



**Social licence and the general public’s attitudes toward research based on linked administrative health data, a qualitative study**

Journal:	<i>CMAJ Open</i>
Manuscript ID	CMAJOpen-2018-0099
Manuscript Type:	Qualitative
Date Submitted by the Author:	25-Jun-2018
Complete List of Authors:	Paprica, P. Alison; Institute for Clinical Evaluative Sciences, ; University of Toronto Dalla Lana School of Public Health, Institute for Health Policy, Management and Evaluation Melo, Magda; Institute for Clinical Evaluative Sciences, Schull, Michael; Institute for Clinical Evaluative Sciences; Sunnybrook Health Sciences Centre, Institute of Clinical Evaluative Sciences, Division of Emergency Medicine, Dept of Medicine
Keywords:	Qualitative research, Health services research, Medical informatics
More Detailed Keywords:	Public involvement in research, Linked administrative health data, Data-intensive health research, Social licence
Abstract:	<p>Background: The research literature, and front-page news about data and privacy, clearly indicate that the public cares about how their data are used. The objective of this qualitative research study was to learn more about the general public’s views about users and uses of linked administrative health data in Ontario, Canada.</p> <p>Methods: A total of eight (8) focus groups were conducted in urban and northern settings in Ontario, Canada in 2015 and 2017. The participants were 63 members of the general public including men and women of varying ages, educational backgrounds and income levels.</p> <p>Findings: Three major themes emerged: (i) the need for assurance about privacy and security, (ii) general support for research based on linked administrative health data with some conditions, (iii) mixed and more negative reactions when there is private sector involvement. Two minor themes were also derived from the data, (i) low knowledge and understanding of how linked administrative health data are used for research and (ii) varying opinions about whether patient consent is needed in cases where research datasets don’t include identifying information.</p> <p>Interpretation: The public generally supports research based on linked administrative health data, but there is no blanket approval. Research and health data holding organizations 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,</p>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

	particularly where there is private sector involvement. Public engagement and involvement can increase transparency and strengthen data-intensive health research.

SCHOLARONE™  
Manuscripts

Confidential

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

P. Alison Paprica, Magda Nunes de Melo, Michael J. Schull

## Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist

Developed from:

Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*. 2007. Volume 19, Number 6: pp. 349 – 357

No. Item	Guide questions/description	Reported on Page #
<b>Domain 1: Research team and reflexivity</b>		
<i>Personal Characteristics</i>		
1. Interviewer/facilitator	Which author/s conducted the interview or focus group?	Facilitator named on Page 2  Involvement of PAP in each focus group described on Page 2 and Page 6  Attendance of MNM and MJS at some focus groups described on Page 6
2. Credentials	What were the researcher's credentials? E.g. PhD, MD	Facilitator's credentials reported on Page 2  Other researchers' credentials reported on Page 1
3. Occupation	What was their occupation at the time of the study?	Page 2
4. Gender	Was the researcher male or female?	Facilitator was female, reported on Page 2
5. Experience and training	What experience or training did the researcher have?	Facilitator experience reported on Page 2
<i>Relationship with participants</i>		
6. Relationship established	Was a relationship established prior to study commencement?	N/A there was no prior relationship

No. Item	Guide questions/description	Reported on Page #
7. Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	The participants knew the facilitator worked for Ipsos and were introduced to, and knew the title and role of, PAP at ICES Page 2
8. Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	Page 2
<b>Domain 2: study design</b>		
<i>Theoretical framework</i>		
9. Methodological orientation and Theory	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	Pages 2-3
<i>Participant selection</i>		
10. Sampling	How were participants selected? e.g. purposive, convenience, consecutive, snowball	Page 2
11. Method of approach	How were participants approached? e.g. face-to-face, telephone, mail, email	Page 2
12. Sample size	How many participants were in the study?	Page 2 and Table 1 on Page 7
13. Non-participation	How many people refused to participate or dropped out? Reasons?	N/A Not tracked or reported
<i>Setting</i>		
14. Setting of data collection	Where was the data collected? e.g. home, clinic, workplace	Page 2
15. Presence of non-participants	Was anyone else present besides the participants and researchers?	Page 2
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	Page 2 and Table 1 on Page 7
<i>Data collection</i>		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Page 2
18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	N/A no repeat interviews carried out

