Personal Privacy and Public Health

Potential Impacts of Privacy Legislation on Health Research in Canada

M. Anne Harris, MSc
Adrian R. Levy, PhD

Kay E. Teschke, PhD

ABSTRACT

Despite variation in Canadian privacy laws between provinces and territories, increasing legislative protection of personal privacy has imposed restrictions on health research across the country. The effects of these restrictions on patient recruitment include increased study costs, durations, and decreased participation rates. Low participation rates can jeopardize the validity of research findings and the accuracy of measures of association by introducing non-response, or participation bias. We constructed simulations to assess potential effects of non-response bias on the accuracy of measures of association in a hypothetical case-control study. Small biases that alter the probability of selecting an exposed case can lead to dramatic inflation or attrition of the odds ratio (OR) in case-control studies. ORs are more unstable and subject to error when the true probability of selecting an exposed case is greater, such that strong positive associations are subject to error even at low levels of bias. Well-powered, population-based epidemiological research is a cornerstone of public health. Therefore, when weighing the benefits of protecting personal privacy, the benefits of valid and robust health research must also be considered. Options might include special legislative treatment of health research, or the use of an "opt-out" (vs. the current "opt-in") construct for consent in confidential research.

Key words: Privacy; legislation; public health; epidemiology

RÉSUMÉ

Les lois sur la protection des renseignements personnels diffèrent d'une province et d'un territoire à l'autre au Canada, mais dans l'ensemble, la protection juridique accrue accordée à la protection de la vie privée impose des restrictions à la recherche en santé dans tout le pays. Comme ces restrictions ont un effet sur le recrutement des patients, elles font augmenter le coût et la durée des études, et diminuer les taux de participation. Or, de faibles taux de participation peuvent compromettre la validité des résultats de recherche et l'exactitude des mesures du degré d'association en introduisant des biais de non-réponse ou de participation. Nous avons construit des simulations pour analyser les effets possibles du biais de non-réponse sur l'exactitude des mesures d'association pour une étude cas-témoin hypothétique. On sait que de faibles biais, s'ils modifient la probabilité de sélectionner un cas exposé, peuvent entraîner une inflation ou une attrition considérable du rapport de cotes (RC) dans une étude cas-témoins. Les RC sont plus instables, et comportent un risque d'erreur plus grand, lorsque la probabilité réelle de sélectionner un cas exposé est plus grande, ce qui fait que des associations fortement positives peuvent être erronées même si les biais sont faibles. La recherche épidémiologique fondée sur la population et dotée d'un bon degré de puissance est la pierre angulaire de la santé publique. Par conséquent, lorsqu'on soupèse les avantages de la protection de la vie privée, il faut tenir compte de l'avantage de pouvoir mener des études de recherche en santé valides et robustes. On pourrait par exemple appliquer un traitement juridique particulier à la recherche en santé ou faire appel à la « participation facultative » (par opposition à la « participation sur demande » en vigueur actuellement) pour obtenir le consentement des patients aux travaux de recherche confidentiels.

Mots clés : protection des renseignements personnels; lois; santé publique; épidémiologie

Department of Health Care and Epidemiology, University of British Columbia, Vancouver, BC
Correspondence and reprint requests: Anne Harris, Department of Health Care and Epidemiology, Faculty of Medicine, University of British Columbia, 5804 Fairview Avenue, Vancouver, BC V6T 1Z3, E-mail: aharris7@interchange.ubc.ca

A trend toward increasing legislative protection of personal privacy has occurred across developed countries, including Canada. Although privacy legislation is intended to protect both individual and collective rights, it has had a major impact on health research by prohibiting or restricting contact of study subjects¹ and/or requiring individual consent to release data for anonymized research.² Privacy acts in five Canadian provinces and in the federal jurisdiction are compared in Table I, selected to provide an example of typical privacy legislation. Early concerns about restrictions on health data collection were expressed in the 1990s.³ More recently, researchers have identified some of the potential effects of privacy legislation on health-related research. Outcomes include study cancellation,⁴ increases in study cost and duration,⁵⁻⁸ and decreases in recruitment and study power.⁹ However, perhaps the most pernicious impact is decreases in participation rates and the attendant risk of introducing a bias. If participation rates differ according to exposure or outcome, systematic non-response can result in biased results and erroneous interpretations.

