

Evaluation of knowledge resources for public health reporting logic: Implications for knowledge authoring and management

Catherine J Staes, BSN, MPH, PhD¹, Rita Altamore, MD², Eun Gyoung Han, MS¹, Susan Mottice, PhD³, Deepthi Rajeev, MS¹, Richard Bradshaw, MS¹

¹Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah

²Washington State Department of Health, Olympia, Washington

³Utah Department of Health, Salt lake City Utah.

Abstract

To control disease, laboratories and providers are required to report conditions to public health authorities. Reporting logic is defined in a variety of resources, but there is no single resource available for reporters to access the list of reportable events and computable reporting logic for any jurisdiction. In order to develop evidence-based requirements for authoring such knowledge, we evaluated reporting logic in the Council of State and Territorial Epidemiologist (CSTE) position statements to assess its readiness for automated systems and identify features that should be considered when designing an authoring interface; we evaluated codes in the Reportable Condition Mapping Tables (RCMT) relative to the nationally-defined reporting logic, and described the high level business processes and knowledge required to support laboratory-based public health reporting. We focused on logic for viral hepatitis. We found that CSTE tabular logic was unnecessarily complex (sufficient conditions superseded necessary and optional conditions) and was sometimes true for more than one reportable event: we uncovered major overlap in the logic between acute and chronic hepatitis B (52%), acute and Past and Present hepatitis C (90%). We found that the RCMT includes codes for all hepatitis criteria, but includes addition codes for tests not included in the criteria. The proportion of hepatitis variant-related codes included in RCMT that correspond to a criterion in the hepatitis-related position statements varied between hepatitis A (36%), acute hepatitis B (16%), chronic hepatitis B (64%), acute hepatitis C (96%), and past and present hepatitis C (96%). Public health epidemiologists have the need to communicate parameters other than just the name of a disease or organism that should be reported, such as the status and specimen sources. Existing knowledge resources should be integrated, harmonized and made computable. Our findings identified functionality that should be provided by future knowledge management systems to support epidemiologists as they communicate reporting rules for their jurisdiction.

MeSH key words: Disease Notification; Knowledge Bases; Decision Support Systems, Clinical; Public Health Practice

Introduction

Timely and complete disease reporting is critical for detecting and controlling emerging health threats, particularly infectious diseases. In each US state, clinicians and hospitals, laboratories, veterinarians, daycare providers and others are required by law to report to public health authorities when they identify a reportable condition, such as anthrax, hepatitis A or lead poisoning. (1-4) Depending on the condition, reporting may lead to public health investigation, immunization, and prophylaxis of susceptible contacts, treatment of infected contacts, implementation of control measures to prevent further spread, and identification of trends and outbreaks. Thus, public health reporting is a key step in the chain of events to initiate control efforts and prevent new instances of disease.

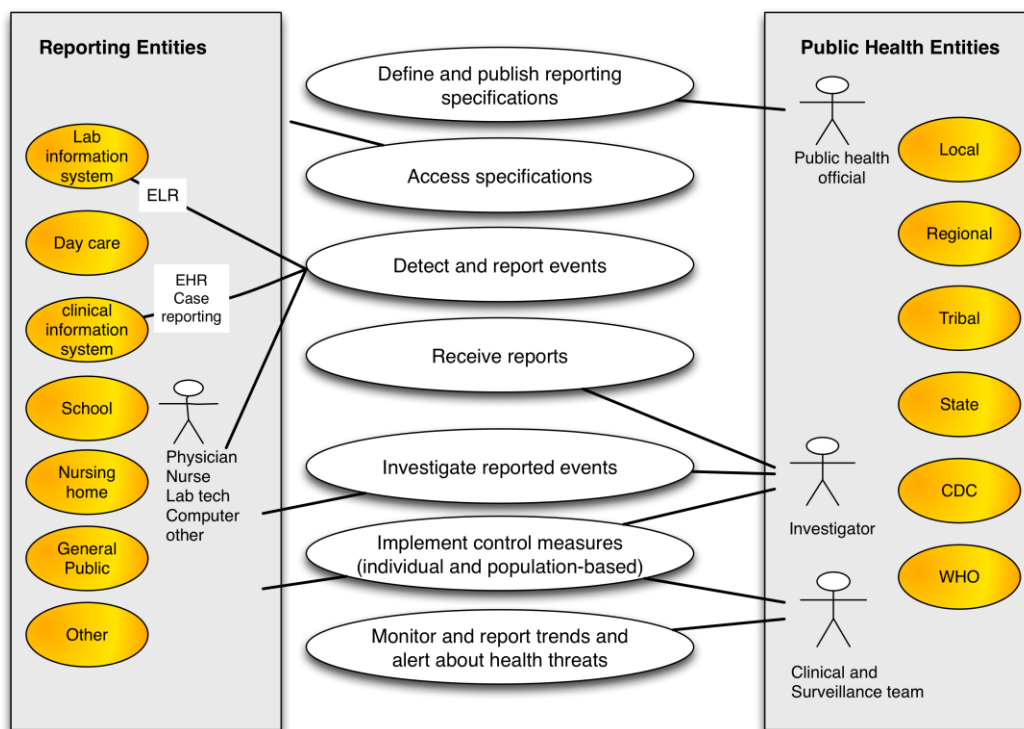


Figure 1: Use case for public health case reporting illustrating the actors and actions involved

