

---

## Improving Aboriginal health data capture: evidence from a health registry evaluation

---

S. L. HARPER<sup>1\*</sup>, V. L. EDGE<sup>1</sup>, C. J. SCHUSTER-WALLACE<sup>2</sup>, M. AR-RUSHDI<sup>3</sup>  
AND S. A. MCEWEN<sup>1</sup>

<sup>1</sup> *Department of Population Medicine, University of Guelph, Guelph, ON, Canada*

<sup>2</sup> *Water-Health Nexus, United Nations University Institute for Water, Environment and Health, Hamilton, ON, Canada*

<sup>3</sup> *Medical Officer of Health, Labrador-Grenfell Health Region, Goose Bay, NL, Canada*

(Accepted 5 November 2010; first published online 7 December 2010)

### SUMMARY

The lack of high-quality health information for accurately estimating burdens of disease in some Aboriginal populations is a challenge for developing effective and relevant public health programmes and for health research. We evaluated data from a health registry system that captured patient consultations, provided by Labrador Grenfell Health (Labrador, Canada). The goal was to evaluate the registry's utility and attributes using modified CDC guidelines for evaluating surveillance systems. Infectious gastrointestinal illness data were used as a reference syndrome to determine various aspects of data collection and quality. Key-informant interviews were conducted to provide information about system utility. The study uncovered limitations in data quality and accessibility, resulting in region-specific recommendations including conversion to an electronic system. More generally, this study emphasized how a systematic and standardized evaluation of health registry systems can help address challenges to obtaining quality health data in often remote areas where many Aboriginal communities are found.

**Key words:** Aboriginal health, gastrointestinal infections, registry evaluation, surveillance system.

### INTRODUCTION

Health registries and surveillance systems that monitor health events are critical to public health practice for estimating the magnitude of a health problem, detecting outbreaks, understanding the natural history of a disease, examining disease distribution and spread, and evaluating control measures [1]. In turn, this information informs the prioritization of public health actions, programme planning, interventions,

and research for disease prevention and health promotion [1]. However, periodic evaluation of a surveillance system is important to ensure that the system is meeting its objectives efficiently and effectively. A widely used framework is the Centers for Disease Control and Prevention's (CDC) Guidelines for Evaluating Public Health Surveillance Systems [2], which could be adapted to evaluate patient-visit registry systems.

Evaluations of health data collection systems used in Aboriginal populations are especially critical due to a recognized lack of good quality health data in some communities [3–5]. Aboriginal populations often live in substandard conditions with less access to the same

\* Author for correspondence: Ms. S. L. Harper, Department of Population Medicine, University of Guelph, Guelph ON, N1G 2W1, Canada.  
(Email: harpers@uoguelph.ca)

quantity and quality of resources as other citizens, which contributes to disparities in health [4, 5]. For example, in Canada, Aboriginal disease mortality and morbidity rates from all causes are substantially higher than for other citizens [6, 7], infant mortality rates in Inuit are more than twice as high [7], Inuit infant mortality rates are up to four times higher [8, 9], and Inuit life expectancies are the shortest of all Canadians [10]. These disparities in rates need to be better understood by investigating environmental and other risk factors for disease; however, relevant health data of adequate quality are limited, resulting in inaccurate and generally underestimated Aboriginal disease rates [11, 12]. This lack of data availability is due in part to gaps in the process of health data collection, which in turn is reflected in generally poor quality of captured data. In some countries health databases and registries do not capture ethnicity/cultural<sup>†</sup> data that would allow for targeted investigations of their Aboriginal population's health [11–13]. Furthermore, even when ethnicity/cultural information is collected in population health databases, the quality of data captured on Aboriginal health can be compromised by lack of uniform reporting [11, 12, 14], incomplete records [15], high monetary costs of patient follow-up [16], and lack of coding validity [3, 17]. The remote location of many Aboriginal communities can impact access to healthcare, limit human resources to provide high-quality healthcare, as well as reduce care-seeking behaviour, which can further compromise the quality and accuracy of captured health data [9, 11, 12, 14]. Consequently, in some countries, national reports of annual disease burden exclude Aboriginal-specific data leaving unfortunate gaps in Aboriginal health statistics and research [15].

