

Social licence and the general public's attitudes toward research based on linked administrative health data: a qualitative study

P. Alison Paprica PhD, Magda Nunes de Melo BScPharm MSc, Michael J. Schull MD MSc

Abstract

Background: Both the research literature and headline news stories indicate that the public cares about how their health data are used. The objective of this study was to learn more about the general public's attitudes toward users and uses of linked administrative health data held by ICES in Ontario, Canada.

Methods: Eight focus groups, with a total of 65 members of the general public, were conducted in urban and northern settings in Ontario, Canada, in 2015 and 2017 using qualitative market research panels established by a market research/public opinion research firm.

Results: Three major themes emerged: (a) the need for assurance about privacy and security, (b) general support for research based on linked administrative health data with some conditions and (c) mixed and more negative reaction when there is private sector involvement. Two minor themes were also derived from the data: (a) low knowledge and understanding of how linked administrative health data are used for research and (b) mixed views on the need to obtain consent when health data do not include identifying information.

Interpretation: The public generally supports research based on linked administrative health data, but there is no blanket approval. Researchers and organizations that hold health data should engage with members of the public to understand and address their concerns about privacy and security and to ensure that research is aligned with social licence, particularly where there is private sector involvement.

Around the world, specialized research centres have developed expertise related to the linkage and analysis of population-wide administrative health data.¹ Canadian examples include ICES in Ontario,² the Manitoba Centre for Health Policy,³ Population Data BC⁴ and the Canadian Institute for Health Information.⁵ These organizations all work with data sets that are created by linking person-level data from different data sets (e.g., prescription drugs, hospital admissions, mortality) then removing or coding identifying information so that research and analyses can be performed while protecting privacy. Recent high-profile reports and initiatives⁶⁻⁸ highlight the potential benefits that could be realized by increasing linkage of, and access to, Canadian health data from these centres and other sites. At the same time, substantial public debate has been generated about questionable practices related to health data, including concerns about private sector access to data under care.data in the United Kingdom⁹ and concerns about privacy and patient consent for My Health Record in Australia.¹⁰ As stewards of health data that cover the entire population, it is the responsibility of organizations like ICES to engage with

the public when considering expanded uses of, and access to, population-wide health data holdings.

A social licence to operate is an informal agreement that is granted by communities and relevant stakeholders to an organization to do certain work.¹¹ Organizations holding a social licence may not even recognize that they have one until it is withdrawn.¹¹ In their analysis of negative public reaction to care.data in the UK, Carter and colleagues note that “the concept of a social licence describes how the expectations of society regarding some activities may go beyond compliance with the requirements of formal regulation; those who do not fulfil the conditions for the social licence (even if formally compliant) may experience ongoing challenge and contestation.”⁹ This

Competing interests: None declared.

This article has been peer reviewed.

Correspondence to: Alison Paprica, alison.paprica@utoronto.ca

CMAJ Open 2019. DOI:10.9778/cmajo.20180099

begs the question, What do we know about the boundaries of social licence for data-intensive health research in Canada? Do members of the general public in Ontario support current and potential expanded uses of what they may reasonably consider to be “their” data, and if yes, under what circumstances? The objective of this qualitative research study was to gain insight into social licence and the general public’s attitudes toward users and uses of linked administrative health data held by ICES in Ontario, Canada.

Methods

Study design

Focus groups were conducted using semistructured discussion guides designed to prompt dialogue among participants (Appendix 1, available at www.cmajopen.ca/content/7/1/E40/suppl/DC1). Each 2-hour focus group had 3 parts: participant reaction to background information about data and ICES, participant views on specific scenarios and research case studies, and time for questions with an ICES representative (P.A.P.). The first set of focus groups in fall 2015 focused mostly on data in general and public sector uses of health data, with some discussion of private sector studies in the last half of each session. In 2016, a decision was made to conduct additional focus groups to learn more about the general public’s views about private sector involvement in research based on linked administrative data. Research case studies designed to represent ones that would interest and involve the private sector were developed, and a second set of focus groups concentrating on the private sector and linked administrative health data was conducted in spring 2017.