No. Item	Guide questions/description	Reported on Page #
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Page 2
20. Field notes	Were field notes made during and/or after the inter view or focus group?	Page 2
21. Duration	What was the duration of the inter views or focus group?	Page 2
22. Data saturation	Was data saturation discussed?	N/A data saturation was not discussed
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	N/A transcripts were not returned to participants
<b>Domain 3: analysis and findings</b>		
<i>Data analysis</i>		
24. Number of data coders	How many data coders coded the data?	Page 2
25. Description of the coding tree	Did authors provide a description of the coding tree?	N/A no explicit coding tree was used
26. Derivation of themes	Were themes identified in advance or derived from the data?	Pages 2-3
27. Software	What software, if applicable, was used to manage the data?	Page 3
28. Participant checking	Did participants provide feedback on the findings?	N/A participants did not provide feedback on the findings
<i>Reporting</i>		
29. Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	Boxes 1-5 on Pages 8-11
30. Data and findings consistent	Was there consistency between the data presented and the findings?	Yes, there was consistency between data presented and findings, see pages 3-6 and 8-11
31. Clarity of major themes	Were major themes clearly presented in the findings?	Pages 3-6 and 8-10
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Pages 3-6 and 11

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

P. Alison Paprica, Magda Nunes de Melo, Michael J. Schull

## Abstract

**Background:** The research literature, and front-page news about data and privacy, clearly indicate that the public cares about how their data are used. The objective of this qualitative research study was to learn more about the general public's views about users and uses of linked administrative health data in Ontario, Canada.

**Methods:** A total of eight (8) focus groups were conducted in urban and northern settings in Ontario, Canada in 2015 and 2017. The participants were 63 members of the general public including men and women of varying ages, educational backgrounds and income levels.

**Findings:** Three major themes emerged: (i) the need for assurance about privacy and security, (ii) general support for research based on linked administrative health data with some conditions, (iii) mixed and more negative reactions when there is private sector involvement. Two minor themes were also derived from the data, (i) low knowledge and understanding of how linked administrative health data are used for research and (ii) varying opinions about whether patient consent is needed in cases where research datasets don't include identifying information.

**Interpretation:** The public generally supports research based on linked administrative health data, but there is no blanket approval. Research and health data holding organizations 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. Public engagement and involvement can increase transparency and strengthen data-intensive health research.

## Introduction

Canada is home to several research centres that have a long-standing role in performing research studies on linked population-wide administrative health data, including the Institute for Clinical Evaluative Sciences (ICES) in Ontario, Canada (1). Recent high profile reports and initiatives (2),(3),(4) highlight the potential benefits that could be realised 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 through questionable practices around use and sharing of personal data, including recent events with Facebook data and Cambridge Analytica (5) and care.data in the United Kingdom (6). As stewards of data that covers the entire population, it is the responsibility of organizations like ICES to engage with the public when considering how to derive additional benefits from population-wide health data holdings.

Carter, Laurie and Dixon-Woods 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

1  
2  
3 compliant) may experience ongoing challenge and contestation” (6). This begs the question, what do we  
4 know about the boundaries of social licence for data-intensive health research in Canada? Do members  
5 of the Canadian general public support current and potential expanded uses of what they may  
6 reasonably considered to be “their” data, and if yes, under what circumstances? The objective of this  
7 qualitative research study was to gain insight into social licence and the general public’s attitudes  
8 toward users and uses of linked administrative health data in Ontario, Canada.  
9  
10

## 11 Methods

12  
13 This qualitative research study included eight two-hour focus groups in 2015 and 2017: four in northern  
14 Ontario (Thunder Bay and Sudbury) and four in Toronto. A total of 63 participants took part in the  
15 research (Table 1). The studies were approved by the Research Ethics Board of the Sunnybrook  
16 Research Institute in Toronto, Ontario, Canada.  
17

18 Purposive sampling was used. The study made use of qualitative market research panels established by  
19 a market research / public opinion research firm (Ipsos) according to the quality guidelines of the  
20 Marketing Research and Intelligence Association. Potential participants from the panels were contacted  
21 by telephone by Ipsos and screened using a recruitment questionnaire to ensure that the perspectives  
22 of northern Ontario residents and urban residents would be reflected, and that participants would  
23 include men and women or varying ages, education and income levels. The screening process also  
24 included questions about level of trust in scientists and level of support for data-intensive health  
25 research to ensure that each focus group included participants with varying views on those topics. As  
26 part of the recruitment process, participants were notified of the purpose of the focus groups, i.e., to  
27 learn about the general public’s views on users and uses of linked administrative health data.  
28 Participants were also informed of the purpose of each focus group, in writing, as part of the process to  
29 obtain their written informed consent immediately before each session, and verbally at the start of each  
30 focus group.  
31  
32  
33  
34

