Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____	_____
Sarita Sadhwani	Date

Validation of Reportable Conditions Trigger Codes (RCTC) content and EHR implementation process.

By

Sarita Sadhwani

Degree to be awarded: MPH

Executive MPH

Sunanda McGarvey, BS, CBAP, CSM	Date
Jon Lipsky, BA, MBA, PMP	Date
Laura Gaydos, PhD Associate Chair for Academic Affairs, Executive MPH program	Date

Validation of Reportable Conditions Trigger Codes (RCTC) content and EHR implementation process

By

Sarita Sadhwani

MPH, Emory University, 2016

MS Life Sciences, Devi Ahilya Vishwavidyalaya, Indore, India, 1996

Thesis Committee Chair: Sunanda McGarvey, BS, CBAP, CSM

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health in Executive MPH program, 2016

Abstract

Validation of Reportable Conditions Trigger Codes (RCTC) content and EHR implementation process

By
Sarita Sadhwani

Abstract

Background: The electronic case reporting (eCR) project is a step towards achieving the vision of bidirectional information flow between public health and health care, it can help bridge the digital gap existing between healthcare and public health. The Reportable Conditions Trigger Codes (RCTC) is the appropriate first step for this initiative, as it can provide timely initiation of electronic initial case reports (eICR) from healthcare to public health when trigger codes are matched to information in a patient's encounter record.

Purpose: The purpose of this research is to validate both the RCTC content and the implementation of standard codes in the EHR processes.
Through research and analysis, this thesis will explore discrepancies between LOINC® codes used within 4 national reference labs, one clinical lab, and the RCTC. The discrepancies will be categorized and reasons for the differences will be defined. In addition the study will evaluate the workflow and interactions between clinical lab Systems and EHRs and study the gaps in the use of LOINC® and SNOMED, and its impact of missing RCTC LOINC® and SNOMED codes on triggering eICR.

Methods: The LOINC® code data used by reference labs was collected, and quantitative analysis of LOINC® codes from selected data sources combined with visual inspection to categorize differences. The workflow process analysis was gathered by conducting a series of interviews with personnel's from an Atlanta hospital system.

Result: Descriptive analysis of validation of LOINC® codes used by four national reference labs for 4 piloted reportable conditions revealed that there were a total of 41 LOINC® codes missing in RCTC, with the highest number of missing codes coming from Quest laboratories, and the highest number of missing code coming from Salmonella. Through inferential analysis the study highlights the four discrepancies patterns in use of LOINC ® codes. Additionally, the analysis of the validation of EHR process, revealed several discrepancies in the use of LOINC® codes and SNOMED-CT codes, and study recommended measurable next steps that can be taken to address the discrepancies.

Conclusion: The study highlighted the discrepancies in use of trigger codes (LOINC®) by reference labs, and suggested the validation of triggers codes against codes actually used by reporters. The study observed gaps through the EHR implementation process analysis and recommended that from a process standpoint the LOINC® codes should be an integral part of EHR, and clinical laboratories should map local codes for lab results reporting to SNOMED codes to facilitate trigger coding at the EHR.

Validation of Reportable Conditions Trigger Codes (RCTC) content and EHR implementation process

By

Sarita Sadhwani

MPH, Emory University, 2016
MS Life Sciences, Devi Ahilya Vishwavidyalaya, Indore, India, 1996

Thesis Committee Chair: Sunanda McGarvey, BS, CBAP, CSM

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in Executive MPH, 2016

Table of Contents

Chapter 1: Introduction

Introduction and background

The Reportable Conditions Trigger Codes (RCTC) are a national set of codes to be implemented in the Electronic Health records (EHR) and matched against encounter information to initiate generation of an electronic Initial Case Report (eICR). They serve to filter encounters that may be reportable from all encounters recorded in an EHR. The electronic case reporting (eCR) is a step towards achieving the vision of bidirectional information flow between public health and health care, it can help bridge the digital gap existing between healthcare and public health. The RCTC is the appropriate first step for this initiative, as it can provide timely initiation of electronic initial case reports from healthcare to public health when trigger codes are matched to information in a patient's encounter record.

In the electronic case reporting flow, the RCTC serves as an initial coarse filter in the EHR or Laboratory Information Systems (LIS) that is paired with a secondary evaluation step that determines if a case meets public health reporting specifications. However, if the coarse filter fails to identify potential cases then those encounters may be completely missed. The relevance of this study is to take the first step in validating the completeness of the RCTC by ensuring that all relevant Logical Observation Identifiers Names and Codes (LOINC®) codes used by four national reference labs and a clinical lab, for four pilot reportable conditions, are included in the RCTC. The RCTC includes more than LOINC® codes, and are actually comprised of the following codes for reportable conditions: Diagnosis codes (ICD 10 and Systematized Nomenclature of

Medicine --Clinical Terms (SNOMED-CT); Test Names from Lab Results Reports (LOINC®) for condition-specific tests; Test Results from Lab Results Reports (SNOMED-CT) for results that are organisms associated with a reportable condition; and Test Orders Placed (LOINC®) for those conditions that are reportable based on suspicion of the condition. Additional codes, such as RxNorm codes may be added in future releases of the RCTC.

A guiding principles in defining RCTC is that there should be one set of trigger codes available to be used by all reporters regardless of the jurisdiction to whom they report. Therefore triggers would include codes for any condition that is reportable in any jurisdiction, and they would be specific only to reportable conditions - which means general test codes (e.g., general culture tests) would not be included in the set of trigger codes.

The RCTC approach is intended to simplify implementation for the reporter by limiting the logic required by the EHR. RCTC is designed to be coupled with secondary evaluation that applies jurisdiction-specific reporting criteria to determine if an encounter is reportable to and to which jurisdiction(s).

Public health uses case reports and Electronic laboratory Reporting (ELR) for reportable diseases/conditions to monitor, control, and prevent the occurrence and spread of reportable conditions. The envisioned electronic case reporting information flow diagram is shown below. The use case is a health condition referral from EHR to public health, where the information starts with a patient visit in the healthcare provider setting. The provider enters the clinical information into the EHR and the patient encounter information is saved. The second swim lane of the flow diagram is represented by the

Electronic Health Record, and is where the RCTC role comes into play. If the encounter information matches a trigger code, the initial electronic case report will be generated. Once the initial case report is sent to the public health community platform, the platform calls RCKMS through a shared service, and reportability is determined based upon jurisdiction requirements. Below highlights the focus of this thesis as 'trigger codes'.

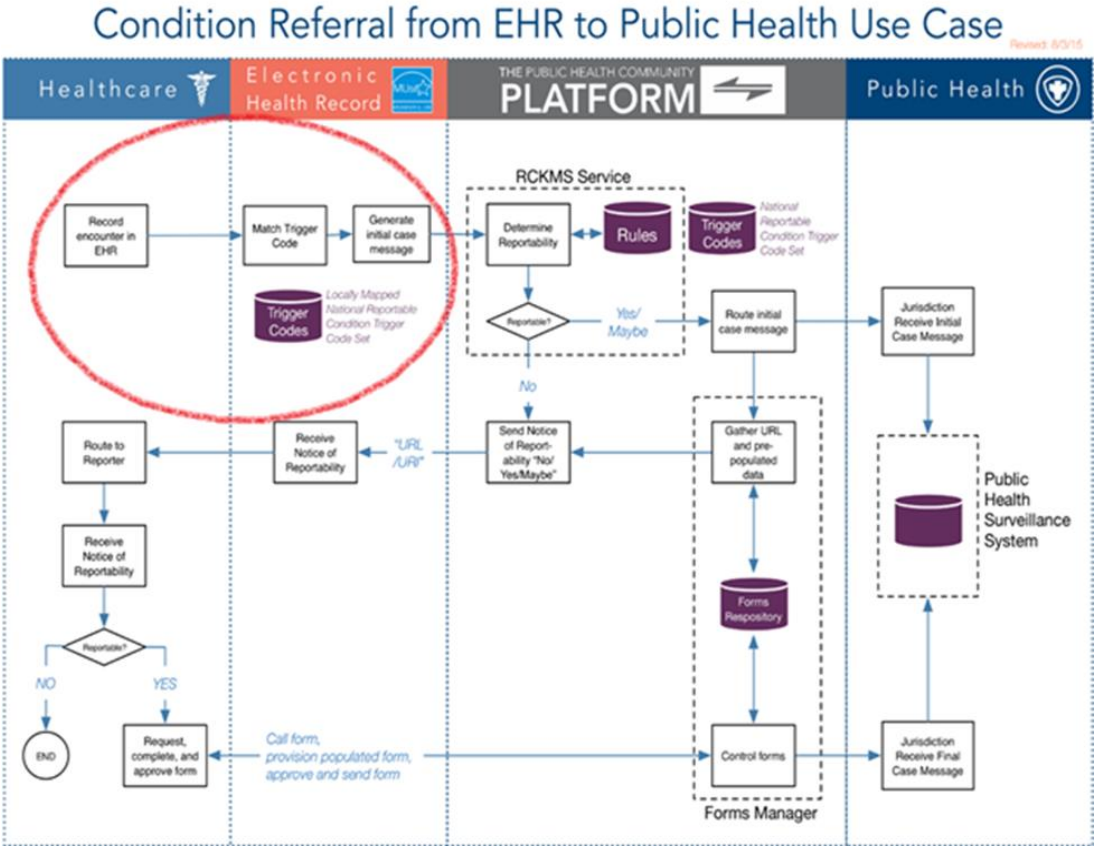


***Figure* 1:** Specifically point out with red dotted circle the area of interest; Trigger codes. Reprinted from CDC/OPHSS/CSELS - RCKMS Content development presentation, August 19th, 2015.