Data collection

Purposive sampling was used to select heterogeneous participants with varying perspectives for each focus group.¹² The target group size of 8–10 people per focus group was selected so that groups would be large enough to include differing perspectives but not so large that group size inhibited some participants from contributing to the dialogue.¹³ The team decided to hold 2 sessions in each location in each study year to decrease the risk that a small number of vocal participants in a single focus group would have a disproportionate or undue impact on the study findings. For practical reasons, we recruited participants for each set of 4 focus groups all at once rather than recruiting participants 1 session at a time. The study team was open to the possibility of conducting additional focus groups if the sample size proved to be insufficient. There was no crossover between focus groups, that is, no person participated in more than 1 focus group. The study made use of qualitative market research panels established by a market research/public opinion research firm (Ipsos) according to the quality guidelines of the Marketing Research and Intelligence Association. Potential participants from the panels were contacted by telephone by Ipsos and screened using a recruitment questionnaire to ensure that the perspectives of northern Ontario residents and urban residents would be reflected and that men and women of varying ages, education and income

levels would be included. The screening process also included questions about level of trust in scientists and level of support for data-intensive health research to ensure that each focus group included participants with varying views on those topics. As part of the recruitment process, participants were notified of the purpose of the focus groups (i.e., to learn about the general public’s views on users and uses of linked administrative health data). Participants were also informed of the purpose of each focus group, in writing, as part of the process to obtain their informed consent in writing immediately before each session and verbally at the start of each focus group. At the end of each session, participants were provided with a cheque for \$75 as a token of appreciation for their time and participation.

Focus groups were moderated by a professional facilitator/researcher (Vanessa Chan, MA, female) who had more than 5 years’ experience with Ipsos performing qualitative research related to social and public affairs issues. This enabled the research team to benefit from the experience of a highly skilled facilitator, provided an environment in which participants would be more likely to feel free to express negative opinions about ICES than if a member of the ICES staff were facilitating, and allowed the research team to focus on observing and understanding the participant discussion. The sessions took place in facilities designed for focus groups, with audio-recording capabilities and space for observation behind a one-way mirror. The discussions followed semistructured discussion guides (Appendix 1), which allowed for free-flowing discussion as well as facilitated discussion of written examples, with prompts on certain questions. Members of the research team (P.A.P. in all focus groups, M.N.M. and M.J.S. in some focus groups) and, sometimes, 1 or 2 additional staff from ICES observed the focus groups from behind a one-way mirror and took independent field notes (P.A.P., M.N.M.) during the sessions. Focus group participants were informed that researchers were in attendance behind the one-way mirror and that sessions were audio-recorded. Audio-recordings were transcribed verbatim by Ipsos. With the exception of statements that were attributed to the facilitator, the transcripts did not attribute the quotes to specific individuals. The purpose of the focus groups was to generate and analyze interactions between participants¹⁴ and thus there was less need to focus on the comments of any one individual.

Data analysis

Preliminary analyses performed after the 2015 and 2017 data collections were completed indicated that the reactions and responses of participants were consistent across the 2 sets of focus groups, despite differences in the specific scenarios and research case studies presented in 2015 and 2017. Accordingly, the data from 2015 and 2017 were analyzed as a whole by the research team. Data were analyzed by P.A.P. and M.N.M. using a qualitative descriptive approach, which is a naturalistic form of inquiry that aims to remain “data-near” while inductively interpreting and thematically grouping and detailing respondent experiences, beliefs and expectations.^{15,16} P.A.P. led the development of the descriptive coding framework on the basis of the verbatim transcripts and field notes taken during

the focus group sessions. The transcripts were read and reread as coding was performed independently by P.A.P. and M.N.M. using a combination of Microsoft Word and Microsoft Excel software. P.A.P. and M.N.M. used an inductive analytic approach to derive themes and subthemes on the basis of the data, and themes and subthemes were socialized and refined through discussion between P.A.P., M.N.M. and M.J.S. M.J.S. reviewed portions of transcripts, with a focus on the key-coded statements that helped define the themes and subthemes. Differences in opinion between P.A.P., M.N.M. and M.J.M. were resolved through iterative discussions. Major themes were those that were prominent in the data from multiple focus group sessions and multiple participants. Each major theme had multiple subthemes. Minor themes were also derived from the data from multiple focus groups and participants, but these were less prominent and indirectly related to the main objective of learning about the attitudes of the general public toward research based on linked administrative health data. Review and coding of transcripts stopped when inductive thematic saturation was achieved,¹⁷ that is, when P.A.P., M.N.M. and M.J.S. agreed that additional coding and thematic analysis would not result in any new codes or themes. The research team was open to the possibility of recruiting additional participants if there were insufficient data to identify

themes; however, on the basis of the finding that themes were strong and pervasive across both the 2015 and 2017 focus groups, no additional participants were recruited.