Defining and publishing reporting specifications is the very first step in the public health reporting process (Figure 1). To implement public health reporting, clinicians, hospital, laboratories and others need information about what, when, how, and where to report. For reporters to ascertain this knowledge, public health authorities must specify reporting requirements and communicate those requirements to the target audience in a usable manner. There are several problems with the current processes. First, the unique reporting requirements for jurisdictions (such as cities, counties, states, and territories) are published in paper-based documents that are mailed/emailed and posted on clinic walls, and the requirements are listed on health department websites. (1, 2, 5-7) The reporting requirements may not be readily accessible or may become out of date, and the specific criteria used to identify reportable events is defined

by the reporter after *interpreting* the requirements. Public health reporting criteria are not provided in computable formats that allow implementation in automated systems. The reporter may or may not interpret the requirements as intended by public health authorities. Second, while websites typically list the name of reportable events in an effort to specify 'what' to report, the lists do not include the clinical and/or laboratory criteria that public health authorities want reporters to use to identify reportable events (1, 2, 5-7) and there is variation in the naming of events and the level of explicitness with which events are specified.(4) Finally, lack of knowledge about reporting requirements and inefficiencies associated with manual processes may contribute to the well-documented problems with delayed and incomplete reporting. (8-12) Recently, there have been major advances in the implementation of standards and clinical information systems, and changes in policies that increase opportunities to automate all or part of the public health reporting process. HL7 standards for vocabulary, messaging, decision support, and knowledge management are actively being developed and implemented in clinical and public health environments.(13) The inclusion of electronic laboratory reporting from healthcare settings to public health as a financial incentive for healthcare organizations to meet their 'meaningful use' requirements further advances the opportunities to automate public health reporting.(14)

Similarly, there have been improvements in the standardization of knowledge about public health reporting requirements, including knowledge represented in Council of State and Territorial Epidemiologists (CSTE) policy documents (e.g., Position Statements) (15), the CDC/CSTE State Reportable Condition Assessment (SRCA)(1), and the recently published Reportable Condition Mapping Tables (RCMT) developed by the Standards Workgroup of the CDC/CSTE Joint Electronic Laboratory Reporting (ELR) Task Force.(16) The CSTE Position Statements were enhanced in response to efforts in 2008 by the federal American Health Information Community to support real time nationwide public health event monitoring and rapid response management. (17) In order to improve the specification of reporting criteria and leverage advances in technology, the Position Statements were redesigned to include clinical and laboratory criteria and case report content for public health reporting of conditions that should be reported locally and included in national surveillance.(15) In 2009 and 2010, 68 new Position Statements concerning public health case reporting of infectious and non-infectious conditions were balloted and approved. (15) The Position Statements define national policy, but they do not address all the criteria required to implement reporting to local and state health departments, do not address emerging health threats, and are not accessible to automated systems. The SRCA provides knowledge on a website that summarizes the list of reportable conditions by state and territory, with links to relevant information on other websites. The information is updated annually and is especially useful and designed for defining the populations under surveillance (i.e., for establishing denominators) to calculate incidence rates of disease. The SCRA does not include specific reporting criteria nor is the knowledge currently in a computable format. The Reportable Condition Mapping Tables (RCMTs) were published in 2011 and represent knowledge about laboratory reporting criteria. (16) The RCMTs contain lists of standard codes for conditions reported nationally, and represents a major effort to harness domain knowledge from experts in standards, laboratory science, and the reportable diseases. The RCMTs provide computable knowledge about laboratory criteria but may or may not represent the criteria for reporting laboratory results in a specific jurisdiction and do not address clinical criteria or the other knowledge required about when, how, and where to report. Each knowledge resource is valuable

and meets a specific need. However, they do not currently meet the need of a laboratory, clinician, or other type of reporter needing a single resource for information about ‘what’s reportable where?’ and ‘how, when, and what should I report?’. The problem is compounded for laboratories and healthcare delivery systems that must report to more than one jurisdiction.

To benefit from improved specification of reporting requirements and leverage automated decision support systems, knowledge such as that found in the position statements needs to evolve through three interacting life cycles (page 19)(18): 1) knowledge generation and validation, 2) knowledge management and dissemination, and 3) clinical decision support implementation and evaluation. The collaborative effort among epidemiologists and the improved representation of reporting specifications in the CSTE position statements reflect this first life cycle. As described by Greenes, the first life cycle concerns knowledge that is often “initially unstructured and unassembled, or even only implicit, and must be extracted (from experts, from databases, or from the literature), organized and synthesized, analyzed for consistency and accuracy, and represented in an unambiguous form that can be computer-interpretable and acted upon (page 19).” (18) The knowledge generated from this effort should be evaluated to ensure its readiness for the knowledge management and dissemination life cycle. During the knowledge management and dissemination life cycle, Greenes describes continuous cycles of: a) curation and content management, b) collaborative authoring and editing, c) versioning and tracking of changes, d) standards-based dissemination, and e) localization and updates.(18) Financial, technical and governance resources are required to support these processes, particularly when the knowledge must be authored by distributed domain experts, undergo formal review and approval, be accessible to multiple and heterogeneous information systems, and may need to be modified (localized) to represent the needs of a subset of users. For example, in the public health reporting use case, there is the need to localize nationally-curated logic to represent unique reporting criteria relevant for a jurisdiction, such as a city, country, state, or territory. All of these life cycle processes and issues are relevant for managing the knowledge associated with public health reporting specifications concerning ‘what’s reportable where’. For example, the CSTE position statements define national policy and the RCMT contains candidate codes for logic, but it is necessary to allow the reporting logic to be ‘localized’ to address local needs and emerging health threats. In addition, knowledge from various sources needs to be harmonized.

Researchers with the Rocky Mountain Center of Excellence in Public Health Informatics have been designing and developing a prototype system to author and manage knowledge about ‘what’s reportable where’ for human review and automated systems. Our goal is to demonstrate a knowledge management tool that allows domain experts (e.g., public health epidemiologists) to author, curate and communicate their specifications in a manner that is unambiguous and computer-interpretable, while acknowledging the business processes associated with public health reporting and the heterogeneous knowledge needs. Therefore, the objectives of the research reported in this paper were to reveal evidence-based requirements for authoring knowledge by:

1. evaluating the tabular reporting logic in the CSTE position statements to assess its readiness for automated systems and identify features that should be considered when designing an authoring interface,

2. evaluating the codes in the RCMT relative to the nationally-defined reporting logic to identify authoring requirements, and
3. describing the high level business processes and knowledge required to support laboratory-based public health reporting.

We focused on the hepatitis-related position statements because they were undergoing revision during the spring of 2011 and allowed us the opportunity to provide and obtain feedback from epidemiologists and CSTE staff. In addition, hepatitis A, B, and C represent conditions that range from immediately reportable with a single laboratory result (i.e., hepatitis A) to those that may require multiple laboratory results performed months apart and at different laboratories (i.e., chronic hepatitis B or C).

