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:

_____ _____

J. Patrick Caneer Date

Innovation and Application of Data Flow Diagram to Inform Stakeholders, Validate Process, Align Operations, and Inform Systems Architecture for the Child Health and Mortality Prevention Surveillance (CHAMPS) Network

By

J. Patrick Caneer

Master of Public Health

Emory University, Rollins School of Public Health

Applied Public Health Informatics

_____

J. Mark Conde
Committee Chair

_____

Timothy Morris
Committee Member

_____

Navit Salzberg, MPH
Committee Member

Innovation and Application of Data Flow Diagram to Inform Stakeholders, Validate Process, Align Operations, and Inform Systems Architecture for the Child Health and Mortality Prevention Surveillance (CHAMPS) Network

By

J. Patrick Caneer

Bachelor of Science
University of Alabama
1997

Thesis Committee Chair: J. Mark Conde

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 Applied Public Health Informatics
2020

Abstract

Innovation and Application of Data Flow Diagram to Inform Stakeholders, Validate Process, Align Operations, and Inform Systems Architecture for the Child Health and Mortality Prevention Surveillance (CHAMPS) Network

By J. Patrick Caneer

Background: The development and implementation of a technical systems architecture for a complex global health childhood mortality surveillance initiative in Sub-Saharan Africa and South Asia lacked a broadly understandable and comprehensive form of technical documentation to communicate the complexity, interdependency, timing, and orchestration of demographic information, specimen collection, clinical and laboratory data, pathology results, photographs, verbal autopsy data, and other case-specific information to stakeholders, program administrators, field workers, clinical staff, laboratorians, epidemiologists, public health professionals, statisticians and information technology professionals.

Key Aims and Methods: The purpose of this thesis project was to develop a new form of informatics diagram to capture the broader surveillance logistics of the network by incorporating processes, study artifacts, standard operating procedures, and data exchange. The key aims of the project were to develop a:

(1) Methodology for the development and refinement of novel data flow diagram for the global case-based public health surveillance program
(2) New type of technical documentation artifact for the global case-based public health surveillance program that could be generalizable for other similar public health surveillance programs

Results: A refined version of the novel data flow diagram artifact and a method for development was established. The resulting technical document met the overall objective of providing a comprehensive depiction of the surveillance data flow that is informative, instructive, and universally engaging. The artifact had the appropriate balance of communicating both a technical and non-technical system workflow consumable by a multitude of international stakeholders.

Conclusions: This project-based thesis presents an innovative technique of technical documentation developed in service of a complex global public health surveillance network that may offer a more comprehensive, versatile, and intelligible illustration technique to capture the flow of surveillance data, potentially disrupting conventional techniques within the industry. The data flow diagram was intentionally designed and honed to communicate the appropriate ratio of operational/functional workflow specificity such that the document could inform the broadest audience of surveillance stakeholders, at the appropriate level magnification. If refined and more broadly adopted, this data flow development methodology and the resulting artifacts could change how public health surveillance networks are designed, communicated, and comprehended.

Innovation and Application of Data Flow Diagram to Inform Stakeholders, Validate Process, Align Operations, and Inform Systems Architecture for the Child Health and Mortality Prevention Surveillance (CHAMPS) Network

By

J. Patrick Caneer

Bachelor of Science
University of Alabama
1997

Thesis Committee Chair: J. Mark Conde

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 Applied Public Health Informatics
2020

Acknowledgements

"A man who carries a cat by the tail learns something he can learn in no other way"
- Mark Twain

My journey in public informatics started with a toe in the water as a non-degree seeking enrollee in an introductory class about applied public informatics. There an enthusiastic evangelist of systems thinking had my attention and later my commitment to enroll and go for it. I did not know what the next five years would have in store for me…

I did not know that going for it would mean that two and half years of classes would need three and half years so that I could maintain my full-time job and try to manage being a husband and father. I did not know that I would switch careers into global heath before I finished my degree and need put things on hold to help craft and implement an international multi-site surveillance network.

There so many people from so many places that I am grateful to have met, learned from, and been inspired by during my public health quest. I would like to thank the faculty and staff at the Rollins School of Public Health, my classmates (across few cohorts), and countless numbers people I have worked with through the CHAMPS program, a project like no other.