In the seven northern communities of Labrador, Canada (Fig. 1), primary healthcare is provided by the Labrador Grenfell Regional Health Authority (LGH), and five of these communities fall within the Inuit region of Nunatsiavut. Patients are provided primary care by resident nurses and visiting physicians and when necessary, patients are medically evacuated by air to the Labrador Health Center (LHC) in Goose Bay. All northern clinics use the

paper-based 'E-Book' health registry system to record information regarding patient visits. The E-Book health registry was originally suggested as a source of data on infectious gastrointestinal illness (IGI) patient visits for a companion study on weather, water quality, and health [18]. However, in order to determine the quality of these data for analysis it was necessary to evaluate and understand the patient health data capture system. The main objectives of this evaluation were to: (1) evaluate the health registry's utility and system attributes using modified CDC guidelines for evaluating surveillance systems, and (2) use IGI as a reference syndrome to provide insight on data quality captured by the E-Book system in two Inuit communities.

## METHODS

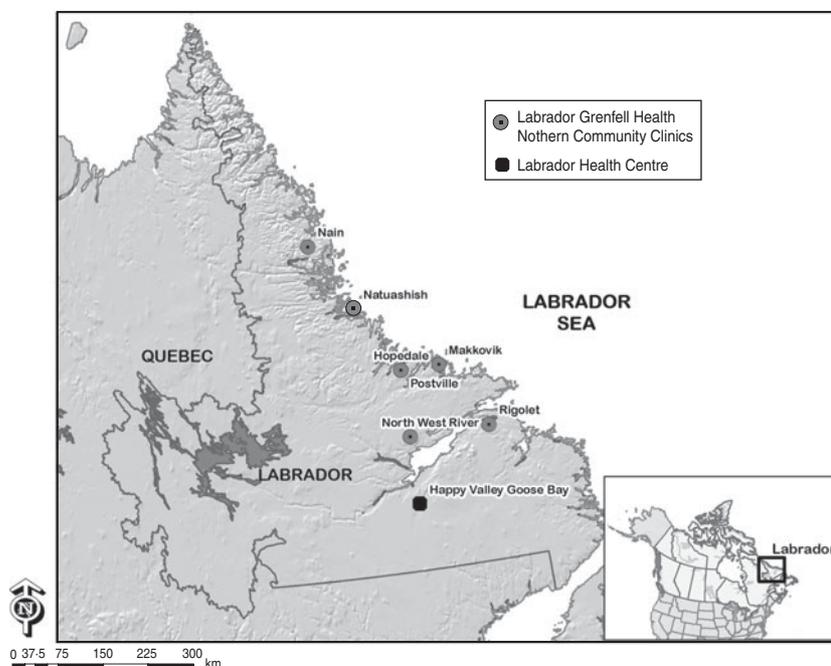
### Study communities

Located on the north coast of Labrador, Nunatsiavut is one of Canada's four Inuit regions. E-Book patient-visit health registries from clinics in two Nunatsiavut communities, Nain and Rigolet, were used for the evaluation because these two communities were participating in a companion study on weather, water quality, and health [18]. Nain (56° N, 61° W) is the most northern community in Labrador with a population of 1034, with mainly Inuit residents (91.8%), approximately equal numbers of men and women, and 27% of the population aged <15 years [19]. Their clinic was staffed by six regional nurses, five personal care attendants (PCAs), one laboratory attendant, four maintenance staff, one clerk typist, and a visiting physician (once a month). The clinic had four holding beds, an incubator, basic trauma and resuscitation equipment, and a defibrillator.

Rigolet (54° N, 58° W) has a population of 269, with approximately equal numbers of men and women. Most residents identify as Inuit (94.3%) and 18.5% of the population are aged <15 years [19]. Rigolet's Groswater Bay Clinic was staffed by two regional nurses, one PCA, and a visiting physician (once every 6 weeks). The clinic had one holding bed, basic trauma and resuscitation equipment, and a defibrillator.

At all northern coastal clinics in Labrador, the LGH uses a paper-based health registry system to record detailed information regarding the reason for each patient visit. Although designated 'E-Book', this is not an electronic record-keeping system, it is strictly

<sup>†</sup> Past literature examining Aboriginal health data often use the term 'ethnicity'; however, here the term 'ethnicity/cultural' is used to recognize that the term "Aboriginal Peoples" refers to organic political and cultural entities that stem historically from the original people of North America, rather than collections of individuals united by so-called "racial" characteristics' [13].



**Fig. 1.** Map of Labrador's seven northern coastal communities where Labrador Grenfell Health has community health clinics (Newfoundland and Labrador, Canada).

paper-based. It was created by the LGH for its own use in their northern community clinics. Detailed patient information is captured in the E-Book system for use at the provincial and regional level, including community name, date of birth, sex, ethnic/cultural origin, date of patient visit, diagnosis description, disease codes (ICD-9), name of person who examined the patient, along with other relevant information.