35 Focus groups were moderated by a professional facilitator/researcher (Vanessa Chan, MA, female) who  
36 has more than five years’ experience with Ipsos performing qualitative research related to social and  
37 public affairs issues. This enabled the research team to benefit from the experience of a highly skilled  
38 facilitator, provided an environment in which participants would be more likely to feel free to express  
39 negative opinions about ICES than if a member of the ICES staff were facilitating, and allowed the  
40 research team to focus on observing and understanding the participant discussion.  
41  
42

43 The sessions took place in facilities designed for focus groups, i.e., with audio-recording capabilities and  
44 space for observation behind a one-way mirror. The discussions followed a semi-structured discussion  
45 guide which allowed for free-flowing discussion as well as facilitated discussion of written examples,  
46 with prompts on certain questions. Members of the research team observed the focus groups from  
47 behind a one-way mirror and took independent field notes during the sessions. Focus group  
48 participants were informed that researchers were in attendance behind the one-way mirror, and that  
49 sessions were audio-recorded.  
50  
51

52 Each focus group started with a general discussion about organizations (in any sector) that collect data  
53 and how that data is used, followed by a brief overview of ICES and the processes ICES uses to protect  
54 privacy. The moderator then facilitated discussions about examples of ICES research studies and  
55 examples of studies with private sector involvement intended to represent the kinds of studies that  
56  
57  
58  
59  
60

1  
2  
3 private sector organizations may be interested in performing. PAP attended all focus groups sessions  
4 and met with participants at the end of each focus group to interact directly with them and answer their  
5 questions.  
6

7 Audio-recordings were transcribed verbatim. Data were analyzed using a qualitative descriptive  
8 approach, which is a naturalistic form of inquiry that aims to remain “data-near” while inductively  
9 interpreting and thematically grouping and detailing respondent experiences, beliefs and expectations  
10 (7),(8) (Sandelowski, 2000, 2010). PAP led the development of the descriptive coding framework based  
11 the verbatim transcripts and field notes taken during the focus group sessions. The transcripts were read  
12 and re-read as coding was performed independently by PAP and MNM using a combination of Microsoft  
13 Word and Microsoft Excel software. PAP and MNM used an inductive analytic approach to derive  
14 themes and subthemes based on the data, and socialized and refined themes and subthemes through  
15 discussion between PAP, MNM and MJS. MJS reviewed portions of transcripts, with a focus on the key  
16 coded statements that helped define the themes and subthemes. Major themes were those that were  
17 prominent based on the data from multiple focus group sessions and multiple participants. In this study  
18 each, major theme had multiple subthemes. Minor themes were also derived from the data from  
19 multiple focus groups and participants, but were less prominent and indirectly related to the main  
20 objective of learning about the attitudes of the general public toward research based on linked  
21 administrative health data. Review and coding of transcripts stopped after thematic saturation was  
22 reached.  
23  
24  
25  
26

## 27 Results

28  
29 Three major themes, each with multiple subthemes, were identified: (i) the need for assurance about  
30 privacy and security, (ii) general support for research based on linked administrative health data with  
31 some conditions and (iii) mixed and more negative reaction when there is private sector involvement.  
32 The research team also identified two minor themes: (i) low knowledge and understanding of research  
33 based on linked administrative health data and (ii) mixed views on the need to obtain consent when  
34 health data don't include identifying information.  
35  
36

37 **Major Theme 1: the need for assurance about privacy and security.** The main concerns about research  
38 based on linked administrative health data related to the security of personal data generally, e.g.,  
39 hacking of Canada Revenue Agency and Canadian banks was noted. Participants responded positively to  
40 information about the ICES process for removing or coding identifying information before data are made  
41 available to researchers, and about the legislated oversight provided by the Information and Privacy  
42 Commissioner of Ontario (IPC) (9). The process of removing names and other direct identifiers was  
43 appreciated, but not seen as sufficient assurance for many participants. Even when fully informed of  
44 privacy and security safeguards, participants noted that risks unavoidably increase when there are more  
45 people and organizations accessing data (Box 1).  
46  
47  
48

