Innovative technology for web–based data management during an outbreak

Shamir N Mukhi[1,2], Tammy L Stuart Chester[2], Justine DA Klaver-Kibria[3], Deborah L Nowicki[4,5], Mandy L Whitlock[4], Salah M Mahmud[4,5] , Marie Louie[6], Bonita E Lee[7]

[1]Canadian Network for Public Health Intelligence, [2]National Microbiology Laboratory – Public Health Agency of Canada, [3]JKK Environmental-Health Consulting, [4]Winnipeg Regional Health Authority, [5]University of Manitoba [6]Provincial Laboratory for Public Health of Alberta, [7]University of Alberta

Abstract

Lack of automated and integrated data collection and management, and poor linkage of clinical, epidemiological and laboratory data during an outbreak can inhibit effective and timely outbreak investigation and response. This paper describes an innovative web-based technology, referred to as Web Data, developed for the rapid set-up and provision of interactive and adaptive data management during outbreak situations. We also describe the benefits and limitations of the Web Data technology identified through a questionnaire that was developed to evaluate the use of Web Data implementation and application during the 2009 H1N1 pandemic by Winnipeg Regional Health Authority and Provincial Laboratory for Public Health of Alberta. Some of the main benefits include: improved and secure data access, increased efficiency and reduced error, enhanced electronic collection and transfer of data, rapid creation and modification of the database, conversion of specimen-level to case-level data, and user-defined data extraction and query capabilities. Areas requiring improvement include: better understanding of privacy policies, increased capability for data sharing and linkages between jurisdictions to alleviate data entry duplication.

Keywords: Outbreak, Laboratory, Surveillance, Epidemiology, Integration, Data, Informatics

Introduction

Infectious disease surveillance at all levels (local, provincial, federal, and international) is essential for timely identification of possible outbreaks. Effective outbreak investigations and measures to reduce the impact and disease transmission are critical and require ongoing collection, management, and analysis of data, in addition to timely aggregation of data across time and jurisdictions (municipal, provincial and territorial).

Seamless data management during an outbreak has been problematic in the past [1]. There are a multitude of challenges related to data management including lack of data sharing agreements, limited resources and technology resulting in manually intensive processes, and lack of shared

1

platforms leading to inability for multiple stakeholders to access real time data. As identified in the Naylor report on the 2003 SARS outbreak in Canada[2], one of the key steps in achieving seamless outbreak management is *"…uniform adoption of highly flexible and interoperable data platforms, that allow sharing of public health information, capture of clinical information from hospitals, and integration into an outbreak management database platform"*. In the spring of 2003, the Public Health Agency of Canada (then, Health Canada) partnered with several Provincial/Territorial and regional public health stakeholders to improve Pan-Canadian public health surveillance, communications, and response through the application of new technologies. This resulted in the creation of the Canadian Network for Public Health Intelligence (CNPHI), a comprehensive framework of applications and resources designed to fill critical gaps in Canada's national public health info-structure [3].

In an attempt to alleviate some of the challenges of data management during outbreaks, a new technology within the CNPHI platform, called Web Data, was developed in 2008. It provides a mechanism for non-technical users to rapidly deploy a secure and adaptable web-based system for managing data. It also has an integrated search and interrogation capabilities and can produce instant data reports.

In the spring of 2009, Web Data technology was put to test to assist with data management during the 2009-2010 H1N1 influenza pandemic. The Winnipeg Regional Health Authority (WRHA), Winnipeg, Manitoba and the Provincial Laboratory for Public Health (ProvLab) in Alberta both employed Web Data technology. The experiences of these organizations represent the epidemiological and laboratory perspectives, respectively.   This manuscript describes the design, development, application and evaluation of the Web Data technology that was employed by these organizations during the H1N1 pandemic.

Design and Development of Web Data

The Web Data technology provides an easy-to-use web-based interface to create customized web-based data collection tools. It allows for user-defined databases and surveys that can be used to collect data in the form of open– and closed-ended questions, both quantitative and qualitative in nature. Surveys can be used to collect a one-time response from users, whereas databases can be used to collect any number of records. The technology consists of the following main components:

1. *Form Builder:* also known as the Designer, allows users to interactively develop forms using a drag and drop approach providing flexibility to non-technical users for rapidly developing secure databases. The form is based on section(s), which can include many fields placed on a grid of rows and columns defined by the user. The Designer includes a pre-defined set of field types categorized as text inputs, selection inputs or display fields. Examples of fields include checkboxes, radio buttons, dropdowns, and a table input. The user may also specify mandatory fields as well as additional descriptive help text for each of the fields as required. Each section can be controlled for read or write access or completely hidden from specific users.

