Appendix 6 (as supplied by the authors): Methods — Extended Version

This study was approved by the Institutional Review Board of McGill University's Faculty of Medicine, as well as by the Health Sciences Research Ethics Board of Queen's University.

Ontario's Grade 8 HPV Vaccination Program

Ontario's grade 8 HPV vaccination program began in September 2007. This publicly funded program offers the three recommended doses of the vaccine, free-of-charge, to all grade 8 girls in the province.[1] The program is primarily delivered through school-based immunization clinics administered by the province's 36 public health units, but eligible girls also have the option of receiving the vaccine at the their local health unit or through their family physician at no cost. During our study period, eligible females had until the end of their grade 8 school year to initiate the vaccine series and until the end of grade 9 to complete it under the publicly funded program. Prior to September 2012, no catch-up program was offered; therefore, females who were not eligible for the program (e.g., completed grade 8 before September 2007) would have had to pay for the vaccine at a cost of approximately $150 per dose.

Data Sources

Data for this study were obtained from Ontario's population-based administrative databases, which are generated by the province's universal health insurance programs and were housed at the Institute for Clinical Evaluative Sciences (ICES). Specifically, we used the following databases: (1) Registered Persons Database (RPDB), Ontario's population registry of insured persons, for information on socio-demographics, (2) Ontario Health Insurance Plan (OHIP) database for information on fee-for-service claims

Appendix to: Smith LM, Kaufman JS, Strumpf EC, et al. Effect of human papillomavirus (HPV) vaccination on clinical indicators of sexual behaviour among adolescent girls: the Ontario Grade 8 HPV Vaccine Cohort Study. *CMAJ* 2014. DOI:10.1503/cmaj.140900.

by physicians, (3) Discharge Abstract Database (DAD) for information on hospitalizations, (4) Same-Day Surgery (SDS) database for information on procedures carried out during same-day surgeries, and (5) National Ambulatory Care Reporting System (NACRS) for information on emergency department visits. These databases have been used extensively in health research, including in the post-marketing evaluation of drug and vaccine effects.[2–6] Details on these databases are available elsewhere (Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140900/-/DC1).[7–11]

We also used the Immunization Records Information System (IRIS) for information on vaccinations, including HPV vaccinations.[11] IRIS databases are maintained by each of Ontario's 36 health units to record and track the immunization status of all school-aged children in their jurisdiction. Although these databases were originally developed for the six designated diseases (diphtheria, tetanus, polio, measles, mumps and rubella) for which immunization is prescribed by the Minister of Health and Long-Term Care (*Immunization of Schools Pupils Act, 1982*), they are currently used for other vaccines as well, particularly those that are publicly funded. Prior to centralizing the IRIS databases ICES, we validated the HPV vaccination data of a medium-sized health unit and found that it captured HPV vaccination status with near-perfect sensitivity (99.8%, 95% confidence interval [CI] 99.3 to 99.9) and high specificity (97.7%, 95% CI 96.3 98.7). Moreover, 98.6% of HPV vaccination dates were accurate.[12] Due to the rigorous and standardized procedures that have developed as a result of the requirements in the *Immunization of Schools Pupils Act*, we expect the HPV vaccine data of other health units to be of similarly high quality.

All data were accessed through the ICES satellite unit at Queen's University. Since residents of Ontario are represented in these databases by a unique encrypted identifier, individual-level record linkage is possible across databases and over time.

Study Population and Cohort Formation

We identified a population-based cohort of all girls eligible for Ontario's grade 8 HPV vaccination program in the first two school years it was offered (i.e., 2007/08 and 2008/09). For the purpose of comparison, we also included girls who were in grade 8 in Ontario in the two years before the program (i.e., 2005/06 and 2006/07), who were therefore ineligible for publicly funded, school-based HPV vaccination. Although we did not have a direct measure of school grade, Ontario school entry practices are such that children typically enter school (Kindergarten) in September of the calendar year during which they turn 5, meaning the vast majority of children in a given grade have the same birth year.[13] Since this means girls in grade 8 typically turn 13 by December 31 of that school year, we identified a cohort of all females born in 1992, 1993, 1994, and 1995 to correspond with grade 8 years of 2005, 2006, 2007, and 2008, respectively. We then restricted the cohort to girls who were alive and residing in Ontario on September 1 of their grade 8 year (cohort entry) and whose immunization records were available at the time of the analysis. Although using birth year to determine grade 8 year misclassifies cohort members who were held back or advanced a grade, we found that this approach correctly identified 96.4% of girls eligible for the program's first two years (i.e., 2007/08 and 2008/09).[14] Cohort members were followed until the earliest of their date of death, occurrence of a study outcome, or March 31 of grade 12 (i.e., March 31 of 2010, 2011, 2012, or 2013, depending on the girl's birth year).