Methods

Systematic evaluation of reporting logic in Position Statements

During the spring of 2011, we evaluated the knowledge that describes public health reporting logic in the CSTE position statements balloted in 2010 for ‘Hepatitis A’, ‘Acute hepatitis B’, ‘Chronic hepatitis B’, ‘Acute hepatitis C’, and ‘Past and present hepatitis C’.(15) Specifically, we evaluated the tabular logic in section VI-B concerning reporting from healthcare and laboratory settings. The tabular logic is organized such that each row in the table represents a unique criterion and each column represents a disjunctive (OR) set of criteria. The status of each criterion is defined in the position statements by the following guiding principles:

- S: This criterion alone is ‘*sufficient*’ to report a case
- N: This *necessary* criterion in conjunction with all other ‘*necessary*’ and any ‘*optional*’ criteria in the same column is required to report a case.
- O: At least one of the ‘*optional*’ criteria in each category in the same column (e.g., clinical, laboratory or epidemiologic findings), in conjunction with all other ‘*necessary*’ criteria in the same column, is required to report a case.
- A: This criterion must be ‘*absent*’

We evaluated the harmonization of criteria across position statements for similar concepts, and evaluated the criteria for ambiguity.(19)

We developed an automated logic verifier tool using an Excel spreadsheet as the user interface for rule definition. The format of the spreadsheet used to input logic mimics the format of the tabular logic found in the position statements (Figure 2). For example, each column is a disjunctive (OR) set of criteria and each row is a unique criterion. We implemented the logic described in the position statements by combining the terms in the following manner: ((Sufficient OR (Necessary AND Optional)) AND Not Present) or in pseudo-code: (S ||(N && O)) && !A).

Evaluation of knowledge resources for public health reporting logic: Implications for knowledge authoring and management

ID	Criterion	Reporting				
		Hepatitis A	Acute B	Chronic B	Acute C	P & P C
Clinical Evidence						
0	Acute onset of jaundice or dark urine				N	
1	Acute onset	N	N	N	N	
2	Jaundice or dark urine	N	N			
3	Jaundice or dark urine or light colored stool		N	N		
<i>Discrete onset of symptoms:</i>						
4	fever	O	O	O	O	
5	headache	O	O	O	O	
6	malaise	O	O	O	O	
7	anorexia	O	O	O	O	
8	nausea	O	O	O	O	
9	vomiting	O	O	O	O	
10	diarrhea	O	O	O	O	
11	abdominal pain	O	O	O	O	
<i>Health record contains diagnosis:</i>						
12	Hepatitis A	S				
13	Acute B		S			
14	Chronic B			S		
15	Acute C				S	
16	Chronic C					S
<i>Death certificate's COD or significant contributor:</i>						
18	Hepatitis A	S				
19	Acute B		S			
20	Chronic B			S		
21	Acute C				S	
22	Chronic C					S
Laboratory Evidence						
23	Elevated ALT > 200 or bili		N	N		
24	Elevated ALT > 400, acute onset			N	N	
25	Hepatitis A IgM (HAV) +	S	N	N		
26	IgM hep B antibody (IgM anti-HBc) +			S	N	N
27	Hepatitis B surface antigen (HBsAg) +			S		N
28	Hepatitis B e antigen (HBeAg) +					S
29	Hepatitis B virus (HBV) DNA +					S
30	Antibodies to hepatitis C virus (anti-HCV) +					S
31	Hepatitis C virus (HCV RIBA) +				S	O
32	Nucleic Acid Test (NAT) for HCV RNA +				S	O
33	(if done) IgM anti-HAV +			A	A	A
Epidemiological Evidence						
34	Contact of lab-confirmed Hep A case 15-50 days prior to onset		N	N		

Figure 2: User interface displaying tabular logic for hepatitis-related 2010 CSTE position statements.

A Java program was used to interface the Excel file with the logic, process the logic rules, and run simulated test cases through the logic rules. The Java program automatically converted the tabular logic into tables that internally represent Boolean logic using bit-vectors. The resulting binary Boolean logic allowed us to analyze simulated case-patient data to determine which test cases would trigger a report. The optimized internal format allowed for efficient comparisons of numerous simulated test cases. We simulated test cases for every conjunctive combination of criteria that would trigger a report according to the tabular logic. This case generation step was repeated for every disease, resulting in a set of report-generating test cases belonging to each disease. A simulated patient who triggered multiple reports did so because of ambiguous reporting logic. The tool ran exhaustive comparisons between the test cases to detect overlapping reports. First, we looked for overlap that would result in the generation of reports for more than

one reportable events. We assumed that each test case was applicable for only one reportable event. For example, in our sample, we assumed that a simulated case-patient had no more than one variant of hepatitis (A, acute B, chronic B, Acute C, or chronic C) and should only meet the criteria for one reportable event. To locate duplicate reports among different reportable events (i.e., ‘false positive’ reports), we used a pairwise disjointness detection algorithm. We tested for individual pairwise overlap for every pair of reportable events. We estimated the induced false positive rate of one event’s reporting criteria on another by calculating the fraction of test cases for the first reportable event that generated ‘false positive’ reports for the second reportable event. Second, we calculated a Total Overlap Ratio using the following formula:

$$\text{Total Overlap Ratio} = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \omega_{i,j}$$

$\omega_{i,j}$: overlap ratio between disease i and j (ith row and jth column in the table).

n: 5 in this case because we were assessing five reportable events.

During two conference calls with the epidemiologists that authored the logic in April and May of 2011, we reported our findings, recommended improvements to draft updated position statements (not published), and obtained general feedback about the spreadsheet interface for testing and viewing the logic. The spreadsheet input format was flexible, a feature that allows the user to experiment with combinations of logic.

Evaluation of codes in the RCMT relative to hepatitis reporting logic

We accessed the reportable condition mapping tables (RCMT) available on the PHIN VADS website on December 12, 2011(20), and downloaded the associated laboratory tests for ‘Viral hepatitis, type A’, ‘Type B viral hepatitis’, and ‘Viral hepatitis C’. The tests are Logical Observation Identifiers Names and Codes (LOINC®) (21). We grouped the LOINC® codes by the laboratory criteria included in the five hepatitis-related position statements balloted in 2011 and describe the proportion of LOINC® codes relevant for each criterion, and those not related to the criteria. In 2011, the laboratory criterion in the hepatitis-related position statements were the same as those in 2010, with the exception of an additional criterion for any ‘Antibodies to hepatitis C virus (anti-HCV) screening-test-positive’ regardless of whether they meet the CDC-defined cutoff. This addition does not change the test codes that should be used to identify reportable events. We did not review the LOINC database for additional relevant codes.