2. *Queries/Reports:* the tool has a simple interface and allows users to develop queries to interrogate their database. Data field(s) can be interactively specified and data exported for further analysis. Customized reports may also be designed if required.

3. *Data Uploader:* in order to support batch data entry, the tool includes an Uploader feature that allows batch uploading of comma delimited data based on a user defined uploader configuration.

4. *Advanced Analysis:* data trending and data analysis can be enabled through seamless integration with the Canadian Early Warning System (CEWS) technology [4] which provides user driven interactive trending, analysis, GIS mapping, reporting and extraction.

5. *Record Transfer:* to foster collaboration and seamless data sharing between multiple implementations of a form used by various stakeholders, the Web Data tool allows sharing of records based on a user defined field sharing configuration. That is, a set of specific fields can be copied or transferred from one record in a specific form into a different form without re-entry of the data.

6. *Cluster Analyzer:* this is a specialized module of the Web Data tool which enables identification of records that meet user defined criteria based on a set of fields. That is, users may define a specific set of fields and their matching preference order to identify a group of records (i.e., clusters). The cluster is then supported by a data table and a chart to visualize the findings.

The technology allows seamless access by multiple users via a secure web-based model. It provides multiple levels of access control including role (reader, writer, administrator), jurisdictional and field level controls.

Deployment of Web Data during the H1N1 Pandemic

The Web Data technology was deployed in two different ways. The WRHA team used the national instance of the Web data technology. Since the instance was not in Manitoba, the WRHA team opted to not include nominal information within the database. The ProvLab team used a local instance of Web Data technology which enabled them to include nominal information, i.e. the physical hardware resided within ProvLab infrastructure.

Both instances were managed by the CNPHI team. WRHA users had to log into the national instance of CNPHI and accessed their form on the same system, while ProvLab users transitioned to local instance of CNPHI and used the technology on it. This provided seamless user management and access to the resources.

3

Evaluation of the application of Web Data during the H1N1 Pandemic

A questionnaire was designed to explore the strengths and limitations of the Web Data technology during its construction and use during the 2009 H1N1 pandemic, as well as to compare and contrast prior data management approaches with Web Data management.

The questionnaire included open and closed ended questions regarding attributes deemed important for data management during an outbreak, the initial data management, the transition to the Web Data technology, and the sequence of events used when employing the new technology for data management. Questions were also asked regarding improvements that were seen as well as successes and limitations of the new Web Data technology.

The questionnaire was sent via email to representatives in ProvLab and WRHA. After receipt of the answers to the questionnaire, follow-up telephone interviews were conducted. For Alberta and WRHA, the representatives who supported and coordinated the implementation of the Wed Data technology provided feedbacks to the questionnaire.

Using the information collected from the questionnaire, a flow chart detailing the steps associated with data-gathering, information sharing, data management, and data technology prior to and following the use of the Web Data technology was constructed. This was done for both the ProvLab and WRHA. The strengths and limitations of the Web Data technology used by WRHA and Alberta ProvLab were identified and compared.

Results

The experience of the Winnipeg Regional Health Authority (WRHA):

The WRHA pandemic response plan was coordinated by the regional health authority. Figure 1 illustrates the flow of information between WRHA and its partners and the changes that occurred (green) following the introduction of Web Data technology.

Influenza A, under the Public Health Act of Manitoba, is included on the notifiable diseases list and is reportable by all laboratories in the province. Cadham Provincial Laboratory, the public health lab for Manitoba, notified Manitoba Health of the first positive pandemic H1N1 2009 laboratory results in early 2009. These positive cases were manually entered into a Manitoba Health H1N1 ACCESS database. Manitoba Health subsequently sorted results and faxed the positive results of Winnipeg residents or non-Winnipeg residents in Winnipeg hospitals to the WRHA Communicable Disease Unit. Demographic and laboratory information was entered into the regional Integrated Public Health Information System (iPHIS). A referral was sent via iPHIS to a community Public Health Nurse, who then completed an H1N1 case investigation form for all community and hospitalized cases in Pandemic Wave 1 and hospitalized cases only in Pandemic Wave 2. Completed investigation forms were then faxed back to the Communicable Disease Unit and filed in a paper chart. Prior to the introduction of the Web Data technology, data from investigation forms for hospitalized patients were captured in an Excel spreadsheet and maintained by the WRHA Population Health Surveillance Unit, as variables required to support

4

surveillance could not be easily or quickly added to iPHIS. All completed forms and updates to forms were faxed to Manitoba Health who then updated their H1N1 ACCESS database.