49 **Major theme 2: support for research based on health administrative data, with conditions.** Generally,  
50 health data were viewed as an asset that should be used for research, and focus groups participants  
51 supported research based on linked administrative health data, with some conditions. Support was  
52 strongest when people saw a public benefit and agreed with the purposes for which studies were  
53 conducted, e.g., focus group participants strongly supported using administrative health data to study  
54 the long-term safety and efficacy of prescription drug products. In contrast, participants expressed  
55  
56  
57  
58  
59  
60



1  
2  
3 concerns when they thought the results of a study could be misused or disadvantage certain groups  
4 (e.g., seniors, people not adhering to their prescription drug medications) (Box 2).  
5

6 **Major theme 3: mixed and more negative reaction when there is private sector involvement.** Some  
7 focus groups participants expressed concerns about private sector involvement in studies based on  
8 linked administrative health data, e.g., the concern that increased pharmaceutical product sales and  
9 profit, not public benefit, would be the primary motive. In contrast, others saw benefits of private  
10 sector involvement including more skilled people being able to use the data and the potential  
11 development of new products and services. Several participants wanted some form of reciprocity when  
12 public data are used in private sector funded studies, e.g., in the form of lower drug prices (Box 3).  
13  
14

15 **Minor theme 1: low knowledge and understanding of research based on linked administrative health**  
16 **data.** Most participants were not aware of studies based on linked administrative health data, despite  
17 regular media coverage of them. Several participants misunderstood the practice of linking  
18 administrative health datasets for studies at the population level with efforts to bring together data  
19 from different healthcare service providers to improve care for individual patients, even after the  
20 moderator provided clarification. In some instances, participants' lack of prior knowledge about  
21 research based on linked administrative health data led to them having concerns about transparency  
22 and trust (Box 4).  
23  
24  
25

26 **Minor theme 2: mixed views on the need to obtain consent when data don't include identifying**  
27 **information.** The subset of focus group participants that expressed views about consent had varying  
28 opinions. Some felt that consent should always be obtained even when study participants in datasets  
29 were not identifiable. Others were direct in stating their views that consent is not necessary if  
30 identifying information is removed before data are used for research (Box 5).  
31  
32

### 33 Interpretation

34 Generally, the participants of the Toronto and northern Ontario focus groups were supportive of  
35 research based on linked administrative health data providing that there was assurance about privacy  
36 and security, but they cared about details including whether there would be a public benefit from a  
37 study, who would have access to health data and whether there could be a potential downside or  
38 negative impact. Repeated misunderstanding about what research based on linked population-wide  
39 data is (i.e., distinct from linking data to inform the clinical care of an individual) suggests that the topic  
40 is hard to understand, and that there is low awareness of research based on linked administrative health  
41 data in the Ontario general public at present. There were mixed views regarding whether consent is  
42 required when health datasets do not contain identifying information.  
43  
44  
45

46 The results of this research study are consistent with literature (10), (11), (12) and the themes identified  
47 in a recent systematic review (13) which included 25 publications from the United Kingdom, the United  
48 States, Canada and other countries. Findings from that systematic review which are reinforced by this  
49 study include: general widespread support for uses of data in health research with some conditions,  
50 concerns about privacy and security, the requirement that there be a public benefit, more trust in public  
51 sector studies compared to private sector studies, and varying views on the need for consent. This study  
52 identified the new subtheme of administrative health data being an asset that should be used for public  
53 benefit, and provides additional information about how public views are influenced by information  
54 about breaches, hacking and violation of trust outside of the health and research sectors. It also begins  
55  
56  
57  
58  
59  
60

1  
2  
3 to identify the types of studies that the public supports provided that appropriate controls are in place,  
4 e.g., studies of the long-term safety and efficacy of a prescription drug product.  
5

