International Journal of Population Data Science

Journal Website: www.ijpds.org

Factors affecting access to administrative health data for research in Canada: a study protocol

Cynthia Kendell[1,2,*], Adrian R Levy[3], Geoff Porter[2], Elaine Gibson[4], and Robin Urquhart[2,3]

Submission History

Submitted:	12/01/2021
Accepted:	24/06/2021
Published:	28/09/2021

[1]Interdisciplinary PhD Program, Dalhousie University, Halifax, Nova Scotia, B3H 4H6
[2]Department of Surgery, Nova Scotia Health, Halifax, Nova Scotia, B3H 2Y9
[3]Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, B3H 1V7
[4]Schulich School of Law, Health Law Institute, Dalhousie University, Halifax, Nova Scotia, B3H 4R2

Abstract

Introduction

In Canada, most provinces have established administrative health data repositories to facilitate access to these data for research. Anecdotally, researchers have described delays and substantial inter-provincial variations in the timeliness of data access approvals and receipt of data. Currently, the reasons for these delays and variations in timeliness are not well understood. This paper provides a study protocol for (1) identifying the factors affecting access to administrative health data for research within select Canadian provinces, and (2) comparing factors across provinces to assess whether and how they contribute to inter-provincial variations in access to administrative health data for research.

Methods

A qualitative, multiple-case study research design will be used. Three cases will be included, representing three different provinces. For each case, data will be collected from documents and interviews. Specifically, interviews will be carried out with (1) research stakeholders, and (2) regulatory stakeholders (10 individuals/group $*$ 2 groups/province $*$ 3 provinces $=$ 60). During within-case analysis, interview data for each stakeholder group will be analyzed separately using constant comparative analysis. Document analysis will occur iteratively, and will inform interview guide adaptation, and supplement interview data. Cross-case analysis will involve systematic comparison of findings across cases.

Discussion

This study represents the first in-depth examination of access to administrative health data in Canada. The main outcome will be an overarching mid-range theory explaining inter-provincial variations in access to administrative health data in Canada. This theory will be strengthened by the inclusion of the perspectives of both researchers and those involved in the regulation of data access. The findings from this study may be used to improve equitable and timely access to administrative health data across provinces, and may be transferable to other jurisdictions where barriers to access to administrative health data have been reported.

Keywords

qualitative research; big data; Canada

*Corresponding Author:
Email Address: cynthia.kendell@ccns.nshealth.ca (Cynthia Kendell)

https://doi.org/10.23889/ijpds.v6i1.1653

Introduction

Background

Administrative health data are generated through the routine delivery of healthcare programs and services [1]. These data are primarily collected and curated to manage health care resources including the evaluation and planning of hospital bed numbers, occupancy and budgets, and payment of physician and drug dispensation claims [2, 3]. However, these data have also come to be recognized as valuable for secondary uses, including clinical, health services, and population health research [4–6].

Administrative health data are typically population-based, capture a wide range of variables, contain data captured over long periods of time, and contain unique individual-level identifiers, making them well-suited to use in health research [5, 7, 8]. The benefits of using administrative health data for research purposes include improved statistical power and reduced non-response or participant bias [9], the opportunity to examine temporal trends, and lower costs compared to primary data collection [7, 8, 10]. Another benefit of using administrative health data for research is the potential for data linkage, which has been defined as "the bringing together from two or more different sources, data that relate to the same individual, family, place or event" [7] (p.767). The linkage of patient-level data contained in multiple databases provides a more comprehensive view of their encounters with the healthcare system than would be possible if a single database was used. Where individual-level identifiers are available, administrative health databases can also be linked to other related databases (e.g. judicial, educational, social services) to study the social determinants of health behaviors and outcomes [5, 8].

Canada has a long history of administrative health data collection and is considered to have some of the most comprehensive administrative health data in the world [2, 3, 11]. While there is some variation in the specific data that are collected across provinces/territories, data that are commonly collected include inpatient hospitalizations, day surgeries, physician billing claims, and prescription medications [2, 3]. Since these are collected at the provincial/territorial level, the majority of provinces have established data repositories that contain various administrative health databases that may be accessed by researchers [12–18].

Despite having what has been deemed an "information rich" environment [19], infrastructure and resources dedicated to facilitating access to administrative health data (i.e. repositories and staff), and a regulatory framework (i.e. ethical guidelines and information legislation) that permits access to health information for research purposes [20–30], there is evidence to suggest that researchers across Canada are not able to access these data in a timely manner [5, 31–38]. Several groups of researchers who have undertaken multi-province studies using linked administrative health data have described delays and substantial interprovincial variations in the timeliness of data access approvals and receipt of data [34–37]. In a 2002 report, Kephart [34] described a multi-province study involving administrative health data where one province was ultimately excluded due to an inability to obtain data access approval after two years. Since then, similar accounts have continued to appear in the literature [34–37]. In 2016, a multi-site study conducted by the Canadian Network for Observational Drug Effect Studies (CNODES) reported that the British Columbia site was not included due to "lengthy timelines" for data access [35]. Elsewhere, variations in the time to obtain required data access approvals across provinces were reported to impact overall project timelines and hinder the ability for work to be carried out in parallel across provinces [36, 37].