Initially, the WRHA Critical Care Program was maintaining an Excel spreadsheet of all hospitalized ICU cases which was updated every morning by overwriting the clinical data from the day before, making new changes difficult to track. This file was emailed to the WRHA Population Health Surveillance Unit every morning, and was managed and stored on one computer. An Access database was later created by the Population Health Surveillance Unit and shared with the WRHA Audits and Quality Analysis team in order to track clinical updates for ICU patients.

These processes were manual and there were multiple layers of duplication of data transfer and entry. During the initial stages of the H1N1 pandemic the WRHA identified a need for a web-based database that was flexible enough to accommodate all fields on the investigation form and allowed for real-time access by multiple users. Web Data technology was used to rapidly develop a secure web-based database for epidemiological, clinical and laboratory information. Initial consultation to setup the system was on June 10th, 2009 with a functional database deployed on June 12th,2009. Data analysts and epidemiologists at the WRHA manually entered non-nominal data from the H1N1 investigation forms for all cases of H1N1 in Winnipeg residents and non-Winnipeg residents in Winnipeg hospitals. The database continues to evolve as needed and allows for interactive data querying. Data were exported for further analysis using Stata (version 11.0) to generate daily and weekly reports.

The Web Data database could not contain personal identifying information due to privacy regulations (see discussion). A unique identifier for each case generated by iPHIS was used to link cases in iPHIS to the web-based database when the two datasets were merged in Stata (version 11.0).

Due to privacy and data sharing regulations Manitoba Health and the WRHA continued to maintain their own databases. Near the end of the second wave, the WRHA was able to export data and send electronic updates to Manitoba Health instead of faxing updated investigation forms. Additionally, First Nations Inuit Health (FNIH) in Manitoba was given read only access to FNIH cases hospitalized in Winnipeg in the Web Data database allowing real time access to information.
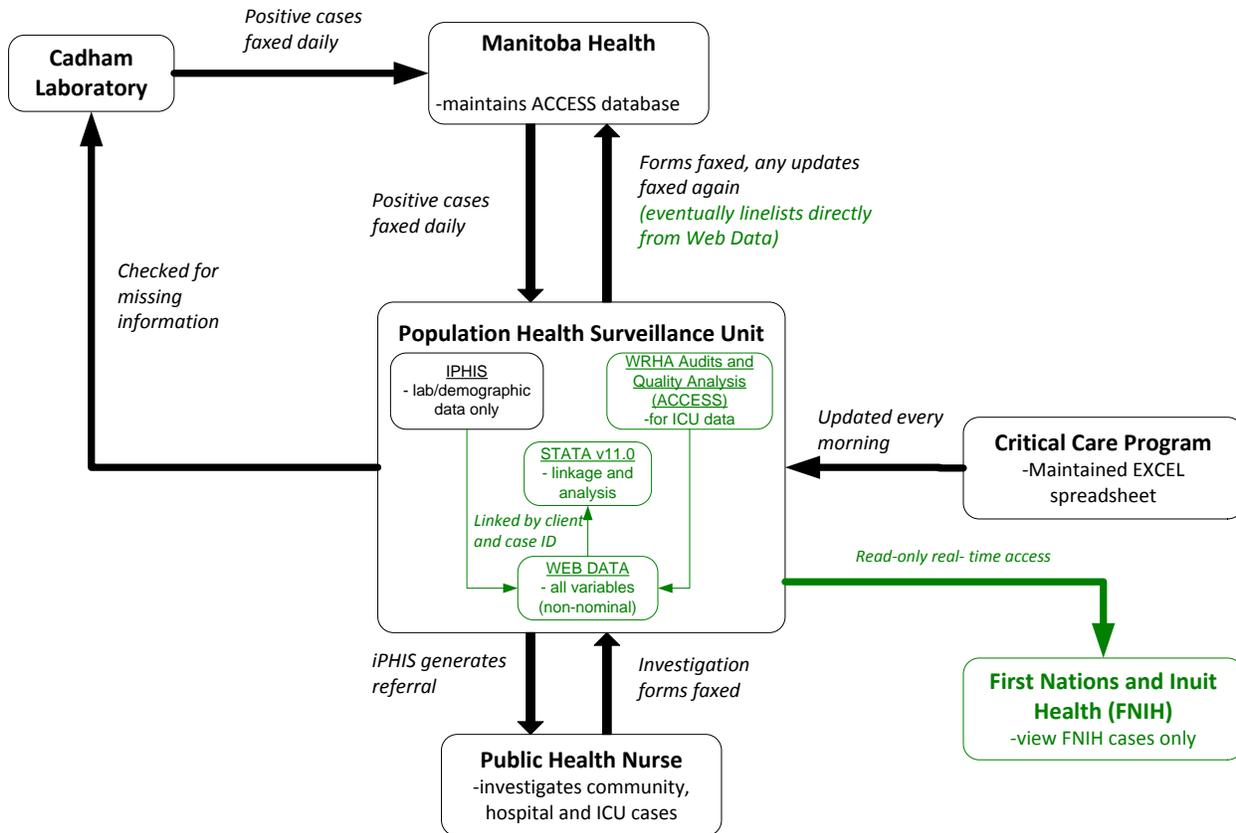
Figure 1: Flow of information before and after (green) the implementation of Web Data technology in the WRHA. Information in green indicates the changes that were made.