6 Given the public's concerns about uses of data generally, social licence for data-intensive health  
7 research is essential. Carter, Laurie and Dixon-Woods note "poorly informed understanding of the social  
8 licence for secondary use of personal medical data, and a failure to recognise that legal authority might  
9 not be enough to secure the social licence, seems to have been at the heart of the controversy  
10 underlying *care.data*" (6). There are indications that social licence for data-intensive health research  
11 varies by jurisdiction. For example, in Denmark, where there is a long-standing history of citizen support  
12 for the use of public data in research, Danish researchers approach patients about participation in  
13 database-based trials directly with little to no involvement of healthcare providers (14), but in Scotland  
14 potential trial participants are generally contacted by someone within the circle of healthcare providers  
15 that patients would reasonably expect to have access to patient data (14). Regarding informed consent,  
16 it is the authors' view that informed consent can contribute to social licence, but doesn't constitute the  
17 complete answer in all circumstances because there are public benefits that can only be realized  
18 through studies based on population-wide non-consented data, e.g., the withdrawal of Vioxx from the  
19 market (15), restrictions on mobile phone use while driving (16) and the identification of the magnitude  
20 of the opioid epidemic (17) all were based on studies of population-wide non-consented data. Further,  
21 consent may not be truly informed in cases where researchers can't describe all the potential future  
22 uses of health data (18).  
23  
24  
25  
26

27 This research study, and the literature, indicate that the general public wants society to realize the  
28 benefits that can be derived from research based on linked administrative health data, but it is  
29 incumbent on the parties involved in research and data sharing to be transparent, and to involve and  
30 engage with members of the public in an ongoing and authentic manner in order to ensure alignment  
31 with social licence. As illustrated by news reports of growing concerns following Cambridge Analytica's  
32 reported misuse of Facebook data, lack of trust in one sector or organization can spread (19), (20) and  
33 have consequences for other practices that rely on data. Public involvement and transparency are  
34 essential to building and maintaining trust. Informational transparency – publicizing information about  
35 what is being done – is a start, but unlikely to yield the benefits that could be realized by involving  
36 patients and the public in governance and decision-making practices to achieve 'participatory  
37 transparency' and 'accountability transparency' (12). As noted in *Consensus Statement on Public  
38 Involvement and Engagement with Data-Intensive Health Research*, a key premise is that the public  
39 should not be characterised as a problem to be overcome (21). Involving the public, and focusing on the  
40 users and uses of health data that they support, can help ensure sustainable and beneficial data-  
41 intensive health research that is aligned with public values (21).  
42  
43  
44  
45

46 This study has limitations. Foremost, results may not be generalizable across or outside of Ontario. It is  
47 possible that participants from other settings, e.g., rural Ontario, remote northern Ontario, specific sub-  
48 populations or other jurisdictions would have different views. Given the multiple references made to  
49 privacy and security concerns outside of the health and research sectors, it is possible that views will  
50 change based on recent public events including Facebook and Cambridge Analytica (5). In addition,  
51 participants' difficulty understanding what research based on linked administrative health data is may  
52 have impacted their ability to understand and respond to the example research studies that they were  
53 presented with. Finally, there are uses of linked administrative health data, (e.g., helping clinical trials  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 recruitment focus on sites with large numbers of eligible patients, artificial intelligence applications) that  
4 were not presented to focus group participants and warrant further study.  
5

6 In summary, consistent with the literature, this qualitative study found that members of the Ontario  
7 public generally support research based on linked administrative health data, but there is no blanket  
8 approval. Researchers and health data holding organizations should engage with and involve members  
9 of the public to ensure that data-intensive health research is trustworthy and within the bounds of social  
10 licence. Focusing on studies that have a clear public benefit, and respecting and addressing public  
11 concerns about privacy and private sector involvement, are likely to increase public support and thereby  
12 increase the impact and the sustainability of research based on linked administrative health data.  
13  
14

15 Contributors: All authors contributed to the design and conception of the study, attended focus groups,  
16 performed critical reviews of drafts of this manuscript, discussed and refined the themes, and approved  
17 the final version submitted for publication. Magda Nunes de Melo led the literature review. Magda  
18 Nunes de Melo and P. Alison Paprica led the work to design private sector example studies,  
19 independently coded transcripts, and performed analyses to identify themes and subthemes. P. Alison  
20 Paprica had the primary responsibility for the descriptive coding framework, and was the lead for  
21 preparation of the manuscript.  
22  
23

24 Acknowledgements: The authors thank Don Willison who provided helpful advice about how to  
25 structure focus groups to maximize the likelihood that participants would understand what research  
26 based on linked administrative health data is before being asked to comment on specific examples of  
27 research studies. Vanessa Chan is acknowledged for her expert facilitation skills and services. We also  
28 thank Mary Tully for her extensive contributions to the examples of research studies used in the 2017  
29 focus groups and Fiona Miller for her advice on qualitative research methods.  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 1 – Participants of Focus Groups