While interprovincial variations in the timeliness of data access may be expected as a result of differences in data access processes, in some cases these variations may be indicative of barriers to timely data access. Currently, the extent to which researchers are experiencing barriers when attempting to gain access to administrative health data in Canada is not well understood. In a 2015 report sponsored by the Canadian Institutes of Health Research (CIHR), a broad range of potential barriers to timely access to health and health-related data were identified; however, minimal evidence was provided to establish the impact of these barriers on data access timelines in Canada, or that researchers in Canada were in fact experiencing barriers to timely data access [5]. More recently, a quality improvement study focused on examining local data access and usage practices was carried out at a single institution in British Columbia [38]. Through interviews with researchers, "lengthy turnaround times" were identified as the primary barrier to accessing and using data for research, although the factors contributing to these lengthy turnaround times were not addressed. Notably, only one study [32] has examined timely access to administrative health data specifically. This study reported wide variations in data access timelines across provinces (from 1 to 18 months) and identified a number of barriers and enablers to timely access. However, there were several important limitations: timeliness data were self-reported by provincial agencies and not confirmed by other means; the starting point for measuring timeliness varied across agencies; and no data were obtained from researchers regarding the experiences with accessing data. Moreover, given that this study was published in 2013, the findings may be of limited applicability to the current Canadian context. Robust evidence is therefore required to assess the extent to which researchers in Canada are experiencing barriers when attempting to access administrative health and the impacts of these on researchers' ability to access data, including the timeliness of data access.

Objectives

The overall aims of this study are to identify the factors (i.e. barriers and enablers) to accessing administrative health data for research within selected Canadian provinces, and to examine how variations in these factors across provinces contribute to reported interprovincial variations in data access timelines across Canada. The specific objectives are to:

1) describe the policies and processes for accessing administrative health data for research purposes in three provinces;

2) explore researchers' experiences with accessing administrative health data for research purposes in each

province, including their ability to obtain data access and the timeliness of data access;

3) explore the perspectives of individuals involved in the regulation of access to administrative health data for research purposes in terms of:

   i. the use of administrative health data for research purposes (e.g. risks and benefits of use),
   ii. the regulation of administrative health data for research purposes (e.g. key considerations, strengths and challenges of the current policies and processes),
   iii. their regulatory role (e.g. gaps in expertise, available supports);

4) compare and contrast (1)–(3) across selected provinces.

Methodology and methods

Methodology

This study will use case study methodology, underpinned by pragmatism. Case study methodology is particularly useful in situations wherein the phenomena being studied cannot be separated from its context [39]. This is relevant to the current study, wherein the factors affecting access to administrative health data for research purposes (i.e. the phenomena) can only be understood in terms of the interactions of the actors involved in research and research oversight, the organizations in which they are situated, and the broader regulatory landscape in which they operate.

There are different approaches to case study, including those published by Stake [40–42], Merriam [43, 44], and Yin [39, 45–47]. All three approaches focus on studying a case, or cases, within a real-life context through the use of multiple data sources, though each proposes different strategies for doing so. Consistent with the pragmatic worldview underpinning this study, we will use a case study approach informed by the works of all three authors, though Merriam [43, 44] and Yin [39, 45–47] will feature more prominently. Since pragmatism embraces both the objective and subjective [48], researchers are not tied to a specific set of methods based on their ontological assumptions, and free to use whichever method (or methods) that are best suited to answering a research question [48].

Study design

This study will employ a multiple case study design (also referred to as "multi-site" [44] or "collective" [40] case study) with embedded units of analysis as set out by Yin [39] (Figure 1). We will use a multiple case study design to enable the examination of variations in data access across jurisdictions, with each of the selected cases representing an example of the social phenomenon being studied [47]. Within each case, there will be two embedded units of analysis, which represent stakeholder groups of particular interest: 1) researchers and research staff who have accessed/sought access to administrative health data, and 2) individuals involved in the regulation and oversight of data access. These two groups differ in terms of their role/mandate, training, interests at stake, incentives, and how their positions within their affiliate institutions. As such, each will have unique insights into the factors affecting access to administrative health data.

Case selection

Three cases will be included in this study, with each representing a "research system". Each research system will encompass a provincial health data repository, as well as the two stakeholder groups described above. To enable the examination of contextual factors, including variations in provincial legislation, cases will be identified in three different provinces. As per guidance from Yin [39], Stake [40], and Merriam [44], cases will be selected in order to maximize the knowledge gained [39, 40, 44]. This means taking into account not only what can potentially be learned from them, but the extent to which they are accessible [40]. More specifically, cases will be chosen based on:

1) The existence of a provincial data repository — The selection of cases from provinces with a provincial data repository will improve comparability across cases by ensuring similar data infrastructure is in place, as well as processes for accessing data.