Based on the survey conducted, the limitations and benefits found by employing Web Data technology for H1N1 case management in the WRHA are as follows:

Benefits

Data Access: At the beginning of the first wave, data from the investigations forms for hospitalized cases were limited to the Population Health Surveillance Unit. With the Web Data database, multiple users could access data in real time. The Web Data database provides secure access that may be audited. In addition remote access to data was easily possible with the internet-based setup. Both features allowed the WRHA to give specific read only access to FNIH.

Data Collection/Extraction: The Web Data database was easily modeled after the provincial surveillance form. Information could be easily extracted for analysis and then transferred to Manitoba Health electronically via a secure file transfer website.

Data Storage/Management: At the beginning of the pandemic, laboratory and demographic information were being entered into iPHIS; data from the investigations forms for hospitalized cases were being entered into an Excel file and another Excel file was used to track clinical updates of patients in the ICU. There was no well-designed system in place to manage pandemic

6

H1N1 data. With the Web Data database, it was possible to rapidly create and modify the working database, as well as adapt to revisions to the case investigation form. Duplicate manual data entry of mandatory data fields was eliminated.

Data Queries/Reports: There was the ability to create and save queries and to easily update case information. Basic reports could also be quickly generated.

Limitations

Data Sharing/Linkage:

Linking epidemiological, clinical and laboratory data was hindered by the lack of pre-defined processes to deal with privacy policies.

Working with, and understanding of, current personal health information privacy policies will enhance data sharing between regional and provincial jurisdictions, and decrease duplication of data entry by having a common shared data management system for multiple stakeholders.

The experience of Alberta Provincial Laboratory (ProvLab):

In the province of Alberta, ProvLab provided centralized testing for respiratory viruses and was the front-line support in confirming H1N1 cases which was critical in the coordination of pandemic response plans. ProvLab employed a web-based platform called DIAL (Data Integration for Alberta Laboratories) [5], a technology also part of the CNPHI framework, specifically designed to enhance the benefits of laboratory data surveillance, and to support easy access to near real-time laboratory data.

Since multiple specimens may be submitted to the ProvLab for a single patient, it was very difficult to have an automated algorithm to translate specimen-based data into patient-based data. Thus, an entirely manual process was initially adopted at ProvLab where a single user generated a patient-based linelist in Excel for public health stakeholders (Figure 2). It is important to note that Figure 2 depicts the flow of information from the laboratory perspective of the ProvLab, i.e., not from a provincial perspective as was illustrated for the WRHA experience, and therefore provincial departments have largely been omitted from the figures.

