

Canadian animal health surveillance network – the road to big data analysis

H KLOEZE¹, A TROKHYMCHUK², T BURNS³, N DE WITH⁴, G DUIZER⁵, H AKWAR⁶, S MUKHI⁶

¹Canadian Food Inspection Agency, Owen Sound, Canada; ²Prairie Diagnostic Services, University of Saskatchewan, Saskatoon, Canada; ³Centre for Coastal Health, Chilliwack, Canada; ⁴British Columbia Ministry of Agriculture, Livestock Health Management and Regulatory Unit, Abbotsford, BC; ⁵Manitoba Ministry of Agriculture, Winnipeg, Canada; ⁶Canadian Network for Public Health Intelligence, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Canada

*harold.kloeze@inspection.gc.ca

Abstract

The Canadian Animal Health Surveillance Network (CAHSN) began as a project in 2006 and has made continuous progress since then. From the outset the goal was to establish a national non centralised surveillance infrastructure based upon the animal health laboratories in Canada. The surveillance component of CAHSN is based upon collation of laboratory test results with data and analytics provided by the Canadian Network for Public Health Intelligence. The implementation required collation of laboratory results but there were serious concerns raised that could potentially derail progress. The concerns involved technical feasibility, privacy, data security and the potential uses of the information and data being shared. Each issue was addressed in order to ensure that the trust developed in the partnership was maintained and enhanced. The CAHSN has evolved and currently has initiatives in disease reporting, syndromic surveillance, disease response and specific surveillance for diseases of importance with new initiatives being developed.

Keywords: *Animal health, surveillance, syndromic surveillance, big data*

Introduction

There is a plethora of available information on animal health related surveillance, in textbooks, peer-reviewed journals, other journals or periodic publications, and through our wired and connected universe of the Internet. In this paper we describe the steps taken to form and actualise a surveillance framework which is non-centralised, non-prescriptive and functions largely as a “collaboration of the willing”.

In Canada responsibility for animal health is shared between the federal and provincial governments, as well as industry. The federal government has responsibility for diseases which are named as reportable under federal legislation, the *Health of Animals Act* and its supporting regulations, with the list of diseases being similar to those diseases included by the World Animal Health Organisation (OIE) as List A diseases. The Canadian Food Inspection Agency (CFIA) is the federal body responsible for responding to these diseases and all suspected occurrences of these diseases must be reported immediately to the CFIA. In Canada no single organisation or governmental body is responsible for the other animal health related diseases. This list includes diseases which are

zoonotic (affect humans but have animal origin or have animal reservoirs), emerging or re-emerging diseases, and endemic diseases not named as reportable which was recognised as a gap in national animal health surveillance in Canada.

This gap was addressed in 2006 via funding from an initiative of Defense Research and Development Canada (DRDC), the current name of this initiative is the Centre for Security Science (1). DRDC provided the funding since the gap revealed a vulnerability for bioterrorism which could result in large scale and potentially irreparable financial damage to Canada’s agricultural sector, with impacts on animal suffering and human welfare. The resulting initiative is the Canadian Animal Health Surveillance Network (CAHSN), which is a federal-provincial-academic animal health network based upon public sector animal health laboratories in Canada (2).

The CAHSN was formed and from the outset had three foci, which together formed an infrastructure capable of detecting and assisting response to serious diseases of animals. The first part is a federal-provincial-academic laboratory network to enhance the detection of animal diseases and to provide surge testing capacity for animal diseases. This has been enabled through the training of over 100 analysts across Canada for our major foreign animal disease threats. The CAHSN includes proficiency testing of analysts, enhancements in laboratory infrastructure to assure that proper biocontainment levels are available in partner laboratories, assistance with quality assurance and regular teleconferences to discuss issues and threats. The laboratory component is led by the CFIA’s National Centre for Foreign Animal Disease in Winnipeg but other laboratories provide assistance and expertise depending upon the issues being faced.

The second part of the CAHSN is the establishment of a national early warning system by formation of a comprehensive surveillance infrastructure which combines animal disease test results from across Canada. The third component is a common scientific informatics platform to allow collection, compilation and analyses of the test results, as well as to allow for a common communication platform for rapid information exchange and alerting. This platform is the Canadian Network for Public Health Intelligence

(CNPHI), which has over 4000 users, primarily involved in various aspects of public health in Canada (3). CNPHI is an initiative of the Public Health Agency of Canada (PHAC) developed and managed by the National Microbiology Laboratory (NML) in Winnipeg. The remainder of this paper will describe the second identified component of CAHSN, the surveillance initiatives.