### E-book evaluation

The CDC's Guidelines provided the framework for the E-Book health registry evaluation [2, 20]. While other excellent frameworks for surveillance system evaluation are available (e.g. Health Canada's Framework and Tools for Evaluating Health Surveillance Systems [21]), the CDC's Guidelines were used in this evaluation because of the clear focus on stakeholder engagement. Briefly, the CDC's guidelines involve five tasks: (i) engaging stakeholders, (ii) describing the system, (iii) focusing the evaluation design, (iv) evaluating system usefulness and attributes, and (v) stating recommendations and sharing lessons learned [2]. The CDC guidelines recommend examining nine system attributes (Table 1) [2], two of which, sensitivity and positive predictive value, were not evaluated because there was no other appropriate source of IGI data in northern Labrador with which

to compare the IGI data captured by the E-Book registry.

Seven key E-Book stakeholders (health officials) from community and regional levels were invited for confidential in-person, on-site, interviews that were audio-recorded in February 2008. The interview guide was pre-tested for content with epidemiologists and academics, and for context with health professionals in the region. The semi-structured interview guide included 40 short answer and 49 Likert scale questions, with the option to elaborate on any response. The interview questions elicited information on: E-Book system operation, regional and local data analysis conducted with E-Book-captured data, E-Book system usefulness; resources required for system operation, public health priorities in the region, gastrointestinal illness in the region, and various system attributes, including simplicity, flexibility, data quality, acceptability, representativeness, timeliness, and stability. All audio recordings of the interviews were transcribed verbatim. Data were examined using a rigorous analytical approach, which included a preliminary exploration of the data by reading interview transcripts and listening to audio recordings, in order to code the data based on the research objectives, as well as emergent and significant themes. Through an iterative process, these codes were then combined and collapsed, until a list of themes from

Table 1. *Attributes used to evaluate the E-Book health registry used by Labrador Grenfell Health, Labrador, Canada [2, 43]*

Surveillance attribute	Definition	Method of evaluation	Qualitative judgement
Simplicity	System design System ease of operation	Qualitative	Major strength Limitation
Flexibility	Ability to adapt to changing information needs or operating conditions	Qualitative	Strength
Data quality	Completeness of data recorded Validity of data recorded	Quantitative	Limitation Limitation
Acceptability	Willingness to participate in the system	Qualitative	Strength
Sensitivity	The proportion of true cases detected by the system	Not evaluated	—
Positive predictive value	The proportion of reported true cases	Not evaluated	—
Representativeness	The ability to describe a health event's distribution in the population by place and time	Quantitative	Strength
Timeliness	Speed between steps in the system	Qualitative	Strength
Stability	Reliability to operate without failure Availability of information to be used by system stakeholders	Qualitative	Strength Severe limitation

each interview was created. This list of themes was then cross-analysed across all interviews, culminating in the creation of a final list of themes which was used to frame the data presentation. To ensure the credibility and accuracy of results, individual follow-up meetings were conducted with each interviewee to discuss and validate study results and conclusions.

Analyses of registry simplicity, flexibility, acceptability, timeliness, and stability attributes were based on information obtained from key-informant interviews, the registry's design, and available LGH documents related to the E-Book registry. Based on this information, each system attribute was then qualitatively ranked as being a 'major strength', 'strength', 'limitation', or 'severe limitation' of the system.

To evaluate system representativeness and data quality, IGI was used as a reference syndrome. A broad case definition of IGI was used to increase sensitivity and case capture; it comprised vomiting and/or diarrhoea not due to chronic conditions such as colitis, diverticulitis, Crohn's disease, and irritable bowel syndrome, nor related to illness from pregnancy, medication use, and/or alcohol/drug use [22]. Anonymized patient-visit data from the E-Book health registry entries that were available for Nain and Rigolet (1 January 2005 to 31 October 2008) were reviewed; the entries related to IGI were manually extracted into an electronic database (Excel; Microsoft Corp., USA).

Transcription accuracy was assessed using a random systematic sample of 10% of the extracted IGI entries (i.e. beginning with a randomly chosen starting point and then performing interval sampling to check 10% of records), with an acceptable error rate of <5%. Moreover, using systematic random sampling, age and sex proportions were extracted from available records of patient visits for all causes from 1 January to 31 October 2008 for Nain ( $n=598$ ) and Rigolet ( $n=166$ ). These proportions were then compared to 2006 Canadian Census data [19] in order to determine if certain segments of the population (by age and sex) used the clinic more often than others.

Ethics approval was obtained from the University of Guelph Research Ethics Board and LGH, and was supported by the Nunatsiavut Government Research Advisory Committee.

