springer Science + Business Media, Inc., 2001.
22. Powell B, Carr Steelman L. The educational benefits of being spaced out: Sibship density and educational progress. *Am Sociol Rev* 1993;58(3):367–81. doi: 10.2307/2095906.
23. Strully K, Mishra G. Theoretical underpinning for the use of sibling studies in life course epidemiology. In: Lawlor D, Mishra G (Eds.), Family Matters: Designing, Analyzing and Understanding *Family Based Studies in Life Course Epidemiology*. Oxford: Oxford University Press, 2009; 39–56.
24. Link BG, Phelan JC. Understanding sociodemographic differences in health – The role of fundamental social causes. *Am J Public Health* 1996;86(4):471–73. PMID: 8604773. doi: 10.2105/AJPH.86.4.471.
25. Rosella L, Manual D, Burchill C, Stukel T. A population based algorithm for the development of diabetes: Development and validation of the Diabetes Population Risk Tool (DPoRT). *J Epidemiol Community Health* 2011; 65(7):613–20. PMID: 20515896. doi: 10.1136/jech.2009.102244.

Received: June 5, 2015
Accepted: October 11, 2015

RÉSUMÉ

OBJECTIFS : Nous avons utilisé une approche axée sur le parcours de vie et maillé des données du Manitoba de la naissance à 18 ans pour faciliter les comparaisons de deux effets importants : l'obtention du diplôme d'études secondaires et le trouble déficitaire de l'attention avec hyperactivité (TDAH). Avec un jeu de variables commun, nous avons cherché à répondre aux questions suivantes : Les indicateurs prédisant l'obtention du diplôme d'études secondaires diffèrent-ils de ceux qui prédisent le TDAH? Quels facteurs sont les plus importants? Les modèles sont-ils bien adaptés à chaque résultat?

MÉTHODE : Les données administratives du Centre d'élaboration et d'évaluation de la politique des soins de santé, au Centre de la politique des soins de santé du Manitoba, ont servi à mener l'un des protocoles d'étude observationnelle les plus robustes : la modélisation multiniveau d'échantillons d'une grande population ($n = 62\ 739$) et de frères et sœurs ($n = 29\ 444$). Les variables incluses étaient les caractéristiques du quartier, des indicateurs de stabilité familiale, ainsi que les états de santé mentale et physique durant l'enfance et l'adolescence.

RÉSULTATS : Les expériences défavorables de l'enfance qui importent pour chaque effet sont différentes. L'instabilité familiale et l'adversité économique ont un effet plus prononcé sur l'abandon des études secondaires avant l'obtention du diplôme, tandis que les problèmes de santé durant l'enfance et au début de l'adolescence ont davantage d'effet sur le TDAH en fin d'adolescence. Les variables incluses dans le modèle ont apporté une précision et une discrimination excellentes.

CONCLUSION : Ces résultats éclairent le rôle de plusieurs variables familiales et sociales et peuvent servir à créer des outils de prédiction fiables et valides pouvant identifier les personnes à haut risque. L'application d'un tel outil à l'échelle d'une population donnerait une idée du fardeau futur de ces effets dans une région ou un pays et permettrait de chiffrer davantage le fardeau du risque dans la population.

MOTS CLÉS : déficit de l'attention avec hyperactivité; niveau d'instruction; risque; études longitudinales