Ethics approval

The studies were approved by the Research Ethics Board of the Sunnybrook Research Institute in Toronto.

Results

Eight focus groups were held with a total of 65 members of the general public in 2015 and 2017 (Table 1). Four were held in northern Ontario (2 in Thunder Bay in 2015 and 2 in Sudbury in 2017) and 4 in Toronto. Each session was 2 hours long. The focus groups in 2015 focused mostly on data in general, and public sector uses of health data, with some discussion of private sector studies in the last half of each session. The focus groups in 2017 discussed the general public’s views about private sector involvement in research based on linked administrative data.

Three major themes were identified: (a) the need for assurance about privacy and security, (b) general support for research based on linked administrative health data with some conditions and (c) mixed and more negative reaction when

Table 1: Participants of focus groups

| Characteristic | Location and date, no. (%) | | | |
|---------------------------|--|--|---|---|
| | Northern Ontario: Thunder Bay Sept. 23, 2015 (n = 17 across 2 focus groups) | Urban: Toronto Sept. 24, 2015 (n = 14 across 2 focus groups) | Northern Ontario: Sudbury Mar. 29, 2017 (n = 18 across 2 focus groups) | Urban: Toronto Mar. 30, 2017 (n = 16 across 2 focus groups) |
| Male | 8 (47) | 5 (36) | 8 (44) | 8 (50) |
| Female | 9 (53) | 9 (64) | 10 (66) | 8 (50) |
| Age, yr | | | | |
| 18–34 | 3 (18) | 5 (36) | 5 (28) | 2 (13) |
| 35–44 | 3 (18) | 4 (29) | 5 (28) | 3 (19) |
| 45–54 | 2 (12) | 3 (21) | 2 (11) | 3 (19) |
| ≥ 55 | 9 (53) | 2 (14) | 6 (33) | 8 (50) |
| Education | | | | |
| High school or lower | 5 (29) | 1 (7) | 6 (33) | 2 (13) |
| Community college | 7 (41) | 4 (29) | 11 (61) | 5 (31) |
| University or higher | 5 (29) | 9 (64) | 1 (6) | 9 (56) |
| Income (before taxes), \$ | | | | |
| ≤ 20 000 | 2 (12) | 0 (0) | 4 (22) | 0 (0) |
| 20 001–39 999 | 8 (47) | 1 (7) | 3 (17) | 4 (25) |
| 40 000–59 999 | 5 (29) | 2 (14) | 4 (22) | 4 (25) |
| 60 000–79 999 | 2 (12) | 1 (7) | 2 (11) | 0 (0) |
| 80 000–99 999 | 0 (0) | 4 (29) | 3 (17) | 5 (31) |
| ≥ 100 000 | 0 (0) | 2 (14) | 2 (11) | 3 (19) |
| Unable to code | | 4 (29) | | |

there is private sector involvement (Box 1). As indicated in Appendix 2 (available at www.cmajopen.ca/content/7/1/E40/suppl/DC1), each major theme had multiple subthemes. The

Box 1: Examples of verbatim statements illustrating major themes

Major theme 1: the need for assurance about privacy and security

- What was the last one [hacking incident], with the government. Anybody recall that? ... CRA, oh Canada Revenue [Agency], yeah, that's what it was. It was a huge one, yeah.
Thunder Bay 2015 — Group 2
- I liked that the personal information is removed, but I have the same thought — like really? [disbelief] It's great on paper, but someone's going to make a connection somewhere or something's going to happen. *Toronto 2015 — Group 1*
- Because it's the digital age, now, too, and anything can be hacked. Look at the guy who just hacked Yahoo from here in Ontario. *Sudbury 2017 — Group 2*
- Well I think because it's health data, it's really important to keep it safeguarded. It's not just some random information. It's personal information. Really personal information.
Toronto 2015 — Group 2
- I found it encouraging that the information and privacy commissioner has an oversight over it and it renews every 3 years. I found that encouraging. Someone's keeping an eye on it. *Sudbury 2017 — Group 1*
- Information can somehow slip out. Especially when there's more people involved. *Thunder Bay 2015 — Group 2*

Major theme 2: general support for research based on linked administrative health data with some conditions