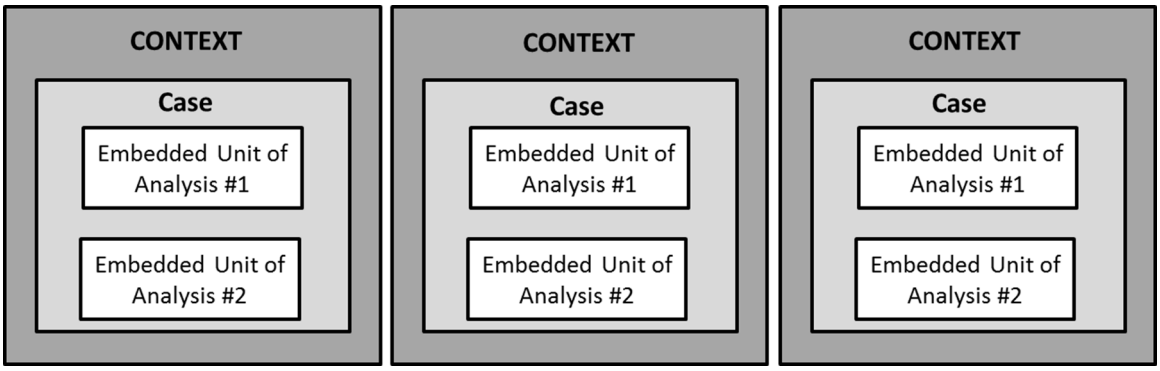
2) Variations in case attributes and contextual factors — The selected cases will represent a range of instances of the phenomenon of interest, consistent with a maximum variation sampling approach [49]. Cases will be identified based on variation in: i) repository structure, organization, and funding; ii) relevant legislation, and iii) reported timeliness of data access. This information will be garnered from publicly available sources (e.g. documents, reports, and websites). Variations in case attributes and context will allow the potential impact of a variety of factors on data access to be examined, and help illuminate the specific circumstances under which findings "hold true"[39]. In addition, common findings across highly variable cases may reveal fundamental aspects of the phenomena being studied [50].

3) Established professional relationships — The existence of established professional relationships between the authors and relevant stakeholder groups in each province will be considered during case selection as this will affect that accessibility of the case (e.g. participant recruitment and access to documents).

Data collection

Case study research is characterized by the use of multiple data collection sources and methods [39, 40, 44, 51]. In this study, interviews will be the primary source of data, providing important historical and contextual information relevant to a case, explaining events and behaviours, and understanding participants' opinions and attitudes [39, 44]. We will also use documentary evidence to provide historical and contextual information relevant to the case that may not otherwise be

Figure 1: Multiple case study design with embedded units of analysis (modified from [39])



obtained [26, 30], and to confirm, corroborate, and supplement information gained from interviews [39].

For each case, we will carry out interviews with key informants (i.e. individuals who have specialist knowledge relevant to the case or specific aspects of the case [52]). Key informants will include members of the two aforementioned stakeholder groups: (1) researchers and research staff with experience accessing administrative health data for research purposes, and (2) those involved in the regulation and oversight of data access.

Interview format and the development of interview guides will be informed by the work of Patton [50] and Rubin and Rubin [53]. Thus, interviews will use a semi-structured approach, taking on the form of a guided conversation rather than a structured interview [39]. Each stakeholder group will be asked a common series of questions related to their own role and experience as well as the facts of the case (i.e. processes for accessing data, required approvals and documentation, actors involved, and relevant policies and legislation), along with a series of questions specific to each group. Researchers and research staff will be asked questions focusing on their experiences when attempting to access administrative health data for research purposes (Supplementary Table 1), while those involved in regulation and oversight of data access will be asked questions focusing on the use and regulation of administrative health data for research purposes (Supplementary Table 2).

Interviews will be carried out in person or via telephone and will be approximately one hour in length. Interviews will be audio-recorded and transcribed verbatim for analysis. Drawing on guidance from Stake [40], notes will be taken throughout interviews to capture impressions, particularly useful quotations, key ideas or concepts to be explored in subsequent interviews, questions that need to be revised or reframed, and to capture recommendations of other sources of evidence to be included or other key informants to be invited to participate.

Documents will be obtained from online sources and requested from key informants. Access to a wide array of documents relevant to each case will be sought, including: administrative documents (e.g. agendas and minutes from research ethics board (REB) and data access committee meetings); policy documents (e.g. provincial legislation and regulations relevant to research involving administrative health data, institutional policies and guidance documents for researchers and oversight bodies); data access documents (e.g. data access forms, data sharing agreements); and evaluations of the provincial health information legislation, and provincial or institutional research reports.

Participants and recruitment

For interviews, we will recruit key informants from two stakeholder groups in each province:

1) *Individuals who have experience accessing or attempting to access administrative health data for research purposes*. Key-informants will include academic researchers (e.g. university faculty or affiliated researchers), research trainees (e.g. graduate students and post-doctoral researchers), and research staff (e.g. research associates, coordinators, and assistants) with experience of accessing or attempting to access administrative health data held by the relevant provincial health data repository within the last 5 years.

2) *Individuals involved in the regulation and oversight of access to administrative health data for research purposes*. Key informants will include individuals affiliated with provincial data repositories who have a role in developing and implementing data access policies and procedures; members of relevant data access committees, privacy review bodies, and REBs; stewards or custodians of databases that are frequently linked to databases held by the provincial data repository; and privacy officers situated within universities, health authorities, and provincial departments of health. Depending on the specific data access processes in place in each province, those who are considered key informants may vary.

