Collation and Analyses of Data to Improve Animal Health Surveillance in Canada

Challenges, Options, Recommendations

November 2014

This work was supported by funding in-kind and in dollars provided by:
AAFC Growing Forward2, CFIA, BMAFF, PDS, MAFRD, OMAFRA and AHL

Table of Contents

Executive Summary

Background and Objectives

- Surveillance is important for its ongoing systematic collection, collation and analyses of information related to animal health, and the timely dissemination of information to those who need to know so that action may be taken.
- The objectives of animal health surveillance are to detect important changes in health, and to provide evidence of health in support of safe trade.
- GF2 funding was allocated to BC, Saskatchewan and Ontario for three separate but related projects to improve surveillance in avian, bovine and porcine sectors respectively.
- While details of the provincial projects differed, leaders of all three agreed to work together to study the advantages and challenges of collating data on the CNPHI platform for national analyses.
- Emphasis was on collation of laboratory data, but CNPHI's ability to manage questionnaire data was also considered.
- The objectives of this interim report (November 2014) were to identify: advantages, challenges, options and recommendations for the collation and analyses of animal health surveillance data in Canada, on the CNPHI platform.

Methods

- Ten production-limiting diseases were studied including three avian, three bovine, and four diseases of swine.
- A list of key destination-data-fields for a data file were identified, into which equivalent data were transferred for the ten diseases, from the respective LIMS of four provincial laboratories (BCMAFF, PDS, MAFRD, AHL), for two time-periods, including: January to March 2013 and June to September 2014.
- The steps required to: collect, collate, standardize and concatenate the data from each source laboratory into a standardized file; were all documented.
- Subsequently, challenges, options and recommendations for managing laboratory data for national animal health surveillance were documented.
- Previously RAIZO, CSHIN and OAHN had demonstrated that clinical-impressions-questionnaires could be useful to obtain surveillance information from veterinary practitioners.
- Accordingly, the questionnaire system available on CNPHI (as of November 2014) was compared to commercial questionnaire software used by CSHIN and OAHN (i.e. FluidSystems™)
- Subsequently, challenges, options and recommendations for questionnaire data were made.

Findings, Challenges and Options

- Many of the findings were already known to experts in this area, but it was useful to document them for a broader decision-making audience.
- Different laboratories used different LIMS platforms with different structures and abilities, making it very challenging to collate equivalent data from them.
- In general LIMS were designed to run laboratories and were not well designed for retrieval of epidemiologically-sound data.
- Coding of parameters was inconsistent within and between LIMS, making it difficult to translate data to a common system of coding, to facilitate reliable epidemiological analyses and counting within a large standardized file.

- While well advanced and having some advantages; as of November 2014, the CNPHI questionnaire system did not yet have the complete utility of commercial questionnaire systems (e.g. FluidSurveys), to be used on animal health surveillance networks (e.g. CSHIN, OAHN). However, it may be significantly improved in the near future.
- Some high-level options for collation of national animal health surveillance data include:
    - Not attempting to collate data into a national system, but making specific requests for information (e.g. counts) from each pertinent jurisdiction as needed (e.g. request each lab to submit a summary report of number tested for disease X in timeframe Y and the number test positive among those). Unfortunately some laboratories cannot meet such requests because their respective LIMS are not designed to do such counts.
    - Create, implement and maintain one huge standardized LIMS to be used by all veterinary diagnostic laboratories in Canada. This would be a very large and complex task and is not likely feasible given the desired independence of laboratories.
    - Improve standardization among LIMS among key data fields and create programs that translate data from each LIMS into a standardized files in CNPHI using standardized structure and coding, to then be used for epidemiologically-sound analyses.
    - Allow jurisdictions and surveillance initiatives to use the questionnaire software of their choice to collect data, but to translate files into standardized structure and coding for uploading to CNPHI for storage and analyses.
    - Improve the CNPHI questionnaire software module so that it meets or exceeds the utility of commercial counterparts for collection, collation, concatenation and analyses of questionnaire data on CNPHI.

Key Recommendations

- A total of 17 recommendations were made in the detailed report.
- Many of them would require significant investment, commitment and cooperation from industry, federal and provincial governments to CNPHI.
- Key high-level recommendations include:
    - Standardize key data fields and coding in veterinary diagnostic LIMS across Canada.
    - Animal health invest in translating and uploading key data from veterinary LIMS into a standardized national files in CNPHI, on an ongoing basis, for ongoing systematic summary and trend analyses including funding CNPHI programmers' creation of an animal health surveillance application similar to the existing CNPHI Public Health Enterics application, but for animal diseases/hazards of concern, identified by species-specific network groups within the network-of-networks approach to national animal health surveillance as supported by the NFAHWC; and generating case count analyses and graphs at the premises/lab-submission level rather than the individual animal (person) level like the Public Health Enterics application currently does .
    - Fund CNPHI staff to build data-mining, data-massage, data-summary and statistical analyses software behind the CNPHI firewall, which is easy to use and easy to obtain training for use by approved federal and provincial users.
    - Fund CNPHI to build / improve highly flexible electronic questionnaire software system in CNPHI that meets or exceeds the utility of commercial questionnaire software.
    - Design and implement standard files in CNPHI to collect, organize, store analyze data from clinical-impressions-questionnaires (commercial or CNPHI), on an ongoing bases, building (concatenating) files over time nationally for respective sectors in the network-of-networks surveillance system.