- For me, I think this is a really great use of the information that's already being collected. It's sitting there, it's waiting to be used. *Sudbury 2017 — Group 1*
- It's already been collected ... it's there and it can be used. So why shouldn't they, if they can get access? It would be so helpful. *Toronto 2015 — Group 2*
- If they're collecting 25 years of data they're going to discover that certain medications are unhealthy or not fit for people, so that's — there's a big benefit in this for sure.
Toronto 2017 — Group 2
- [Indicating concern about potential for misuse of data] And then they combine all that together, and they say, okay, well, this person has got this and this and this. Wasting medication or treatment or whatever on this person, beyond this age is useless. Let's just let this person die.
Thunder Bay 2015 — Group 1

Major theme 3: mixed and more negative reaction when there is private sector involvement

- I'd rather not have a private company because I feel like they're just out to make money. *Sudbury 2017 — Group 2*
- People who are really good at this kind of work always tend to work for the private sector because the money's better ... someone with a fresh idea can come in and see something that's totally different. *Toronto 2015 — Group 2*
- I guess I just think maybe they [the private sector] could fund their own research. I'm not sure the taxpayers should pay for it. But I guess, as you said, if they're giving us an appropriate price or a better drug being released, then I guess it's okay.
Toronto 2017 — Group 1

research team also identified 2 minor themes: (a) low knowledge and understanding of research based on linked administrative health data and (b) mixed views on the need to obtain consent when health data do not include identifying information (Box 2). No major differences were noted between the 2015 and 2017 focus groups or between the views of northern participants and Toronto urban participants.

Major theme 1: the need for assurance about privacy and security

The main concerns about research based on linked administrative health data related to the security of personal data generally (e.g., the hacking of the Canada Revenue Agency). Participants responded positively to information about the ICES process for removing or coding identifying information before data are made available to researchers, and about the legislated oversight provided by the Information and Privacy Commissioner of Ontario.¹⁸ The process of removing names and other direct identifiers was appreciated, but many participants did not see it as sufficient assurance. Even when fully informed of privacy and security safeguards, participants noted that risks unavoidably increase when there are more people and organizations accessing data (Box 1 and Appendix 2).

Major theme 2: general support for research based on health administrative data, with some conditions

Generally, health data were viewed as an asset that should be used for research, and focus group participants supported

Box 2: Examples of verbatim statements illustrating minor themes

Minor theme 1: low knowledge and understanding of research based on linked administrative health data

- Is this actually happening today, where they're collecting a lot of data? *Toronto 2017 — Group 2*
- We don't know them (ICES) so how can we trust them. We've never heard of them until today, so we can't possibly trust them. *Thunder Bay 2015 — Group 1*
- [Indicating that the participant has confused research based on linked administrative health data with integration of data to inform individual patient care] It's like one-stop shopping. Once you get into the system, all your information is in one place, for your services or programs or health care, whatever that you may need to link up to, to help you in your health.
Thunder Bay 2015 — Group 2

Minor theme 2: mixed views on the need to obtain consent when health data do not include identifying information

- I don't think consent is needed as much to gather data when it's nameless and faceless. *Toronto 2017 — Group 2*
- So the first thing is no one really tells you when you go to the doctor that your data will be shared, right? That's number one. We don't know. They haven't gotten anyone's consent.
Toronto 2015 — Group 2
- And I think if it's something the company's doing because they want stats on how their drugs are being used, then I think people should be consenting to it. I don't need to make them even more profitable than they are without my consent.
Sudbury 2017 — Group 2

research based on linked administrative health data, with some conditions. Support was strongest when people saw a public benefit and agreed with the purposes for which studies were conducted (e.g., focus group participants strongly supported using administrative health data to study the long-term safety and efficacy of prescription drug products). In contrast, participants expressed concerns when they thought the results of a study could be misused or disadvantage certain groups (e.g., seniors, people not adhering to their prescription drug medications) (Box 1 and Appendix 2).

Major theme 3: mixed and more negative reaction when there is private sector involvement

Some focus group participants expressed concerns about private sector involvement in studies based on linked administrative health data (e.g., the concern that increased pharmaceutical product sales and profit, not public benefit, would be the primary motive). In contrast, others saw benefits of private sector involvement including more skilled people being able to use the data and the potential development of new products and services. Several participants wanted some form of reciprocity when public data are used in studies funded by the private sector (e.g., in the form of lower drug prices [Box 1 and Appendix 2]).