As illustrated in Figure 2, patient demographics, and pertinent health information, along with specimenswere submitted to the ProvLab for influenza/respiratory virus testing. When a specimen tested positive for H1N1, a single user would manually enter the patient demographic information of the positive specimen into an Excel spreadsheet. Each day the user had to check for duplicate specimens from a single patient at least once to maintain a patient-based dataset. On a daily basis, this excel spreadsheet was sent to Alberta Health and Wellness by the single user as a province-wide patient-based linelist. The linelist was also employed to generate daily reports for Public Health Professionals. At the same time, appropriate personnel from Alberta Health and Wellness and other regional public health stakeholders, also maintained their own ACCESS database using information from notifiable laboratory reports for H1N1 cases that were reported by ProvLab.
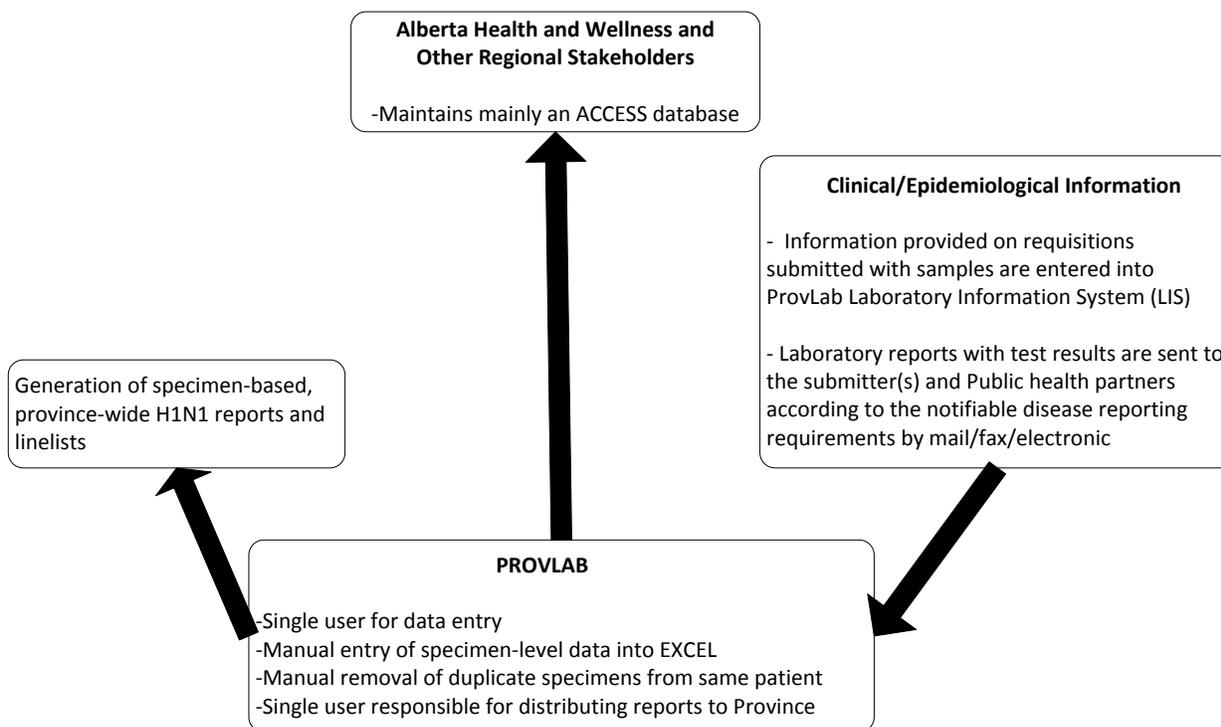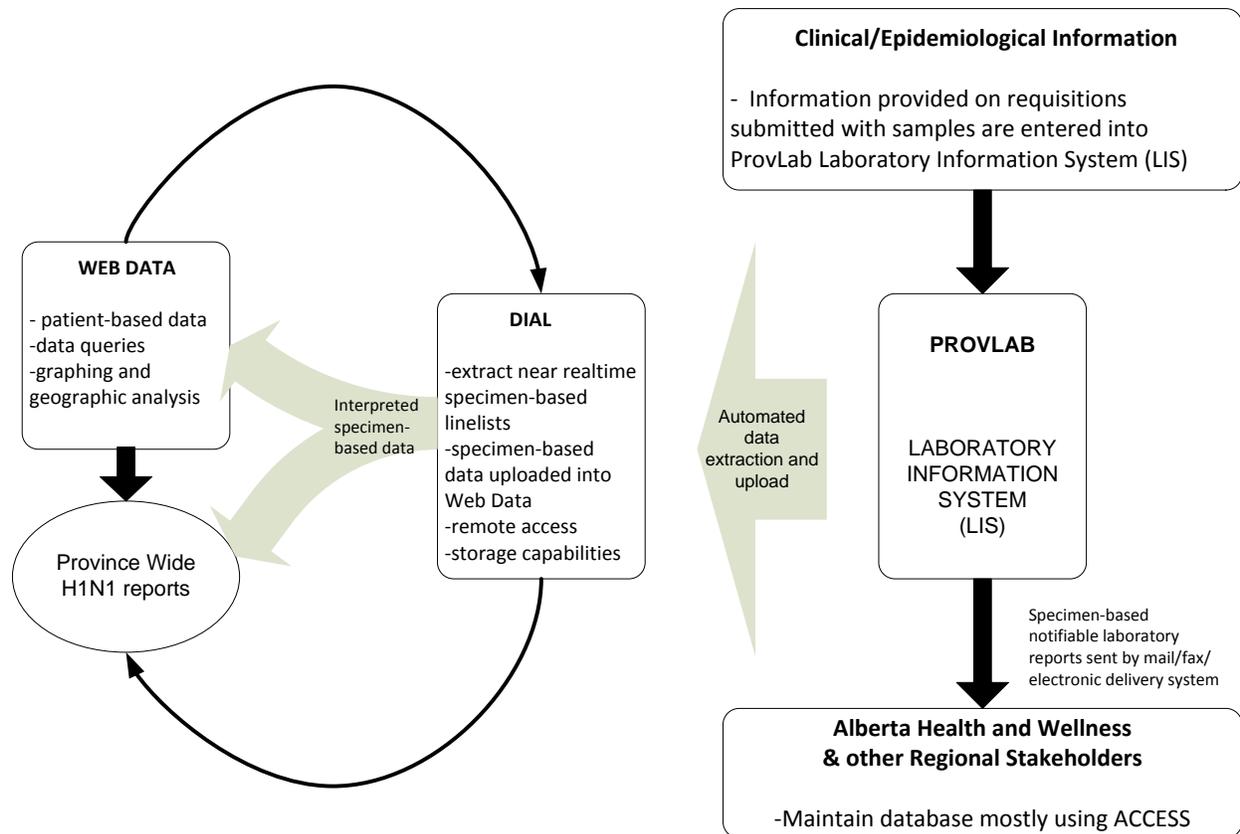
Figure 2: ProvLab H1N1 case management using the previous system.