Problem:

There are several components needed to achieve eCR, including: trigger codes to filter clinical encounters and identify a reportable condition; standard data extractions from EHRs that create an HL7 electronic initial case report; data transaction and transport protocols; and security protocols. EHRs are responsible for implementing trigger codes and matching them to encounter data to initiate an eICR, the completeness and applicability of trigger codes is paramount to successful implementation of eCR. Therefore, it is critical that the trigger codes are both inclusive of codes of interest to public health, and validated against codes actually used by reporters. Not only do healthcare providers need to use standard codes, but they need to use them consistently. Additionally the healthcare and clinical lab systems should use the data standards appropriately during the workflow that take place among EHR, Clinical labs, Reference Labs, and public health while transferring the cases information.

Purpose:

The purpose of this research is to validate both the RCTC content and the implementation of standard codes in the EHR processes. The validation of RCTC content is done by evaluating LOINC® trigger codes in the RCTC for 4 pilot jurisdiction reportable conditions (Chlamydia, Pertussis, Salmonellosis, and Gonorrhea) against codes currently in use by four national reference labs. The validation of EHR implementation process is done by reviewing the workflow among the EHR, clinical labs, and public health labs.

For validation purpose the study is limited to LOINC® codes only, since national reference lab exclusively uses LOINC® codes to identify the tests that they ran in the results report that is returned to the party that ordered the tests to be run. This study is designed to answer the following research questions:

1) What are the gaps in the use of LOINC® codes and the RCTC vision for implementing trigger codes within the EHR?

    a) What is the impact on completeness of eCR for reportable conditions when LOINC® codes are not used as envisioned by EHRs?

2) What are the discrepancies between LOINC ® codes used for triggering electronic case reporting (eCR) using RCTC and LOINC® codes used by national reference labs?

    a. What are the patterns observed in the discrepancies?

3) Are there measurable next steps that can be taken to address discrepancies between the national references labs included in the thesis research?

Significance statement

LOINC® and SNOMED CT are the two most complete coding systems representing lab test type and result information, respectively. Therefore, these two information coding systems were specifically recommended for use in coding laboratory information in electronic health records by the U.S. Department of Health and Human Services Office of National Coordinator for Health Information Technology. (Dhakal, S., Burrer, S. L., Winston, C. A., Dey, A., Ajani, U., Groseclose, S. L., 2015)

There is a need to evaluate trigger codes (RCTC) used for eICR report generation for completeness and consistency. If the trigger codes are not aligned with codes used by the LIS or EHR, then a match will not occur. This study starts the validation process by reviewing LOINC® codes used by four national reference labs and or four pilot conditions against the RCTC LOINC ® codes. LOINC® codes represent 31.59 % of codes in the RCTC, and were selected as a starting point for the validation process because they were commonly used by the national reference labs. National reference labs were selected because their code mapping is publicly available, and they represent 68.29 % of lab reporting for the pilot conditions that are a part of this study. Automated disease reporting supports the case reporting meaningful use stage-3 objective that is part of the Health Information Technology for Economic and Clinical Health (HITECH) act, enacted as part of the American Recovery and Reinvestment Act (ARRA).

The analysis of the EHR process can help understand the usage of LOINC®/SNOMED codes and RCTC vision. This study starts the validation of EHR process by analyzing laboratory work flow and the interaction between Clinical Lab systems and EHRs.

There is an opportunity for healthcare providers to improve the exchange of laboratory data by using standardized trigger codes. The usage of codes may be direct or implemented via a mapping between each institutional specific laboratory code and name in their system's term dictionary to LOINC® names or codes. Evaluating trigger code use by laboratories and EHR can help improve the completeness of trigger codes in RCTC, and in turn improve the completeness of case reporting.

Terms Definition:

Term	Definition
ELR- Electronic Lab Reporting	**ELR is the electronic transmission from laboratories to public health of laboratory reports which identify reportable conditions.**
eCR- Electronic Case Reporting	**Electronic reporting of reportable data from EHR to public health. The "triggering" of an eCR report by an EHR based on the presence of clinical data in the patient's record that matches a particular set of coded values (trigger codes).**
eICR	**Electronic initial case reporting.**
RCTC- Reportable Conditions Trigger Codes	**Reportable Conditions Trigger Codes - codes implemented in the health care system to match against encounter information and initiate an eICR. Universal code set for all reportable conditions in all jurisdictions (lab order, result, diagnosis) for piloted Pertussis, Salmonellosis, Gonorrhea, and**

	Chlamydia, published within PHIN VADS.
PHINVADs	**Public Health Information Network Vocabulary Access and Distribution System.**
EHR	**Electronic Health Record**
LOINC®	**Logical Observation Identifiers Names and Codes.**
RCKMS	**The Reportable Condition Knowledge Management System is an authoritative, real-time portal to enhance disease surveillance by providing comprehensive information to public health reporters about the "who, what, when, where, and how" of reporting.**
SNOMED-CT	**Systematized Nomenclature of Medicine--Clinical Terms**
LIMS	**Laboratory Information System**
MURR Interface	**Meaningful use interface used for reporting of reportable conditions to public health.**
PHAs	**Public Health Agencies**

Chapter 2: Review of Literature

Introduction

This literature review evaluates the workflow and interactions between Clinical Lab Systems and EHRs, Specifically, this review examines the gaps in the use of LOINC® and SNOMED, and its impact of missing RCTC LOINC® and SNOMED codes on triggering eICR. Although there is a body of research on EHR as a great source to advance quality measurement and secondary uses of clinical data in EHRs, research that also incorporates the role of the EHR in eCR is largely absent. The EHR is the first line of defense to capture potential cases based on information recorded as a part of a patient encounter and matched to trigger codes.

The electronic case reporting and use of trigger codes (using RCTC) is a project which is currently under development by CDC/CSTE, but there is lack of literature on the topic. Therefore, grey literature is used for understanding the process of eCR and RCTC, and the source of grey literature is accessed from CDC resources.

Work Flow analysis among EHR, lab, and outpatient's lab.

The workflow analysis for an Atlanta Hospital System was derived from interviews with a senior lab computer analyst, infection prevention personnel, and personal experience from working at the lab as a Medical Technologist.

The workflow analysis looks at interactions between the EHR, the clinical lab, outpatient lab, reference lab, and public health department, with the focus being on those points in the flow where LOINC® codes are implemented.

The chosen Atlanta Hospital system - the Northside Hospital system includes three not-for-profit hospitals, located in Atlanta, Forsyth and Cherokee, with a total of 852 licensed beds. The main campus located in Atlanta is in charge of taking care of operations of hospital information technology systems.

Background information for process analysis was gathered through interview on 7-22-16 at 1:30 pm with IT personnel at hospital system. The question that were asked to gather requirement were:

1. Are you using data standards such as LOINC® for ordering lab tests?
2. Are you using SNOMED for reporting of lab test results?
3. Would you be able to provide the LOINC® and SNOMED used at your hospital system for 4 pilot conditions of eCR?
4. Is the hospital system reporting lab results to public health electronically?
5. Is the LIMS capable of reporting results to the EHR?
6. Is the hospital system following guidelines of Meaningful Use Stage-3 for lab reporting?
7. To whom does the hospital system report outpatient results?
8. Who is responsible for reporting of reportable conditions for outpatient labs?
9. Who is responsible for reporting of lab tests sent out to reference labs?

The question that were asked to Infection prevention personnel were:

1. What is the role of infection prevention in reporting of reportable conditions to public health?
2. How is the reporting of reportable conditions to public health performed?
3. Is reporting done manually or electronically?

The workflow diagram is below:

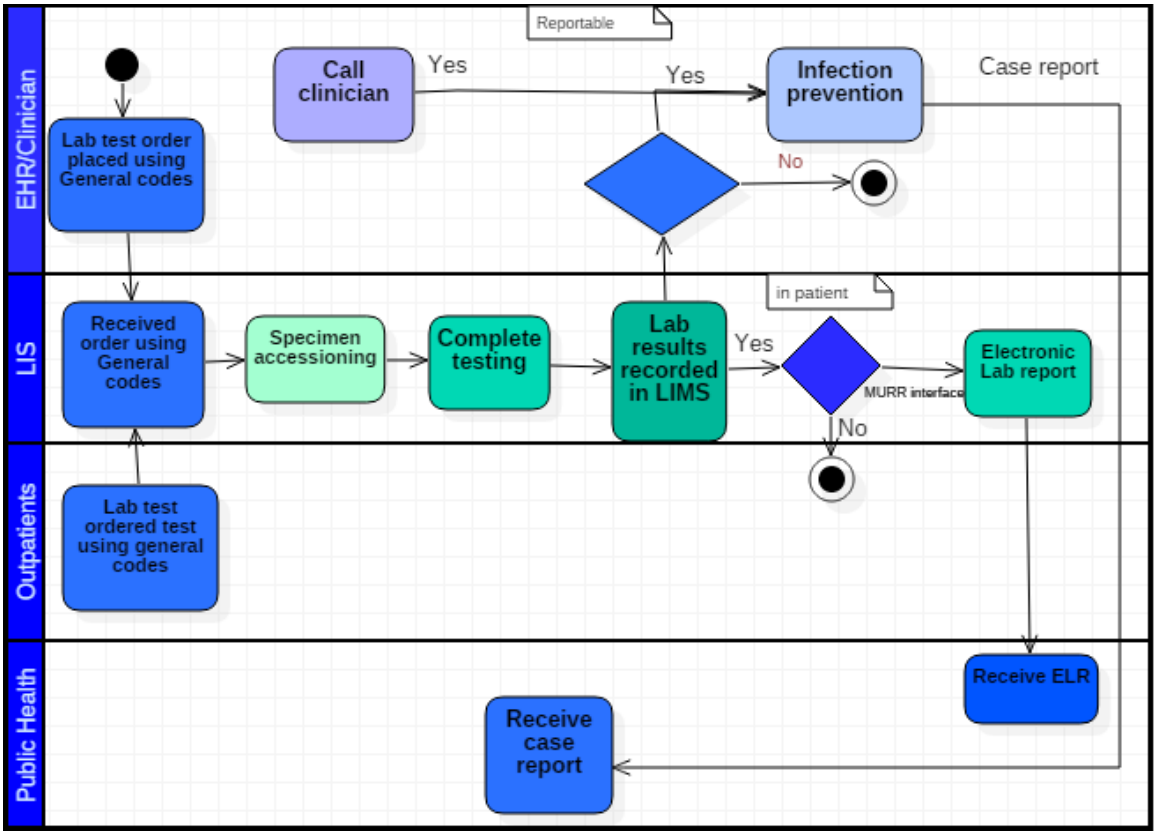


***Figure* 2: The workflow process interaction diagram among EHR, clinical Lab, outpatients, and public health.**

The steps in the workflow process are listed below:

- The clinicians or healthcare provider uses medical record number/patient id to places order in EHR.
- Lab receives the order depends upon the source it comes from, the orders comes through EHR or from outpatient's lab.
- The clinical lab uses medical record number to look up the order placed, if the order matches with specimen received, the specimen accessioned by the laboratory (i.e., enters lab testing process) .
- The specimen accessioning is performed using the patient id and general test code for test name to order and receive the specimen.