High-level business process analysis

To improve our understanding of the knowledge required to support laboratory-based public health reporting, we evaluated business processes associated with reporting knowledge in both the laboratory and public health setting. To understand reporting processes at a laboratory, we directly observed the work processes of the compliance officers tasked with public health reporting at a major reference laboratory. We diagrammed the information gathered, and validated and improved the generalizability of the documented processes by conducting structured interviews with compliance officers at the reference laboratory and the central laboratory for a multi-hospital healthcare network. To understand the business process of reporting from the perspective of a public health agency, we conducted interviews using storyboards with

epidemiologists from three public health programs in Utah. The objective of the interviews was to understand the processes related to handling incoming reports. In addition, to understand the parameters for criterion that may need to be authored by epidemiologists, we identified situations when reporting logic may need to change and reviewed more detailed reporting criteria being developed by the epidemiology program.

Results

Systematic evaluation of reporting logic in Position Statements

Using the automated tool to systematically evaluate the tabular reporting logic in the five hepatitis-related position statements (Figure 2), we identified problems with overlapping and overly-complicated logic. First, tabular logic was true for more than one reportable event. For example, half (52%) of the test cases simulated to meet the criteria for chronic hepatitis B also generated a report for acute hepatitis B, and vice versa. Most (90%) of the test cases simulated to meet the criteria for acute hepatitis C generated a report for 'Past and Present hepatitis C', and vice versa. There was no overlap between viral hepatitis type A, B, and C.

Second, the automated tool uncovered tabular logic that was unnecessarily complex. Sufficient conditions often superseded Necessary and Optional conditions. For example, the logic for hepatitis A defined in columns 2 and 3 in Figure 2 are not necessary given the criteria in column 1. A positive test for IgM antibodies to hepatitis A was sufficient to report a case to public health therefore the more complicated logic was irrelevant. Similarly, for acute hepatitis B, columns 3 through 6 in the tabular logic were subsumed by the 'sufficient' hepatitis B laboratory criterion in columns 1 and 2. When a criterion is 'sufficient' for reporting, it is not necessary to include other sets of criteria that contain the already-'sufficient' criterion. When we removed the superseded conditions from the logic for the hepatitis variants, and reanalyzed the pairwise overlap in the logic, the number of cases detected and the total overlap ratio did not change, indicating that the superseded conditions did not contribute to case finding.

Manual review of the tabular logic (Figure 2) revealed additional problems: a) tabular logic occasionally violated the guiding principles (e.g., Optional criteria for acute hepatitis B did not always occur with a Necessary criterion); b) some criterion were ambiguous and included more than one concept (e.g., hepatitis A logic included the criterion: 'elevated ALT >200 or bili'); c) apparently similar criterion varied across different position statements; and, d) laboratory criteria needed more definition to be actionable (e.g., temporal parameters are required to make the hepatitis C criterion of 'elevated ALT >400, acute onset' actionable).

When we reviewed the logic and our findings with the position statement authors, we identified features of the analyzer tool that should be considered when developing tools to author and manage reporting logic in the future. The position statement authors indicated that the spreadsheet input for the logic of multiple reportable events provided new insights into problems with the logic and lack of harmonization of criteria. They mentioned that they had never before seen all of the rules on one page and they found it useful. We were able to provide the epidemiologists with a spreadsheet to formulate and modify logic by inserting N, O, S, and A for applicable cells, and then retest the impact on overlapping logic within and between reportable diseases.

Evaluation of knowledge resources for public health reporting logic: Implications for knowledge authoring and management

Evaluation of associated laboratory test codes in the RCMT

The proportion of hepatitis variant-related codes included in RCMT that correspond to a criterion in the hepatitis-related position statements varied between hepatitis A (36%), acute hepatitis B (16%), chronic hepatitis B (64%), acute hepatitis C (96%), and past and present hepatitis C (96%).

Among the 22 codes for hepatitis A virus (HAV), only 8 (36%) were for IgM antibodies, the reporting criterion in the position statement; the remaining codes were for HAV IgG antibodies (n=5), HAV antibodies not further specified (Ab) (n=7), HAV RNA (n=1), and hepatitis panels used to order tests (n=2).

Among the 129 codes for hepatitis B virus (HBV), 89 (69%) were for criteria found in one or both of the HBV-related position statements. The remaining codes were either not related to any reporting criterion (n= 37 IgG or antibody tests comprising 29% of the total), potentially useful tests that may not yet have been considered by the position statement authors (n= 2 codes for rRNA or core Ag), and a hepatitis panel used to order tests (n=1).

Among the 90 codes for hepatitis C virus (HCV), 86 (96%) were for criteria found in both of the HCV-related position statements. The two hepatitis C related position statements use the same set of tests to trigger a report, but the acceptable values differ. Specifically, test result values for antibodies to HCV (anti-HCV) screening tests must exceed the cut-off ratios defined by CDC to meet the criteria for acute HCV results, whereas the result values are not required to meet these criteria for establishing 'past or present hepatitis C'. The remaining codes were potentially useful tests that may not yet have been considered by the position statement authors (n= 3 codes for rRNA or Ag) and a hepatitis panel used to order tests (n=1).

There were no laboratory criteria in the five position statements without corresponding LOINC codes.

Business process analysis relevant to knowledge authoring needs

Consistent with Figure 1, public health reporting starts with the process of a public health agency communicating reporting requirements to a laboratory and other reporting entities. The Utah Department of Health and the county health departments publish reporting requirements on their websites, including pdf files that may be downloaded, and call laboratories by phone when necessary. While reporting requirements have been fairly stable over time, we identified a variety of situations when reporting requirements did or could change thus requiring the ability to add new, remove, or update reporting requirements. For example, new reportable events were *added* when:

- a. a new condition emerged (e.g., west nile virus H1N1 influenza,). Thus authoring tools would require the ability to add a new condition and any associated lab tests. Note that the condition may not yet exist in SNOMED-CT and laboratory tests may not yet be in LOINC. There is often an urgency associated with publicizing this new requirement.
- b. a problem with an existing reportable condition emerges, requiring an expansion to the organisms under surveillance and a new event with its own reporting specifications. For example, the H1N1 influenza strain was reported separately from the "seasonal" strain routinely reported among hospitalized persons.