Minor theme 1: low knowledge and understanding of research based on linked administrative health data

Most participants were not aware of studies based on linked administrative health data, despite regular media coverage of them. Several participants misunderstood the practice of linking administrative health data sets for studies at the population level and confused it with efforts to bring together data from different health care service providers to improve care for individual patients, even after the moderator provided clarification. In some instances, participants' lack of prior knowledge about research based on linked administrative health data led to them having concerns about transparency and trust (Box 2 and Appendix 2).

Minor theme 2: mixed views on the need to obtain consent when health data do not include identifying information

The subset of focus group participants who expressed views about consent had varying opinions. Some felt that consent should always be obtained even when study participants in data sets are not identifiable. Others were direct in stating their views that consent is not necessary if identifying information is removed before data are used for research (Box 2 and Appendix 2).

Interpretation

Generally, the participants in the Toronto and northern Ontario focus groups were supportive of research based on linked administrative health data providing that there was assurance about privacy and security, but they cared about details including whether there would be a public benefit from

a study, who would have access to health data and whether there could be a potential downside or negative impact. Repeated confusion about the nature and purpose of research based on linked population-wide data (i.e., distinct from analyses in which data are linked to inform the clinical care of an individual) suggests that the topic is hard to understand and that there is low awareness of research based on linked administrative health data among members of the general public in Ontario at present. There were mixed views regarding whether consent is required when health data sets do not contain identifying information.

The results of this research study are consistent with literature^{19,20,21} and the themes identified in a recent systematic review²² that included 25 publications from the UK, the United States, Canada and other countries. Findings from that systematic review that are reinforced by this study include the following: general widespread support for uses of data in health research with some conditions, concerns about privacy and security, the requirement that there be a public benefit, more trust in public sector studies than in private sector studies, and varying views on the need for consent. This study identified the new subtheme of administrative health data being an asset that should be used for public benefit, and it provides additional information about how public views are influenced by information about breaches, hacking and violation of trust outside of the health and research sectors. It also begins to identify the types of studies that the public supports provided that appropriate controls are in place (e.g., studies of the long-term safety and efficacy of a prescription drug product).

Given the public's concerns about uses of data generally, social licence for data-intensive health research is essential. Carter and colleagues note that "poorly informed understanding of the social licence for secondary use of personal medical data, and a failure to recognise that legal authority might not be enough to secure the social licence, seems to have been at the heart of the controversy underlying care.data."⁹ There are indications that social licence for data-intensive health research varies by jurisdiction. For example, in Denmark, where there is a long-standing history of citizen support for the use of public data in research, Danish researchers approach patients about participation in database-based trials directly with little to no involvement of health care providers,²³ but in Scotland potential trial participants are generally contacted by someone within the circle of health care providers that patients would reasonably expect to have access to patient data.²³ Regarding informed consent, it is the authors' view that informed consent can contribute to social licence, but it does not constitute the complete answer in all circumstances because there are public benefits that can be realized only through studies based on population-wide nonconsented data (e.g., the withdrawal of Vioxx from the market,²⁴ restrictions on mobile phone use while driving²⁵ and the identification of the magnitude of the opioid epidemic²⁶ all were based on studies of population-wide nonconsented data). Further, consent may not be truly informed in cases where researchers cannot describe all the potential future uses of health data.²⁷

This research study, and the literature, indicate that the general public wants society to realize the benefits that can be derived from research based on linked administrative health data, but it is incumbent on the parties involved in research and data sharing to be transparent and to involve and engage with members of the public in an ongoing and authentic manner to ensure alignment with social licence. As illustrated by news reports of growing concerns following Cambridge Analytica's reported misuse of Facebook data,²⁸ lack of trust in a sector or organization can spread^{29,30} and have consequences for other practices that rely on data. Public involvement and transparency are essential to building and maintaining trust. Informational transparency — publicizing information about what is being done — is a start, but it is unlikely to yield the benefits that could be realized by involving patients and the public in governance and decision-making practices to achieve “participatory transparency” and “accountability transparency.”²¹ As noted in the *International Consensus Statement on Public Involvement and Engagement with Data-Intensive Health Research*, a key premise is that the public should not be characterized as a problem to be overcome.³¹ Involving the public, and focusing on the users and uses of health data that they support, can help ensure sustainable and beneficial data-intensive health research that is aligned with public values.³¹