## RESULTS

For the period 1 January 2005 to 31 October 2008, the E-Books contained a total of 48 620 clinic records in Nain, and 16 656 clinic records in Rigolet; 541 and 175 records were related to IGI, respectively. The overall proportion of IGI consultations of all patient visits from 1 January 2005 to 31 October 2008 was 1.1% in Nain and 1.0% in Rigolet (Table 2). Since the health registries were anonymized, it was not possible to exclude patients that visited the clinic more

Table 2. Annual counts of infectious gastrointestinal illness (IGI)-related patient visits and total patient visits to the clinics in Nain and Rigolet (Labrador, Canada) from 1 January 2005 to 31 October 2008

Year	Nain (population = 1034)*			Rigolet (population = 269)*		
	IGI patient visits	Total patient visits	Proportion of IGI visits	IGI patient visits	Total patient visits	Proportion of IGI visits
2005	130	16 370	0·8 %	33	5202	0·6 %
2006	107	12 697	0·8 %	47	4280	1·1 %
2007	188	11 921	1·6 %	41	4365	0·9 %
2008†	116	10 500	1·1 %	54	3510	1·5 %
2005–2008†	541	51 488	1·1 %	175	17 357	1·0 %

\* Statistics Canada, 2006 [19].

† Clinic records from January 1 to 31 October 2008.

than once for the same episode of IGI. A detailed description of IGI patient-visit trends by season, sex, and age are presented elsewhere [18].

### System evaluation

Seven LGH employees who regularly worked with the E-Book system at regional and community levels were interviewed (response rate 100%). The qualitative rankings derived for each system attribute are summarized in Table 1. Results from stakeholder interviews, document review, and data quality analysis are summarized by system attributes below.

*Purpose and objectives of the system.* Based on E-Book documents and interviews, no record could be found, nor could interviewees remember the original purpose of the E-Book health registry. While each stakeholder's description of the registry's purpose varied, the stakeholders interviewed generally described the E-Book registry's intention as a disease classification system that summarized and described the types of patient visits. However, there was much confusion around the objectives of the system where, particularly at the community level, interview respondents were unsure of how the data captured by the system were used. All respondents reported that the detailed data requested through this system should be further utilized for health statistics and feedback on the overall health of each community.

*System design and scope.* Fig. 2 illustrates the patient-visit data capture system. The attending nurse entered patients' information on a patient chart. From this chart, the PCA or clerk copied patient-visit records into the E-Book, ascribing the international disease

code (ICD-9) associated with the diagnosis or complaint. Month-end summaries of total patient counts were sent to LHC where count data were transferred from paper to an electronic file. These data eventually contributed to annual reports that were disseminated to community clinics and senior LGH management to inform decision-making processes, including clinic workload and budgetary requirements.

*Simplicity.* The structure of the E-Book health registry system was described as relatively simple (Fig. 2); however, some stakeholders reported that registry operation was not easy due to disease-coding procedures. Most reported that the ICD-9 coding scheme could be unclear, confusing, inconsistent, outdated, and a burden on their time. Furthermore, some PCAs and clerks who coded data most often had little or no healthcare or ICD training, making coding even more difficult.

*Flexibility.* Some concerns were expressed regarding the flexibility of the system to capture new or emerging diseases and new case definitions due to the limitations of, and dependency on, an available short-list of ICD-9 codes that were generally used. Furthermore, while the system had few supply costs, personnel costs were considered high because of the time-consuming transcription, making the system vulnerable to personnel funding cuts. All stakeholders agreed that conversion to an electronic system could bring more flexibility to the system.

*Data quality.* Except for the ICD-9 entries, over 99% of fields were complete based on IGI patient-visit records for both communities. One community did not ascribe ICD-9 codes to any patient visit. The