- After lab test are performed and the result are entered into LIMS, the test result are not reported using LOINC® or SNOMED. Instead lab technologist uses standard general code per organization policy. For example Salmonella species, instead of using code used for Salmonella.
- If the lab test result is reportable conditions, the Meaningful use lab interface (MURR) triggers ELR that goes to public health. The reportability of lab result is determined by SNOMED-CT codes that is specific for reportable conditions.
- The case report is not generated electronically, the Infection prevention personnel pulls epidemiology lab report and conduct their own surveillance and if the finding matches with reportability criteria the next event get initiated.
- The Infection prevention personnel is responsible for reporting of reportable conditions for both inpatients and outpatients to public health.
- The notification of case report is done through website: https://sendss.state.ga.us notification, by manually filling out form.
- Interviews with Infection prevention personnel, revealed that most of the time they are entering clinical information of patients, and labs are already reported to public health through ELR.
- Interview with IT personnel revealed that the mapping of local codes to standard codes for reportable conditions are done on case to case basis, and the mapping is done automatically at meaningful use lab interface (MURR). The MURR interface generate the ELR for reportable condition to send to public health. The mapping of local codes to standard codes for reportable conditions at MURR interface is performed only for in-house patients only.

- Labs do not use EHR, and the lab personnel are not aware of the interoperability standards use, they are aware of patient medical record number and general code being use by hospital.
- The hospital lab IT personnel said, ‘the hospital is responsible for reporting of reportable conditions for only in house patients, and the lab use general code provided by reference lab for ordering the special test that are not performed at facility.”
- The lab test that are sent out to reference lab are ordered using their general code.

Summary of LOINC® and SNOMED-CT usage with in workflow:

The RCTC anticipates that triggers fire at the following points in the flow:

- Test order placed (LOINC®), reporting required based on suspicion (for example anthrax), includes lab orders where the suspicion of the condition is, itself, reportable. For that orders placed, coded in LOINC® , includes at least one test specific to a reportable condition and orders placed value sets must be used to match codes against tests ordered rather than tests performed.
- Test name from lab results report (LOINC®), test name specific for reportable condition, this criteria includes laboratory test names coded in LOINC® -specific to reportable condition. Result could be non-specific (positive, detected, numeric) and is important in the context of the test name resulted. This excludes generic tests, for example general bacterial cultures.
- Test result from lab results report (SNOMED-CT), result value that represent reportable conditions, included laboratory values, such as organisms found in generic tests, coded in SNOMED-CT. These results are relevant for non-specific

tests such as cultures where the lab test performed (lab test name) is not specific to a reportable condition.

The analysis shows the standard codes being introduced at the following points:

After reviewing the work flow, it was found that the LOINC® and SNOMED codes are introduced only at the point of ELR, after the lab entered results into the LIMS. If the results are matched to reportable conditions of trigger codes at MURR interface, the ELR get generated.

Role of electronic case reporting and ELR in public health surveillance:

Presently, public health case reporting from providers is manual, paper-based, and labor intensive which can result in reportable conditions not being reported to local and state public health departments.(Dixon,2014 ; Rajeev, D., Staes, C. J., Evans, R. S., Mottice, S., Rolfs, R., Samore, M. H., … Huff, S. M., 2010)

There are many challenges in fully implementing public health electronic case reporting due to technical, policy, and funding challenges in supporting interoperability. Recent efforts including ELR and eCR have focused on sharing data between health care and public health entities.

ELR use has increased the volume and timeliness of reporting compared with the traditional faxed reports. ELR helps identify reportable conditions determined by confirmatory testing and supports case reporting at the state or local level. ELR is used by laboratory providers to help them meet state reportable diseases laws mandating that providers report cases of specified diseases to the health department. ELR supports overall public health surveillance by helping improve the timeliness and accuracy of case

reporting and confirmation to state and local health departments. It also supports national public health surveillance by improving the timeliness and accuracy of notifiable disease data voluntarily shared by states with CDC.

ELR has several drawbacks: it is limited to conditions that have laboratory tests; labs lack data elements important to public health, such as; patient demographics, location, and clinical data; and lab reporting lacks the timeliness needed to inform public health of conditions reportable upon suspicion. These drawbacks indicate the need for electronic case reporting. (Rajeev, D., 2010)

As the use of EHRs increases in the USA, the opportunities for electronic case reporting also increase. An electronic case report contains laboratory results similar to an ELR message, but also additional information about patient demographics, clinical findings, and other relevant data that can be extracted from the EHR. In the USA, each state requires that a specific set of conditions be reported to public health authorities by a clinician who diagnoses the condition, in some cases in the absence of laboratory confirmation.

Role of LOINC® and SNOMED standardized coding support ELR and Case Reporting:

In the field of public health, the data sharing takes place among several entities, including a health department or between: local health departments, local and state health departments, state health departments, health providers and health departments, state health departments and the Centers for Disease Control and Prevention (CDC). The capability to share data meaningfully, though depends upon the existence of a common, clearly defined set of data elements. Additionally, data elements must be collected in a

standardized way that is conductive to being machine-processable. LOINC® and SNOMED are two coding standards developed to promote standardization. (ONC)

Since Electronic Laboratory Reporting (ELR) and eCR both have the potential to enable more accurate, timely, and cost-effective reporting, both have been systematically promoted as a public health priorities. Electronic Laboratory Reporting (ELR) has been promoted as a public health priority and its inclusion as a meaningful use objective for public health has served as a catalyst to accelerate its adoption. Use of structured, unique, and national available coding systems such as LOINC®, SNOMED for both ELR and electronic case reporting improves the computational characteristics of data. There are several coding strategies available, the recommendation by Office of the U.S. National Coordinator for Health Information Technology has suggested incorporating LOINC® for laboratory orders and (SNOMED CT) codes for laboratory results to standardize and eCR.(Dhakal , 2015)

Chapter 3: Method and Analysis:

This chapter describes the methodology used by study. The study followed four process steps:

1. Evaluating the process flow and interactions between Clinical and Lab(s) with a focus on where in the flow LOINC® codes are implemented.
2. Collecting the LOINC® codes used by reference labs.
3. Creating an analysis framework for use with the limited data set of this study, but extensible to a broader data set for future studies.

4. Analyzing collected LOINC® codes using the analysis framework created for this study.

1) Process flow and interactions between Clinical and Lab(s)

The process flow focused on where standards LOINC® codes get implemented. The workflow was gathered by conducting a series of interviews with a senior lab computer analyst, infection prevention personnel, and a medical technologist. The workflow analysis looks at interactions between the EHR, the clinical lab, outpatient lab, reference lab, and public health department, with the focus being those points in the flow where LOINC® codes are implemented.

After conducting interviews the workflow process diagram was created using StarUML (open-source Unified Modeling Language (UML) and Model Driven Architecture (MDA) tool).

2) Data Collection:

The methodology describes the framework and process used to identify and categorize discrepancies between the RCTC and national lab LOINC® codes and determine reasons for the discrepancies.

Data sources and data preparation

This study involved data gathering from several disparate data sources. The data then had to be cleansed, and prepared for analysis. The data integration was aided by use of a simple data model to organize data from the disparate sources. The data sources included:

- LabCorp – https://www.labcorp.com/.EdosPortlet/TestMenuLibrary?libName=File+Library&compName=LOINC
- Quest – http://www.questdiagnostics.com/testcenter/TestCenterHome.action
- Mayo – http://www.mayomedicallaboratories.com/test-catalog/loinc-codes.php
- ARUP – http://www.aruplab.com/Testing-Information/resources/LOINC-codes.xlsx
- Atlanta Hospital lab: The LOINC and SNOMED data was obtained through emails.

Data preparation.

Once obtained, the raw data was cleansed of various errors. For some data sources, the first challenge was to make the data usable for further evaluation. This was the case with LabCorp where the LOINC® codes pulled from their website were only available as a locked pdf document, which was non-editable. This locked pdf file had to be uploaded to an online PDF converter (online2pdf.com) which was used to unlock the PDF. Concept codes and concept names were then extracted and saved to an MS-Excel CSV (Comma Separated Values) format. Next, duplicate concept codes were eliminated. The duplicates were general tests that were run for multiple conditions.

The Quest laboratory LOINC® codes were pulled using each conditions name search on their website. The retrieved LOINC® codes were saved in CSV format to be used for analysis. The LOINC® codes for ARUP and Mayo were easily obtained from their website and were downloaded and saved in CSV format.