- Implement a farm premises ID system linked with valid longitude latitude location data and require its use on each and every laboratory submission with linkage to all testing and results data. Alternatively, require the submitter to record the longitude and latitude in decimal degrees (to three decimal places using their smart-phone or GPS), for the location of where the animal(s) were at the time the sample(s) were collected and record that information on all respective laboratory submissions for capture in LIMS systems.

Glossary and Acronyms

Term	Explanation
AAFC	Agriculture and Agri-Food Canada
AHA	Animal Health Act
AHL	Animal Health Laboratory
BCMAFF	British Columbia Ministry of Agriculture Food and Fisheries
BI	Business Intelligence - a brand of data-mining software
CAHSN	Canadian Animal Health Surveillance Network
CFIA	Canadian Food Inspection Agency
CNPHI	Canadian Network of Public Health Intelligence
Cognos	Cognos - a brand of data-mining software
CSHIN	Canadian Swine Health Intelligence Network a national network of swine health surveillance modeled after RAIZO with national conference calls among participating regions and provinces for swine (see Network-of-Networks, RAIZO, OAHN)
DSP	Disease Surveillance Plan
FAD	Foreign Animal Disease
FluidSurveys	FluidSurveys - a brand of survey questionnaire software
GF2	Growing Forward 2 - an AAFC funding program
LIMS	Laboratory Information Management System
MAFRD	Manitoba Agriculture Food and Rurural Development
MAPAQ	Le ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec
Network-of-Networks	phrase describing combining national sector specific networks with regional or provincial networks of expert committees monitoring clinical and laboratory animal health information and communicating findings (see also RAIZO, OAHN, CSHIN)
NFAHWC	National Farmed Animal Health and Welfare Council
OAHN	Ontario Animal Health Network - a provincial animal health surveillance system for Ontario, modeled after RAIZO, involving provincial quarterly conference calls among sector specific expert committees, with participation in CSHIN and other sector specific national calls contributing to the national network-of-networks (see RAIZO, CSHIN, network-of-networks)
OMAFRA	Ontario Ministry of Agriculture Food and Rural Affairs
SEAC	Surveillance and Epidemiology Advisory Committee
PDS	Prairie Diagnostic Services
PHAC	Public Health Agency of Canada
RAIZO	Reseau d'Alerte et d'Infromation Zoosanitaire - a system of sector specific expert committees each with representation from private veterinarians, laboratory, veterinary college research and provincial government - who monitor clinical and laboratory diagnoses and discuss via quarterly provincial conference calls by sector, contributing to the Network-of-Networks. (see OAHN, CSHIN, Network-of-Networks)

Background

- Surveillance is defined as the systematic ongoing collection, collation and analyses of information related to animal health, and the timely dissemination of information to those who need to know, so that action can be taken.
- Its objectives are: a) to detect important changes in health status in a timely manner, and b) to provide evidence of health status in support of safe trade.
- Animal health surveillance is important to the welfare of: animals, the agri-food economy, trade, food safety and public health.
- Veterinary practitioners and diagnostic laboratories serve as important sources of animal health surveillance data
- Notwithstanding the challenges of collating and standardizing data from various sources (e.g. provincial laboratories, clinical-impressions-surveys from species-specific networks), there are distinct advantages to analyzing appropriately collated and standardized data sets to attain national health surveillance information. Some examples include:
    - Obtaining baseline counts and documenting normal variability of submissions and test-results for various diseases and hazards of concern, by region and host animal species and production types across Canada.
    - Being able to demonstrate and quantify national surveillance coverage, providing evidence of quality and safety in support of trade and stable prices for all.
    - Being able to detect changes in health beyond normal variability in counts among host species and types, in space (e.g. by region), and over time (e.g. this quarter vs. last quarter, or vs. the same quarter last year).
- There have been calls to improve animal health surveillance in Canada
- Accordingly, GF2 provided funding to do so.
- Subsequently, three separate but related projects were proposed and accepted for GF2 funding, based in BC, Saskatchewan and Ontario, with an emphasis on poultry, cattle and swine surveillance respectively.

Objectives

- The projects were independent and thus differed in their respective design and scope.
- Accordingly, each project had its own objectives and responsibilities for reporting back to respective overseers.
- However, at least one objective was common to all three projects, which was:
    - To study challenges and opportunities, and to make recommendations for uploading relevant data (e.g. from veterinary diagnostic laboratories and other sources), into the CNPHI platform as part of the CAHSN, for ongoing collection, collation and analyses, to improve animal health surveillance in Canada.
- More complete reports on the individual projects will follow at a later date.
- The objective of this interim report was:
    - To summarize methods, findings, challenges, opportunities and recommendations (as of November 2014), concerning the collection, collation and analyses of selected animal health laboratory data and clinical impressions questionnaire data, on the CNPHI platform as part of the CAHSN, to improve animal health surveillance in Canada.