I especially want to thank my committee members, Mark Conde, Tim Morris, and Navit Salzberg for their support, guidance, and encouragement. Your influence in my achievement of this milestone are so much more than this paper.

The most important acknowledgement of all goes to my wife and sons who supported and encouraged me to persevere. I think we did it. I promise I will be a lot more fun going forward…

Table of Contents

Chapter 1 Introduction

1.1. Introduction and Rationale

Public health surveillance is defined by the Centers for Disease Control as "the ongoing systematic collection, analysis, and interpretation of health-related data essential to the planning, implementation, and evaluation of Public Health practice, closely integrated with the timely dissemination of these data to those who need to know. The final link in the surveillance chain is the application of those data to prevention and control" [17]. Collection methods of public health surveillance data continue to evolve [2] with the "technology" of the times with an ever-increasing dependence on electronic data capture, transfer, storage, analysis, dissemination, and representation (e.g., visualization).

The relatively young interdisciplinary field of public health informatics continues to play an essential role in bridging the information gap between the surveilled and those looking to investigate, evaluate, monitor, and impart public health interventions[2,6,7,15]. When describing the application of public health informatics (by informaticians) in the field of surveillance, Krishnamurthy and St. Louis state that "Informaticians use disciplines such as information science, computer science, communications theory, psychology, neuroscience, and systems engineering to understand and address the information requirements of an organization" [7].

The Collaborative Requirements Development Methodology (CRDM) [11], developed by the Public Health Informatics Institute (PHII), is a commonly utilized informatics methodology applied in the field of public health informatics in recent decades. The

CRDM methodology relies heavily on a classic operational or systems engineering approach whereby subject matter experts and/or participants within a business process are engaged in a detailed workflow or task flow analysis. This approach begins by first considering the general context of the business (or public health operation). Next, the current state of how the work is being done is elicited from stakeholders and documented in a series of task flow diagrams that illustrate the primary tasks and decision points performed when completing a specific task. Once agreement is reached on the current state, a second phase of the methodology calls for a critical evaluation of the current state to identify inefficiencies or opportunities to improve the current state workflows resulting in a new enhanced future state. The final phase of the CRDM is to translate the future state workflows into a comprehensive body of functional system requirements. The functional system requirements in conjunction with the future state workflows serve as the reference documentation to perform a system evaluation for a technology system(s) that may best serve the newly designed future state of the business process.

In early 2015, the Bill & Melinda Gates Foundation announced their commitment of an initial $75-million-dollar investment in the Child Health and Mortality Prevention Surveillance Network (CHAMPS), granted to the Emory University Global Health Institute (EGHI) [5]. The Public Health Informatics Institute (PHII), a division of the Task Force for Global Health, was one of the foundational partners working in association with EGHI to engineer, implement, and operate the CHAMPS network. The CHAMPS project was an ambitious and broad surveillance project intended to bring together multiple complicated surveillance streams (or disciplines), including verbal autopsy, child clinical medical record abstraction, maternal clinical record abstraction, postmortem minimally invasive tissue sampling (MITS) and specimen collection, and

laboratory testing results, into a comprehensive case file of evidence to aid interdisciplinary medical and public health experts in the determination of a definitive cause of death for children under the age of 5 in multiple countries of sub-Saharan Africa and South Asia[14]. While pilot programs had been instituted in a number of locations to perform some of the individual components of the surveillance goals of CHAMPS [4, 8,9], there had yet to be a single program or initiative that had attempted to bring all such data elements and streams together to come to a definitive cause of death for individual cases, nor a singular effort attempted at such a geographically significant scale with the potential of influencing health outcomes for such a sizable region of the globe. Lastly, no such program had ever attempted to bring this volume and type of data from multiple low-resource countries into a single standardized comprehensive surveillance data repository.