Discussion

Issues affecting implementation

The World Health Organisation defines public health surveillance as follows: *Public health surveillance is the continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice. Such surveillance can: serve as an early warning system for impending public health emergencies; document the impact of an intervention, or track progress towards specified goals; and monitor and clarify the epidemiology of health problems, to allow priorities to be set and to inform public health policy and strategies* (4). While this is not exactly transferable to the animal health realm it serves as a better guide for our work in CAHSN than more traditional animal health surveillance definitions. The focus of CAHSN is not traditional surveillance but is early warning surveillance, and response, on a national or regional basis. In much the same way that population or herd medicine moved the unit of interest away from individual animals to the herd level, CAHSN attempts to move the focus to the regional or national level.

CAHSN, through the CNPHI informatics platform, is able to enact a close to real time data transmission of laboratory results from the various laboratories in Canada. From the outset this data feed was envisioned to allow for data analyses, cross country reports on diseases of interest, syndromic surveillance and other novel analytics using large data sets. When initial attempts were made to implement this data collection various issues were raised indicating that there needed to be a considered approach to implementation that addressed all concerns of our partners.

The first issue raised was that of technical feasibility. The disease test results reside in different Laboratory Information Management Systems (LIMS), each with their own nomenclature and their own data structures. In addition, the testing done by each laboratory is unique since each laboratory does testing that is pertinent to their clients. The data can be retrieved from each LIMS, but it is often a laborious process. The first system we implemented was a national laboratory test results aggregation for Bovine Spongiform Encephalopathy (BSE). This system was designed, implemented and maintained in part as a way of proving that the basic technology worked, that data could be reliably transferred from each LIMS to CAHSN in near real time (once daily at present) and that reports could be developed. At this point the BSE system contains well over 300,000 test results with daily additions, this system, while important in and of itself, has also functioned as a

demonstration of the agile and organic technology that allowed for data collation from disparate sources.

The second issue raised was that of privacy. Canada, and each province, have legislation designed to protect the privacy of citizens. There are prescribed ways to investigate this issue and to protect privacy. This was accomplished through a dedicated project, funded by the CFIA that investigated the impact of the future data collection on privacy. A legal opinion was obtained from the government's Department of Justice and a Privacy Impact Assessment (PIA) was conducted as per federal policy and the results were shared with the partners. To ensure that only required data was collected and that privacy concerns were minimised, a group of federal and provincial animal health epidemiologists devised a "minimum data set" that ensured effective analysis was possible. This work has been peer reviewed and published (5) and continues to form the basis for the data collections. The data set does contain geographic references which are non-specific and indicate an region where the premises is located but names of owners, specific locations or addresses and other contact information are not part of the data set.

A third concern was that of data security. Partners were concerned whether the transfer and storage of their information would be safe. The CFIA project around privacy has a second component which was a security review comprised of a Statement of Sensitivity, Concept of Combined Operations, and a Threat Risk Assessment on the data transfer protocols (these products form the basis of the IT security review process in CFIA). This was revisited after five years in 2016.

The final major concern raised was that of purpose of the data collection and use of the data. Upon completion of the PIA and security assessment, an Information Sharing Agreement was finalised. This agreement laid out the data to be collected (based on the "minimum data set"), the purposes the data may be used for and the publication arrangements regarding reports derived from the data. This agreement is not binding and does not force data sharing; it simply provides a framework. The agreement is bilateral between CFIA and each partner to allow the partners to adopt this on their terms and timeframe. At this time there are 10 signed agreements and the final two are in the signature process. The bilateral arrangements were chosen because obtaining signatures on a comprehensive agreement covering all parties can be exponentially more challenging.

The final result of the processes above was development of trust, which is the basis for our partnership and allowed our progress to date. Various concerns were addressed as can be evidenced by the discussion above. Discounting concerns about security, or privacy may have saved time but would have fractured any developed trust.