Locations and Dates	Total and Sex	Age	Highest Level of Education	Total household income (before taxes)
Northern Ontario: Thunder Bay – two (2) focus groups on Sep. 23, 2015  Urban: Toronto – two (2) focus groups on Sep. 24, 2015	31 participants  Male: 13 Female: 18	18 to 34: 8 35 to 44: 7 45 to 54: 5 55 and older: 11	High School or lower: 6 Community College: 11 University or higher: 14	\$20,000 or less: 2 \$20,001 to \$39,999: 9 \$40,000 to \$59,999: 7 \$60,000 to \$79,999: 3 \$80,000 to \$99,999: 4 \$100,000 or more: 2 Unable to Code: 4
Northern Ontario: Sudbury – two (2) focus groups on March 29, 2017  Urban: Toronto – two (2) focus groups on March 30, 2017	32 participants  Male: 15 Female: 17	18 to 34: 8 35 to 44: 6 45 to 54: 5 55 and older: 13	High School or lower: 7 Community College: 15 University or higher: 10	\$20,000 or less: 3 \$20,001 to \$39,999: 7 \$40,000 to \$59,999: 8 \$60,000 to \$79,999: 2 \$80,000 to \$99,999: 8 \$100,000 or more: 4

Confidential

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

Subtheme – generic skepticism about the ability of any organization (in any sector) to ensure privacy

- Everybody likes to keep their information quiet. It depends. I don't trust the banks and how they keep their information, they're always getting hacked, people are getting hacked. Thunder Bay 2015-G1
- 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 huge one, yeah. Thunder Bay 2015-G2
- I'm not really a Facebook person. I know the internet and your stuff can be stolen. Thunder Bay 2015-G2

Subtheme – importance of de-identification and other privacy and security measures for health data

- 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-G2
- I mentioned at the start that if this was a guarantee that my privacy would be secure then I would be for it [research based on linked administrative health data] – Toronto 2015-G1
- The anonymity of those individuals, the privacy of those people should be maintained. I think that's what's critical here – Toronto 2015-G1

Subtheme – external oversight of health data was valued

- It stands out that it's an outside party that's doing the oversight, it's not internal to the institute, it's external - the IPC [Information and Privacy Commissioner of Ontario]. So that's a good thing – Thunder Bay 2015-G2
- It's also on-going oversight and renewal every three years. So it's just not one step. Continuously renewing and renewing and renewing and adapting to the data. Toronto 2015-G1
- PHIPA [the Personal Health Information and Protection Act]. It just reassures me that somebody's got some form of control on how the information is gathered and how the information is distributed. Thunder Bay 2015-G1

Subtheme – incomplete confidence in processes intended to ensure privacy and security of personal health information

- You can build a profile of individual without necessarily ever knowing their name, simply by picking up one piece of data here and one piece of data there. They're not necessarily been identified, but they still serve to build a notion of who that person is. Thunder Bay 2015-G1
- 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-G1
- I believe that if someone ...if a hacker or there's some computer problem, right? Which has happened before, many, many times – personal information can get out there. Toronto 2015-G1

Subtheme – increased risk if more individuals and organizations are provided with access to health data

- I don't like the idea that it's branching out and then branching out, and then who knows where it's going to branch out to. Toronto 2015-G1
- Information can somehow slip out. Especially when there's more people involved. Thunder Bay 2015-G2
- You've already got people [at ICES] with specialized training on privacy and security issues, you've already got that level of protection. Rather than just open everything up, and private sector having to wade through it. Thunder Bay 2015-G2

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

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

Subtheme –linked administrative health data is an asset that should be used for research

- 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-G1
- You can't go forward unless you have information. And the way to get information on a broad scale is through data collection. Thunder Bay 2015-G2
- I don't have a problem with any data being collected or used, I think that's the way of it, I think that's what we have to do. Data is collected to make changes, that's what it's for. Thunder Bay 2015-G2
- 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-G2

Subtheme – importance of public benefit

- 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-G2
- I can see the benefit to use it so they can follow trends and everything. It's kind of something I assumed they did already. Thunder Bay 2015-G1
- It just shows how research can be beneficial. I mean even though the deaths were relatively small, several hundred in more than ten thousand people, it really served them well to figure out that that certain drug had double the risk as opposed to other antibiotics. It was beneficial in the long run. Toronto 2015-G1
- I do know people who don't take their illnesses as seriously as they should and then don't take their medication. So in that aspect, yeah, I can see a good benefit. Sudbury 2017-G2