In the summer of 2015, PHII informaticians initiated the CRDM with CHAMPS stakeholders to elicit the fundamental business processes and functional requirements necessary to meet the informatic needs of the surveillance network. However, application of the CRDM for this initiative faced a significant challenge; the CHAMPS surveillance network was not a pre-existing "business process", rather it was a vision of a network conceptualized in a grant proposal. Therefore, there was not a current state, nor a group of subject matter experts experienced with or within the current state of CHAMPS for informaticians to reference and engage during the CRDM. This lack of a current state was further confounded during the tactical planning of the project due to the silos of subject matter expertise among those responsible for the respective components of surveillance data sources required (e.g., verbal autopsy, histopathology, laboratory, etc.). A lack of common and cohesive understanding of the where, when, and how these components fit within the whole system complicated the ability to achieve

consensus vision for enterprise-level data flow and subsequent technical design. In this context, the CRDM was a valuable method for capturing hyper-focused topics with small groups of related subject matter experts that could provide insight into specific components of the larger network. However, the volume of individual hyper-focused business process workflows distracted from the ability to identify critical pathway(s) for achievement of the overall surveillance objectives. Informaticians, as the common denominator of each of the surveillance streams, were responsible for developing of all aspects of the network and needed to devise a new approach to adjust the project perspective to the appropriate level of magnification.

This method-based thesis will describe the techniques developed, tested, and applied by informaticians in the field that were iteratively refined resulting in what is now referred to as a data flow diagram. The data flow diagram was able to bring together a number of critical concepts and components at a high enough level of context to provide an overall picture of the surveillance data collection and workflows while at the same time providing sufficient level of specificity regarding information exchanges, critical workflows, essential system interactions, case reporting forms, and standard operating procedures. This innovation in technical documentation and method facilitated the communication of complexity, interdependency, timing, and orchestration of the surveillance deliverables to the stakeholders, program administrators, field workers, clinical staff, laboratorians, epidemiologists, public health professionals, statisticians and information technology professionals involved in the development and operation of the surveillance network.

This paper is an in-depth description of the specific aims, deliverables, and informatics challenges encountered by informaticians and the CHAMPS implementation partners in

2015-2017 as they began developing the CHAMPS surveillance protocol, architecture and design of systems, and strategies for implementation in low-resource countries. Through an overview of CRDM, this paper will explore how this fundamental informatics methodology was applied at the onset of the project to attempt to inform potential system needs. Finally, this paper will address how the CRDM was not fully meeting the needs of the programmatic and informatics stakeholders to reach a clear consensus on the surveillance network requirements, thus creating a need for a more evolved approach to meet the complex challenges of CHAMPS Network design.

1.2 Review of the literature

1.2.1 Overview of Collaborative Requirements Development Methodology

The CRDM, developed by the Public Health Informatics Institute (PHII), is a commonly utilized informatics methodology applied in the field of public health informatics in recent decades. The CRDM methodology relies heavily on a classic operational or systems engineering approach whereby subject matter experts and/or participants within a business process are engaged in a detailed workflow or task flow analysis. This approach begins by first considering the general context of the business (or public health operation). Next, the way that work is being done, or the current state, "is documented in a series of task flow diagrams which elicit the primary tasks and decision points engaged in completing a specific task. Once agreement is reached on the current state, a second phase of the methodology calls for a critical evaluation of the current state to seek inefficiencies or opportunities to improve the current state task flows resulting in a new enhanced future state. The final phase of the CRDM is to translate the future state workflows into a comprehensive body of functional system requirements. The functional

system requirements in conjunction with the future state workflows serve as the reference documentation to perform a system evaluation for a technology system(s) that may best serve the newly designed future state of the business process.

1.2.2 Example Documentation Artifacts of the CRDM

1.2.2.1 Business Process Matrix

A business process matrix is an organized listing of related business processes offering details of each goal, objective, rule, and trigger in a matrix format. An example of a business process matrix is provided in Figure 1.