Evaluation of knowledge resources for public health reporting logic: Implications for knowledge authoring and management

- c. new control measures were available so case finding was important and prevention strategies need to be evaluated (e.g., varicella zoster).

Reportable events need to be able to be *updated* when the health department:

- a. seeks a denominator for situational awareness requesting that all tests performed, not only ‘positive results’, be reported (e.g., during a large cryptosporidium outbreak) ;
- b. relaxes constraints associated with current reportable events to improve case finding and for situational awareness (e.g., requesting all persons diagnosed with influenza be reported, rather than only those that are hospitalized).
- c. wants more timely reports for an emerging problem associated with an existing reportable condition, thus requesting preliminary as well as final results.
- d. increases specificity for a reportable event by specifying that only organisms resistant to antibiotics be reported.
- e. Needs to add synonyms to the name of a clinical condition, or make other changes based on improved understanding about associated organisms.

Reportable events were *removed* when priorities changed and receiving reports was no longer indicated (e.g., MRSA, Kawasaki, Rheumatic fever, CMV). Finally, the epidemiologists indicated other needs, such as the ability to change the reporting time frame, where reports may be sent, and the methods available for receiving reports, but these are beyond the focus of this paper.

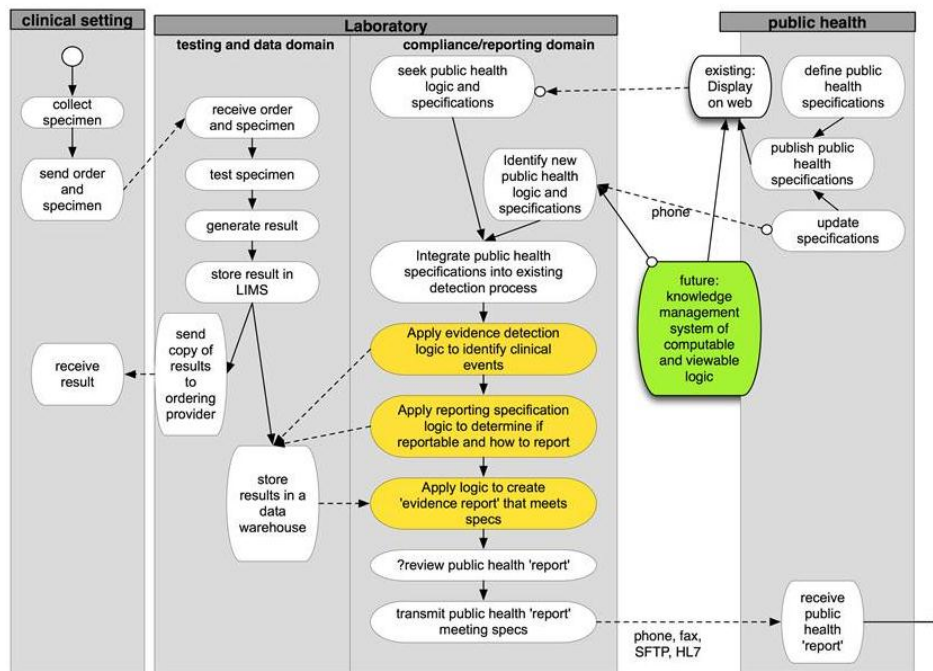


Figure 3: Example business processes and logic required to communicate reporting requirements and implement public health reporting within a laboratory

Evaluation of knowledge resources for public health reporting logic: Implications for knowledge authoring and management

To manage public health reporting in the laboratories, we identified three distinct processes that each use a different set of knowledge (Figure 3). First, the laboratories identified the test results associated with reportable events (i.e., using ‘evidence detection logic’). This knowledge was implemented using the laboratories ‘local’ codes that had been mapped to reportable events. Potentially, this could be handled by the code sets provided by RCMT, but not all laboratories have mapped their local codes to LOINC[®] codes and new laboratory tests may not yet have an assigned LOINC[®] code. Second, the laboratories applied constraints, such as age, pregnancy status, test results, or hospitalization status, to identify reportable events relevant for a given jurisdiction (i.e., using ‘reporting specification logic’). For example, the age and blood lead level criteria for reporting ‘childhood lead poisoning’ varies by state. Finally, after establishing that an event is reportable in a given jurisdiction, the laboratories created a report and transmitted the information using knowledge concerning how, when, and what to include in a report relevant for a given jurisdiction.

Finally, we identified two sets of knowledge to complete the reporting process at the public health agency. First, the health department requires ‘evidence response logic’ to triage incoming reports, handle duplicate reports from multiple sources (e.g., reports from laboratory and healthcare sources for the same person), and prioritize those reports that require investigation. The response to an incoming report may be ‘reject the evidence’, ‘hold for more evidence’ (e.g., repeated hepatitis B surface antigen results required to establish chronicity), ‘respond to evidence in an automated manner’, or ‘respond to evidence with action by public health personnel’. Finally, health departments require logic to classify events for surveillance, which is similar but different than the reporting logic communicated to reporting entities.