Measurement and Analysis

The Regression Discontinuity Design – To address our objectives, we used the regression discontinuity design (RDD), a quasi-experimental approach that was specifically created to evaluate the causal effects of interventions.[15–19] The RDD is used in situations when assignment to an intervention (e.g., HPV vaccine program) is determined by the value of an observed continuous factor (e.g., birth date), referred to as the "forcing variable", being on one side of a fixed eligibility cut-off or the other, causing the probability of receiving the intervention (e.g., HPV vaccine) to increase discontinuously at this cut-off. In terms of Ontario's grade 8 HPV vaccination program, assignment to the intervention was based on whether individuals were in grade 8 before or after the September 2007 program implementation date (i.e., born December 31, 1993 or earlier *vs.* January 1, 1994 or later), causing the probability of receiving the vaccine to jump at the eligibility cut-off (Appendix 2A, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140900/-/DC1). RDD analyses are used to measure any corresponding discontinuous change in the probability of the outcome at the same eligibility cut-off (Appendix 2B), which is interpreted as the causal effect of the intervention. Correspondingly, a null effect is reflected by *continuity* in the outcome across the cut-off (Appendix 2C).

The major advantage of the RDD rests on the notion that the eligibility criteria and implementation date, which determine the assignment cut-off, are based on administrative decisions, meaning the exact location of the eligibility cut-off is random with respect to the characteristics of cohort members. Consequently, individuals falling directly on either side of the cut-off are comparable with respect to all measured and

unmeasured confounders; the only factor that differentiates them is their probability of receiving the vaccine. This type of design is particularly valuable in studies of vaccine effects because individuals who opt for vaccination tend to have different health beliefs and behaviours than those who do not. Since health beliefs and behaviours are strongly associated with health outcomes and are difficult to identify and quantify, traditional methods of analysis that directly compare vaccinated and unvaccinated individuals are prone to confounding bias.[20–23] Conversely, by controlling for this type of observed and unobserved confounding, the RDD facilitates reliable causal inference.[15–17]

Forcing variable and cut-off – As mentioned above, our study design exploits the fact that girls were eligible for the HPV vaccination program based on when they were in grade 8. Since school grade was estimated based on birth date, females born January 1 1992 to December 31, 1993 (corresponding with the 2005/06 and 2006/07 grade 8 calendar years) were ineligible for the HPV vaccination program, whereas females born January 1, 1994 to December 31, 1995 (corresponding with the 2006/07 and 2007/08 grade 8 years) were eligible for this program. Accordingly, the forcing variable was based on birth date and December 31, 1993 *vs.* January 1, 1994 defined either side of the eligibility cut-off. For the purposes of analysis, the forcing variable was collapsed into three-month intervals (referred to as "birth year quarters"), meaning cohort members born October 1, 1993 to December 31, 1993 were directly on the ineligible side of the cut-off and cohort members born January 1, 1994 to March 31, 1994 were directly on the eligible side. Cohort members born earlier/later than those dates were represented with increasing distance from the cut-off on the ineligible/eligible sides (Appendix 3, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140900/-/DC1).

Exposure Ascertainment – Two levels of exposure were analyzed. First, to evaluate the impact of the vaccination program, exposure was based solely on program eligibility. Therefore, cohort members who were in grade 8 in the 2005/06 and 2006/07 school years were classified as *ineligible* and those in grade 8 in the 2007/08 and 2008/09 were classified as *eligible.* This approach is analogous to an "intention-to-treat" (ITT) definition of exposure. Second, to assess the impact of vaccination, actual HPV vaccine receipt was also taken into account. A girl was classified as *vaccinated* if she received three doses of the vaccine between September 1 of grade 8 and August 31 of grade 9, which is the program vaccination period; otherwise, she was considered *unvaccinated.* The use of three doses for the primary exposure definition was based on the fact that this vaccine is administered as a three-dose series in Ontario. However, we also conducted sensitivity analyses based on receipt of at least one dose to assess whether the act of vaccination may have been sufficient to induce disinhibition. Similarly, we defined HPV vaccination status based on two doses in light of recent evidence that suggests two doses provide adequate protection.[24,25]