All known data inconsistencies due to use of different file formats and extra attributes used by reference lab for their own requirements not applicable to this study were removed. LOINC® codes not related to reportable conditions were removed. An attribute for the data Source was added at the end of each record to indicate where it came

The common concerns related to data noise such as incorrect attribute values, duplicate records and incomplete data records were removed.

This study used R, the statistical computing language, to accomplish a majority of the data mining and analysis. R's sqldf package was used to create several data frames, and the CSV files were loaded into these data frames. SQL statements (all SQL queries used for analysis are included in Appendix below) were coded against the data frames to analyze the data, similar to what is done for relational database tables.

4) Framework used by study

The framework analysis has been used with the limited data set of this study, but extensible to a broader data set for future studies.

Study framework and rationale for using framework.

The study used R and the sqldf package to create an extensible framework, built to support the 4 reportable conditions included in this study, but designed to support validation of the remaining reportable conditions (approximately 120 conditions) in the future. The validation of all reportable conditions, especially if data sources are expanded to include clinical labs and public health labs, in addition to reference labs, would result

in a much larger data set than exists for this initial study, but would still be supported by the framework and would be queryable using SQL.

The data sources by condition (Salmonella, Neisseria, Pertussis and Chlamydia) were saved as CSV files and loaded into the R framework to determine discrepancies and identify missing RCTC Concept Codes;

- The Salmonella CSV file is stored internally in a data frame referenced as df_Sal.
- The Neisseria CSV file is stored internally in a data frame referenced as df_Nei.
- The Pertussis CSV file is stored internally in a data frame referenced as df_Per.
- The Chlamydia CSV file is stored internally in a data frame referenced as df_Chl.

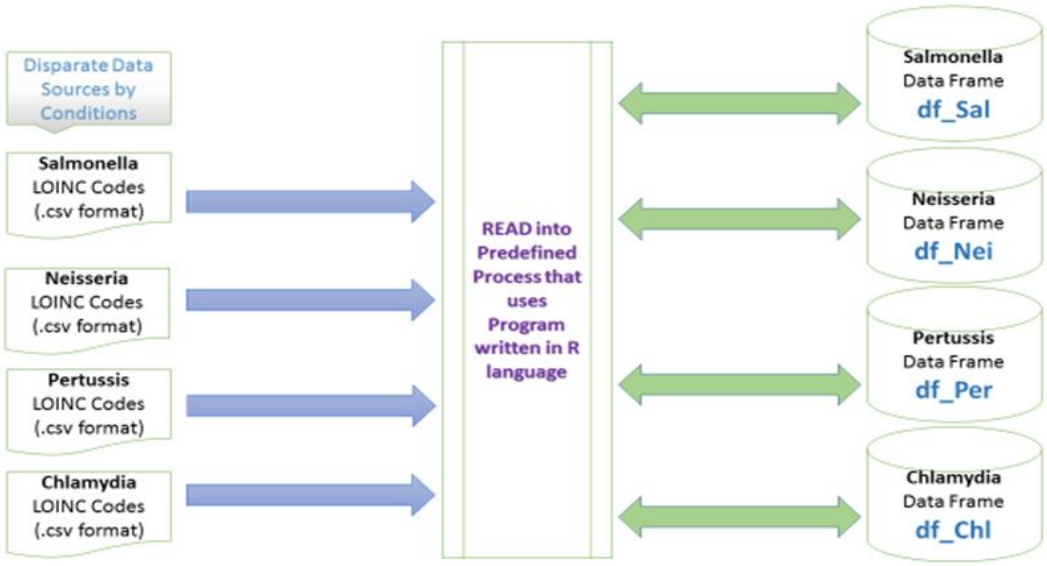


***Figure* 3: Create data frames for each condition using csv files.**

The above 4 data frames were sorted by condition name and summarized as follows:

- 92 Unique Codes for Salmonella

- 101 Unique Codes for Neisseria
- 85 Unique Codes for Pertussis
- 377 Unique Codes for Chlamydia

The LOINC® codes were retrieved from different sources (LabCorp, Quest, Mayo, ARUP) and split for each condition in a separate condition CSV files. The Reporting Conditions Mapping Table (RCMT) and RCTC codes for the 4 conditions were retrieved from PHIN VADS website and were appended in the appropriate condition CSV files. These condition CSV files were read into 4 data frame and then sorted by unique data source. This results showed 6 unique data sources listed below -

1. ARUP
2. LabCorp
3. Quest
4. Mayo
5. RCMT
6. RCTC

After the above step, the 4 conditional data frames were split into 6 frames organized by data source. The example for ARUP codes is shown in figure 7, and was repeated for each of the data sources.

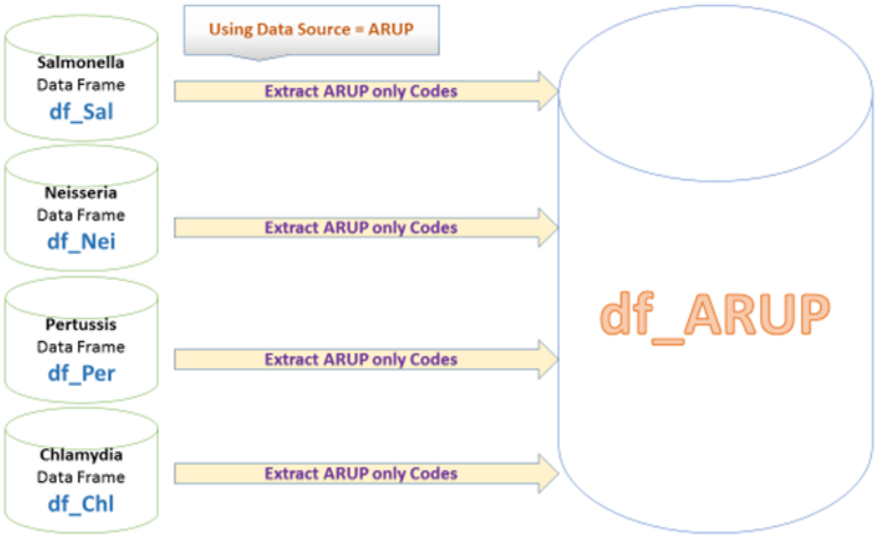


***Figure* 4: Creating data frame for each lab by extracting the concept code from conditions.**

After splitting, each of the data frames (df_ARUP, df_LabCorp, df_Quest, df_Mayo and df_RCMT) were compared against the df_RCTC to identify missing RCTC codes. The missing RCTC codes were loaded into a separate data frame referenced as df_Missing_RCTC.

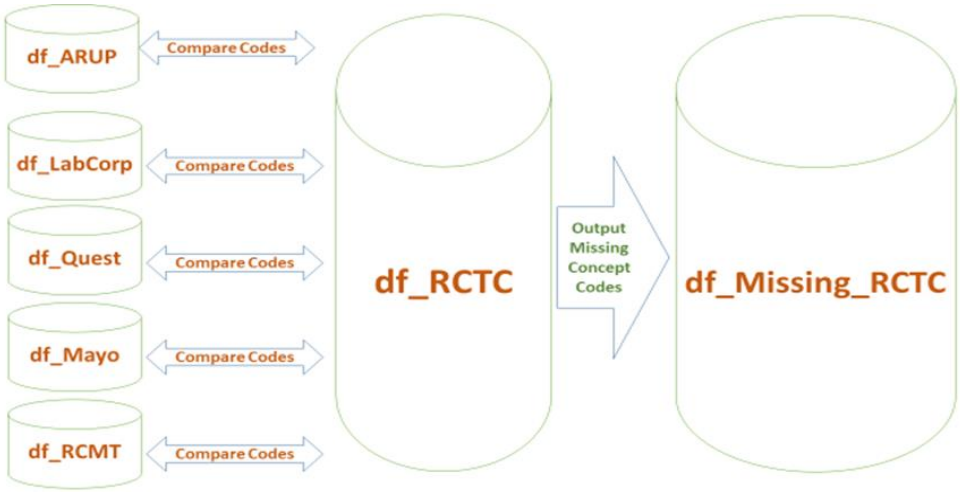


***Figure* 5: Creating Missing_RCTC data frame of concept code for each labs ( ).**

The attributes in the df_Missing_RCTC include Concept Code, Frequency, Data Source, Concept Name and Condition Name. The df_Missing_RCTC were exported to a CSV file RCTC_Missing_Codes.csv.

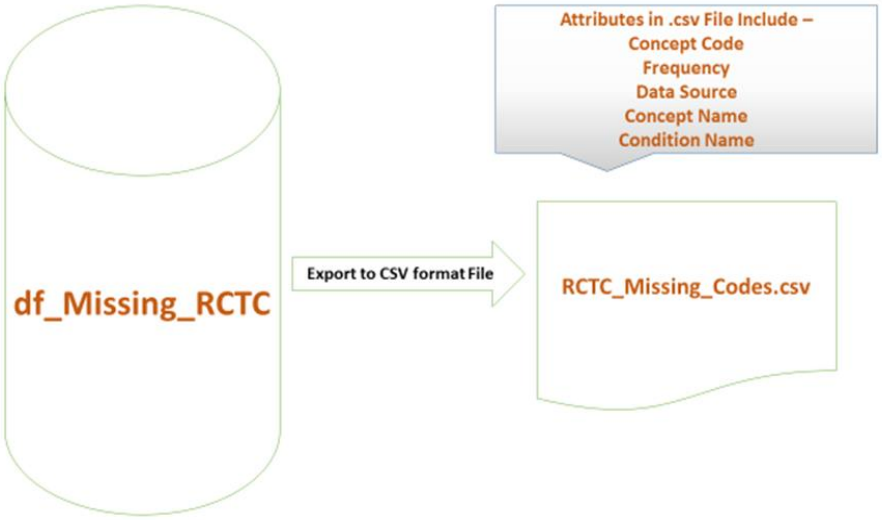


***Figure* 6: Creating csv file of dataframe missing_RCTC.**

4) Analyzing collected LOINC® codes using the analysis framework created for this study

The RCTC_Missing_Codes.csv file was loaded into MS-Excel and used to build pivot tables, charts and the final discrepancy tables that support the results described in chapter 4.