We will identify participants using two strategies. First, we will use purposive sampling to ensure that study participants include those individuals from whom the most can be learned (i.e. key informants) [50]. Based on the research team's knowledge of the research and regulatory landscapes in Canada, as well as information obtained from online searches and relevant published literature and reports, the team will compile a preliminary list of potential participants. As individuals are recruited and interviewed, we will also use snowball sampling [50] , whereby participating key informants

will be asked to identify other potential participants in their province.

Individuals who are selected for study inclusion will receive an email invitation to participate. A modified Dillman method [54] will be used to increase response rates. A consent form will be sent to all participants several days prior to the interview, and written consent will be obtained prior to the start of the interview. For each case, we will aim to recruit 10 key informants from each stakeholder group, totaling 20 interviews per case and 60 interviews for the entire study (10 individuals/group * 2 groups/province * 3 provinces = 60). For each case, recruitment will stop once theoretical saturation is reached [49, 55, 56].

Data analysis

As is common in qualitative research, data collection and analysis will occur concurrently [44]. For analysis, we will use an approach consistent with that of Merriam [44]. Since we are using a multiple case study design, analysis will occur on two "levels". Each case will first be analyzed separately as a standalone study (i.e. "within-case analysis"), followed by the comparison of findings across cases (i.e. "cross-case comparison").

Within-case analysis

Within case analysis has several components, including the development of detailed case descriptions, analysis of individual data sources, and the integration of multiple data sources. These activities will occur concurrently and iteratively.

Case descriptions will be developed for each of the three cases, and will focus on providing the "facts of the case", including: (1) a description and brief history of the provincial data repository, (2) a summary of policies (statutes as well as institutional policies) and guidelines relevant to research involving the secondary analysis of personal health information, (3) a description of the various approval processes that researchers must navigate to access to administrative health data for research purposes, and (4) other relevant contextual factors, such as changes in leadership, organizational restructuring, sources of funding, etc. Detailed case descriptions help the researcher understand the context in which the phenomena of interest occurs, providing a critical frame of reference for interpreting findings [39, 44], and also assists the reader in determining the extent to which findings are transferrable [57].

Interview transcripts will be analyzed using the constant comparison method [49], which uses a process of coding and categorization as a means of developing a theory or explanation. Analysis will be informed by Strauss and Corbin's approach [58], which uses open, axial, and selective coding. To ensure consistency throughout the coding process, a coding framework (i.e. "codebook") will be developed by two authors (CK and RU). The authors will independently code a minimum of four interviews (additional if necessary), two for each stakeholder group, and then meet to discuss and refine as needed. Three additional transcripts will be coded by one author (CK) using the revised codebook and additional code refinement may occur if necessary. Once refinement is complete and the codebook has been reviewed by RU, the remaining transcripts will be coded by CK. Qualitative data analysis software (NVivo, QSR International) will be used to facilitate coding. The authors (CK and RU) will meet as needed throughout the analysis process to review the codes and emerging theory.

Analysis of documentary evidence will begin with a preliminary review of all documents obtained for each individual case. A database of documents will be created, containing document title, date of creation, document type, author and/or institution to which it pertains, and brief summary of content (2–3 sentences). This process of reviewing and cataloging will assist with "triaging" documents in order to identify those that are most directly relevant to the case [39]. Three types of information will be extracted from relevant documents: (1) facts pertaining to the case, which will be incorporated into the development of case descriptions, (2) inferences or conclusions drawn on the basis of available documents, which will be explored via other data collection methods, and (3) information contained within case documents that corroborates or contradicts emergent findings from other data sources which will facilitate triangulation [39].

The integration of findings is embedded within the analytic process, occurring as the researcher moves back and forth between interviews and documents. While interview data will comprise the majority of study data and play the largest role in theory development, documentary evidence will shape theory development in several ways. Documentary evidence will inform the development of case descriptions, influence the conduct of interviews (via the addition of new interview questions and probes), and assist in the ongoing refinement of emergent theory. Documentary evidence may also be incorporated into text descriptions of themes resulting from interview data where appropriate. For example, if interview findings indicate that a specific policy lacks clarity, excerpts from that policy may be incorporated into the description of the theme. As described by Baxter [59], each individual data source is a piece of the "puzzle" and contributes to the overall understanding of the phenomenon being studied.

Cross-case comparison

During this stage of analysis, findings from individual cases will be compared and synthesized. Since each individual case is treated as a separate study, synthesizing findings across cases is similar to synthesizing findings across multiple studies [39]. The output of cross-case synthesis may vary depending on the study, and can include a unified description of the cases; categories and themes or typologies developed using all cases; or the development of a general theory that spans all cases [44]. In this study, cross-case analysis will focus on (1) systematically comparing findings for each individual case for each objective, and (2) developing a general theory on access to administrative health data that applies to all three cases included in the study.