Methods

- With respect to the common objective of investigating collation animal health surveillance data on the CNPHI platform; members of all three projects agreed to work together in a synergistic manner involving frequent conference calls, consensus agreement on key decisions, and strategic assignments for contacting and working with people at specific data sources (laboratories, questionnaires) and CNPHI.
- Through consultation with various relevant stakeholders, the group agreed on a list of three production-limiting-diseases / hazards, in each of poultry and cattle and four in swine (total ten diseases) , that were used as examples to collate data from laboratories.
- Similarly the group agreed on a list of key data fields relevant to surveillance (and likely available from source laboratories), that were used as destination fields in files, into which data from analogous data fields were uploaded from selected laboratories.
- While many laboratories were consulted, the group agreed to focus on uploading data from four provincial-level laboratories including BCMAFF, PDS, MAFRD and AHL.
- Two separate trials were conducted uploading data, including one set for the three-month period of January to March 2013 and then data for the period of July to September 2014.
- Laboratory data for the identified data fields (for each of the selected diseases, from each of the three species, from each of the participating laboratories), was reformatted in a step-by-step process on personal computers of analysts involved in the study. Then used to produce summary reports. That step-by-step reformatting process (outside CNPHI) was documented to record and communicate its challenges (see Appendix 1).
- Since one project was also involved in collecting clinical impressions data from veterinarians using commercial questionnaire data (as part of the network-of-networks approach to surveillance), part of the group also studied the challenges and opportunities for using CNPHI for the collection, collation and analyses of questionnaire data.

Findings, Advantages and Challenges

- Table 1 lists project team members

Table 1: Project Team Members

Name	Email	Affiliation(s)	Project Role
Anatoliy Trokhymchuk	anatoliy.trokhymchuk@pds.usask.ca	University of Saskatchewan	Saskatchewan Lead
Bruce McNab	bruce.mcnab@rogers.com	University of Guelph	Ontario Lead
Carl Ribble	carl.ribble@gmail.com	Centre for Coastal Health, University of Calgary	British Columbia Lead
Harold Kloeze	Harold.Kloeze@inspection.gc.ca	Canadian Food Inspection Agency	Canadian Animal Health Surveillance Network Lead
Stefan Iwasawa	stefan.iwasawa@viu.ca	Centre for Coastal Health	BC Research Assistant
Theresa Burns	theresa_burns@hotmail.com	Centre for Coastal Health	BC Lead Researcher
Tyler Stitt	tyler.stitt@viu.ca	Centre for Coastal Health	BC Researcher

- Table 2 lists laboratory contacts and systems.

Table 2: Laboratory Contacts and Systems

Province	SEAC Member	LIMS Contact	Contact Email	LIMS
British Columbia	Nancy de With	-	Nancy.Dewith@gov.bc.ca	VADDS
British Columbia	-	-	*Victoria.Bowes@gov.bc.ca*	VADDS
British Columbia	-	Linda Thomson	Linda.Thomson@gov.bc.ca	VADDS
Alberta	Delores Peters	-	delores.peters@gov.ab.ca	
Saskatchewan	-	-	anatoliy.trokhymchuk@pds.usask.ca	
Saskatchewan	Wendy Wilkins	-	wendy.wilkins@gov.sk.ca	
Manitoba	*Glen Duizer	-	Glen.Duizer@gov.mb.ca	
Manitoba	Dale Douma	-	Dale.Douma@gov.mb.ca	
Manitoba	-	David Hunt		
Ontario	*Bruce McNab	-	bruce.mcnab@rogers.com	In house system
Ontario	Tim Pasma	-	Tim.Pasma@ontario.ca	
Ontario	-	Bev McEwen		In house system
Quebec	Luc Bergeron	-	Luc.Bergeron@mapaq.gouv.qc.ca	
New Brunswick			jim.goltz@gnb.ca	In house system (limited scope)
Nova Scotia			SPEARMJG@gov.ns.ca	VADDS
PEI			edobbin@upei.ca	In house system (currently upgrading)
Newfoundland/ Labrador			laurarogers@gov.nl.ca , hughwhitney@gov.nl.ca	
Yukon	Carolyn Cooper	-	Carolyn.Cooper@gov.yk.ca	
Yukon	Mary Vanderkop	-	Mary.Vanderkop@gov.yk.ca	
CIFA	Andre Vallieres	-	Andre.Vallieres@inspection.gc.ca	
CIFA	Cheryl James		Cheryl.James@inspection.gc.ca	
CIFA	Krista Howden		Krista.Howden@inspection.gc.ca	
CIFA	Margaret Morrison		Margaret.Morrison@inspection.gc.ca	

- Table 3 lists the target data fields, including some repetition (e.g. location data) required due to the variability of data collected by various LIMS.

Table 3: List of Target Data Fields
Data Field
Laboratory (source laboratory)
Case ID or Submission Number (assigned at lab)
Veterinarian Town or Postal Code (of the submitting veterinarian)
Producer / Owner Town or Postal Code (of the owner of the animal)
Premises ID (unique ID of the premises of the animal(s) tested)
CCSD (Consolidated Census Subdivision code for the premises)
Province of the animal(s) tested
Species
Animal type (production type)
Body System
History
Date Submitted (received) at the laboratory
Sample ID
Specimen (e.g. lung tissue, rectal swab, whole blood etc.)
Lab Test Type (e.g. PCR, Antibody ELISA, culture & isolation etc.)
Hazard / Disease for which the test is testing (e.g. PRRS, PED, ILT etc.)
Test Result Level
Test Interpretation (e.g. positive, suspect, negative, unknown)
Diagnostic Code
Final Diagnoses

- Table 4 lists the diseases by species for which data were requested from labs.