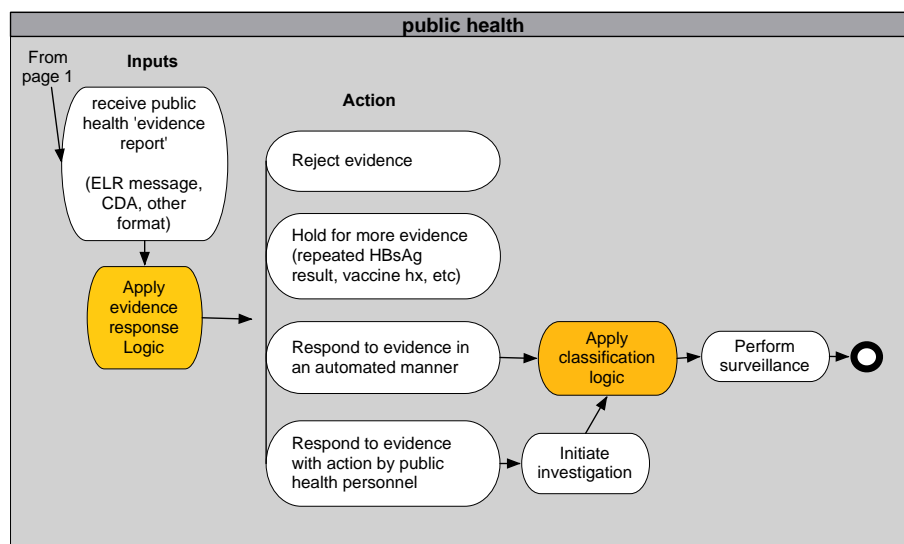


Figure 4: Example business processes and logic required to manage incoming reports at the health department and apply logic for surveillance

During the course of this research, epidemiologists with the Utah Department of Health defined laboratory findings they wished to receive by defining criteria using ‘lab findings by method’ (e.g. HAV IgM) as shown in the position statements, and including the specimen source, whether preliminary or final results are expected, and the results to be reported.

Discussion

Our evaluation revealed strengths and limitations of two important sources of knowledge for public health reporting, and identified features that should be considered when designing and implementing a knowledge management system to support public health reporting. Each resource, and the information provided on health department websites, contains critical information that needs to be integrated and shared in both a human-readable and computer-interpretable format. Some of the knowledge may be curated on a national level, but epidemiologist need to be able to modify requirements based on the needs of their jurisdiction. Systematic evaluation of the tabular reporting logic in the hepatitis-related CSTE position statements revealed problems but also specific opportunities for improvement. The CSTE should be applauded for their efforts to structure reporting logic and use a collaborative process to gather input from domain experts distributed throughout the US and from federal, state and local public health agencies. To date, this knowledge engineering task to create reporting logic has been performed using text-based position statement documents that summarize one reportable event at a time. The knowledge is then verified by experts who review and vote on the text-based documents. This task would likely benefit from a tool that allows authors and reviewers to: a) create and test various combinations of logic using standard concepts, and b) view the logic for a variety of diseases in one view to improve logic harmonization. Better tools are needed to support the early steps in the lifecycle of ‘knowledge generation and validation’ (18) and eliminate wasted effort creating and negotiating overly complicated and redundant logic. Better yet, the tools should go on to support the knowledge management and dissemination phase.

While the reporting logic in the position statements was structured, we identified several problems that limited the readiness of the logic to be adopted for automated systems. The tabular logic were sometimes overly complicated given the presence of simpler, sufficient logic and would come true for more than one reportable event. The overlapping logic appears to be unintended for acute and chronic hepatitis B, but unavoidable for acute and ‘past and present’ hepatitis C. This situation will require that reporting systems and public health receiving systems have the ability to reconcile reports for the two reportable events associated with the same person. To be fully ready for automated systems, the logic needs to be unambiguous and computer-interpretable. While narrative logic may be easier to create because domain experts can use their own words and purposefully express ambiguity, the tabular logic was one step closer to the computationally useful form of logic needed for automated systems. Additionally, in comparison to the narrative logic, the tabular logic was more structured, unambiguous, and explicit, and was easier to a) compare tabular logic across a set of reportable conditions, and b) expose problems with narrative ‘AND/OR’ logic statements and misplaced modifiers. Finally, it is possible to automatically generate narrative text from the tabular logic to allow domain experts to verify the tabular logic in a narrative form.

Use of the prototype tool with the epidemiologists revealed features that may be useful for future systems. The tool presented the user with a user-friendly interface that mimics the original source of the knowledge (the position statement). The interface was a spreadsheet that could be manipulated in any spreadsheet application, such as MS Excel or OpenOffice. This provided freedom and flexibility to efficiently compute new rules on the fly and test different logical propositions. Additionally, the tool allowed the user to verify overlapping reporting rates and interference among other diseases' reporting logic, and may be used to test logic within any of the other 68 disease-specific CSTE position statements. We hypothesize that the views and output generated by the tool could save time during the logic development process if it were available to the authors. Unfortunately, the tool requires manual entry of the logic. It would be best to implement this functionality within a knowledge management system that supports rule authoring, thus allowing authors to define, test, and share the logic in one application.

The RCMT represents a major effort by experts in standards, public health, and laboratory science to identify and standardize the code sets that could be used for detecting reportable events. We identified variation in the specificity of codes for the criteria included in the five hepatitis-related position statements. We suspect that similar issues would be found for other conditions. It is unreasonable to expect that one set of codes would satisfy all uses. Therefore, we recommend that LOINC codes be classified to the level commonly used for reporting criteria (e.g., IgM, IgG, total antibody, etc). This would allow epidemiologists to select criterion and automate the process of identifying the relevant LOINC codes. Automatically classifying the LOINC codes has been addressed in the past(21), but is not yet solved.

Limitations

Our investigation has limitations. First, our study is limited to the 'what' in 'what's reportable where', and does not address the full spectrum of reporting specifications concerning 'when', 'how' and 'where' to report. Lack of standardization in these criteria have been identified by others (5) and our team (unpublished data), but this is not the focus of this paper. The specifications for 'when' and 'how' are not composed at the national level and specified in position statements. That knowledge must be generated in the context of a given state or other jurisdiction. Second, we limited our investigation to the hepatitis-related position statements. While it is possible that the problems we identified are unique to hepatitis, we found similar problems with other position statements when graduate informatics students evaluated 26 other position statements as part of class assignments during 2009 and 2010. Despite the limited focus on hepatitis, we believe that the problems and requirements we identified are generalizable to other reportable events and position statements.