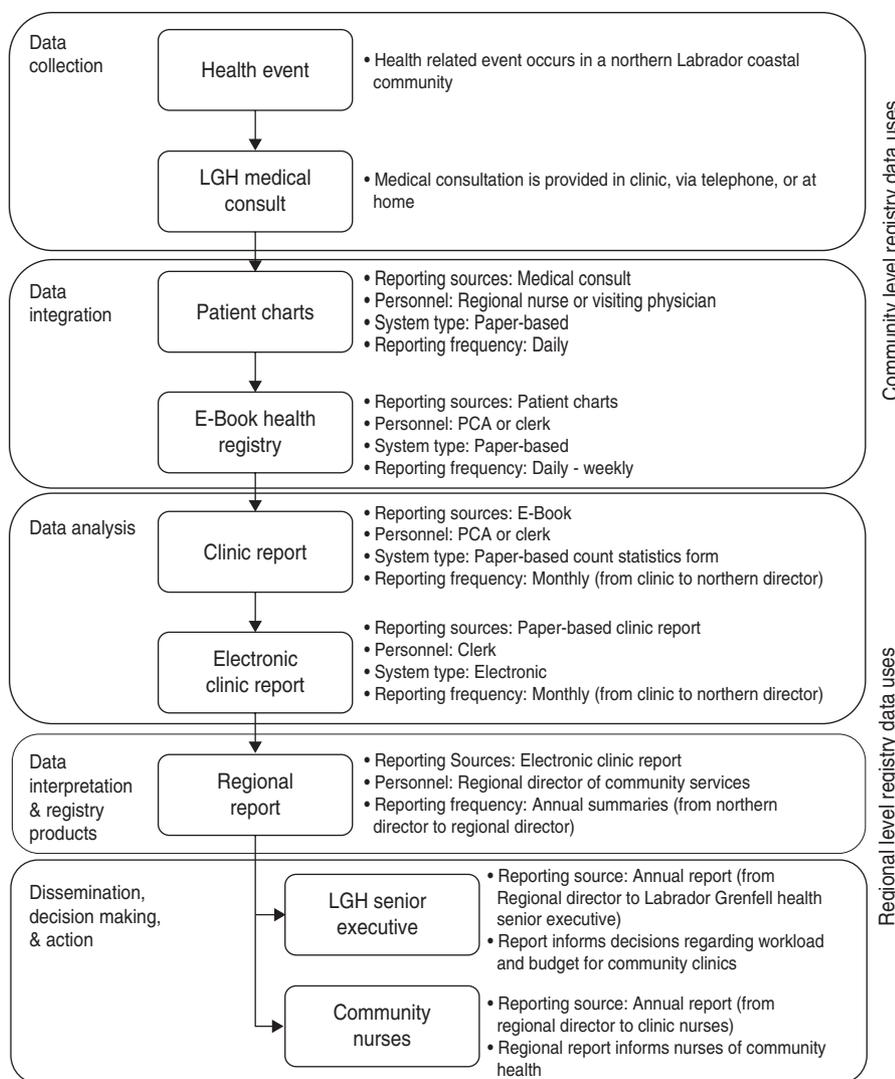


Fig. 2. Flowchart displaying the flow of data captured by the E-Book health registry used by Labrador Grenfell Health (LGH), Newfoundland and Labrador, Canada.

validity of IGI data (2005–2007) was examined for one community by matching the written visit description against assigned ICD-9 codes, which revealed that 18.3% of IGI diagnosis descriptions did not match the assigned ICD-9 code. With regard to all entries in the E-Books (e.g. for all diagnoses) for this community, 6.2% of outpatient department numbers were incorrect, some entries were illegible (0.2–11.7% per month), and others were out of chronological order (11–42.2% per month).

*Acceptability.* All seven northern community clinics participated in the E-Book health registry system; however, the level of compliance was reported to vary by community. System components that were reported by stakeholders to have low acceptability

included E-Book registry security, usefulness, time burden, and data dissemination to stakeholders. In particular, most community-level respondents felt that the frequency and type of data analysis were not acceptable, and that the time required to record the extensive amounts of data into the registry did not seem justified considering its use and minimal feedback.

*Representativeness.* It was reported by stakeholders that all patient visits were recorded in the E-Book registry, and thus, the E-Book registry captured data representative of all patient visits. Determining if patient visits are actually representative of health events occurring in the community over time and place is more difficult, but knowledge of population

demographics can provide some insight on system representativeness [2, 20]. Thus, in one community, comparisons between the proportions of age and gender in the population were made with patient visits and revealed that females were over-represented. A slightly lower proportion of community members aged <20 years visited the clinic, whereas those aged >75 years represented a higher proportion than found in the community.

*Timeliness.* Data collected by the system were analysed annually to inform year-end recruitment and budgetary decisions. For these purposes, as well as examining long-term health trends, the time from data capture to data analysis was considered acceptable by stakeholders. However, since data were analysed on an annual basis, it was not useful for short-term health monitoring, including outbreak detection.

*Stability.* Overall, the stakeholders reported that the system operated with few interruptions and attributed this to its paper-based nature. The availability of E-Book information for use by its stakeholders was also examined: monthly statistical counts (e.g. total patient visits, total deaths, total births) conducted at local clinics were used to inform annual recruitment and budgetary decisions and considered by regional stakeholders as easy to access. Conversely, the availability and accessibility of data by registry stakeholders for other uses, such as examining community health trends, was considered quite poor by all community stakeholders and some regional stakeholders.