Cross-case comparison will be aided by the use of tables to display the facts and findings of each individual case study as recommended by Yin [39]. For example, a table presenting a specific set of characteristics for each case will be created to facilitate cross-case comparison of the cases themselves. Tables will also be created to facilitate cross-case comparison

of findings for each research objective. These will use a matrix format in order to identify findings (i.e. using "+") that occurred across two or more cases versus those that were unique to a single case (as exemplified in [60]). The use of these tables will help clearly identify key similarities and differences between cases and their contexts, facilitate the comparison of specific case findings, and shed light on the relationship between the two (i.e. how variations in study findings are related to variations in cases and their contexts). The end result of this process will be the development of a robust theory that can be applied across cases to explain interjurisdictional variations in data access.

Discussion

This study will be the first in-depth study of access to administrative health data in Canada. As such, we expect that it will address several important knowledge gaps. First, this study will provide detailed descriptions of the processes for accessing administrative health data for each of the provincial data repositories included in this study. Explicating these processes will be integral to understanding how variations in process contribute to interprovincial variations in data access, and on a more practical level, may serve as a resource for researchers seeking access to data. Second, this study will provide empirically derived evidence regarding the extent to which Canadian researchers are, in fact, experiencing barriers when attempting to access administrative health data, and what these are. To date, researchers' experiences accessing data are not well documented within the literature. Third, a comprehensive taxonomy of the multi-level factors affecting access to administrative health data in Canada will be developed, which does not currently exist within the literature. Finally, the main outcome of this study will be an overarching mid-range theory explaining interprovincial variations in access to administrative health data in Canada. Currently, the specific barriers and enablers to data access, and how these contribute to observed interprovincial variations in data access, are not well understood.

While this study is focused on examining access to administrative health data in Canada, it has relevance beyond the Canadian context. The use of administrative health data for generating evidence to inform decision-making at various levels of health systems is becoming more widespread, however, researchers in many jurisdictions outside Canada struggle with issues related to data access [61–64]. The findings of this study may help support efforts to improve data access in these jurisdictions, particularly in those employing a similar governance framework to that in Canada (i.e. an institution-based research ethics model intersecting with information legislation regulating the access, use, and disclosure of personal health information for research purposes).

This study has a number of strengths. First, the use of case study methodology provides a framework for interprovincial comparisons that is not available via other qualitative approaches. A second strength is the focus on a limited number of carefully selected cases, which allows the researcher to gain an in-depth understanding of the case(s) as well as insight into the interaction of significant factors characteristic of the phenomenon of interest [44]. In the context of this study, we will move beyond simply identifying and cataloging the barriers and enablers to data access to understanding how the interaction of various factors may create and/or mitigate barriers to data access. This includes understanding how contextual factors may affect data access, providing an opportunity to gain insight into how the current COVID-19 pandemic may be affecting data access. Strengths also include the use of multiple data collection methods and the inclusion of a wide range of stakeholders, which will ensure a thorough examination of the factors affecting access to administrative health data and improve study rigour by facilitating the triangulation of data (i.e. from different sources, and different stakeholder groups) [39]. Finally, as a result of having existing professional relationships in each of the selected provinces, we expect that we will have improved access to the cases and that key informants will be willing to share their experiences and perspectives with us.

Limitations

A limitation of this study is the lack of comprehensive quantitative data to examine data access timelines in each province. Although provincial data repositories may capture or be able to calculate the time from the receipt of a researcher's application to the date approval is granted, this represents only a portion of the total data access timeline, excluding other aspects of the overall process of data access (e.g. application preparation, obtaining REB approval, contracts and data sharing agreements, dataset preparation, etc.), the time to access data held by other data holders, as well as applications that remain incomplete. To obtain complete timeliness data, the research team would need to request detailed information from researchers and all relevant data holders in an attempt to assemble data access timelines for individual studies. This resource intensive process could only be undertaken for a small sample of studies in each province, which would ultimately provide limited insights into the overall timeliness of data access in a province, or into the factors affecting data access. With that being said, the primary focus of this study is not to measure timeliness, but rather to gain an understanding of the contexts and conditions under which researchers are able to gain timely access to administrative health data for research, which case study methodology is well-suited to address.

The inclusion of a limited number of cases may also be perceived as a study limitation due to lack of generalizable findings [65], however, the traditional notion of generalizability (i.e. "statistical generalisation" [66]) is not an appropriate criterion for assessing the value of case study research. Instead, the aim of this study is "transferability" [44] (also referred to as "naturalistic generalization" or "case-to-case transfer"[66]), which occurs when findings observed in one context can be applied in a different context. The onus is on the reader to determine the extent to which the findings from a particular case or cases may apply in another context, while it is the responsibility of the researcher to provide sufficient detail to enable this process [44]. In this study, transferability of findings will be facilitated through the process of "particularization" (i.e. by focusing on the details and uniqueness of each of the selected cases [40, 67]) and the use of "thick description"

(i.e. rich, literal descriptions of each case and its context) to describe the context in which each case is embedded [44].