The Web Data technology was rapidly deployed within ProvLab during the pandemic due to the immediate need for managing case data in a structured manner allowing for web-based secure sharing and collation (see figure 3). The first H1N1 case was reported on April 28, 2009 in Alberta. The first meeting to discuss possible implementation of the Web Data technology as part of the pandemic response from the laboratory perspective was held in the afternoon of May 4, 2009. The database was designed with stakeholders' input and set up quickly. Data was entered and the Web Data technology was deployed and became functional with the first case-based line list extract distributed to public health partners on May 11, 2009. The application continues to be used for influenza case-based data in the province by both laboratory and public health partners.

Initially, specimen-based data was manually entered into Web Data to provide patient-based data. However, as the number of positive H1N1 specimens increased, modified specimen-based linelists generated using DIAL were batch uploaded (i.e., multiple specimens at one time) into the Web Data technology removing the need for manual data entry. A review process was developed to compare H1N1 positive specimen test results to patient-based information in Web Data. This was needed to identify duplicate positive specimens submitted from a single patient. This allowed for the maintenance of a web enabled patient-based database of all influenza A positive specimens in Alberta from the start of the pandemic. The queries function of the Web Data technology allowed easy access and extraction of patient-based linelist for public health stakeholders.

8

Data connectivity was also established between the Web Data and DIAL so all the analytical functionalities of the DIAL platform could be used to analyze patient-based H1N1 data. Data analysis went from non-existent to excellent for patient-based data, so that laboratory quality issues and public health questions could be answered quickly. Secure and audited access to DIAL and Web Data based case management databases were available remotely through the Internet. Separate web databases were set up for the province of Alberta, the Northwest Territories and Nunavut to suit specific user needs in compliance with the Health Information Act.

Figure 3. ProvLab perspective of H1N1 case management using Web Data supported by DIAL.