Business Process	Goal	Objective	Business Rules	Trigger(s)
	The major health goal that the business process supports. The goal is the end state to be achieved by the work of the health agency and should be defined in terms of the benefits provided to the community/ population or individual/client.	A concrete statement describing what the business process seeks to achieve. The objective should be specific to the process such that one can evaluate the process or reengineer the process and quantify performance measures. A well-worded objective will be SMART (Specific, Measurable, Attainable/Achievable, Realistic and Time-bound).	A set of criteria that defines or constrains some aspect of the business process. Business rules are intended to assert business structure or to control or influence the behavior of the health agency (business).	Event, action, or state that initiates the first course of action in a business process. A trigger may also be an input, but not necessarily so.
Billing and Accounts Receivable	Assurance that the fiscal process supports the strategic goals of public health and complies with all legal and policy requirements.	Process accounts receivable in a timely, efficient and accurate manner to assure cash flow, compliance with legal requirements, and alignment with budgeted public health activities.	GAAP; OMB Circ A-133 (Single Audit); U.S. Government Auditing Standards; City/county charter/law; State law.	Services provided to customer.
Communicable Disease and Clinical Intervention & Treatment	Early identification, treatment and resolution of health condition; Promotion and protection of population's health.	Complete accurate and timely screening, diagnostic and treatment processes; Assure individual compliance with and completion of recommended course of treatment.	Public and private medical providers; State and local health departments; Diagnostic labs and other entities; Hospitals/acute care system; Standardized and best practice communicable disease control.	Client presents with symptoms or risk factors; Periodicity schedule for screening; Referral for services; Alert (population risk); Lab report; Surveillance reports; Clinical diagnosis.
Community Health Assessment	Assessment of the health status and needs of the community.	Compile and analyze data as requested by stakeholders.	Statistical methods; Data collection protocols; Reporting guidelines.	Public Health requirements to track core set of indicators; Identification of data gaps by stakeholders; Continued health surveillance; Health emergency identified by Public Health staff or others.

Figure 1 Example of a Business Process Matrix [12]

1.2.2.2 Context Diagram

A context diagram provides a high-level visualization of various entities involved in a business process or series of business processes (as outlined in a business process matrix) and each entity is connected by lines that document the exchange of information or deliverables among the entities. Figure 2 provides an example context diagram for placing a food order.

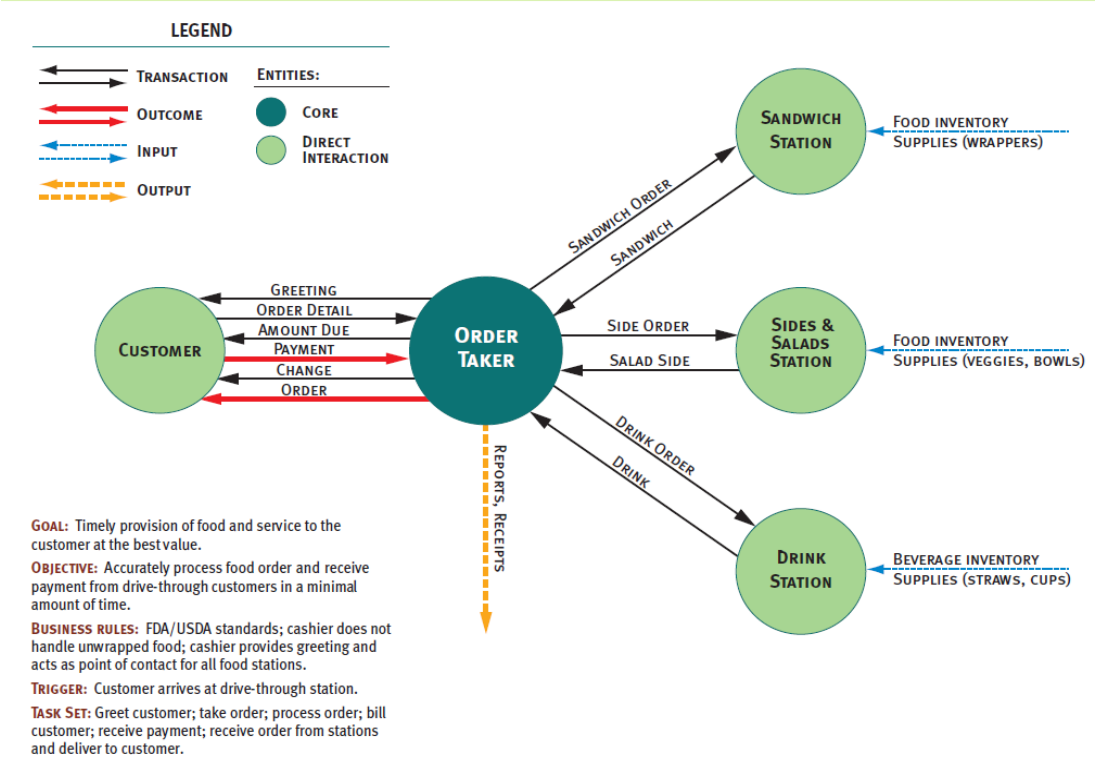


Figure 2 Example of a Context Diagram [12]