Conclusion

Administrative health data represent a powerful research tool that may be leveraged to address the informational needs of healthcare decision-makers and subsequently inform healthcare delivery and health policy. The potential of these data to trigger positive change within healthcare systems is undermined when researchers are unable to access these data in a timely manner, or where barriers to access exist. The findings from this study may be used to improve equitable and timely access to administrative health data across provinces. For example, the findings of this study may be used to improve harmonization of policies and processes across jurisdictions, identify areas where researchers and regulatory/oversight bodies may require additional supports to facilitate improved access to administrative health data, or inform the development of a national framework guiding access to administrative health data specifically for research purposes.

Acknowledgements

The authors would like to acknowledge Margaret Jorgensen for providing administrative support for this study. Financial support for this study is provided by the Cancer Outcomes Research Program, Department of Surgery, Nova Scotia Health.

Statement on conflicts of interest

The authors have no financial conflicts of interest to disclose. CK, AL, GP, and RU are users of administrative health data for research purposes. CK is a current member of the Data Access Committee at Health Data Nova Scotia. AL is the former Director of Health Data Nova Scotia and financially supports its operation using peer-reviewed funding.

Ethics statement

This study received ethics approval from the Nova Scotia Health Research Ethics Board (Reference number: 1025301).

Supplementary Tables

Supplementary Table 1: Interview guide for researchers and research staff.
Supplementary Table 2: Interview for regulatory stakeholders.

References

1. University of Manitoba. Term: Administrative Health Data. 2013; Available from: http://mchp-appserv.cpe.umanitoba.ca/viewDefinition.php?definitionID=102210
2. Lucyk K, Lu M, Sajobi T, Quan H. Administrative health data in Canada: lessons from history. BMC Med Inform Decis. 2015;15(1):69. https://doi.org/10.1186/s12911-015-0196-9
3. Cadarette SM, Wong L. An introduction to health care administrative data. Can J Hosp Pharm. 2015;68(3). https://doi.org/10.4212/cjhp.v68i3.1457
4. Canadian Institute for Health Information (CIHI). Better Information for Improved Health: A Vision for Health System Use of Data in Canada. Ottawa:ON: CIHI, 2013.
5. The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation. Accessing Health and Health-Related Data in Canada. Ottawa:ON: Council of Canadian Academies, 2015.
6. Organisation for Economic Co-operation and Development. Health Data Governance: Privacy, Monitoring and Research-Policy Brief. 2015.
7. Holman CD, Bass JA, Rosman DL, Smith MB, Semmens JB, Glasson EJ, et al. A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. Aust Health Rev. 2008;32(4):766–77. https://doi.org/10.1071/AH080766
8. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health. 2011;32(1):91–108. https://doi.org/10.1146/annurev-publhealth-031210-100700
9. Jones J. The effects of non-response on statistical inference. J Health Soc Policy. 1996;8:49–62. https://doi.org/10.1300/J045v08n01_05
10. Brook EL, Rosman DL, Holman C. D'Arcy J. Public good through data linkage: measuring research outputs from the Western Australian Data Linkage System. Aust N Z J Public Health. 2008;32(1):19–23. https://doi.org/10.1111/j.1753-6405.2008.00160.x
11. Quan H, Smith M, Bartlett-Esquilant G, Johansen H, Tu K, Lix L. Mining administrative health databases to advance medical science: geographical considerations and untapped potential in Canada. Can J Cardiol. 2012;28(2):152–4. https://doi.org/10.1016/j.cjca.2012.01.005
12. Population Data BC. 2020 [cited 2020 December 20]; Available from: https://www.popdata.bc.ca/.
13. Manitoba Centre for Health Policy. University of Manitoba; 2020 [cited 2020 December 22]; Available from: http://umanitoba.ca/faculties/health_sciences/medicine/units/chs/departmental_units/mchp/.
14. ICES. 2020 [cited 2020 December 22]; Available from: https://www.ices.on.ca/.
15. NB Institute for Research, Data, and Training. University of New Brunswick; 2020 [cited 2020 December 22]; Available from: https://www.unb.ca/nbirdt/.

16. Health Data Nova Scotia. Dalhousie University; 2020 [cited 2020 December 22]; Available from: https://medicine.dal.ca/departments/department-sites/community-health/research/hdns.html.

17. Projects. University of Prince Edward Island; 2020 [cited 2020 December 20]; Available from: http://chcresearch.ca/projects/.

18. Newfoundland and Labrador Center for Health Information. 2020 [cited 2020 December 22]; Available from: https://www.nlchi.nl.ca/.

19. Roos LL, Gupta S, Soodeen R-A, Jebamani L. Data quality in an information-rich environment: Canada as an example. Can J Aging. 2004;24(Suppl 1):153–70. https://doi.org/10.1353/cja.2005.0055

20. Freedom of Information and Protection of Privacy Act, RSBC 1996, c. 165.

21. Health Information Act, RSA 2000, c. H-5.

22. Personal Health Information Act, CCSM 1997, c. 51.

23. Personal Health Information Protection Act, SO 2004, c. 3.

24. Personal Health Information Privacy and Access Act, SNB 2009, c P-7.05.

25. Personal Health Information Act, SNS 2010, c. 41.

26. Health Information Act , RSPEI 2014, c.31

27. Personal Health Information Act, SNL 2008, c P-7.01

28. Health Information Act, SNWT 2014, c. 2.

29. Health Information Privacy and Management Act, SY 2013, c. 16.

30. Access to Information and Protection of Privacy Act, SNWT (Nu) 1994, c 20.

31. Canadian Institutes of Health Reserach. International Review Panel Report 2005-2010. Ottawa, ON: Canadian Institutes of Health Reserach, 2011.