Chapter 4: Results

Introduction

This chapter presents the results from the evaluation of LOINC® codes in the Reportable Conditions Trigger Codes (RCTC), and LOINC® codes used by national reference labs and hospital lab. The results describe the following aspects of the study:

1. The discrepancies between LOINC® code usage, whether it be within a lab, across reference labs, or between the labs and RCTC.
2. Descriptive analysis, followed by inferential analysis of discrepant codes.
3. Workflow analysis among EHR, lab, and outpatient's lab.

Findings

The analysis was conducted to find LOINC® codes that are used by ARUP, LabCorp, Quest, Mayo and Atlanta Hospital lab that are of interest to public health, but are missing in RCTC. Based on findings presented in tables 1 through 4, it was evident that there were a total of 41 LOINC® codes missing in RCTC across the 4 Notifiable conditions (Chlamydia, Gonorrhea, Pertussis, and Salmonellosis), with the highest number of missing codes coming from Quest laboratories , and the highest number of missing code coming from Salmonella.

Summary of missing LOINC® codes in RCTC by reference labs are included in figures below :(Chart 1)

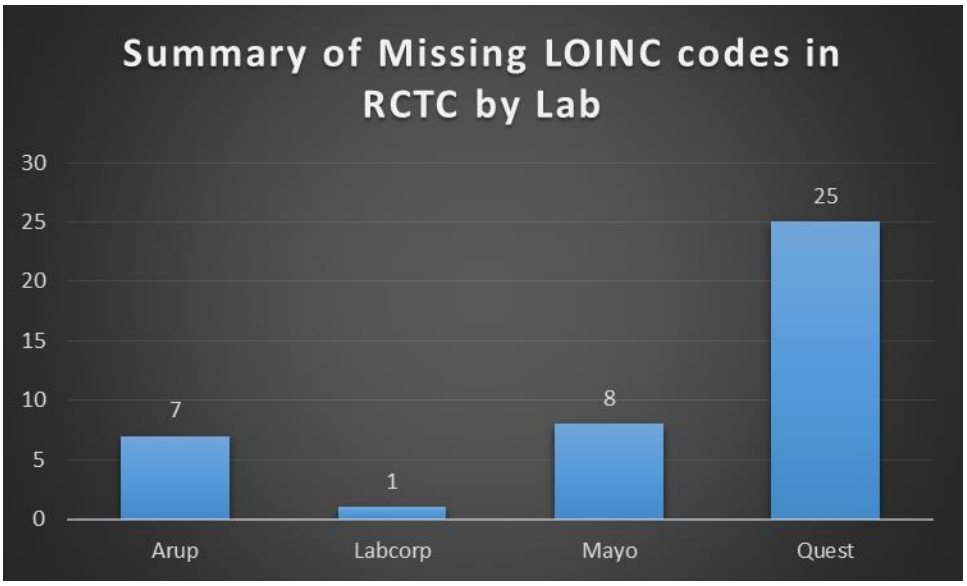
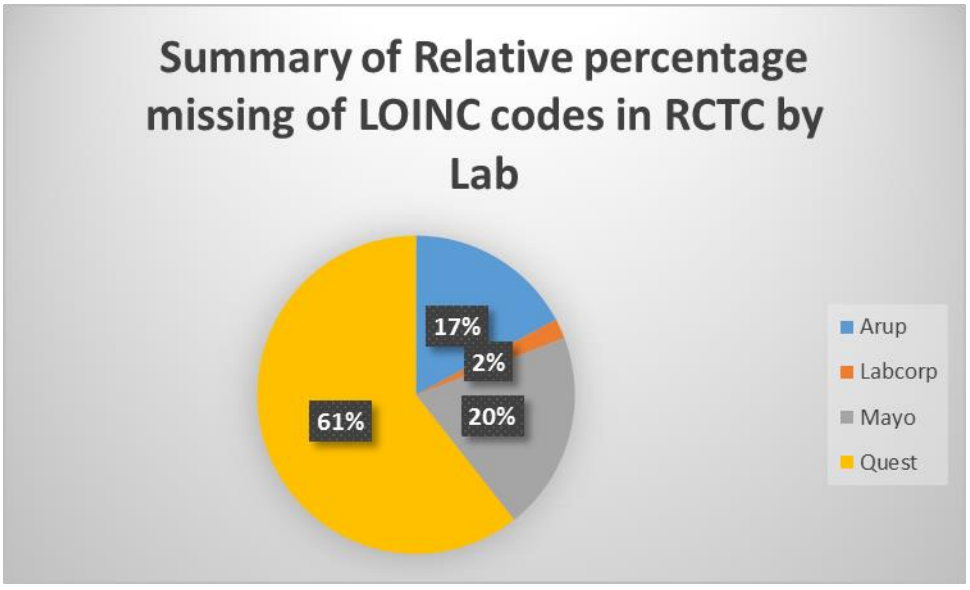


Chart 1: Summary of missing LOINC® codes in RCTC by lab.

Quest laboratories have 3 times more RCTC count of missing concept codes then Arup and Mayo.



Pie Graph 1: Summary of Relative percentage of missing LOINC® codes in RCTC by Lab.

The 61% of the missing RCTC codes are from Quest, while 17% are from Arup, 2 % from LabCorp, and 20% from Mayo.

Summary of missing LOINC® codes in RCTC by conditions name are included in figures below; **(Chart 2)**

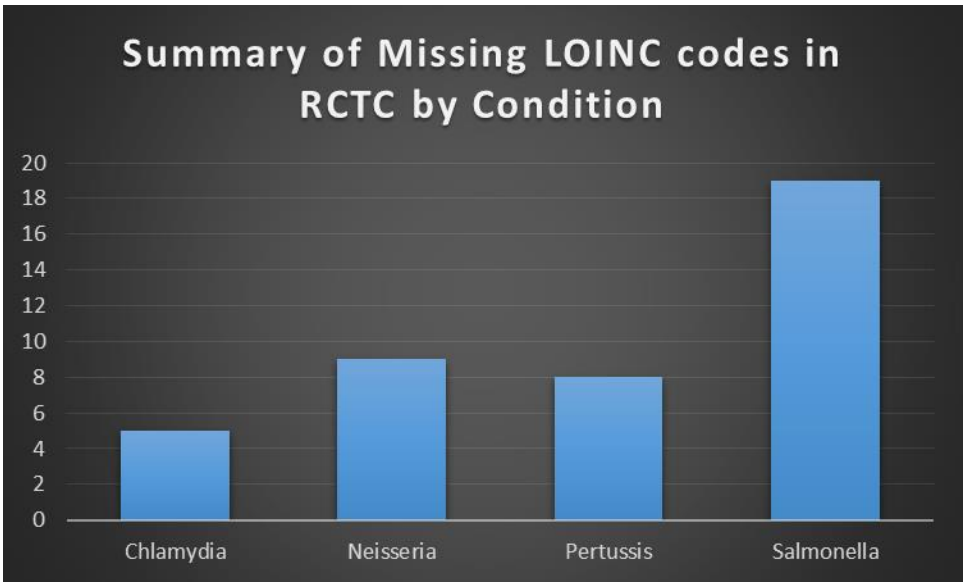
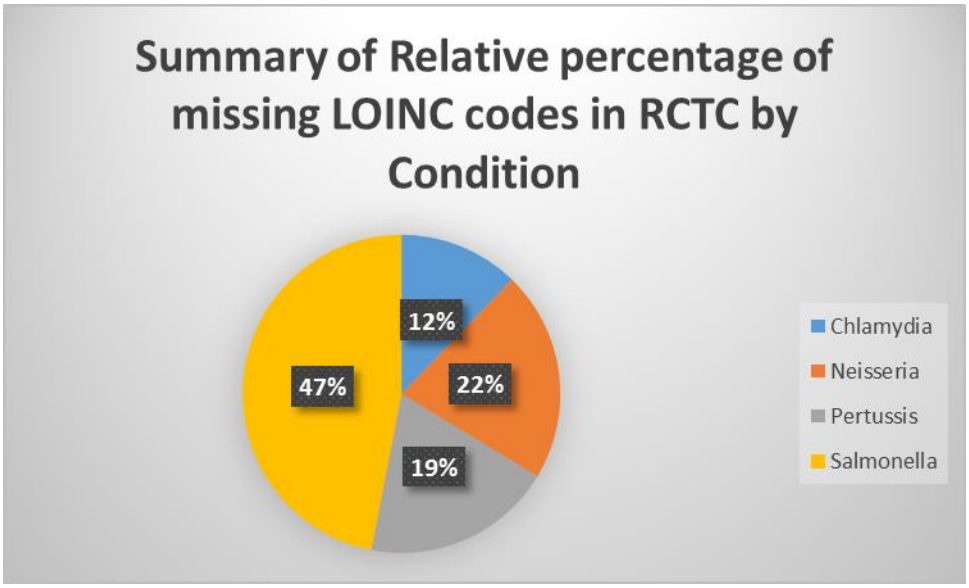


Chart 2: Summary of missing LOINC® codes in RCTC by condition name.

Salmonella has twice as many missing concept codes than Neisseria and pertussis, and 3 times as many as Chlamydia.



Pie Graph 2: Summary of Relative percentage of missing LOINC® codes by condition name.