Table 4: Diseases by species/sector for which data were requested	
Industry	Diseases of interest
Bovine	IBR
	Bovine Viral Diarrhea
	calf scours
Porcine	Porcine Reproductive and Respiratory Syndrome (PRRS)
	Porcine Epidemic Diarrhea (PED)
	Influenza
	Streptoccus suis
Poultry	Mycoplasma gallisepticum (Mg)
	Infectious Laryngotracheitis (ILT)
	REO Virus

- Many of the findings were already known to experts in this area, but it was useful to document them for a broader decision-making audience.
- A key finding was that LIMS systems and structures differed significantly between labs such that some key data fields did not exist in completely analogous manners between some systems, regardless of coding used within such fields.
- Tests applied, critical thresholds and case definitions varied between labs and within labs over time.
- Coding varied over time deliberately as procedures and tests evolve, and varied inadvertently due to inconsistent coding within and between LIMS (e.g. many different codes used to

represent a positive test result: positive, Pos, +, pos, 4+, 1, p, P, pstv, yes, y, organism name…. etc.)

- The capture of key information varied within and between systems, including inconsistent or incomplete recording on submission forms. (e.g. premises ID or animal production type not filled in on the submission form)
- Precise location or premises ID information was often missing in LIMS case data.
- Systematic errors existed in coding (e.g. breeds coded as if they were species)
- Given the variability in structure and coding within and between LIMS, final collated files were incomplete and inconsistent.
- Most LIMS were designed to run labs and were not designed to harvest surveillance information in an epidemiologically sound manner (e.g. some systems created a new submission number for billing purposes, within the same epidemiological case, leading to double counting of cases).
- Counting the denominator was extremely challenging (i.e. submissions that would have found evidence of the hazard-of-interest if it had been there).
- Some LIMS simply could not be queried for surveillance or epidemiological purposes, such that data had to be transferred to another system before it could be queried (e.g. BCMAFF, MAFRD).
- Significant time and effort was required to reformat data from the various LIMS, outside CNPHI. This included significant risk of error (e.g. potential misalignment of record-rows, invalidating whole blocks of data).
- Calf scours was included as a test disease in the study because it was a common bovine syndrome. Unfortunately it proved extremely difficult to impossible to pull numerator and denominator data for calf scours from various LIMS.
- The steps required to reformat the data were documented in the Appendices of this report (see Appendix 1)
- CNPHI staff demonstrated CNPHI's ability to create applications that collate and analyze pubic health data to generate useful counts, trends, bar-graphs and maps, for specific diseases and time-windows (e.g. national PH enterics CNPHI application)
- It would be a significant step forward if CNPHI could build something similar for CAHSN.
- While CNPHI programmers can build specific applications for clients (e.g. the PH enterics system), they need a specific a proposal for desired output AND access to, and full understanding of, the data sources (e.g. LIMS), that will be used AND funding to do so.
- Also, during such a build, CNPHI programmers need specific direction from client(s) describing aspects the client wants to be able to change on their own within the application (without having to go back to CNPHI programmers)
  For example, a client may want ability:
  - to change the time-window under study
  - to change the bar period (e.g. day, week, month, quarter, year) over the total time-window under study (e.g. 2013-2014)
  - to change the cut-off of for alert (e.g. 1.5 vs. 2 vs. 3 SD)
  - to alter categories of outputs e.g. from one county, to groups of counties, or one province, to groups of provinces, or national (assuming appropriate data were provided).
  - to change queries to select records of specific tests (hazards/diseases eg. PRRS, influenza H3N2) or groups of tests or syndromes (e.g. bovine respiratory diseases).
-

- Currently, since funding has not yet been provided to CNPHI to do so, there are no animal health specific applications written in CNPHI analogous to the CNPHI PH Enterics application. Nevertheless, participants working on this project in BC were able to collate data from three laboratories to provide an example of the type of reports that could be produced for CAHSN (see Appendix 2 for an example)
- Unfortunately (as of November 2014), the questionnaires used at the time by CSHIN and OAHN (written in FluidSurveys), could not be reproduced exactly in CNPHI. However, it is expected that CNPHI programmers could produce various electronic questionnaires that could be of use to the networks-of-networks approach supported by NFAHWC, if specific requests/proposals for questionnaire design and ongoing questionnaire data storage and analyses are made to CNPHI programmers.

Discussion and Options

- Greater geographic coverage, greater representativeness and more statistical power can be achieved by collating animal health surveillance data from across the country
- The NFAHWC has recently endorsed a network-of-networks approach to animal health surveillance in Canada including use of clinical impression and laboratory data. It would be useful if the CAHSN could support such an approach through use of the CNPHI platform, but funding for additional programing in CNPHI would be required to achieve this.
- One must collect and collate data that are consistently coded in a meaningful manner, providing accurate information on: the host animal(s) tested, their location, the date, the samples tested, the diseases/hazards for which tests were conducted, the tests performed, and the tests results.
- Tests may be clinical or laboratory based.
- Most health data is collected using imperfect clinical and laboratory tests, conducted on units that have been voluntarily presented for examination or testing due to some concern within the context of availability and affordability at that time and place.
- Once appropriately collated and coded, then the data may be analyzed to identify important changes in counts of test results in hosts, space or time.
- However, if there are errors or inconsistencies in the data, its collation or coding, then analyses cannot be accurate. Erroneous counts, interpretations and conclusions will be made, leading to the distribution of misinformation, causing inappropriate actions and trade.
- Accordingly, extreme care must be applied when collecting, collating, analyzing and interpreting data from different or inconsistent systems.
- Different jurisdictions use different labs, with different case submission forms, different test procedures, different data systems, different database structures, different coding, and different consistency of use.
- Even within a lab, much or all of the above change and evolve over time.
- It is virtually impossible to count cases or detect changes in counts of cases without unique premises ID or longitude/latitude location data to allow premises to be distinguished (counted as separate cases) from a similar case on another premises. Since precise location or premises ID information was often missing in case submission data, case counts were unreliable.
- Also, since any mapping or spatial-clustering-analyses requires location data in longitude and latitude; either valid longitude and latitude data must be linked to premises ID data or valid longitude and latitude data must be recorded at the time of sample collection and recorded accurately in all LIMS records associated with those samples.