Use Case development

Based on our findings, we developed a high-level use case for authoring and accessing public health reporting specifications (figure 5). The use case includes key actors involved in the process of authoring and accessing reporting specifications, and the major system functionalities that would be required to meet the system goals. We included major existing knowledge resources (including the CSTE Position Statements and the Public Health Information Network Vocabulary Access Distribution Service (PHIN VADS)) that may provide sharable content for public health reporting specifications. The actors and required functionality were identified iteratively as we gathered information during the business process analysis and evaluation of

Evaluation of knowledge resources for public health reporting logic: Implications for knowledge authoring and management

current resources described above. The goal is to allow public health authorities to develop and access sharable knowledge (typically curated at a national level), while allowing public health authorities in state or other jurisdictions to use the sharable knowledge and author reporting specifications that meet their needs (i.e., ‘localizing content’). As shown in Figure 5, public health authorities need to author as well as access specifications, but their role becomes specialized at the jurisdiction level (e.g. state level) as the specifications must meet the jurisdiction-specific legal reporting rules. Because reporting specifications vary by jurisdiction, any efforts to standardize the detection logic at the national level must support the capability for localization at the jurisdictional level (e.g., city/county, state, territory). The custodian of the reporting specifications (e.g., a state epidemiologist or state health officer) could designate a domain expert to perform two functions: 1) author the jurisdiction-specific reporting actions concerning when, where, how and why to report, and 2) link this information to reporting criteria. The reporting criteria may be adopted from the sharable source, or may need to be localized to, for example, modify age, hospitalization, or laboratory criteria to meet local needs. The Public Health Information Network Vocabulary Access Distribution Service (PHIN VADS)) may be used to provide sharable content for public health reporting specifications.

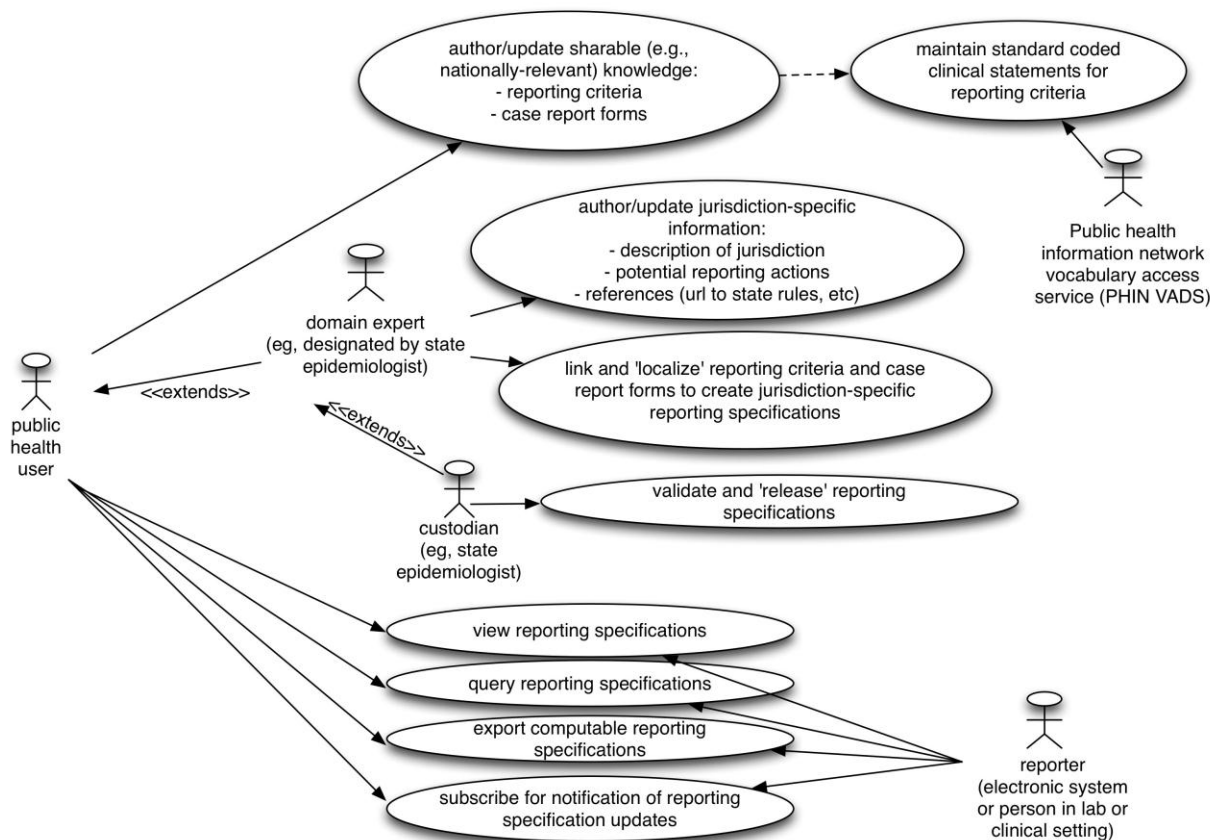


Figure 5: High level Use Case for authoring and accessing reporting specifications using national resources for sharable content while incorporating jurisdiction-specific content required by reporting systems.

Recommendations

We identified numerous features that should be requirements for systems that manage the knowledge associated with ‘what’s reportable where’. The requirements we identified compliment those identified using user-centered design and other methods reported elsewhere. In particular, knowledge authoring and management tools should allow epidemiologists to:

- systematically assess reporting logic to identify logic that is overcomplicated, or overlapping with other logic for the same or a different reportable event.
- specify laboratory reporting criteria at the level of granularity described in the position statements (typically, a combination of laboratory finding and method), allowing for linkage to relevant knowledge from the RCMT (i.e., the corresponding combinations of relevant LOINC® and SNOMED codes).
- identify problematic logic and experiment with improved logic before balloting and publishing reporting requirements
- track the provenance of knowledge and show the history of decisions that were made as the knowledge was developed over time. For example, a feature should be included that allows authors and reviewers to comment on the knowledge and the system should track the users and dates associated with the comments
- collaboratively author reporting criteria that meet national specifications (e.g., as reflected in the position statements), but allow the logic to be localized to meet state or other jurisdictional needs.

In addition, when researchers validate and publish reporting logic, as for example, Klompas et al have done for acute hepatitis B (22), then this further validated logic should be incorporated into the knowledge management system for review and adoption by others.