32. Meagher N, McGrail K. Data Access Review Times: A Study. Population Data BC, 2013.

33. Katz A, Enns J, Wong ST, Williamson T, Singer A, McGrail K, et al. Challenges associated with cross-jurisdictional analyses using administrative health data and primary care electronic medical records in Canada. Int J Popul Data Sci. 2018;3(3). https://doi.org/10.23889/ijpds.v3i3.437

34. Kephart G. Barriers to Accessing and Analyzing Health Information in Canada. Canadian Institute for Health Information, 2002.

35. Azoulay L, Filion KB, Platt RW, Dahl M, Dormuth CR, Clemens KK, et al. Incretin based drugs and the risk of pancreatic cancer: international multicentre cohort study. BMJ. 2016;352:i581. https://doi.org/10.1136/bmj.i581

36. Groome PA, McBride M, Jiang L, Kendell C, Decker K, Grunfeld E, et al. Lessons learned: it takes a village to understand inter-sectoral care using administrative data across jurisdictions. Int J Popul Data Sci. 2018;3(3):1–13. https://doi.org/10.23889/ijpds.v3i3.440

37. Butler A, Smith M, Jones W, Adair CE, Vigod SN, Lesage A, et al. Multi-province epidemiological reserach using linked administrative data: a case study from Canada. Int J Popul Data Sci. 2018;3(3). https://doi.org/10.23889/ijpds.v3i3.443

38. Ho HKK, Görges M, Portales-Casamar E. Data access and usage practices across a cohort of researchers at a large tertiary pediatric hospital: qualitative survey study. JMIR Med Inform. 2018;6(2). https://doi.org/10.2196/medinform.8724

39. Yin RK. Case Study Research: Design and Methods. 4th ed. Thousand Oaks, CA: Sage; 2009.

40. Stake R. The Art of Case Study Reserach. Thousand Oaks, CA: Sage; 1995.

41. Stake R. Qualitative Case Studies. In: Denzin N LY, editor. The Sage Handbook of Qualitatative Research (3rd ed). Thousand Oaks, CA: Sage; 2005.

42. Stake R. Multiple Case Study Analysis. New York, NY: Guilford Press; 2006.

43. Merriam SB. Case Study Research and Case Study Applications in Education. San Francisco, CA: Jossey-Bass; 1998.

44. Merriam SB. Qualitative Research: A Guide to Design and Implementation (3rd ed.). San Francisco, Ca: Jossey-Bass; 2009.

45. Yin RK. Enhancing the Quality of Case Studies in Health Services Research. Health Services Research. 1999;35(5):1209–24.

46. Yin RK. Applications of Case Study Research. 2nd ed. Thousand Oaks, CA: SAGE Publications; 2003.

47. Yin R. Case study research: Design and Methods. 2nd ed ed. Newbury Park, CA: Sage; 1994.

48. Tashakkori A, Teddlie C. Mixed Methodology: Combining Qualitative and Quantitative Approaches. Thousand Oaks: SAGE Publications; 1998.

49. Glaser B, Strauss A. Discovery of Grounded Theory: Strategies for Qualitative Work. Chicago, IL: Aldine Transaction; 1967.

50. Patton MQ. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks, CA: SAGE Publications; 2002.

51. Creswell JW. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: SAGE Publications; 2013.

52. Payne G, Payne J. Key Informants. Key Concepts in Social Research. London: SAGE Publications; 2004.

53. Rubin H, Rubin I. Qualitative Interviewing: The Art of Hearing Data. Thousand Oaks, CA: Sage Publications; 1995.

54. Dillman DA, Smyth JD, Christian LM. Mail and Internet Surveys: The Tailored Design Method. 3rd ed. New York, NY: John Wiley and Sons; 2009.

55. Urquhart C. Grounded Theory for Qualitative Research: A Practical Guide. Thousand Oaks: Sage; 2013.

56. Given LM. 100 Questions (and Answers) About Qualitative Research. Thousand Oaks: Sage; 2016.

57. Stake R. Case Studies. In: Denzin NK, Lincoln YS, editors. Handbook of Qualitative Research Thousand Oaks, CA: Sage Publications; 1994.