- Notwithstanding the challenges, valuable information is hidden within laboratory and clinical questionnaire data.
- The CAHSN has taken steps to standardize laboratory testing for at least some important FADs and has partnered with the CHPHI to provide a secure electronic health data storage and retrieval platform.
- More recently, evolution of programs such as RAIZO, CSHIN, OAHN and veterinary-farm-call-data projects in Ontario have used a mixture of laboratory and clinical information to gain information of use to decision making.
- Rather than attempting to force standardization on data sources (labs and clinics), those initiatives achieved greater success by tailoring data-collection to make it easy for each source to submit their data. Then data-clean-up and coding were standardizing after the raw data were collected and collated.
- As of November 2014 key components of the CNPHI system were:
    - Creating and maintaining a secure site, with only approved users having access to only the sections and data for which they have been approved.
    - Building, implementing and maintaining project-specific information management systems, upon detailed consultation with the specific client group (e.g. Public Health Enterics application)
    - Deliberately not allowing data to be altered once in the system (at least not without extremely high system security clearance), so as to insure the integrity of the data uploaded to the system.
- While the above aspects of CNPHI are logical and commendable, they also limit its use without additional programming and development of specific applications by CNPHI programmers.
- Currently, if one wishes to standardize coding after data has been uploaded to CNPHI then there are two options:
    I. Download the data onto another system (which may or may not be secure) and use search-and-replace if-then functions in some other data management system to translate and standardize the coding. Then re-upload the standardized data to CNPHI. Such a process is cumbersome and defeats the security of CNPHI, because overall security is only as secure as the weakest link in the chain.
    II. Alternatively, CNPHI programmers could write specific translation/data standardization programs in CNPHI for data coming in from each lab. Such programs would have to be continuously refined as coding evolves in each source LIMS. That would require a significant commitment by CNPHI and respective LIMS staff. Note that currently in both i) and ii), users must then download the clean standardized files to analyze the data (e.g. counts etc.) in software outside CNPHI. This again defeats the security of CNPHI if the system(s) on which the data are analyzed are not secure. But applications of fairly routine analyses (e.g. case count tables, bar graphs, maps of counts at the provincial level) could be developed by CNPHI for animal health similar to the current Public Health Enterics applications in CNPHI.
- In the longer term, some options for consideration include:
    a) Limited transfer of raw data to centralized system(s), but rather harvesting of specific counts and information locally by local experts of respective lab data systems, with resultant reports of information communicated and stored centrally for discussion by expert committees. Some laboratories cannot meet such requests because their respective LIMS are not designed to do such counts. However, such labs could upload more raw data to CNPHI and CNPHI programmers develop applications to produce

relevant premises-level counts from such lab data to be collated in CNPHI with counts produced by other labs for overall analyses in CNPHI of animal health data similar to the PH Eneritcs CNPHI application but at the premises-case level.

b) Implement one huge national LIMS used by all government and university veterinary diagnostic laboratories with standardized coding throughout the system. This national LIMS must be designed in a way that in addition to running individual labs, it is readily queried to pull epidemiologically sound data. Development, implementation and maintenance of such a system would be a monumental task, and is not likely to be achieved especially given laboratories' need an desire for some independence.

c) Implement Regional LIMS systems. For example, could BCMAFF, PDS and MFRD all use the new LIMS of PDS? Could this be done in a way that reports sent to submitters look like they are coming from the lab to which they made submissions, but all data from all three labs are coded in a standard manner and stored in a common back-end of the regional LIMS?

d) Continue to maintain and evolve respective provincial LIMS, but move to as much standardization in case definitions and coding as possible within and between various LIMS formats. This is a large task requiring much cooperation recoding and retraining at each laboratory, including ongoing diligence.

e) Continue provincial LIMS, but translate exported files of specific fields needed for national collation into standard coding at the respective lab, before they upload their file to the national system (e.g. CNPHI)

f) Continue provincial LIMS, but export and upload files with the specific data fields needed for national collation, but translate them into standardized coding after the arrive at the national data center (e.g. CNPHI), but before they are concatenated into one large high quality data file in CNPHI for storage and subsequent analyses.

g) A combination of a), d), e) and f) above.

Recommendations

- The following recommendations are not trivial. Many of them would require significant investment, commitment and cooperation from industry, federal and provincial governments.