Based on the survey conducted, the limitations and benefits of using Web Data technology for H1N1 case management in the ProvLab are as follows:

Benefits:

Data Access: Data could be accessed remotely by authorized users via the internet.

Data Collection/Extraction: Previous manual entry of data was practically eliminated except for a manual check of the transformed patient-based data generated from specimen-based linelists before batch upload into the Web Data technology. Validated lab data were updated onto DIAL automatically twice daily. Batch data support allowed for uploading of a large amount of specimen data at one time. With Web Data manual entry errors were largely reduced.

9

Data Storage/Management: Data automatically transformed from specimen to case-based data. The database could accommodate an evolving H1N1 case definition and provided a relational database for patients with multiple specimens.

Data Queries/Reports: Data storage and management, as well as patient-level data allowed for excellent capabilities for generating queries and reports.

Data Sharing: All laboratory results are stored in the ProvLab laboratory information system with restricted user access. However, the Web Data technology and DIAL can be accessed remotely and securely. These applications allow authorized regional jurisdictions easy access to relevant data.

Improvements required

Data Linkage: The case management database was used exclusively in the laboratory setting for managing cases with multiple specimens and it was not linked to clinical and epidemiological information. As a result clinical and epidemiological data could not be used to define specimen-testing priorities at the height of the pandemic.

Discussion

Both the Winnipeg Regional Health Authority and Alberta ProvLab had positive experiences with implementation of the Web Data technology. The main benefits included:

Improved and Secure Data Access.
Web Data allowed for multiple users to access the database in real time, such as laboratory personnel, epidemiologists, and public health officials. The database was secure and was accessible via the Internet with user audit functions.

Improved Efficiency and Error Reduction with Batch Data Uploader
The single user manual input of data was inefficient and errors were more easily made. Automated batch uploading allowed for very efficient data uploading, removed duplication of efforts, and reduced data entry errors. There was also less strain on human resources during pandemic.

Electronic Collection and Transfer of Data
The electronic collection and transfer of data was efficient and reduced errors associated with manual data entry. Updates could be automatically sent to the appropriate public health officials in a timely manner.

Rapid Creation and Modification of the Database
Over the course of the pandemic, the epidemiology of H1N1 changed, requiring the modification of the H1N1 case definition on investigation forms. Web Data technology provided flexibility and allowed for rapid creation, or adjustment of existing variables as needed. It could

accommodate clinical, epidemiological and laboratory data input needs in a timely fashion and could respond to changes in processes and algorithms.

Conversion of Specimen- to Case-Level Data
Case based data were needed to understand the epidemiology of the pandemic, for public health decision making and setting laboratory-testing priorities at the ProvLab.

Data Extraction and Queries
The ability to create, save and sort within queries was possible.

Although the WRHA and the ProvLab had different approaches to, and applications of, the pandemic case management tool, the limitations to Web Data and the areas requiring improvement were similar:

Compliance with Privacy Policies/Laws
Protection of the privacy of personal information, which encompasses one's health information, is paramount in both Alberta and Manitoba [6, 7]. Typically, personal information (e.g., personal health number, name, date of birth) cannot be stored on physical hardware outside the jurisdiction where it was collected. It is important therefore, that government policies and technological advances be established to maintain privacy such that data sharing and jurisdictional collaboration occurs as efficiently as possible during pandemic response. One possible solution to this problem is to deploy a local instance within a jurisdiction.

Data Sharing and Linkage between Jurisdictions
It was difficult to share information across various stakeholders that employed unlinked data systems. It was clear that each group collected only a part of the data required to paint a complete picture of the outbreak and epidemiology of the pandemic required for a collaborative response. Having a single database accessible to each authorized professional jurisdiction to input their data would allow for the best use of resources. Having access to other geographical jurisdictions would allow for best practices and epidemiological insight to become more apparent.

Duplication of Data Entry
Duplication of data entry also stems from privacy policy. Each jurisdiction (e.g., WRHA or Critical Care Unit) maintains its own database in order to protect the personal information of individuals. In WRHA, an additional layer of coding using duplicated data entry (i.e., the linkage of iPHIS to Web Data database via Stata) was applied with the usage of Web Data to protect privacy.