Outcome Ascertainment – Our primary outcome was a composite measure of incident non-HPV-related STIs and pregnancy (Appendix 4, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140900/-/DC1). We also assessed each endpoint separately. We excluded anogenital warts (an STI caused by HPV) from our measure of STIs because a decrease in this endpoint is an intended effect of the program and the vaccine. To ensure fixed follow-up time with equal probability of the outcomes for all cohort members, outcomes were ascertained between September 1 of grade 10 and

March 31 of grade 12. A case was defined as incident if there was no indication of that event (STI or pregnancy) in the previous 365 days.

Baseline Characteristics – To describe the study cohort, we identified a number of baseline characteristics relating to socio-demographics, vaccination history, health service use, and medical history.

Statistical Analyses – To evaluate the program impact (i.e., intention-to-treat effect), linear regression was used to model the association between program eligibility and the outcome. To evaluate the vaccine impact, two-stage linear regression was used to estimate the association between program eligibility and the outcome *and* the association between program eligibility and HPV vaccine exposure. In the two-stage analysis, the estimate of interest was the ratio of coefficients from the two regressions, which represents the absolute impact of HPV vaccination on the outcome. Similarly, one- and two-stage log-binomial regressions were used to estimate the relative impact of program eligibility and vaccination on the outcomes of interest. In all analyses, cohort members born in 1993 or 1994 (i.e., closest to the cut-off) were weighted twice as heavily as those born in 1992 and 1995 because individuals closest to the cut-off are the most comparable. Moreover, analyses were conditioned on birth timing (i.e., birth quarter) because we found that, across birth years, females born early (or late) in the year were the most comparable (Appendix 5, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140900/-/DC1). Similar effects of relative age have been found in other studies as well.[26–28]

Sensitivity analyses were executed to test the robustness of our results to our various assumptions. For example, we assessed the impact of using different weights for

birth year. Also, as previously mentioned, vaccination status was re-defined based on receipt of at least one and a least two doses. In addition, exposure and outcome ascertainment windows were altered to ascertain vaccine exposure in grade 8 (since this is when most girls are vaccinated) and outcomes in grades 9 to grade 12. Furthermore, we conducted sensitivity analyses that controlled for neighbourhood income quintile, hepatitis B vaccination, and a recent sexual health-related outcome (i.e., pregnancy, diagnosis of an STI, or cervical cancer screening) in addition to birth quarter.

Data management was carried out using SAS statistical software version 9.3 (SAS Institute Inc., Cary, North Carolina), and statistical analyses were executed using Stata version 13.1 (StataCorp, College Station, Texas).

References

1. Ministry of Health and Long-Term Care. Ontario's HPV Vaccination Program. 2013. (Accessed March 13, 2014, at http://www.health.gov.on.ca/en/ms/hpv/.)

2. Lipscombe LL, Gomes T, Levesque LE, Hux JE, Juurlink DN, Alter DA. Thiazolidinediones and cardiovascular outcomes in older patients with diabetes. *JAMA* 2007;298(22):2634-43.

3. Lipscombe LL, Levesque L, Gruneir A, Fischer HD, Juurlink DN, Gill SS, et al. Antipsychotic drugs and hyperglycemia in older patients with diabetes. *Arch Intern Med* 2009;169(14):1282-9.

4. Kwong JC, Tanuseputro P, Zagorski B, Moineddin R, Chan KJ. Impact of varicella vaccination on health care outcomes in Ontario, Canada: effect of a publicly funded program? *Vaccine* 2008;26(47):6006-12.

5. Juurlink DN, Stukel TA, Kwong J, Kopp A, McGeer A, Upshur RE, et al. Guillain-Barre syndrome after influenza vaccination in adults: a population-based study. *Arch Intern Med* 2006;166(20):2217-21.

6. Smith LM, Brassard P, Kwong JC, Deeks SL, Ellis AK, Levesque LE. Factors associated with initiation and completion of the quadrivalent human papillomavirus vaccine series in an ontario cohort of grade 8 girls. *BMC Public Health* 2011;11:645.