Non-response or participant bias

Non-response bias can occur when study participants differ systematically from non-participants in ways related to the association under examination.¹⁰ Refusal to participate often comprises a small proportion of non-response,^{1,2,6} with a more substantial proportion attributed to the inability to contact potential subjects, for example when patients die or move before contact can be made.^{6,11} McCarthy et al.⁵ studied the effect of Minnesota legislation requiring study subjects to consent to the use of the information contained in their medical records for research purposes. They enrolled 19% of subjects, compared to 93% in a neighbouring jurisdiction where study-specific consent was not required.⁵ Approximately 43% of the non-response comprised failure to contact (unreturned consent forms and follow-up calls).⁵ In Canada, Tu et al.⁶ assessed the practicality of requiring patient consent prior to enrolment into the Canadian Stroke Network registry. The overall participation rate was 39% with significantly lower participation among women, older

TABLE I**Comparison of Privacy Legislation in Canadian Jurisdictions**

	British Columbia	Alberta	Manitoba	Ontario	Québec	Canada (Federal)
Separate legislation governs health information	N	Y	Y	Y	N	N
Name of relevant legislation (reference)	Freedom of Information and Protection of Privacy Act ²¹	Health Information Act ²²	Personal Health Information Act ²³	Personal Health Information Protection Act ²⁴	Act Respecting Access To Documents Held By Public Bodies And The Protection Of Personal Information ²⁵	Personal Information Protection and Electronic Documents Act ²⁶ and The Privacy Act ²⁷
Forbids release of information to contact of individuals for purposes of research	Y	N	N	N	N	N
Stipulates requirements for researchers applying for personal information	Y	Y	Y	Y	Y	Y
States that information must be released to researchers if requirements met	N	N	N	N	N	N

persons, and those with increased comorbidity and mortality.⁶ Similarly low participation rates have been observed in registries in the United States following enforcement of consent requirements at enrolment.⁷

Restrictions on contact of study subjects can also reduce participation. For example, clinical trials,⁸ studies of disease causation¹ and descriptive studies of health outcomes following treatment¹² can all require contacting potential subjects identified through medical records or registries. Although prohibited in British Columbia (see Table I), most jurisdictions in North America, Europe, and Australia do not prohibit such contact outright, but often require additional consent stages^{1,13,14} that can substantially reduce enrolment.² In one study from the United Kingdom which employed a consent form prior to sending survey forms, the return rate of consent forms was 25%, meaning the pool of potential subjects was very small compared to the number initially contacted, and the upper limit on final participation rates was 25%.¹⁵

While there are several ways in which overall participation rates can drop, and non-response increase, the ideal epidemiologic study has a high participation rate and no differences between participants and non-participants.¹⁶ The higher the rate of non-response (no matter its source) in a study, the greater the probability that differences between participants and non-participants will influence study results leading to biased analyses and therefore conclusions.¹⁰ Estimating the impact of non-response is impossible when researchers are

The information in Table I of this article reports on laws as of January 2008, yet all legislation is subject to change. For example, on May 29, 2008, the British Columbia Legislative Assembly passed Bill 24 (e-Health Act and "consequential amendments" to the Freedom of Information and Protection of Privacy Act). Sections of both acts would now permit release of information to researchers for the purposes of contacting potential study participants. This proposed change reflects legislators' recognition of the impediments to public health research in that province. The new legislation retains certain conditions, and how it will be applied in practice will be of intense interest to researchers. This example demonstrates that policy changes affecting research are not inexorable, and highlights the importance of evaluating the potential effects of privacy-protecting policy on research in the public interest.