The 47% of the missing RCTC concept codes are from Salmonella, while 22% are from Neisseria, 19 % from Pertussis, and 12% from Chlamydia.

Frequency table (Pivot Table) of RCTC missing by Concept Code and Condition Name

The frequency of LOINC concept codes for each reportable condition (Chlamydia, Gonorrhea, Pertussis, and Salmonellosis) not found in RCTC is presented in Pivot Table 1 below. The frequency value includes codes across all 4 reference labs included in the study. The frequency table (table 1) of LOINC codes not found in RCTC, generated by criteria selected by Condition name and Concept code.

Table 1

Frequency of RCTC missing by Concept Code and Condition Name

Concept Code	Condition name -Chlamydia	Condition name - Neisseria	Condition name - Pertussis	Condition name -Salmonella	Grand Total
13284-5				2	2
Salmonella typhi H D Ab				1	1
Salmonella typhi H D Ab				1	1
13285-2				2	2
Salmonella typhi O D Ab				1	1
Salmonella typhi O D Ab				1	1
13916-2			1		1
Bordetella parapertussis Ag			1		1
17562-0				1	1
Salmonella typhi/paratyphi Abs Interp				1	1
20423-0		1		1	2
Beta lactamase organism identified				1	1
BETA LACTMASE		1			1
21070-8		1		1	2
Antibiotic XXX				1	1
MIC SENSITIVITY (BREAK POINT)		1			1
22517-7				2	2
Salmonella paratyphi A H Ab				1	1
Salmonella paratyphi A H Ab				1	1
22521-9				2	2
Salmonella paratyphi B H Ab				1	1
Salmonella paratyphi B H Ab				1	1
23667-9		1		1	2
Bacteria identified				1	1

ID BY PROBE		1			1
29723-4			3		3
Bordetella parapertussis by PCR			1		1
Bordetella parapertussis DNA			2		2
31208-2 (LOINC code used for Specimen source of Unspecified specimen)	1	3	3	1	8
B. pertussis/parapertussis Source			1		1
Bordetella pertussis and Bordetella parapertussis: Molecular Detection: PCR			1		1
C. trachomatis Source	1				1
N. gonorrhoeae by TMA		1			1
SOURCE:		1			1
Specimen source			1	1	2
Specimen source:Prid:Pt:XXX:Nom:		1			1
35347-4				1	1
Microscopic observation				1	1
42588-4			1		1
Bordetella pertussis and Bordetella parapertussis: Molecular Detection: PCR			1		1
45187-2		1		1	2
Antibiotic XXX				1	1
KIRBY-BAUER (DISK DIFFUSION)		1			1
55617-5		1		1	2
Antibiotic XXX				1	1
E-TEST		1			1
57769-2				2	2
Salmonella O: Type Vi Ab				1	1
Salmonella typhi O Vi Ab				1	1
59464-8	1				1

Chlamydia IgM Panel Interpretation	1				1
612-2				1	1
Bacteria identified				1	1
6912-0	1				1
Chlamydophila pneumoniae Ab.IgA	1				1
6914-6	1				1
Chlamydophila pneumoniae Ab.IgM	1				1
6916-1	1				1
Chlamydophila psittaci Ab.IgG	1				1
74384-9		1			1
N. gonorrhoeae by TMA		1			1
Grand Total	5	9	8	19	41

Summary of Descriptive Analysis of all above 3 Pivot tables.

Based on the results of the descriptive analysis and frequency table of missing RCTC by concept code and concept name, the inferences of missing RCTC LOINC® codes were drawn and the RCTC missing codes discrepancies were further categorized.

Inferential Analysis of result obtained from Descriptive analysis:

In this step the missing LOINC® codes in RCTC were assessed and inferences were drawn. For validation purposes the LOINC® codes were verified using LOINC.org as the gold standard. After verifying the codes, the discrepancies that were found, and any pertinent observations were categorized by type, and the associated implications were suggested.

LOINC discrepancy and observation categories (Table 2)

The discrepancies/observations were organized by the following categories:

- Use of different LOINC® codes for similar tests due to missing specimen types
- Use of general LOINC® codes instead of SNOMED for description of specimen source
- Use of non-specific general LOINC® codes for pathogen confirmation methods
- Missing relevant LOINC® codes

Table 2

LOINC discrepancy and observation categories

Discrepancies / observations type	**LOINC code**	**Concept Name**	**Implication**
Use of different LOINC codes for similar tests due to missing specimen types.	42588-4 29723-4	Bordetella parapertussis DNA in Nasopharynx. Bordetella parapertussis DNA in Unspecified specimen.	It's an observation found in missing RCTC code. It's not uncommon for reference labs to use the nonspecific LOINC codes in event of missing specimen type by healthcare providers.

Use of general LOINC codes instead of SNOMED for description of specimen source	31208-2	Specimen source [Identifier] of Unspecified.	Reference labs have used this general code when specimen source information is missing. Instead of LOINC the labs should have used SNOMED, since the recommendation is that SNOMED - CT should be used to code the specimen source.
Use of non-specific general LOINC codes for pathogen confirmation methods	35347-4 31208-2	Microscopic observation in unspecified specimen	All the specified LOINC codes in this category are used for either preliminary, secondary confirmation and or

	23667-9 21070-8 20423-0 45187-2 55617-5	Bacteria identified in unspecified specimen Id by probe Other Antibiotic [Susceptibility] by Minimum inhibitory concentration (MIC) Beta lactamase organism identified in Isolate Other Antibiotic [Susceptibility] by Disk diffusion (KB)	for susceptibility testing. These code use cannot provide any specific information towards the identification of reportable condition pathogens.

		Other Antibiotic [Susceptibility] by Gradient strip	
Missing relevant LOINC codes	The list of missing codes in RCTC is attached as a separate table (Table 3) in the description of designated category.		May requires addition to RCTC values sets.

Each Discrepancy category is discussed below.

Use of different LOINC® codes for similar tests due to missing specimen type

The study observed the use of two separate LOINC codes for the same test. For example, LOINC code '42588-4' was used for lab test 'Bordetella parapertussis DNA in Nasopharynx', whereas the LOINC code '29723-4' was used for 'Bordetella parapertussis DNA in Unspecified specimen'. Since reference labs rely on the information provided by healthcare providers, in the event of a missing specimen type, the reference lab has no other choice than to use nonspecific LOINC® codes.

Use of general LOINC codes instead of SNOMED for description of missing specimen source

The inclusion of specimen source is important in the electronic initial case report for reportable conditions, as certain tests are of interest to public health only when run on a sterile specimen, for example, Neisseria meningitis. This condition is reportable only if the 'specimen collection site is normally sterile body site', such as blood, CSFs, synovial fluid, pleural, or pericardial fluid. (CSTE position statement, 09-ID-42, 2009)

The recommendation is that SNOMED - CT should be used to code the specimen source, but clinical labs and reference labs are instead using LOINC® codes. The validation analysis shows that reference labs are using the general LOINC® code '31208-2: Specimen source [Identifier] of Unspecified specimen' when the source of the specimen is missing.

Use of nonspecific general LOINC® codes for pathogen confirmation methods

The labs are using nonspecific general LOINC® codes for determination or confirmation of lab tests for pathogen identification and confirmation. Examples of these LOINC codes are:

1) LOINC code 35347-4 is used for Microscopic observation in unspecified specimen.
2) LOINC code 31208-2 is used for Bacteria identified in unspecified specimen.
3) LOINC code 23667-9 is used for Id by probe for Neisseria identification.
4) LOINC code 21070-8, is used by Quest for MIC SENSITIVITY (BREAK POINT) for Neisseria. Minimum inhibitory concentration (MIC) are used for antibiotic sensitivity that is performed according to specimen source and organism submitted, but Quest has used it for Neisseria.

5) LOINC code 20423-0 is used for identification of an organism using beta lactamase test as a secondary confirmation test, Quest has used the nonspecific secondary confirmation test as a confirmatory test for both Neisseria identification and Salmonella identification. For Neisseria gonorrhea the confirmatory test is to identify the pathogen using a bacterial culture method. Beta lactamase is just one of the methods that labs are using as a secondary method of identification of a Neisseria gonorrhea colony that grew on the culture. The beta lactamase test is not considered confirmatory test for any organism.

6) LOINC Code 45187-2 is used for antibiotic susceptibility testing and the method used is Kirby bauer (disk diffusion), which is a general LOINC code that can be used for any organism antibiotic susceptibility testing. But Quest has specified using it for Neisseria.

7) LOINC code 55617-5 is used for antibiotic susceptibility testing and the method used is Gradient strip (E-test), which is a general LOINC code that can be used for any organism antibiotic susceptibility testing. But Quest has specified using it for Neisseria.

Missing relevant LOINC® codes:

Several LOINC codes, for example for Salmonella and Bordetella parapertussis, simply seem to be missing in RCTC and should be vetted and considered for addition by SMEs.

Table 3:

The list of missing relevant LOINC codes

Concept_Code	Data_Source	Local Concept_Name for Lab	Condition_Name
17562-0	Arup	Salmonella typhi/paratyphi Abs Interp	Salmonella
29723-4	LabCorp	Bordetella parapertussis DNA	Pertussis
42588-4	Mayo	Bordetella pertussis and Bordetella parapertussis: Molecular Detection: PCR	Pertussis
13284-5	Mayo	Salmonella H: Type d	Salmonella
13285-2	Mayo	Salmonella O: Type D	Salmonella
22517-7	Mayo	Salmonella H: Type a	Salmonella
22521-9	Mayo	Salmonella H: Type b	Salmonella
57769-2	Mayo	Salmonella O: Type Vi	Salmonella
13916-2	Quest	Bordetella parapertussis Ag	Pertussis
29723-4	Quest	Bordetella parapertussis DNA	Pertussis
13284-5	Quest	Salmonella typhi H D Ab	Salmonella
13285-2	Quest	Salmonella typhi O D Ab	Salmonella
22517-7	Quest	Salmonella paratyphi A H Ab	Salmonella
22521-9	Quest	Salmonella paratyphi B H Ab	Salmonella
57769-2	Quest	Salmonella typhi O Vi Ab	Salmonella

The result of workflow analysis among EHR, lab, and outpatient's lab:

The gaps that were observed in use of LOINC codes in workflow that can affect the RCTC vision of automatically initiating an electronic initial case reporting (eICR) are:

1) The healthcare providers are not using national data standards when placing lab orders in EHR.