1. Central Secure Data
   Continue to support the concept of, and contribute to the development of a secure, national animal health surveillance data system (e.g. CNPHI/CAHSN) for safe: controlled access, collection, collation, storage, data-standardization, analyses, summary of veterinary diagnostic laboratory and clinical data, from multiple sources across Canada. Since the core aspects of CNPHI are maintained by funding from public health for the benefit of public health (including zoonotic diseases); animal health sectors must be reassured by public health officials that CNPHI will be maintained even if it is used by CAHSN to manage animal health data for non-zoonotic diseases (e.g. animal production-limiting diseases that have no direct impact on public health). It is also recommended that animal health sectors fund CNPHI to develop application for animal health and including a CAHSN "front-end" to the CNPHI system, so that when animal health people log-onto the system they see primarily CAHSN headings, rather than primarily CNPHI headings. Notwithstanding this adjustment to the front-end-look of system for CAHSN vs. CNPHI users, the "back-end" structure should remain common to both.

2. Consistency Within LIMS
   Ensure consistent coding within respective animal health LIMS systems across Canada, especially for key data fields required for uploading information centrally.

3. Consistency Between LIMS For Key Data Fields
   Standardized coding between all animal health LIMS in Canada among the key data fields needed for uploading and collation to a central data system (e.g. CNPHI).

4. Fund the Build of More Animal Health Tools/Applications in CNPHI for CAHSN
   Fund CNPHI to expand its available tools to also include providing a secure environment in which approved users can store, manage, edit, standardize, and analyze their data, using tools with which they are familiar, all behind the CNPHI firewall and security systems (e.g. build an animal health application similar to the CNPHI public health enterics application but for animal health hazards at the premises/submission level identified by networks, in the networks-of-networks system supported by the NFAHWC).

5. Start With Data Already In CNPHI
   Encourage CNPHI and CAHSN staff to first achieve, validate and generate useful reports for #4 for data from BCMAFF and MAFRD which are already being uploaded into CNPHI. Then expand to include PDS data.

6. Expand System by System
   Once # 5 has been achieved for BCMAFF and MAFRD to justify expansion to PDS, then have AHL commit to uploading the subsequent year of data so that #4 and #6 can be achieved for AHL concatenated with BCMAFF, PDS and MAFRD data.

7. Validate and Continue
   Validate data and case counts from CNPHI with respective BCMAFF, PDS, MAFRD and AHL data

with respective LIMS and laboratory staff. Refine and continue on an ongoing basis and expand to include remaining provincial LIMS systems.

8. Switch to CNPHI / CAHSN Questionnaire Systems as They Improve
   If and when CNPHI either sets up FluidSurveys within CNPHI, or develops an equivalent survey system within CNPHI, then various networks in the networks-of-networks can switch to CNPHI for respective clinical-impressions-questionnaires for data collection, collation, concatenation and analyses behind the CNPHI security wall.

9. Use Best Current System for Collection But CNPHI /CAHSN for Collation
   Until #8 is achieved, consider continuing to use FluidSurveys within networks to collect data. However, also systematically upload, concatenate and store flat files of survey data on CNPHI that were collected using FluidSurveys. This should include concatenating clinical impressions data over time (various quarters) and space (e.g. OAHN and RAIZO and CSHIN data)

10. Level of Organization:
    Even once we have a clean, complete file of consistently and accurately coded data, properly collated and concatenated from all participating laboratories; there are then challenges in standardizing the analyses and interpretation.
    A key decision that will influence many aspects will be deciding upon the "level of organization" at which "cases" are classified and counted as "positive", "negative" "suspicious", or "unknown". It is recommended that this "level of organization" be at the "species, premises and time-window level" (ie. what, where, when), where a "premises" is a land parcel with one or more animals of a given species. There may be more than one production type of that species on the premises and there may be more than one species on the premises. The time window may default to the date the samples arrived at the lab or if known, within a reasonable range of time (e.g. within 3 weeks of the first positive samples from that premises). Thus a case to be counted as one case becomes the presence of one or more animals of a given species, on the premises, on a given date (or applicable time window) that is positive to one or more tests specific to the hazard/disease in question. Thus the presence of one hazard found in two species on a premises counts as two cases. The presence of one hazard found in two different production types of one species on one premises at one point in time counts as one case. The presence of one hazard found in one animal of one species with many animals of that species on the premises counts as one case at the premises level. The presence of many animals of one species found with the one hazard at one point in one or more groups or barns on one premises, still only counts as one case at the premises level. The presence of one hazard found in at least one animal on each of two different physical premises, even of the same company, counts as two cases. All of the above influences how we roll-up lab data from multiple samples and multiple tests in one submission to classify and count it as a case or not.

11. High-Level Case Definition:
    Another key question is are we counting cases of clinical disease, or production limiting, or cases of infection, or exposure (with or without clinical disease or production limitations), or presence of the hazard (biological or chemical agent) only within animals, or presence of the hazard within animals or in the environment of the premises, or evidence of the hazard was present (and may still be (e.g. evidence of an immune response to the natural hazard but not vaccine)? Among other things these decisions have implications for data collected including the

type of sample tested and the type of test (e.g. was it an environmental swab or a swab taken directly from an animal, does the test detect live infectious hazards (e.g. virus isolation) or possibly dead hazards or parts of hazards (e.g. PCR) or reaction to hazards (eg. serum antibody)? It is proposed that a positive case be defined as: one or more animals, of one species, on one premises, at one point in time (or applicable time window), that is test positive to one or more tests based on: nucleic acid test or antigen detection test, or antibody detection test, or chemical identification or culture and isolation or pathonumonic lesion, specific to the biological, chemical, radiologic or physical hazard in question, but not including response to vaccination.