58. Strauss A, Corbin J. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park, CA: Sage Publications; 1990.

59. Baxter P, Jack S. Qualitative case study methodology: study design and implementation for novice researchers. Qual Rep. 2008;13(4):544–59. https://doi.org/10.46743/2160-3715/2008.1573

60. Urquhart R, Porter GA, Sargeant J, Jackson L, Grunfeld E. Multi-level factors influence the implementation and use of complex innovations in cancer care: a multiple case study of synoptic reporting. Implement Sci. 2014;9. https://doi.org/10.1186/s13012-014-0121-0

61. Milea D, Azmi S, Reginald P, Verpillat P, Francois C. A review of accessibility of administrative healthcare databases in the Asia-Pacific region. J Mark Access Health Policy. 2015;3:28076. https://doi.org/10.3402/jmahp.v3.28076

62. Iversen A, Liddell K, Fear N, Hotopf M, Wessely S. Consent, confidentiality, and the Data Protection Act. BMJ. 2006;332(7534):165–9. https://doi.org/10.1136/bmj.332.7534.165

63. Moore H, Blyth C. Optimising the use of linked administrative data for infectious diseases research in Australia. Public Health Res Pract. 2018;28(2):e2821810. https://doi.org/10.17061/phrp2821810

64. Downs JM, Ford T, Stewart R, Epstein S, Shetty H, Little R, et al. An approach to linking education, social care and electronic health records for children and young people in South London: a linkage study of child and adolescent mental health service data. BMJ Open. 2019;9(1):e024355. https://doi.org/10.1136/bmjopen-2018-024355

65. Flyvberg B. Case Study. In: Denzin NK LY, editor. The SAGE handbook of qualitative reserach (4th ed). Thousand Oaks, CA: Sage; 2011.

66. Firestone WA. Arguments for generalizing from data as applied to qualitative research. Educ Res. 1993;22(4):16–23. https://doi.org/10.2307/1177100

67. Stake R. The case study method in social inquiry. Educ Res. 1978;7(2):5–8. https://doi.org/10.2307/1174340

Supplementary table 1: Interview guide for researchers and research staff

Participant Background

1. Describe your role (e.g., researcher, research staff, graduate student/trainee, other).
2. How long have you been in this role?
3. Briefly describe the research involving administrative heath data that you have been directly involved with (e.g., content area, databases used, single/multi-jurisdictional, etc.)
4. is your involvement in the data access process?

Facts Related to the Case

5. What is involved in accessing administrative health data for research via [repository X]?
   - Required reviews and approvals
   - Application preparation and submission
   - Relevant policies and legislation
   - Data-related costs
6. Do these things change if you are linking to external datasets? Doing a multi-province study? If so, how?

Experiences Accessing Data

7. How would you describe your experiences accessing administrative health data for research purposes via [repository x]?
   - Challenges
   - Barriers and enablers
8. How long does it take to get access to data to your study data?
   - Is that satisfactory?
   - Why do you think it takes that amount of time?

Specific Factors Affecting Data Access/Timely Data Access

9. In your opinion, what are the main factors affecting researchers' ability to access administrative health data for research purpose in [province x]?
   - Resources
   - Supports available to assist with accessing data
   - Provincial information legislation
   - Your personal knowledge and experience
10. Regarding the overall process of accessing administrative health data for research in [province X]:
    - Is there a clear pathway?
    - Is the process transparent?
    - Are the oversight/approval bodies involved responsive?
    - Are there "bottlenecks" in the process? If so, where?
11. Does the feedback provided to you during the data access process reflect an understanding of the research and related risks? If not, please explain.
12. Are there certain databases or types of data that are harder to get access to than others? If so, please explain.

Closing

13. Do you have any final comments about accessing administrative health data for research purposes in [province x] that you would like share?

Supplementary table 2: Interview guide for researchers and research staff

Participant Background

1. Relevant to data access, how would you describe your current role (e.g., data custodian, REB member, privacy committee member, other)?
   – What are your main responsibilities?
   – What is the main aim of objective of your role?
2. How long have you been in this role?
3. Please describe your background and how you came be in this role.

Note: All subsequent questions will be adapted based on the individual's role.

Facts related to the case

4. For researchers who require access to administrative health data for research purposes
   –What are the required reviews and approvals relevant to [regulatory body x]?
   –What happens once the researcher submits an application to [regulatory body x]?
   –What happens once to [regulatory body x] decides to approve or reject an application?

Perspectives on the use and regulation of administrative health data for research

5. What are your thoughts on the use of administrative health data for research?
6. What are your thoughts on the processes involved in gaining access to administrative health data for research?
   – What is the workload involved for researchers? Those in regulatory/oversight roles?
   – Can current processes be streamlined?
7. What are your thoughts on the policies that are in place?
   – Are they documented and accessible to you?
   – Are the requirements clear?

Factors affecting researchers' access to administrative health data for research

8. What things impact your decision/the decision of [regulatory body x] to approve an application?
9. Once approval has been granted, what are some things that impact the time it takes for the researcher get access to the requested dataset?
10. What do you think are the main factors that impact researchers' ability to access administrative health data for research?
11. What do you think are the main factors impacting the timeliness with which researchers are able to access administrative health data?

Regulatory Role

12. How would you describe your level of knowledge/expertise with regard to:
   – The methodological aspects of research involving administrative health data?
   – Ethical/legislative requirements related to disclosures of health information for research?
   – Issues related to privacy and confidentiality?
13. What organization/institutional supports are currently available to you to support you in your role?
14. What additional supports would assist you in your role?

Closing

15. Do you have any final comments about accessing administrative health data for research purposes in [province x] that you would like share?