## DISCUSSION

This study illustrates the usefulness of a systematic evaluation of a health registry system in order to gain clarity on its usefulness, strengths, and gaps. Further, the reliance of a companion study that used these health data highlighted the importance of understanding the underlying data for purposes such as public health activities and research. We successfully adapted the CDC's Guidelines [2] to further engage system stakeholders and examine data quality using a reference syndrome, thus making it possible to provide informed recommendations to LGH related to specific attributes of their health registry system as implemented in Nunatsiavut. This study demonstrated how the CDC's framework for evaluating surveillance systems can be adjusted so that it is

applicable as a systematic and standardized approach to evaluating health registry systems in general.

Understanding the source, quality, and representativeness of health data is critical to public health programming and other research applications. As observed in other evaluations using a reference syndrome [23–25], these descriptive analyses provided LGH with a suite of information about the registry itself, as well as IGI patient-visit trends in Nunatsiavut. Descriptive data analysis is critical in all evaluations [15] because it enriches the results by providing grounded and useful information for healthcare providers, policy makers, and health researchers, which is particularly needed in the context of Aboriginal health. In addition, our evaluation of the E-Book system and descriptive analyses on the reference syndrome (IGI) provided insight on data quality and informed our decisions regarding the use of the system's health data for a comprehensive study linking weather, water, and human health in Nunatsiavut [18].

The evaluation of this region's health registry system revealed that the original purpose of the E-Book system was apparently lost, forgotten, or not communicated, and that the understanding of how the data were used was particularly unclear and vague for all stakeholders. To improve the registry, we recommended that the system's purpose, goals, and objectives be clarified and communicated appropriately, because these are considered critical components of any registry system [20]. This finding emphasizes the importance of regular evaluation of any health registry in use, to help ensure that the process is meeting the stakeholders' needs, and ultimately health system goals. For remote Aboriginal communities such as those in this study, evaluations of this type might enhance the usability and quality of Aboriginal health data being captured, an issue that is of particular concern.

The E-Book registry evaluation uncovered concerns about data quality and accessibility of coded health data for all stakeholders. Paper-based systems, such as the E-Book registry, are more prone to human error, and can result in illegible, missing, and incomplete entries [26]. For example, the E-Book registry had some entries in which visit descriptions did not match the assigned ICD-9 code, outpatient department numbers were incorrect, entries were illegible, and visits were recorded out of chronological order. The E-Book stakeholders recommended a conversion from the existing paper-based system to an electronic

system for enhanced capture of their Aboriginal health data. Converting to an electronic system could increase data quality [27–30], allow improvement via data editing (program software to check data validity), consolidate health information into one entry [31], and facilitate the effective and efficient assignment of disease codes [32, 33]. Moreover, conversion to an electronic system would improve the timeliness, accessibility, and usability of captured data [27, 28, 33–35], as well as allow greater flexibility in responding to health events in a community [27, 29, 30]. The short-term financial and human resource costs of conversion from a paper-based to an electronic system can be high (e.g. the initial cost of custom software development [29, 30]). However, open software can create innovative systems and help reduce these costs substantially [36]. Furthermore, in the long term it is likely to be cost-effective because the increased timeliness, accessibility, and usability of captured data allows public health interventions to be implemented earlier, resulting in decreased direct medical costs, mortality, and morbidity [29, 30, 33]. In general, considering the remote locations of many Aboriginal communities and the demonstrated success of other electronic health registries, an electronic data collection system that provides an excellent patient-visit registry and community-level health statistics in near real time would greatly enhance public health surveillance in remote Aboriginal populations.

The ICD-9 coding frustrations identified here were also reported by other studies [32, 33, 37], and might explain some of the quality issues of coded data in the E-Book system and in other systems [26, 38]. Updating to ICD-10 codes might alleviate some of these coding frustrations. Other studies reported that ICD-10 codes provide clearer disease definitions [39], resulting in higher quality data capture [39, 40], provide a more intuitive coding system [39], and allow more detailed diagnoses [41] than ICD-9 codes. Still, ICD training for nurses, PCAs, and clerks in these communities in order to make coding easier to understand and more efficient [37, 42] is seen as critical in avoiding coding errors and compromising data quality [32, 37]. Furthermore, training courses should be designed to be accessible and appropriate, especially in the context of training in remote communities. Involvement of multiple stakeholders in the design of training modules is needed to ensure cultural respect, local relevance, appropriate language(s), suitable learning platforms and pedagogy

(e.g. in-person vs. online training), and accessibility and affordability to all end-users and trainees. Initially, updating coding systems, as well as developing and implementing training represent increased costs; however, over the long term it has been reported to be cost-effective [33]. Standardized training on system operation, including disease coding, could further improve the quality of Aboriginal health data captured by registry systems.