Subtheme: concerns about how the study findings could be used or misused

- That's broad and scary ....this is Big Brother watching you, asking me if I'm taking my medication? I see insurance companies loving this. Sudbury 2017-G1
- 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-G1

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

Subtheme: concerns about private sector motives, particularly pharmaceutical companies' motives

- I do like the idea of collecting data for the greater good, but that data, if it's for the profit of somebody, I have a problem with that. Sudbury 2017-G1
- What is the information used for? Again, what are they using it for? To try to help people...Or are they just looking at it at a way- okay, we can sell more cancer drugs. Thunder Bay 2015-G1
- I've taken some statistics courses before and one of the things I remember from the course is that you can take data, and you can interpret it any way you want depending on what you're looking for. And companies do that all the time. Toronto 2015-G1
- I'd rather not have a private company because I feel like they're just out to make money and it's not really about doing the research of helping diabetics [or] helping asthma patients. Sudbury 2017-G2
- If it's a pharmaceutical company, they have a vested interest in selling that drug and making money, so inevitably there's a bit of cynicism. Toronto 2017-G2

Subtheme – benefits of involving the private sector

- 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-G2
- An aspect I liked was the collaboration [private sector using government-generated data]. We need society to collaborate more for the common good... to improve the life of Ontarians. Toronto 2017-G2
- I find that there's going to be multiple uses for that kind of information... this is a collaborative way that they can help companies and encourage them to come up with products we all need. Sudbury 2017-G1

Subtheme: expect public benefit in exchange for ability to use public data in private sector studies

- If they're going to use the public data, then they really have to give a good price on the medication. Toronto 2017-G1
- A pharmaceutical company is taking data that's funded by our tax dollars and being used to basically turn a dollar at the end of the day, I feel that's a misuse of that data. I do believe, though, that if that data came at a premium to a private institution with a fund that went into our healthcare system, I might not be so adverse to it. So if they paid into the Ministry of Health, I would have no issue with them utilizing the data that was funded by our tax dollars. Sudbury 2017-G1
- If they're giving us an appropriate price or a better drug being released, then I guess it's okay. Toronto 2017-G1

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Box 4: 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-G2
- 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-G1
- [Indicating that the participant has misunderstood what research based on linked administrative health data is] I had my parents both in long-term healthcare and I did have a problem at one point where certain information was not shared between the pharmacy and the doctor and home care. And so I was constantly calling to make sure things were done properly. Oh, we don't have a record the doctor said. It became a very big problem. Toronto 2015-G1
- [Indicating that the participant has misunderstood what research based on linked administrative health data is] 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 healthcare, whatever that you may need to link up to, to help you in your health. Thunder Bay 2015-G2

Box 5: Minor theme 2 – variation in opinions about the need for people to provide consent for use of data when datasets do not contain identifying information

- Either way it doesn't matter to me. Because the people that didn't give consent are being anonymous anyway. Toronto 2017-G1
- I don't think consent is needed as much to gather data when it's nameless and faceless. Toronto 2017-G2
- How come we didn't get permission that our data is in a study? I can understand where that discomfort is coming from. Sudbury 2017-G1
- It comes back to the consent. I think it would be strange, let's say I had an infection and go to the doctor, and my information isn't just being used in healthcare, but it's being moved external to a drug company. And I get the purpose is to get funding to have more adequate treatment, but I hate not knowing. Like you don't go into a doctor's office thinking this information of mine is going to be used. Toronto 2017-G2