Conclusion

The logic in the hepatitis-related position statements was insufficient to discriminate reportable events and may be overly complicated given the presence of simpler logic to trigger reports. However, the tabular logic can be assessed and improved using automated tools, is one step closer to the computable form, and represents a major effort within the ‘knowledge generation and validation’ phase. We recommend classifying the LOINC codes in the RCMT to the criteria included in the position statements to allow epidemiologists to select the codes they prefer to use for reporting in their jurisdiction. We suspect variation will decrease over time as users understand the variation they are imposing on reporting facilities. In addition, we recommend using knowledge management tools to author, verify, improve, and authenticate logic, and continually incorporate improved logic that has been validated in clinical systems.

Acknowledgements

This research was supported by CDC Grants COE1: 5P01HK000030 and 5P01HK000069. P.I. Matthew Samore, MD, University of Utah. In addition, coauthors (EGH) were partially funded through support from the National Library of Medicine (NLM Training Grant No.T15LM007124). We acknowledge Lisa Ferland and Monica Huang at CSTE, the authors of

the position statements, and the CSTE/CDC Case Report Standardization Workgroup for their support with this investigation.

Corresponding author

Catherine Staes, BSN, MPH, PhD
Dept of Biomedical Informatics, University of Utah
Catherine.staes@hsc.utah.edu

Conflict of Interest

There are no conflicts of interest.

References

1. State Reportable Conditions Website [database on the Internet]. Council of State and Territorial Epidemiologist. 2011 [cited December 4, 2011]. Available from: <http://www.cste.org/dnn/ProgramsandActivities/PublicHealthInformatics/StateReportableConditionsQueryResults/tabid/261/Default.aspx>.
2. Utah Administrative Code. Rule R386-702. Communicable Disease Rule. Effective February 1, 2008. (Available from: <http://www.rules.utah.gov/publicat/code/r386/r386-702.htm>) [Accessed March 14, 2008].
3. Roush S, Birkhead G, Koo D, Cobb A, Fleming D. Mandatory reporting of diseases and conditions by health care professionals and laboratories. *Jama*. 1999 Jul 14;282(2):164-70.
4. Jajosky R, Rey A, Park M, Aranas A, Macdonald S, Ferland L. Findings from the Council of State and Territorial Epidemiologists' 2008 Assessment of State Reportable and Nationally Notifiable Conditions in the United States and Considerations for the Future. *Journal of Public Health Management & Practice*. 2011;17(3):255-64.
5. M'Ikanatha N M, Welliver DP, Rohn DD, Julian KG, Lautenbach E. Use of the Web by state and territorial health departments to promote reporting of infectious disease. *Jama*. 2004 Mar 3;291(9):1069-70.
6. Florida Department of Health, Notifiable Disease Reporting 2011 [cited 2011 September 16]; Available from: http://www.doh.state.fl.us/disease_ctrl/epi/topics/surv.htm.
7. Reportable Diseases in Utah (Available from: <http://health.utah.gov/epi/report.html>) [Accessed March 14, 2008].
8. Konowitz PM, Petrossian GA, Rose DN. The underreporting of disease and physicians' knowledge of reporting requirements. *Public Health Rep*. 1984;99(1):31-5.
9. Overhage JM, Grannis S, McDonald CJ. A Comparison of the Completeness and Timeliness of Automated Electronic Laboratory Reporting and Spontaneous Reporting of Notifiable Conditions. *Am J Public Health*. 2008 February 1, 2008;98(2):344-50.
10. Schramm MM, Vogt RL, Mamolen M. The surveillance of communicable disease in Vermont: who reports? *Public Health Rep*. 1991 Jan-Feb;106(1):95-7.
11. Silk BJ, Berkelman RL. A review or strategies for enhancing the completeness of notifiable disease reporting. *J Public Health Management Practice*. 2005;11(3):191-200.
12. Staes CJ, Gesteland P, Allison M, Mottice S, Rubin M, Shakib J, Boulton R, Wuthrich A, Carter ME, Lecaster M, Samore MH, Byington CL. Urgent Care Physician's Knowledge

- and Attitude about Public Health Reporting and Pertussis Control Measures: Implications for informatics. *J Public Health Manag Pract.* 2009;15(6):1-8.
13. HL7. Health Level Seven International. 2011 [cited 2011 December 12]; Available from: <http://www.hl7.org/index.cfm>.
 14. CMS. EHR Incentive Program. Baltimore: Centers for Medicare and Medicaid Services; 2011 [cited 2011 December 10]; Available from: <https://http://www.cms.gov/EHRIncentivePrograms/>.
 15. CSTE. 2010 Position Statements. Council of State and Territorial Epidemiologists; 2010 [cited 2011 September 16]; Available from: <http://www.cste.org/dnn/AnnualConference/PositionStatements/2010PositionStatements/tabid/422/Default.aspx>.
 16. CDC. Reportable Condition Mapping Tables (RCMTS) Another step toward standardizing electronic laboratory reporting (ELR). 2011 [updated July 22, 2011; cited 2011 September 30]; Available from: <http://www.cdc.gov/ehrmmeaningfuluse/rcmt.html>.
 17. CSTE. CSTE Official List of Nationally Notifiable Conditions. Atlanta: Council of State and Territorial Epidemiologists; 2007 [updated 2007; cited 2011 October 5]; Available from: <http://www.cste.org/PS/2007ps/2007psfinal/EC/07-EC-02.pdf>.
 18. Greenes RA, editor. *Clinical Decision Support: The road ahead*. Boston: Elsevier, Inc; 2007.
 19. Cimino J. Desiderata for controlled medical vocabularies in the twenty-first century. *Methods Inf Med.* 1998;37(4-5):394-403.
 20. CDC. PHIN Vocabulary Access and Distribution System. Atlanta: CDC; 2011 [cited 2011 December 11]; Available from: <https://phinvads.cdc.gov/vads/SearchVocab.action>.
 21. Steindel S, Loonsk JW. Introduction of a hierarchy to LOINC to facilitate public health reporting. *AMIA Annu Symp Proc.* 2002:737-41.
 22. Klompas M, Haney G, Church D, Lazarus R, Hou X, Platt R. Automated identification of acute hepatitis B using electronic medical record data to facilitate public health surveillance. *PLoS ONE.* 2008;3(7):e2626 (available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2440348/?tool=pubmed>).