12. Roll-Up of Data to Case Level of Organization:
    Individual, multiple or pooled samples may be collected from single or multiple animals of one species, within one submission to a lab, to make a "case" diagnoses. Single or multiple tests may be conducted in series or in parallel on individual, split, sub- or pooled samples, from those submitted to the lab. Data from that testing must be rolled up and interpreted at the species/premises / time case definition level described in #12 and #13. Some results (e.g. culture and isolation) may take much longer to attain than others (e.g. PCR). A decision must be made whether this roll-up and interpretation of multiple-test data results to the case level-of-organization, is done by the source laboratory before uploading to CAHSN/CNPHI or by CAHSN/CNPHI representatives after more raw (multiple separate test and sub-sample data results) have been uploaded to CASHN/CNPHI. There are advantages and disadvantages to both approaches.
    Roll-up at the lab would require more work by lab staff and would be susceptible to variability within and between labs over time. But it may be argued that the local lab staff are in the best position to interpret those data components and rolling it up to an interpretation at the case level of organization. In contrast uploading of more raw component test data gives CAHSN/CNPHI staff more opportunity to study sources of variability and test various methods of interpretation and potential case definitions (eg. test interpretation in series vs. in parallel, using different cut-offs eg SN ratios). But this latter approach would place greater responsibility on CAHSN/CNPHI data interpretation which could be wrong. These decisions also influence the formatting required for electronic data files for uploading to and storage at CAHSN/CNPHI. It is recommended that these decisions are made in a transparent manner and consistently followed for CAHSN/CNPHI data combined from various LIMS.

13. Numerator vs. Denominator:
    Additional decisions will be required to determine if combined data and analyses will focus on just "numerator" type data, (i.e. case positives), or denominator and numerator data to gain information on how many units were tested and among those how many were positive. In general collecting data to measure denominators is much more difficult than just counting positives for numerators. To get counts for the denominator for a specific hazard one must consider how many submissions had the opportunity to be tested for the hazard and would have been found positive if the hazard was there. This requires different filters be applied to different records and requires more complete information to do so. For example, looking for a GI disease, bacterial culture of intestine may be relevant and count in the denominator, but bacterial culture of lung should not likely be counted as a test that had opportunity to find the GI hazard. In some cases submitters may put limits on how much they want to spend, so relevant tests were not done. This all requires good knowledge of lab procedures used at each lab, thorough information being recorded, and appropriate filtering of records to be included

and excluded from counts.

14. Missing Data:
    As noted above, desired data are not always recorded on the submission form or are missing for any other reason. Such missing data will often reduce the ability to filter records appropriately for counting. Policies on missing data will be required.

15. Rapid Evolution of Testing and Coding:
    New tests and refinement of protocols and coding at labs are continuously being updated and changed. Data analysts need be kept aware of such changes and their impacts on: computer programing used for data uploads from labs to CAHSN/CNPHI, specific case definitions and filters used to count records as in or out of numerators or denominators of interest etc. This will require significant ongoing effort by participating labs and CAHSN/CNPHI staff.

16. Counting Undiagnosed Cases:
    For the current projects we are looking at uploading data pertinent to at least three production limiting diseases for each of poultry, bovine and porcine. But such an approach may prevent surveillance systems from detecting new, emerging and as yet un-recognized hazards. To address this, some jurisdictions monitor trends in counts of undiagnosed cases. It is recommended that CAHSN/CNPHI consider how that may be achieved in Canada.

17. Premises ID or Longitude and Latitude
    Given the above recommendations for level-of-organization and high-level case definition, it is essential that lab data can distinguish between the premises involved in the testing. This requires some sort of unique code for each separate premises in the data and it requires that premises-ID be accurately recorded on every single submission to all participating laboratories. Quebec and Manitoba currently achieve this. Also, since any mapping or spatial-clustering-analyses requires location data in longitude and latitude; either valid longitude and latitude data must be linked to premises ID data, or valid longitude and latitude data must be recorded at the time of sample collection and recorded accurately in all LIMS records associated with those samples. Therefore, an alternative to a premises ID system could be that people submitting samples to a lab be required to record the longitude and latitude in decimal degrees (to three decimal places using their smart-phone or GPS), for the location of where the animal(s) were at the time the sample(s) were collected and record that information on all respective laboratory submissions for capture in LIMS systems. To work properly, either the premises ID system or the longitude/latitude system must be fully supported by industry and governments. It is very difficult for labs to enforce such systems on their own without the full support of industry and government or mandatory regulations.

Appendices

Appendix 1: Data Translation From Source to Destination Files

The following is a description of the steps required to standardize data collected from various laboratories.

Appendix 2: Example Report Collating Data From Multiple Laboratories

Disease Surveillance Report: Reovirus

This report presents data from submissions to the British Columbia (BC) Animal Health Centre (AHC), Manitoba (MB) Veterinary Diagnostic Services Laboratory (VDSL) and Ontario (ON) Animal Health Laboratory (AHL) between Jul. 1, and Sep. 30, 2014. This report is a draft for data review and discussion purposes only. No conclusions should be drawn from numbers or figures presented below.