2) At the time of accessioning, the lab receives the orders placed by physicians, but the lab tests recorded in the LIMS do not employ LOINC® codes.

3) Reporting of lab results in the LIMS are not done using LOINC®/ SNOMED.

4) Reporting of reportable conditions to public health is done manually.

The analysis suggested that the hospitals are not using LOINC® codes when placing lab test orders in the LIMS, instead local codes are being used for this purpose. After the lab tests are performed at the reference labs, the results of the lab tests are sent back to the clinical lab using fax. The lab test results are resulted using clinical terminology, for example use of clinical terminology 'Salmonella species isolated', instead of using SNOMED codes.

The interview findings revealed that hospitals are not required to report outpatient test results to public health since the meaningful use guidelines are limited to in-house patients only per ONC guidelines for Meaningful Use stage 2 for clinical lab.

The ONC guidelines for Meaningful Use stage 2 for clinical lab is, "Certification Requirements for Lab Stage 2 Core Objective & Measure: ELR 170.314(f) (4) standard Electronic Reportable Laboratory Results. Inpatient setting only — transmission of

reportable laboratory tests and values/results. EHR technology must be able to electronically create reportable laboratory tests and values/results for electronic transmission in accordance with HL7 - 170.205(g),SNOMED - 170.207(a)(3), and LOINC 170.207(c)(2) standards."

The ELR data results submission for Stage 2 which continue to apply for the Mod Period (2015-2017), as well as Stage 3. (ONC)

Almost all hospitals have outpatient labs where they receive specimens from ambulatory healthcare providers, but according to the lab objective for meaningful use stage 2, the hospital labs are only required to send reportable condition reports for inpatients lab results only. At present the LOINC and SNOMED coding is not intrinsically required in the lab to EHR processing and is therefore not done, resulting in a process interaction that does not support matching of recorded codes against Trigger Codes in the manner envisioned.

The workflow analysis revealed several gaps in the use of LOINC® codes: not using LOINC codes when placing lab orders, not employing LOINC® codes for scheduling lab tests in the lab, failure to use LONIC® or SNOMED for lab result recording or reporting, and hospitals not having a meaningful use incentive to report outpatient test results to public health.

Chapter 5: Discussion

Introduction

This chapter discusses the importance of validating RCTC content and assessing the EHR implementation process, for feasibility of its applicability towards triggering eCR.

Summary of Study

This study highlights the disparities in the use of triggers codes (LOINC® and SNOMED codes) within EHR, through work flow analysis of processes that takes place amongst healthcare providers, EHR, clinical lab, outpatients (ambulatory healthcare providers and nursing homes) and public health. The study conducted several rounds of interviews with entities such as Atlanta hospital system IT staff, clinical labs and infection prevention staff to understand the processes. Based upon the analysis conducted, this study was able to show that

- The healthcare providers are not using LOINC® codes while ordering the lab tests.
- The clinical labs were not using LOINC® codes when accessioning specimen in LIMS system.
- The outpatient labs were not using LOINC® codes for lab test ordering for ambulatory healthcare providers and nursing homes.

Additionally, the clinical labs do not use SNOMED codes when recording lab test results in the LIMS, causing the SNOMED codes to get assigned for reportable condition only at MURR interface while sending ELR to public health. The eCR relies on both LOINC® and SNOMED codes to be implemented and to trigger at multiple points during the recording of patient encounter information. This means that the trigger codes need to

match what the lab uses and returns to the EHR, and the key interaction points between the EHR and the lab needs to use standard codes. The disparity found in use of standard codes by EHR can potentially affect the RCTC vision of initiating the eCR for reportable conditions.

The second part of the study highlights the patterns observed in discrepancies between LOINC ® codes for triggering electronic case reporting (eCR) using RCTC and LOINC® codes used by national reference labs. To study the patterns of discrepancies the study conducted validation analysis of four piloted reportable LOINC codes collected form 4 national reference labs. Based upon analysis conducted, the study recorded and analyzed the missing LOINC® codes that were not found in RCTC. The study observed the four discrepancies patterns in use of LOINC ® and categorized the discrepancies as; reference labs are using different LOINC® codes for similar tests due to missing information of specimen types, using LOINC® codes for specimen source description instead of SNOMED, using general LOINC ® codes for methods used for pathogen confirmation, and missing LOINC® codes in RCTC that are used by reference labs. The study recommended measurable next steps that can be taken to address the discrepancies. These measurable steps are communication with labs, and additions of missing LOINC® codes to RCTC and RCKMS value sets.

Limitations

The study limitations were that the data was only collected from reference labs, and the process analysis was limited to one Atlanta hospital lab system. The study requested data from public health laboratories for ELR, but was unable to obtain data

within the time constraints of the study. Although the study established the process flow for one hospital lab system, it would have been more beneficial if the data was collected from more hospitals, additional vendors, health care providers of different sizes, and in rural and urban settings. The study was conducted on only four piloted conditions and inclusion of additional conditions could enhance the validation results. Other limitations are; reviewing the discrepancies with other lab experts and vocabulary SMEs, extending data collection to other clinical labs (not just the clinical –lab flow, but the codes they use), exploring local to standard code mapping (SuperScripts, LIC, Intermountain).

Conclusion

While this study was a smaller study, it uncovered coding discrepancies and issues in the process of establishing the value of further assessment. The study analyzed a total of 1080 LOINC codes across 4 reportable conditions collected from 4 reference labs for validation of RCTC LOINC codes. The study only validated LOINC® codes since the LOINC system was used more commonly than SNOMED system in reference labs. The study also found that the hospitals are using SNOMED for reportable conditions, and LOINC® codes are not used by hospital labs. Both, the reference labs and Atlanta hospital system have differences in use of choice of data standards. The reason for the differences may be due to nature of work. The reference labs are more geared towards conducting laboratory testing and are not directly involved in reporting results to public health but instead send the test results directly to the ordering authorities. In contrast, the hospital labs are involved with both conducting lab tests and reporting of lab test results to public health. The differences in choice of data standards for laboratory testing ordering and reporting may inhibit effective analysis of validation of RCTC codes.

Increased use of completeness of coded data by EHR, clinical, lab, reference labs are needed for further validation of RCTC and eCR project success.

Implications and Recommendations

Public health officials have identified the need for advancing electronic reporting of reportable conditions for years through various initiatives. Utilizing existing systems of electronic reporting and adding additional layers of analysis can enhance electronic reporting efforts. The additional recommended layer is, to include the process of evaluation of the completeness of LOINC® trigger codes in the RCTC for Jurisdiction reportable conditions (Chlamydia, Pertussis, Salmonellosis, and Gonorrhea) by comparing them against codes currently in use by reference labs, public health labs, and clinical labs.

The eCR initiative is an ongoing project and the RCTC volume will go up gradually as each reportable condition is added. The recommendation for future efforts would be to incorporate the evaluation and validation of all trigger codes (listed in RCTC guidelines) for reportable conditions. Future projects should include public health labs, and more clinical labs in addition to the national reference laboratories included in this study.

The validation process is not a one-time effort, it should be an ongoing process, and since the use of LOINC® codes will continue to increase, it means each institution needs a process to maintain their local LOINC® mappings. The reference laboratories often conduct rare tests that hospital clinical labs do not conduct so there will always be new LOINC® codes to be included. For example now there is a new test for Zika virus

that might have new LOINC® codes assigned along with SNOMED-CT and ICD-10 to it.

The recommendation is to keep the RCTC and eCR flexible and up-to-date to accommodate the LOINC® codes used for upcoming new lab tests. Since the study could not validate the public health lab and clinical lab trigger codes, extending this research would be beneficial for to examine the completeness of trigger codes.

The study recommendation from a process standpoint is that the LOINC® codes should be an integral part of EHR - Computerized physician order entry (CPOE) portal to facilitate the triggering of the first case encounter. Other recommendations would be to map local codes used by clinical laboratories for lab results reporting to SNOMED codes to facilitate trigger coding at the EHR.

The study was conducted on smaller dataset (1080 records) for 4 piloted reportable conditions, and the study was still able to find significant findings in discrepancies between LOINC® codes used within 4 national reference labs and clinical lab. The study implies that the eCR project should seriously look at extending the study and ensure the success of reportable condition trigger codes for initiating electronic case reporting.