7. Iron K. Moving toward a better health data system for Ontario. *ICES Investigative Report*. Toronto: Institute for Clinical Evaluative Sciences, 2006.

8. Iron K, Zagorski BM, Sykora K, Manuel DG. Living and dying in Ontario: An opportunity for improved health information. *ICES Investigative Report*. Toronto: Institute for Clinical Evaluative Sciences, 2008.

9. Improving health care data in Ontario. *ICES Investigative Report*. Toronto: Institute for Clinical Evaluative Sciences, 2005.

10. Juurlink D, Preyra C, Croxford R, Chong A, Austin P, Tu J, et al. Canadian Institute for Health Information Discharge Abstract Database: A Validation Study. Toronto: Institute for Clinical Evaluative Sciences, 2006.

11. Ministry of Health and Long-Term Care. Data sources and population health indicators limitations (An appendix to the Initial Report on Public Health). 2009. (Accessed March 6, 2014, at http://www.health.gov.on.ca/en/public/publications/pubhealth/init_report/pdfs/appendix4.pdf.)

12. Smith LM, Lévesque LE, Nasr M, Perry AG. Validity of the Immunization Record Information System (IRIS) database for epidemiologic studies of the human papillomavirus (HPV) vaccine. *Canadian Journal of Clinical Pharmacoepidemiology* 2010;17(1):e90-e127.

13. Teach in Ontario. Accessed July 4, 2014 at http://www.teachinontario.ca/tio/en/structure.htm)

14. Lévesque LE, Smith LM, Perry AG, Nsar M, Hogan ML, Martin A, et al. The Ontario Grade 8 HPV Vaccine Cohort Study: A Feasibility and Validity Evaluation. *Ontario Vaccine Sciences Symposium*. Toronto, ON, 2011.

15. Cook TD. ‘‘Waiting for Life to Arrive’’: A history of the regression-discontinuity design in Psychology, Statistics and Economics. *Journal of Econometrics* 2007;142:636-54.

16. Hahn J, Todd P, Van der Klaauw W. Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica* 2001;69(1):201-09.

17. Imbens GW, Lemieux T. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 2008;142:615-35.

18. Lee H, Munk T. Using Regression Discontinuity Design for Program Evaluation. Section on Survey Research Methods - Joint Statistical Meeting; 2008; 1675-1682 (Denver, Colorado).

19. Zuckerman IH, Lee E, Wutoh AK, Xue Z, Stuart B. Application of regression-discontinuity analysis in pharmaceutical health services research. *Health Serv Res* 2006;41(2):550-63.

20. Fine PE, Chen RT. Confounding in studies of adverse reactions to vaccines. *Am J Epidemiol* 1992;136(2):121-35.

21. Chen RT, Davis RL, RHodes PH. Special Methodological Issues in Pharmacoepidemiology studies of vaccine safety. In: Strom BL, editor. *Pharmacoepidemiology*. Fourth ed: John Wiley & Sons, Ltd, 2005:455-85.

22. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol* 2006;35(2):337-44.

23. Nelson JC, Jackson ML, Weiss NS, Jackson LA. New strategies are needed to improve the accuracy of influenza vaccine effectiveness estimates among seniors. *J Clin Epidemiol* 2009;62(7):687-94.

24. Dobson SR, McNeil S, Dionne M, Dawar M, Ogilvie G, Krajden M, et al. Immunogenicity of 2 doses of HPV vaccine in younger adolescents vs 3 doses in young women: a randomized clinical trial. *JAMA* 2013;309(17):1793-802.

25. Herweijer E, Leval A, Ploner A, Eloranta S, Simard JF, Dillner J, et al. Association of varying number of doses of quadrivalent human papillomavirus vaccine with incidence of condyloma. *JAMA* 2014;311(6):597-603.

26. Buckles KS, Hungerman DM. Season of Birth and Later Outcomes: Old Questions, New Answers. *The review of economics and statistics* 2013;95(3):711-24.

27. Hancock DJ, Adler AL, Cote J. A proposed theoretical model to explain relative age effects in sport. *Eur J Sport Sci* 2013;13(6):630-7.

28. Morrow RL, Garland EJ, Wright JM, Maclure M, Taylor S, Dormuth CR. Influence of relative age on diagnosis and treatment of attention-deficit/hyperactivity disorder in children. *CMAJ* 2012;184(7):755-62.