Comparison to other Web Data like tools

Internet has become an integral part of health care with 77% of health department staff reporting use of the internet to search for health information at least once or twice a day [8]. In 2004, while 84% of 56 health departments reported the use of web-based reportable disease lists in the United States, only 9% had secure web-based capability for case reporting [9]. A more recent

survey still revealed substantial variations in the adaption of electronic system in United States with some web-based manual electronic laboratory reporting system [10].

The concept of building rapid databases is also available though publicly available Epidata software, which is an open-source software that allows users to build custom data collection tool. When compared to Web Data technology, however, Epidata is limited by single user data entry, desktop focus and is not ideal for large datasets [11].

The description of other similar electronic web-based system is limited and there is little data on the deployment and user feedback to allow direct comparison with the Web Data technology [12-14].

Conclusion

The Web Data technology provides an innovative platform to rapidly implement, with minimal technical expertise, a secure web-based database during an outbreak. It provides an easy-to-use interface for data entry and mechanisms to seamlessly link data, query and export data, and to adapt to change. In order to maximize its effectiveness and use, there needs to be simultaneous work on resolving data sharing and privacy issues before an outbreak.

Acknowledgements

Correspondence

Dr Shamir N Mukhi, 204.771.4698, shamir.nizar.mukhi@phac-aspc.gc.ca

References

[1] Martin SM, Bean NH. 1995. Data Management Issues for Emerging Diseases and New Tools for Managing Surveillance and Laboratory Data. Emerg Infect Dis. 1(4), •••. http://dx.doi.org/10.3201/eid0104.950403
[2] Learning from SARS: Renewal of Public Health in Canada. A report of the National Advisory Committee on SARS and Public Health. October 2003.
[3] Mukhi SN, Aramini J, Kabani A. 2007. Contributing to communicable disease intelligence management in Canada. Can J Infect Dis Med Microbiol. 18(6), 353-56.
[4] Aramini J, Mukhi SN. Canadian Application of Modern Surveillance Informatics. In: Lombardo J.S. and D. L. Buckridge Disease Surveillance: a public health informatics approach, pp. 315-328. John Wiley & Sons Inc. Publication.

[5] Mukhi SN, May-Hadford J, Plitt S, Preiksaitis J, Lee B. DIAL: A platform for real-time Laboratory Surveillance, Online Journal of Public Health Informatics, ISSN 1947-2579, http://ojphi.org Vol.2, No. 3, 2010.

[6] Government of Manitoba. 2007. Freedom of Information and Protection of Privacy (FOIPP) Act. Available at http://web2.gov.mb.ca/laws/statutes/ccsm/f175e.php Accessed July 2010.

[7] Government of Alberta. 2003. Freedom of Information and Protection of Privacy (FOIPP) Act. Available at: http://www.oipc.ab.ca/pages/PIPA/ReadAct.aspx Accessed July 2010.

[8] Turner AM, Petrochilos D, Nelson DE, Allen E, Liddy ED. 2009. Access and use of the Internet for health information seeking: a survey of local public health professionals in the northwest. J Public Health Manag Pract. 15(1), 67-69. http://dx.doi.org/10.1097/01.PHH.0000342946.33456.d9

[9] M'ikanatha NM, Welliver DP, Rohn DD, Julian KG, Lautenbach E. 2004. Use of the Web by state and territorial health departments to promote reporting of infectious disease. JAMA. 291, 1069-70. http://dx.doi.org/10.1001/jama.291.9.1069

[10] CDC. 2009. Status of State Electronic Disease Surveillance Systems --- United States, 2007. MMWR. 58, 804-07.

[11] Achonu C. Epidata: A potential tool for pandemics and large scale outbreaks, http://www.apheo.ca/resources/projects/epidata/CDSN%20epidata%20presentation%2028oct10.pdf, accessed on May 12th, 2011.

[12] Snacken R, Manuguerra JC, Taylor P. 1998. European Influenza Surveillance Scheme on the Internet. Methods Inf Med. 37, 266-70.

[13] Kant L, Krishnan SK. 2010. Information and communication technology in disease surveillance, India: a case study. BMC Public Health. 10, S11. http://dx.doi.org/10.1186/1471-2458-10-S1-S11

[14] Mao Y, Wu Z, Poundstone K, Wang C, Qin Q, et al. 2010. Development of a unified web-based national HIV/AIDS information system in China. Int J Epidemiol. 39, S79-89. http://dx.doi.org/10.1093/ije/dyq213