Only two of the seven northern community clinic E-Book registries were evaluated for data quality. Further studies would be needed to evaluate the quality of data captured in the other remote communities in terms of comparability. While using a reference syndrome to assist in the evaluation provided useful information on IGI-related patient visits for LGH, as well as the companion study [18], the validity of coded data for other health events captured by the E-Book registry was not investigated. However, by conducting key-informant interviews from across the region, results regarding the utility of the system, as well as the usability of the registry data, are probably representative of the E-Book system operating in all of the northern clinics, and for all reasons for patient visits.

Standardized and systematic evaluations of health databases are often inexpensive [26], quick, effective, and simple but important starting points in addressing ways to improve the quality of Aboriginal health data captured by existing registries. Evaluations could also indicate that new or enhanced registry systems need to be implemented, creating opportunities to develop new methodologies and techniques to collect health data that are socially, culturally, and geographically relevant in remote Aboriginal settings. Results from evaluations can also be used to advocate political and resource support to make necessary improvements in system efficacy and efficiency, as well as improving the quality of data captured. Higher quality health data are essential to improve our understanding of the magnitude of Aboriginal health problems and will allow for the prioritization of actions, programme planning, and research for disease prevention and health promotion in these populations, both now and in the future [1]. Thus, the focus should be on improving the evidence base through systematic data collection and reporting, which will ultimately result in a better understanding of Aboriginal health, a message that could well be applied to the broader Canadian and international contexts, especially when considering

the missing/incomplete data found globally in Aboriginal-specific health statistics.

## ACKNOWLEDGEMENTS

We thank Maureen Baikie, Peggy Baikie, Gwen Elliott, Kathy Elson, and Norma Forsey from Labrador Grenfell Health for their support in this project. We particularly thank the communities of Nain and Rigolet for their involvement and support for this project. We also thank Mary Denniston, John Lampe, and Gail Turner from the Nunatsiavut Government for their continued support. Many thanks to Ashlee Cunsolo Willox for her support and guidance with qualitative methods and analysis, and to Rob Meyers for creating the Labrador map. Financial support was provided by ArcticNet, Nasivvik, and the Public Health Agency of Canada.

## DECLARATION OF INTEREST

None.

## REFERENCES

1. **Teutsch SM, Thacker SB.** Planning a public health surveillance system. *Epidemiological Bulletin Pan American Health Organization* 1995; **16**: 1–16.
2. **CDC.** Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. *Morbidity and Mortality Weekly Report* 2001; **50**: 1–35.
3. **Smylie J, Anderson M.** Understanding the health of Indigenous peoples in Canada: key methodological and conceptual challenges. *Canadian Medical Association Journal* 2006; **175**: 602–605.
4. **Gracey M, King M.** Indigenous health part 1: Determinants and disease patterns. *Lancet* 2009; **374**: 65–75.
5. **King M, Smith A, Gracey M.** Indigenous health part 2: The underlying causes of the health gap. *Lancet* 2009; **374**: 76–85.
6. **Caron NR.** Getting to the root of trauma in Canada's Aboriginal population. *Canadian Medical Association Journal* 2005; **172**: 1023–1024.
7. **MacMillan HL, et al.** Aboriginal health. *Canadian Medical Association Journal* 1996; **155**: 1569–1578.
8. **Luo Z, et al.** Birth outcomes in the Inuit-inhabited areas of Canada. *Canadian Medical Association Journal* 2010; **182**: 235–242.
9. **Harris SB, et al.** Disease patterns among Canadian Aboriginal children: Study in a remote rural setting. *Canadian Family Physician* 1998; **44**: 1869–1877.
10. **Adelson N.** The embodiment of inequity: Health disparities in Aboriginal Canada. *Canadian Journal of Public Health* 2005; **96**: S45–S61.
11. **Auer AM, Andersson R.** Canadian Aboriginal communities and medical service patterns for the management of injured patients: a basis for surveillance. *Public Health* 2001; **115**: 44–50.
12. **Auer AN, Andersson R.** Canadian Aboriginal communities: a framework for injury surveillance. *Health Promotion International* 2001; **16**: 169–177.
13. **ITK.** *Inuit in Canada: A Statistical Profile*. Ottawa, Canada: Inuit Tapiriit Kanatami, 2008.
14. **Guthrie JA, et al.** HIV and AIDS in Aboriginal and Torres Strait Islander Australians: 1992–1998. *Medical Journal of Australia* 2000; **172**: 266–269.
15. **Mak DB, Watkins RE.** Improving the accuracy of Aboriginal and non-Aboriginal disease notification rates using data linkage. *BMC Health Services Research* 2008; **8**.
16. **Clough AR.** Some costs and challenges of conducting follow-up studies of substance use in remote Aboriginal communities: An example from the Northern Territory. *Drug and Alcohol Review* 2006; **25**: 455–458.
17. **Kozyrskyj AL, Mustard CA.** Validation of an electronic, population-based prescription database. *Annals of Pharmacotherapy* 1998; **32**: 1152–1157.
18. **Harper SL.** Weather, water, and infectious gastrointestinal illness in the context of climate change in Nunatsiavut, Canada (dissertation). Guelph, ON, Canada: University of Guelph, 2009.
19. **Statistics Canada.** 2006 Community profiles (<http://www12.statcan.ca/english/census06/data/profiles/community/Index.cfm?Lang=E>). Accessed 7 April 2009.
20. **CDC.** Guidelines for evaluating surveillance systems. *Morbidity and Mortality Weekly Report* 1988; **37**: 1–18.
21. **Health Canada.** Framework and tools for evaluating health surveillance systems, 2004.
22. **Majowicz SE, Horrocks J, Bocking K.** Demographic determinants of acute gastrointestinal illness in Canada: A population study. *BMC Public Health* 2007; **7**.
23. **Gazarian M, et al.** Evaluation of a national surveillance unit. *Archives of Disease in Childhood* 1999; **80**: 21–27.
24. **Matsui T, et al.** Evaluation of national Tsutsugamushi disease surveillance – Japan, 2000. *Japanese Journal of Infectious Diseases* 2003; **55**: 197–203.
25. **Takahashi T, et al.** Evaluation of a public health *Salmonella* surveillance system in King County, Washington. *American Journal of Infection Control* 2004; **32**: 7–11.
26. **Dixon J, et al.** Assessment of the reproducibility of clinical coding in routinely collected hospital activity data: A study in two hospitals. *Journal of Public Health Medicine* 1998; **20**: 63–69.
27. **McDonald CJ, et al.** A framework for capturing clinical data sets from computerized sources. *Annals of Internal Medicine* 1997; **127**: 675–682.
28. **Doyle TJ, Glynn MK, Groseclose SL.** Completeness of notifiable infectious disease reporting in the United