Limitations

This study has limitations. Foremost, results may not be generalizable across or outside of Ontario. It is possible that participants from other settings (e.g., rural Ontario, remote northern Ontario or other jurisdictions) or specific subpopulations would have different views. It is also possible that increasing the number of focus groups or selecting different participants may have resulted in differences in the themes and subthemes that were identified by the study team. Second, the discussion guides were informed and reviewed by people who were not on the research team, but they were not pilot tested or validated. The team decided not to track or attribute quotes to specific individuals within focus groups and briefly discussed the implications of this decision. Given the multiple references made to privacy and security concerns outside of the health and research sectors, it is possible that views will change on the basis of recent public events such as the one involving Facebook and Cambridge Analytica.²⁸ In addition, participants' difficulty understanding the nature and purpose of research based on linked administrative health data may have affected their ability to understand and respond to the sample research case studies with which they were presented. Finally, there are uses of linked administrative health data (e.g., helping clinical trial recruitment focus on sites with large numbers of eligible patients, artificial intelligence applications) that were not presented to focus group participants and warrant further study.

Conclusion

This qualitative study found that members of the Ontario public see data as an asset that should be used, and they generally support research based on linked administrative health

data, but there is no blanket approval. Researchers and organizations holding health data should engage with and involve members of the public to ensure that data-intensive health research is trustworthy and within the bounds of social licence. If researchers focus on conducting studies that have a clear public benefit and respect and address public concerns about privacy and private sector involvement, public support is likely to increase, enhancing the impact and the sustainability of research based on linked administrative health data.

References

1. International Population Data Linkage Network Member Directory [homepage]. Available: https://www.ipdln.org/member_directory (accessed 2018 Oct. 16).
2. About ICES. Toronto: ICES; 2018. Available: <https://www.ices.on.ca/About-ICES/> (accessed 2018 June 8).
3. Manitoba Centre for Health Policy [homepage]. Winnipeg: University of Manitoba; 2018. Available: http://umanitoba.ca/faculties/health_sciences/medicine/units/chs/departamental_units/mchp/ (accessed 2018 Oct. 16).
4. Population Data BC [homepage]; updated 2018 Sept. 12. Available: <https://www.popdata.bc.ca/> (accessed 2018 Oct. 16).
5. Ottawa: Canadian Institute for Health Information [homepage]. Available: <https://www.cihi.ca> (accessed 2018 Oct. 16).
6. The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation CoCA. *Accessing health and health-related data in Canada*. Ottawa: Council of Canadian Academies; 2015.
7. *Investing in Canada's future: strengthening the foundations of Canadian research*. Advisory Panel on Federal Support for Fundamental Science; 2017.
8. Canada's strategy for patient-oriented research (SPOR). Ottawa: Canadian Institutes of Health Research; 2018. Available: <https://www.researchnet-recherchenet.ca/rnr16/vw/OpprntnyDtls.do?prog=2794&view=currentOpps&org=CIHR&type=EXACT&resultCount=25&sort=program&next=1&call=1&masterList=true> (accessed 2018 June 8).
9. Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. *J Med Ethics* 2015;41:404-9.
10. Kemp K, Baer Arnold B, Vaile D. My health record still isn't safe enough to proceed. It needs more than a band-aid fix. *ABC News* 2018 Aug. 1. Available: <https://www.abc.net.au/news/2018-08-02/my-health-record-still-not-safe/10063026> (accessed 2018 Oct. 16).
11. Rooney D, Leach J, Ashworth P. Doing the social in social license. *Soc Epistemology* 2014;28:209-18.
12. Patton M. *Qualitative design and data collection. Qualitative evaluation and research methods*. Beverley Hills (CA): Sage; 1990:169-86.
13. Freeman T. 'Best practice' in focus group research: making sense of different views. *J Adv Nurs* 2006;56:491-7.
14. Smithson J. Using and analysing focus groups: limitations and possibilities. *Int J Soc Res Methodol* 2000;3:103-19.
15. Sandelowski M. Whatever happened to qualitative description? *Res Nurs Health* 2000;23:334-40.
16. Sandelowski M. What's in a name? Qualitative description revisited. *Res Nurs Health* 2010;33:77-84.
17. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant* 2018; 52:1893-907.
18. Toronto: Information and Privacy Commissioner of Ontario [homepage]; 2018. Available: <https://www.ipc.on.ca/> (accessed 2018 June 8).
19. Moat KA, Bullock H. Panel summary: Integrating data across sectors for public service improvement in Ontario. Hamilton (ON): McMaster Health Forum; 2016.
20. Wellcome Trust. *The one-way mirror: public attitudes to commercial access to health data*. London (UK): Ipsos MORI; 2016.
21. Aitken M, Cunningham-Burley S, Pagliari C. Moving from trust to trustworthiness: experiences of public engagement in the Scottish Health Informatics Programme. *Sci Public Policy* 2016;43:713-23.
22. Aitken M, de St Jorre J, Pagliari C, et al. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Med Ethics* 2016;17:73.
23. *Best brains exchange — using administrative, electronic medical record and patient health record data in clinical trials*. Ottawa: Canadian Institutes of Health Research; 2016. Available: <http://www.cihr-irsc.gc.ca/e/13931.html> (accessed 2018 Oct. 16).
24. Mamdani M, Rochon P, Juurlink DN, et al. Effect of selective cyclooxygenase 2 inhibitors and naproxen on short-term risk of acute myocardial infarction in the elderly. *Arch Intern Med* 2003;163:481-6.
25. Min ST, Redelmeier DA. Car phones and car crashes: an ecologic analysis. *Can J Public Health* 1998;89:157-61.