Disease overview

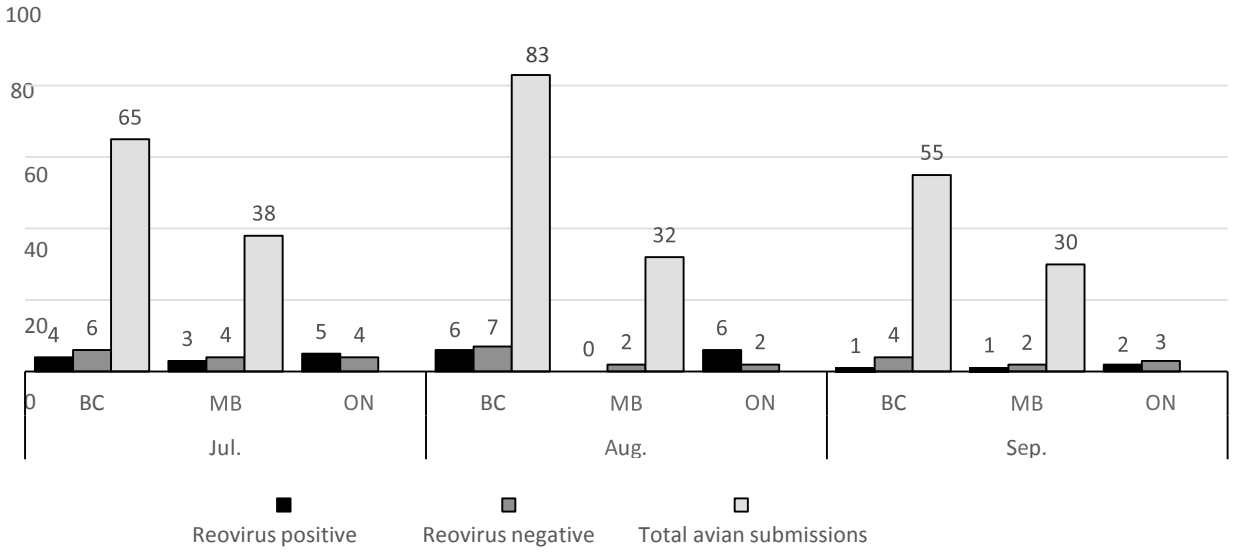
- Avian Reovirus infects chickens, turkeys, Muscovy ducks and other avian species.
- Ubiquitous virus of variable pathogenicity. Transmitted *in ovo* to developing embryos, and horizontally between chicks.
- Clinical disease occurs primarily in meat-producing chickens between 4 and 16 weeks of age, with a peak incidence near 7 weeks of age.
- Clinical features include acute lameness and swollen and inflamed joints (often hock joints); severe cases may result in rupture of gastrocnemius tendon.
- Post-mortem findings include arthritis and tenosynovitis, erosion of the articular cartilage, excess clear fluid in synovial capsules turning turbid in the presence of bacteria or mycoplasmas.
- The BC AHC uses conventional polymerase chain reaction (PCR) to test for Avian Reovirus. The MB VDSL and ON AHL use real- time PCR.

Surveillance Data

Table 1. Estimated size of BC, MB and ON supply-managed poultry flock by commodity.

	Broiler	Broiler Breeder	Layer	Turkey
BC provincial flock (n)	~98 million	~348,000	~2.4 million	~2.6 million
MB provincial flock (n)	~29 million	~1.3 million	~2.0 million	~1.7 million
ON provincial flock (n)	~195 million	~7.9 million	~7.6 million	~9.3 million

Figure 1. Total poultry submissions and Reovirus PCR tests to BC AHC and MB VDSL, and Reovirus PCR tests to ON AHL: Jul. 1, to Sep. 30, 2014.



Commentary

In BC, MB and ON, all tests were from chickens. In ON, 1 test result not shown in Figure 1 was reported as inconclusive. Total avian submission numbers were not available for ON.

There are a number of important data limitations that must be considered when interpreting the information provided above. These include the following;

1. At this time in the Canadian Animal Health Surveillance Network (CAHSN) data system it is not possible to distinguish multiple samples submitted from the same premises from multiple cases (birds from different farms). Premises ID would allow quick and effective determination of whether samples originate from the same or different flocks.
2. At this time in CAHSN it is not possible to differentiate commercial from backyard flock submissions.
3. Poultry total submission numbers for this report include submissions classified as chickens and turkeys.
4. Whole birds, tissues and swabs are included in total submission numbers.
5. Environmental samples and fluff are excluded from total submission numbers.
6. Samples collected as part of active surveillance programs are excluded from total submission numbers.
7. Samples submitted for serological testing are excluded from total submission numbers because serological testing is routinely performed in poultry flocks to determine vaccine response rather than disease status.
8. Data includes submissions to the BC AHC, MB VDSL and ON AHL. Given the nature of the poultry production system, it is possible that some samples from birds might have been sent for testing out of province.

Acknowledgements

Funding for this project has been provided by *Growing Forward 2*, a federal-provincial-territorial initiative. We acknowledge the support of the following provincial laboratories for participating in this endeavor: BC, SK, MB, ON, and QC.

This report was compiled by the Centre for Coastal Health for the BC Ministry of Agriculture.