- States: an analytical literature review. *American Journal of Epidemiology* 2002; **155**: 866–874.
29. **Effler P, et al.** Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. *Journal of the American Medical Association* 1999; **282**: 1845–1850.
  30. **Overhage JM, Grannis S, McDonald CJ.** A comparison of the completeness and timeliness of automated electronic laboratory reporting and spontaneous reporting of notifiable conditions. *American Journal of Public Health* 2008; **98**: 344–350.
  31. **Hogan WR, Wagner MM.** Accuracy of data in computer-based patient records. *Journal of the American Medical Informatics Association* 1997; **4**: 342–355.
  32. **Surjan G.** Questions on validity of international classification of diseases-coded diagnoses. *International Journal of Medical Informatics* 1999; **54**: 77–95.
  33. **O'Neill L, Klepack W.** Electronic medical records for a rural family practice: a case study in systems development. *Journal of Medical Systems* 2007; **31**: 25–33.
  34. **McGlinchey EA, et al.** Ability to perform registry functions among practices with and without electronic health records. *Journal of General Internal Medicine* 2008; **23**: 241–241.
  35. **Health Canada.** Toward electronic health records (<http://www.hc-sc.gc.ca/ohih-bsi/>). Accessed 4 March 2009.
  36. **McDonald CJ, et al.** Open source software in medical informatics – why, how, and what. *International Journal of Medical Informatics* 2003; **69**: 175–184.
  37. **O'Malley KJ, et al.** Measuring diagnoses: ICD code accuracy. *Health Services Research* 2005; **40**: 1620–1639.
  38. **Romano PS, Mark DH.** Bias in the coding of hospital discharge data and its implications for quality assessment. *Medical Care* 1994; **32**: 81–90.
  39. **Kokotailo RA, Hill MD.** Coding of stroke and stroke risk factors using International Classification of Diseases, revisions 9 and 10. *Stroke* 2005; **36**: 1776–1781.
  40. **Jensen AO, et al.** Accuracy of ICD-10 discharge diagnoses of skeletal related events and bone metastases secondary to prostate cancer and breast cancer in a Danish population-based hospital registry. *Pharmacoepidemiology and Drug Safety* 2008; **17**: S35.
  41. **Henderson T, Shepherd J, Sundararajan V.** Quality of diagnosis and procedure coding in ICD-10 administrative data. *Medical Care* 2006; **44**: 1011–1019.
  42. **Stausberg J, et al.** Reliability diagnoses coding with ICD-10. *International Journal of Medical Informatics* 2008; **77**: 50–57.
  43. **Jhung MA, et al.** Evaluation and overview of the national electronic injury surveillance system-cooperative adverse drug event surveillance project (NEISS-CADES). *Medical Care* 2007; **45**: S96.