TABLE II**Effects of Altering the Proportion of Sampled Cases Classified as Exposed in a Hypothetical Case-control Study**

Effect of Bias on Odds Ratio	Bias	2x2 Table		Odds Ratio	Error
Unbiased	No Bias: 1.0	E+	Case 220 Control 120	True OR: 2.5	No Error: 1.0
		E-	280 380		
		Tot	500 500		
Inflation: exposed cases overrepresented	2.0	E+	Case 440 Control 120	23.2	9.3
		E-	60 380		
		Tot	500 500		
	1.5	E+	Case 330 Control 120	6.1	2.4
		E-	170 380		
		Tot	500 500		
	1.2	E+	Case 260 Control 120	3.4	1.3
		E-	240 380		
		Tot	500 500		
Attenuation: exposed cases under-represented	0.9	E+	Case 200 Control 120	2.1	0.8
		E-	300 380		
		Tot	500 500		
	0.8	E+	Case 180 Control 120	1.8	0.7
		E-	320 380		
		Tot	500 500		
	0.6	E+	Case 120 Control 120	1.0	0.4
		E-	380 380		
		Tot	500 500		

Bias was measured by the ratio of the proportion of exposed cases in the biased vs. unbiased samples. Error was measured by the ratio of the biased OR estimate divided by unbiased or true OR.

Figure 1. Relationship between the magnitude of an introduced bias and the magnitude of error for different true odds ratios (OR) in a hypothetical case-control study

Bias was introduced by manipulating the proportion of cases who were exposed, and is measured by the ratio of the proportion of exposed cases in the biased vs. unbiased samples. Error was measured by the ratio of the biased OR estimate divided by unbiased or true OR.

*N.B.: The y-axis employs a logarithmic scale.

unable to collect data on non-participants. However, several studies have addressed this issue by: simulating the effect of privacy rules;⁶ performing before/after study designs;⁸ and making comparison to jurisdictions with different privacy rules.⁵ Those studies have generally found that the characteristics of non-respondents, and thus potential direction and magnitude of bias, vary between studies. For example, age has been found to be positively associated both with participation^{15,17-19} and with non-participation.^{2,14} Similar ambiguity has been observed for sex^{14,19} and health status.¹⁹ These contradictory findings exemplify the fact that determinants of participation and non-response (and thus potential biases) are complex, often unexpected, and not easily accommodated with post-hoc analytic methods.

Simulations examining the impact of non-response

Previous simulations by Jones¹⁰ demonstrated that the probability of obtaining a

correct estimate declines sharply as participation rates decrease, even at low levels of bias. However, we might also wish to understand the magnitude of inaccuracy introduced by different levels of non-response bias. To address this issue, we constructed two simulations based on parameters that would be commonly observed in a case-control study, a common epidemiologic study design in which the odds of exposure in disease cases are compared to the odds of exposure in healthy controls via a ratio (the odds ratio, OR).

Simulation A: Effect of Non-response Bias on Observed Odds Ratios

An estimate of the magnitude of bias in a case-control study is the ratio comparing the proportion of sampled case subjects who are exposed in a biased vs. unbiased sample. When this ratio is 1, no bias exists. When the ratio exceeds 1, exposed cases are overrepresented resulting in an incorrectly inflated OR. Conversely, when the

ratio is less than 1, exposed cases are under-represented resulting in an incorrectly attenuated OR. The degree of error introduced can be assessed by the ratio of the observed OR over the true or unbiased OR.

Consider a hypothetical case-control study in which 500 cases and 500 controls are to be recruited. In this example, the true, unbiased OR is 2.5. Table II illustrates the observed OR estimates when the proportion of cases who are exposed is altered due to different levels of non-response. The upshot is that impact of relatively small non-participation biases, on the order of 10-15%, can have a measurable effect on OR estimates, ranging from substantial attenuation to substantial inflation. Furthermore, the magnitude of error is greater than the magnitude of the bias, that is, our analyses inflate the biases in our data.