Products and technical progress

At this time a novel surveillance application has been developed on the CNPHI platform to collate data and allows for interactive and rapid data interrogation and analysis. This system, referred to as Animal Health Syndromics, includes close to real time data transmission from the three public sector laboratories in Western Canada that conduct testing for clinical submissions. These laboratories are the provincial laboratories in Manitoba and British Columbia as well as Prairie Diagnostic Services in Saskatchewan which is a partnership of the government of Saskatchewan and the University of Saskatchewan with Directors from the four western Canadian provinces. The province of Alberta does not have a public animal health diagnostic laboratory. This provides a data flow of all clinical laboratory results from these laboratories. In calendar year 2015 this comprised some 50,000 submissions which consisted of 947,301 recorded tests. This has allowed the development of diverse products. The products are not controlled by a central administration but are developed by our partners or through projects and are all collaborative, the major ones are described here.

Disease specific surveillance: The BSE system mentioned previously collects all BSE test results from the eight testing laboratories in near real time. The data is enhanced with downloadable reports, charts and tables which were designed in consultation with the system users. A novel Quality Assurance system is being developed based upon this data collection to allow investigation of data errors such as incorrect dates. There are also data collections for other Transmissible Spongiform Encephalopathies and one is under development for Rabies. The basic framework of CAHSN is very useful for data aggregation when multiple laboratories test for a disease.

Syndromic surveillance: Syndromic surveillance has been an objective of CAHSN from the outset but was hampered by lack of data. The basis for our analytics is to use laboratory submission data, collected in close to real time, and to develop syndromes based upon either test requested or sample type submitted. The concept is that if a veterinarian submits lung tissue from a beef farm and requests Bovine Herpesvirus 1 and Parainfluenza-3 testing, they are investigating a respiratory syndrome. Various analytics are applied to this such as Cusums or moving averages for signal generation.

Disease reporting: The British Columbia Ministry of Agriculture has commissioned the Centre for Coastal Health (an epidemiology group) to develop specific disease reports using CAHSN data. This initiative started with the development of poultry disease specific reports for Infectious Laryngotracheitis, *Mycoplasma gallisepticum* and avian reovirus. The project has been very successful and has been renewed for three more years and will include three swine and bovine diseases. This project has not only added value

to the data collected but has provided a reporting template. One very important lesson from this project was to consult extensively both to select diseases to report on and to refine the case definition for reporting. Diseases were selected that were of interest but would not raise any concerns regarding trade. The case definitions were developed in consultation with the laboratories to ensure that the testing performed was understood and accurately reflected in the reports.

Emergency response: The availability of the data and the existing partnership allows for rapid collection and collation of laboratory information in a disease response situation. This was proven in 2009 with the detection of pandemic Influenza A/H1N1/09 virus in swine in Alberta. Within 10 days we were able to initiate rapid data collection using CNPHI platform, which continued for the next two years (6). The CAHSN partnership proved to be a vital component in Canada's response to the animal health aspects of this disease incursion through data collection as well as coordinated laboratory testing and shared test protocols.

New projects: Independent work on antimicrobial susceptibility testing (AMS) data as recorded in CAHSN has led to projects on a provincial level in British Columbia. We are currently developing a project to further this work and extend to western Canada as well as investigate the possibility of integration with human origin data. A second project under development will further our "big data" analytics through development of new indicators using Bayesian techniques.

Conclusion

The CAHSN has been under development for a decade. The development has been stepwise and concerns and issues were addressed to ensure trust was developed. While our focus from the outset was surveillance, and specifically "big data" analytics based upon an aggregated laboratory data set, the development initially was much more involved in building a framework or foundation which addressed all identified concerns of our partners. The value of the network is now becoming evident with the number of new and innovative projects that are possible because of the framework developed and the data that is being collected in near real time. This development would not be possible without the support and hard work of all CAHSN partners.

References

1. **DRDC.** www.drdc-rddc.gc.ca/en/research-centres.page
2. **Kloeze H et al.** *Transbound Emerg Dis* 57, 414-419, 2010
3. **Mukhi S et al.** *Can J Infect Med Microbiol* 18, 353-356, 2007
4. **WHO.** www.who.int/topics/public_health_surveillance/en/
5. **Kloeze H et al.** *Transbound Emerg Dis* 59, 264-268, 2012
6. **Kloeze H et al.** *Can Vet J* 54, 501-503, 2013