Appendix 1

SQL queries used in study

SQL Code

```
**Sets the current working directory
setwd("C:/Big Data/R Sample")

**Install the sqldf package if not already preinstalled
if("sqldf" %in% rownames(installed.packages())==FALSE){install.packages("sqldf")}
library(sqldf)/

** Create 4 condition based dataframes by reading in the csv files
df_Sal <- read.csv('Salmonella.csv',header=TRUE)
df_Nei <- read.csv('Neisseria.csv',header=TRUE)
df_Per <- read.csv('Pertussis.csv',header=TRUE)
df_Chl <- read.csv('Chlamydia.csv',header=TRUE)

**Sort each of the 4 dataframes to get the unique Concept codes
sort(unique(df_Sal$Concept_Code))
sort(unique(df_Nei$Concept_Code))
sort(unique(df_Per$Concept_Code))
sort(unique(df_Chl$Concept_Code))

**Sort each of the 4 dataframes to get the unique Data Sources
sort(unique(df_Sal$Data_Source))
sort(unique(df_Nei$Data_Source))
sort(unique(df_Per$Data_Source))
sort(unique(df_Chl$Data_Source))
```

**Create dataframe for each of the Data Sources (4 national labs and RCTC) to include Concept code, frequency, Concept Name and Condition Name by combing data from each of the 4 condition based dataframes created initially.

```
df_Arup <- sqldf("SELECT Concept_Code, COUNT(*) AS Frequency, 'Arup' as Data_Source,
        Concept_Name,'Salmonella' AS Condition_Name
        FROM df_Sal
        WHERE Data_Source = 'Arup'
        GROUP BY Concept_Code
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'Arup' as Data_Source,
        Concept_Name,'Neisseria' AS Condition_Name
        FROM df_Nei
        WHERE Data_Source = 'Arup'
        GROUP BY Concept_Code
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'Arup' as Data_Source,
        Concept_Name,'Pertussis' AS Condition_Name
        FROM df_Per
        WHERE Data_Source = 'Arup'
        GROUP BY Concept_Code
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'Arup' as Data_Source,
        Concept_Name,'Chlamydia' AS Condition_Name
        FROM df_Chl
        WHERE Data_Source = 'Arup'
        GROUP BY Concept_Code")


df_Labcorp <- sqldf("SELECT Concept_Code, COUNT(*) AS Frequency, 'Labcorp' as
Data_Source,
```

```
            Concept_Name,'Salmonella' AS Condition_Name
            FROM df_Sal
            WHERE Data_Source = 'Labcorp'
            GROUP BY Concept_Code
            UNION
            SELECT Concept_Code, COUNT(*) AS Frequency, 'Labcorp' as
Data_Source,
            Concept_Name,'Neisseria' AS Condition_Name
            FROM df_Nei
            WHERE Data_Source = 'Labcorp'
            GROUP BY Concept_Code
            UNION
            SELECT Concept_Code, COUNT(*) AS Frequency, 'Labcorp' as
Data_Source,
            Concept_Name,'Pertussis' AS Condition_Name
            FROM df_Per
            WHERE Data_Source = 'Labcorp'
            GROUP BY Concept_Code
            UNION
            SELECT Concept_Code, COUNT(*) AS Frequency, 'Labcorp' as
Data_Source,
            Concept_Name,'Chlamydia' AS Condition_Name
            FROM df_Chl
            WHERE Data_Source = 'Labcorp'
            GROUP BY Concept_Code")


df_Mayo <- sqldf("SELECT Concept_Code, COUNT(*) AS Frequency, 'Mayo' as
Data_Source,
          Concept_Name,'Salmonella' AS Condition_Name
          FROM df_Sal
```

```
WHERE Data_Source = 'Mayo'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'Mayo' as Data_Source,
Concept_Name,'Neisseria' AS Condition_Name
FROM df_Nei
WHERE Data_Source = 'Mayo'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'Mayo' as Data_Source,
Concept_Name,'Pertussis' AS Condition_Name
FROM df_Per
WHERE Data_Source = 'Mayo'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'Mayo' as Data_Source,
Concept_Name,'Chlamydia' AS Condition_Name
FROM df_Chl
WHERE Data_Source = 'Mayo'
GROUP BY Concept_Code")

df_Quest <- sqldf("SELECT Concept_Code, COUNT(*) AS Frequency, 'Quest' as
Data_Source,
Concept_Name,'Salmonella' AS Condition_Name
FROM df_Sal
WHERE Data_Source = 'Quest'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'Quest' as Data_Source,
```

```
Concept_Name,'Neisseria' AS Condition_Name
FROM df_Nei
WHERE Data_Source = 'Quest'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'Quest' as Data_Source,
Concept_Name,'Pertussis' AS Condition_Name
FROM df_Per
WHERE Data_Source = 'Quest'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'Quest' as Data_Source,
Concept_Name,'Chlamydia' AS Condition_Name
FROM df_Chl
WHERE Data_Source = 'Quest'
GROUP BY Concept_Code")

df_RCMT <- sqldf("SELECT Concept_Code, COUNT(*) AS Frequency, 'RCMT' as
Data_Source,
Concept_Name,'Salmonella' AS Condition_Name
FROM df_Sal
WHERE Data_Source = 'RCMT'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'RCMT' as Data_Source,
Concept_Name,'Neisseria' AS Condition_Name
FROM df_Nei
WHERE Data_Source = 'RCMT'
GROUP BY Concept_Code
```

```
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'RCMT' as Data_Source,
        Concept_Name,'Pertussis' AS Condition_Name
        FROM df_Per
        WHERE Data_Source = 'RCMT'
        GROUP BY Concept_Code
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'RCMT' as Data_Source,
        Concept_Name,'Chlamydia' AS Condition_Name
        FROM df_Chl
        WHERE Data_Source = 'RCMT'
        GROUP BY Concept_Code")

df_RCTC <- sqldf("SELECT Concept_Code, COUNT(*) AS Frequency, 'RCTC' as
Data_Source,
        Concept_Name,'Salmonella' AS Condition_Name
        FROM df_Sal
        WHERE Data_Source = 'RCTC'
        GROUP BY Concept_Code
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'RCTC' as Data_Source,
        Concept_Name,'Neisseria' AS Condition_Name
        FROM df_Nei
        WHERE Data_Source = 'RCTC'
        GROUP BY Concept_Code
        UNION
        SELECT Concept_Code, COUNT(*) AS Frequency, 'RCTC' as Data_Source,
        Concept_Name,'Pertussis' AS Condition_Name
        FROM df_Per
```

```
WHERE Data_Source = 'RCTC'
GROUP BY Concept_Code
UNION
SELECT Concept_Code, COUNT(*) AS Frequency, 'RCTC' as Data_Source,
Concept_Name,'Chlamydia' AS Condition_Name
FROM df_Chl
WHERE Data_Source = 'RCTC'
GROUP BY Concept_Code'')
```

**Create dataframe for the missing RCTC codes by comparing the Concept Code between RCTC dataframe and the dataframe for each of the 4 national labs created above.

```
df_RCTC_Missing <- sqldf("SELECT * FROM df_Arup
        WHERE Concept_Code NOT IN
            (SELECT Concept_Code FROM df_RCTC)
            UNION
            SELECT * FROM df_Labcorp
            WHERE Concept_Code NOT IN
            (SELECT Concept_Code FROM df_RCTC)
            UNION
            SELECT * FROM df_Quest
            WHERE Concept_Code NOT IN
            (SELECT Concept_Code FROM df_RCTC)
            UNION
            SELECT * FROM df_Mayo
            WHERE Concept_Code NOT IN
            (SELECT Concept_Code FROM df_RCTC)
            UNION
```

```
SELECT * FROM df_RCMT
WHERE Concept_Code NOT IN
(SELECT Concept_Code FROM df_RCTC)
ORDER BY Data_Source, Condition_Name")
```

**Write the RCTC Missing dataframe values into a csv format file.

```
write.csv (df_RCTC_Missing, file = "RCTC_Missing-0706.csv", row.names=FALSE)
```

References:

ARUP laboratories LOINC codes, Retrieved from: http://www.aruplab.com/testing/cpt-loinc-codes

Deepthi Rajeev, Catherine J Staes, R Scott Evans, Susan Mottice, Robert Rolfs, Matthew H Samore, Jon Whitney, Richard Kurzban, Stanley M Huff J Am Med Inform Assoc. 2010 Jan-Feb; 17(1): 34–41. doi: 10.1197/jamia.M3299

Dhakal, S., Burrer, S. L., Winston, C. A., Dey, A., Ajani, U., & Groseclose, S. L. (2015). Coding of Electronic Laboratory Reports for Biosurveillance, Selected United States Hospitals, 2011. Online J Public Health Inform, 7(2), e220. doi:10.5210/ojphi.v7i2.5859

Dixon, B. E., Vreeman, D. J., & Grannis, S. J. (2014). The long road to semantic interoperability in support of public health: experiences from two states. J Biomed Inform, 49, 3-8. doi:10.1016/j.jbi.2014.03.011

Electronic Case Reporting: eICR and Trigger Implementation Discussion, May 17th, 2016, CDC, Retrieved from: https://www.cdc.gov/ehrmeaningfuluse/docs/vendors_collaboration_initiative_webinars/2016-05-17-vendorcall_eicr_and_trigger_codes_final-508.pdf

LabCorp laboratories LOINC codes, Retrieved from: https://www.labcorp.com/wps/portal/!ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os_hACzO_QCM_IwMLXyM3AyNjMycDU2dXQwN3M6B8JG55AwMCuv088nNT9QtyI8oBPK_DWA!!/dl2/d1/L0lDU0NTQ1FvS1VRIS9JSFJBQUlnb0FNeUtibTZtL1lCSkp3NDU0a3N1eWx3ISEvN19VRTRTMUk5MzBPR1MyMElTM080TjJONjY4MC92aWV3SXRlbQ!!/?itemId=3832482

Mayo laboratories LOINC codes, Retrieved from: http://www.mayomedicallaboratories.com/test-catalog/appendix/loinc-codes.html

ONC, CMS's regulatory requirements for meaningful use, Retrieved from: ttps://www.healthit.gov/sites/default/files/meaningfulusetablesseries2_110112.pdf

Quest laboratories LOINC codes, Retrieved from: http://www.questdiagnostics.com/home/physicians/test-center/loinc-codes.html

RCMT, PHINVADS, Retrieved from: http://www.cdc.gov/ehrmeaningfuluse/rcmt.html

RCTC, PHINVADS, Retrieved from: https://phinvads.cdc.gov/vads/SearchVocab.action