26. Gomes T, Juurlink DN, Dhalla IA, et al. Trends in opioid use and dosing among socio-economically disadvantaged patients. *Open Med* 2011;5:e13-22.
27. O'Doherty KC, Christofides E, Yen J, et al. If you build it, they will come: unintended future uses of organised health data collections. *BMC Med Ethics* 2016;17:54.
28. McMahon T. Facebook broadens estimate of data misuse to 87 million people, including more than 600 000 Canadians. *The Globe and Mail* [Toronto] updated 2018 Apr. 5. Available: <https://www.theglobeandmail.com/world/article-facebook-moves-to-improve-data-collection-transparency-as-zuckerberg/> (accessed 2018 June 8).
29. Smee B. Facebook's data changes will hamper research and oversight, academics warn. *The Guardian* [London] 2018 Apr. 25. Available: <https://www.theguardian.com/technology/2018/apr/25/facebooks-data-changes-will-hamper-research-and-oversight-academics-warn> (accessed 2018 June 8).
30. Curry B. Ethics committee urged to look beyond Facebook with privacy study. *The Globe and Mail* [Toronto] 2018 Apr. 16. Available: <https://www.theglobeandmail.com/politics/article-ethics-committee-urged-to-look-beyond-facebook-with-privacy-study/> (accessed 2018 June 8).
31. Aitken M, Tully M, Porteus C, et al. Consensus statement on public involvement and engagement with data-intensive health research [abstract]. *IJPD* 2018;3:246.

Affiliations: ICES (Paprica, Nunes de Melo, Schull), Toronto, Ont.; Institute for Health Policy, Management and Evaluation (Paprica, Schull), University of Toronto, Toronto, Ont.; Sunnybrook Health Sciences Centre (Schull), Toronto, Ont.

Contributors: All authors contributed to the design and conception of the study, attended focus groups, critically reviewed drafts of the manuscript, discussed and refined the themes and subthemes, and approved the final version submitted for publication. Magda Nunes de Melo led the literature review. Magda Nunes de Melo and Alison Paprica led the work to

design private sector example studies, independently coded transcripts and performed analyses to identify themes and subthemes. Alison Paprica had the primary responsibility for the descriptive coding framework and was the lead for preparation of the manuscript. All of the authors gave approval of the final version for publication and agreed to be accountable for all aspects of the work.

Acknowledgements: The authors thank Don Willison, who provided helpful advice on how to structure focus groups to maximize the likelihood that participants would understand what research based on linked administrative health data is before being asked to comment on specific research case studies. Vanessa Chan is acknowledged for her expert facilitation skills and services. The authors also thank Mary Tully for her extensive contributions to the research case studies presented in the 2017 focus groups and Fiona Miller for her advice on qualitative research methods.

Disclaimer: This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results and conclusions reported in this article are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred. The Marketing Research and Intelligence Association (MRIA) is a Canadian not-for-profit association representing the market intelligence and survey research industry. The authors have no affiliation with the MRIA. The market research/public opinion research firm (Ipsos) that recruited participants and provided the facilities and facilitator for the focus groups is a member of the MRIA.

Supplemental information: For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/7/1/E40/suppl/DC1.