Simulation B: Magnitude of Error vs. Magnitude of Bias

Figure 1 shows the relationship between the magnitude of a bias affecting the proportion of sampled cases who are exposed and the magnitude of error in the odds ratio, again in a hypothetical case-control study enrolling 500 cases and 500 controls. The error is exponentially related to bias, with an asymptote occurring at the degree of bias that produces a sample of exposed cases only. That is, as the bias towards sampling exposed cases increases, the number of unexposed cases approaches 0, and both the estimated odds ratio and the error ratio approach infinity. This simulation also shows that the greater the true odds ratio, the more labile the error ratio at relatively low levels of bias (because the number of unexposed cases rapidly approaches 0) (Figure 1). This relationship might be counter-intuitive given the oft-repeated mantra that larger true effect sizes are more resilient in the face of bias – this is true only for non-differential biases.

These simulations demonstrate that the introduction of relatively small biases due to increases in non-participation can have a dramatic effect on conclusions drawn, either attenuating or inflating the observed effect size. These simulations manipulated the distribution of cases, but similar results could be produced by manipulating the distribution of controls. Clearly,

researchers must be cautious when interpreting the results of studies with low participation rates, and might wish to seek methods of alleviating the potential biases introduced by non-response.

CONCLUSIONS: IMPACTS ON PUBLIC HEALTH

Internationally, privacy legislation has imposed restrictions on contacting individuals or using data about their care for research purposes. One effect has been to increase non-response in epidemiological studies. While some of this lower participation can be attributed to informed refusal,¹ a greater portion results from logistical challenges such as failure to contact eligible subjects. While impediments to recruitment can inflate individual study costs and durations,⁸ they can also hinder scientific progress by producing studies subject to non-response bias.⁶ Impediments to health research could result in losses of societal benefits by delaying or prohibiting research into disease etiology, diagnosis and treatment, or health services delivery. In the application of public policy, the potential harms of legislation must be weighed against its benefits. Possible options for policy-makers include:

1. Study public opinions regarding research and privacy. A recent population-based survey done in the United Kingdom²⁰ found that the majority of respondents were in favour of allowing the national cancer registry to collect and use identifiable data without requiring prior consent.
2. Jurisdictions in Canada without separate, parallel privacy legislation governing health research could institute such laws. This would allow the application of different restrictions to those on marketing research, for example.
3. Since a large portion of non-response comprises failure to contact, registries could establish an "opt-out" (rather than the current "opt-in") policy for

anonymous and confidential research into certain questions of substantial public health impact. The default would allow inclusion of patients who die or leave the hospital before their wishes may be known, but those who wish to opt-out may do so.

4. Governments might consider protecting research in health information legislation, preventing impediments to health research if requirements (e.g., ethics approval) are met. This would allow researchers legal recourse if data are not released when all criteria are met.

REFERENCES

1. Iversen A, Liddell K, Fear N, Hotopf M, Wessely S. Consent, confidentiality, and the Data Protection Act. *Br Med J* 2006;332:165-69.
2. Yawn BP, Yawn RA, Geier GR, Xia Z, Jacobsen SJ. The impact of requiring patient authorization for use of data in medical records research. *J Fam Pract* 1998;47:361-65.
3. Westrin CG, Nilstun T. The ethics of data utilisation: A comparison between epidemiology and journalism. *Br Med J* 1994;308:522-23.
4. O'Grady KA, Nolan TM. Privacy: Bad for your health? *Med J Aust* 2004;180:307-8.
5. McCarthy DB, Shatin D, Drinkard CR, Kleinman JH, Gardner JS. Medical records and privacy: Empirical effects of legislation. *Health Serv Res* 1999;34:417-25.
6. Tu JV, Willison DJ, Silver FL, Fang J, Richards JA, Laupacis A, et al. Impracticability of informed consent in the Registry of the Canadian Stroke Network. *N Engl J Med* 2004;350:1414-21.
7. Armstrong D, Kline-Rogers E, Jani SM, Goldman EB, Fang J, Mukherjee D, et al. Potential impact of the HIPAA privacy rule on data collection in a registry of patients with acute coronary syndrome. *Arch Intern Med* 2005;165:1125-29.
8. Wolf MS, Bennett CL. Local perspective of the impact of the HIPAA privacy rule on research. *Cancer* 2006;106:474-79.
9. Chen DT, Worrall BB, Brown RD, Jr., Brott TG, Kissela BM, Olson TS, et al. The impact of privacy protections on recruitment in a multicenter stroke genetics study. *Neurology* 2005;64:721-24.
10. Jones J. The effects of non-response on statistical inference. *J Health Soc Policy* 1996;8:49-62.
11. Worrall BB, Chen DT, Meschia JF. Ethical and methodological issues in pedigree stroke research. *Stroke* 2001;32:1242-49.
12. Melton LJ, III. The threat to medical-records research. *N Engl J Med* 1997;337:1466-70.
13. Walley T. Using personal health information in medical research. *BMJ* 2006;332:130-31.

14. Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. Patterns of consent in epidemiologic research: Evidence from over 25,000 responders. *Am J Epidemiol* 2004;159:1087-94.
15. Angus VC, Entwistle VA, Emslie MJ, Walker KA, Andrew JE. The requirement for prior consent to participate on survey response rates: A population-based survey in Grampian. *BMC Health Serv Res* 2003;3:21.
16. Madigan MP, Troisi R, Potischman N, Brogan D, Gammon MD, Malone KE, et al. Characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women. *Int J Epidemiol* 2000;29:793-98.
17. Sogaard AJ, Selmer R, Bjertness E, Thelle D. The Oslo Health Study: The impact of self-selection in a large, population-based survey. *Int J Equity Health* 2004;3(1):3.
18. Kjoller M, Thoning H. Characteristics of non-response in the Danish Health Interview Surveys, 1987-1994. *Eur J Public Health* 2005;15:528-35.
19. Boshuizen HC, Viet AL, Picavet HS, Botterweck A, van Loon AJ. Non-response in a survey of cardiovascular risk factors in the Dutch population: Determinants and resulting biases. *Public Health* 2006;120:297-308.
20. Barrett G, Cassell JA, Peacock JL, Coleman MP. National survey of British public's views on use of identifiable medical data by the National Cancer Registry. *Br Med J* 2006;332:1068-72.
21. Government of British Columbia. 1996. Freedom of Information and Protection of Privacy Act. Available online at: http://www.qp.gov.bc.ca/statreg/stat/F/96165_01.htm (Accessed November 30, 2007).
22. Government of Alberta. 2001. Health Information Act. Available online at: <http://www.oipc.ab.ca/hia/DetailsPage.cfm?id=89> (Accessed November 30, 2007).
23. Government of Manitoba. 1997. Personal Health Information Act. Available online at: <http://web2.gov.mb.ca/laws/statutes/ccsm/p033-5e.php> (Accessed November 30, 2007).
24. Government of Ontario. 2004. Personal Health Information Protection Act. Available online at: http://www.e-laws.gov.on.ca/DBLaws/Statutes/English/04p03_e.htm#BK58 (Accessed November 30, 2007).
25. Government of Québec. 1982. Act Respecting Access To Documents Held By Public Bodies And The Protection Of Personal Information. Available online at: http://www2.publicationsduquebec.gouv.qc.ca/dynamicSearch/telecharge.php?type=2&file=/A_2_1/A2_1_A.html (Accessed November 30, 2007).
26. Government of Canada. 2000. Personal Information Protection and Electronic Documents Act. Available online at: http://www.privcom.gc.ca/legislation/02_06_01_01_e.asp (Accessed November 30, 2007).
27. Government of Canada. 1985. Privacy Act. Available online at: <http://laws.justice.gc.ca/en/P-21/index.html> (Accessed November 30, 2007).

Received: February 27, 2007

Accepted: February 19, 2008