## Reference List

- 1 Institute for Clinical Evaluative Sciences. Institute for Clinical Evaluative Sciences 2018 [cited 2018 Jun 8]; Available from: URL: <https://www.ices.on.ca/About-ICES/>
- 2 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, Canada: Council of Canadian Academies; 2015. Report No.: R864.A24 2015.
- 3 Advisory Panel on Federal Support for Fundamental Science (Canada). Investing in Canada's Future: Strengthening the Foundations of Canadian Research. Advisory Panel on Federal Support for Fundamental Science; 2017.
- 4 Canada's Strategy for Patient-Oriented Research (SPOR). Canadian Institutes of Health Research 2018 [cited 2018 Jun 8]; Available from: URL: <https://www.researchnet-recherchenet.ca/rnr16/vwOpprtntyDtls.do?prog=2794&view=currentOpps&org=CIHR&type=EXACT&resultCount=25&sort=program&next=1&all=1&masterList=true>
- 5 McMahon T. Facebook broadens estimate of data misuse to 87 million people, including more than 600,000 Canadians. The Globe and Mail 2018 [cited 2018 Jun 8]; Available from: URL: <https://www.theglobeandmail.com/world/article-facebook-moves-to-improve-data-collection-transparency-as-zuckerberg/>
- 6 Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. J Med Ethics 2015 May;41(5):404-9.
- 7 Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-40.
- 8 Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84.
- 9 Information and Privacy Commissioner of Ontario. Information and Privacy Commissioner of Ontario 2018 [cited 2018 Jun 8]; Available from: URL: <https://www.ipc.on.ca/>
- 10 Moat KA, Bullock H. Panel summary: Integrating data across sectors for public service improvement in Ontario. Hamilton, Canada: McMaster Health Forum; 2016 Mar 5.
- 11 Ipsos MORI. The one-way mirror: Public attitudes to commercial access to health data. London: Wellcome Trust 2016.
- 12 Aitken M, Cunningham-Burley S, Pagliari C. Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. Science and Public Policy 2016;43(5):713-23.

- 1  
2  
3 13 Aitken M, Jorre JdS, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and  
4 linkage of health data for research purposes: a systematic review and thematic synthesis of  
5 qualitative studies. BMC medical ethics 2016;17(1):73.  
6  
7  
8 14 Best Brains Exchange - Using Administrative, Electronic Medical Record and Patient Health Record  
9 Data in Clinical Trials. 5-17-2016. (Unpublished Work)  
10  
11 15 Mamdani M, Rochon P, Juurlink DN, Anderson GM, Kopp A, Naglie G, et al. Effect of selective  
12 cyclooxygenase 2 inhibitors and naproxen on short-term risk of acute myocardial infarction in the  
13 elderly. Archives of Internal Medicine 2003;163(4):481-6.  
14  
15 16 Min ST, Redelmeier DA. Car phones and car crashes: an ecologic analysis. Can J Public Health 1998  
16 May;89(3):157-61.  
17  
18 17 Gomes T, Juurlink DN, Dhalla IA, Mailis-Gagnon A, Paterson JM, Mamdani MM. Trends in opioid  
19 use and dosing among socio-economically disadvantaged patients. Open Medicine 2011;5(1):e13-  
20 e22.  
21  
22 18 O'Doherty KC, Christofides E, Yen J, Bentzen HB, Burke W, Hallowell N, et al. If you build it, they  
23 will come: unintended future uses of organised health data collections. BMC medical ethics 2016  
24 Sep 6;17(1):54.  
25  
26 19 Smee B. Facebook's data changes will hamper research and oversight, academics warn. The  
27 Guardian 2018 April 25 [cited 2018 Jun 8]; Available from: URL:  
28 [https://www.theguardian.com/technology/2018/apr/25/facebooks-data-changes-will-hamper-](https://www.theguardian.com/technology/2018/apr/25/facebooks-data-changes-will-hamper-research-and-oversight-academics-warn)  
29 [research-and-oversight-academics-warn](https://www.theguardian.com/technology/2018/apr/25/facebooks-data-changes-will-hamper-research-and-oversight-academics-warn)  
30  
31  
32 20 Curry B. Ethics committee urged to look beyond Facebook with privacy study. The Globe and Mail  
33 2018 April 16 [cited 2018 Jun 8]; Available from: URL:  
34 [https://www.theglobeandmail.com/politics/article-ethics-committee-urged-to-look-beyond-](https://www.theglobeandmail.com/politics/article-ethics-committee-urged-to-look-beyond-facebook-with-privacy-study/)  
35 [facebook-with-privacy-study/](https://www.theglobeandmail.com/politics/article-ethics-committee-urged-to-look-beyond-facebook-with-privacy-study/)  
36  
37  
38 21 Aitken M, Tully M, Porteus C, Denegri S, Cunningham-Burley S, Banner N, et al. Consensus  
39 Statement on Public Involvement and Engagement with Data-Intensive Health Research. 2018.  
40 (Unpublished Work)  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60