1.2.2.3 Task flow (or workflow) diagram

A workflow diagram depicts a business processes that support a clear objective with defined beginning and ending (or endpoints). Key features are event triggers that indicate what event or issue must occur to start the business process, process steps, and critical decision points within the process. Figure 3 is an early rendition of the CRDM workflow model.

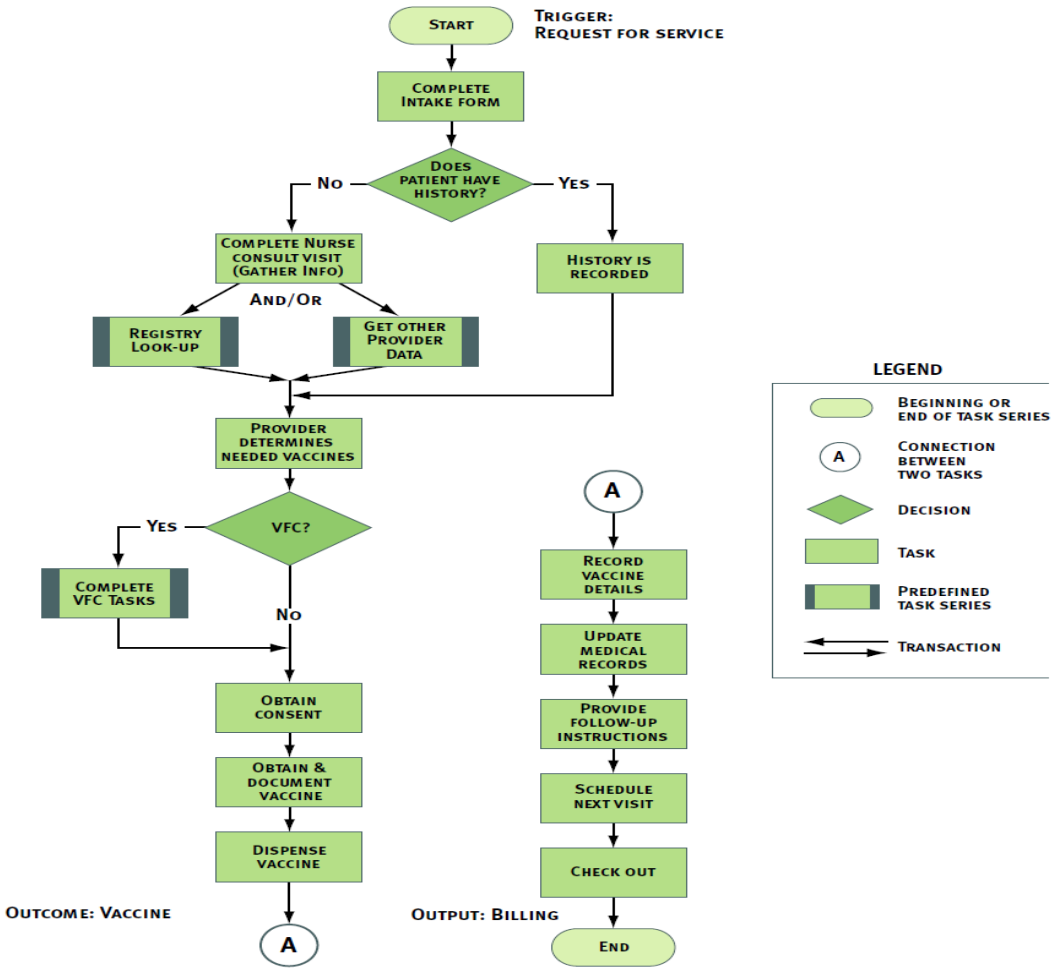


Figure 3 Example of a Task flow diagram [12]

Figure 4 is a more recent evolution of the workflow that includes swim lanes to indicate transference of steps/decisions among roles and enumerated descriptions of the workflow process steps and objectives.

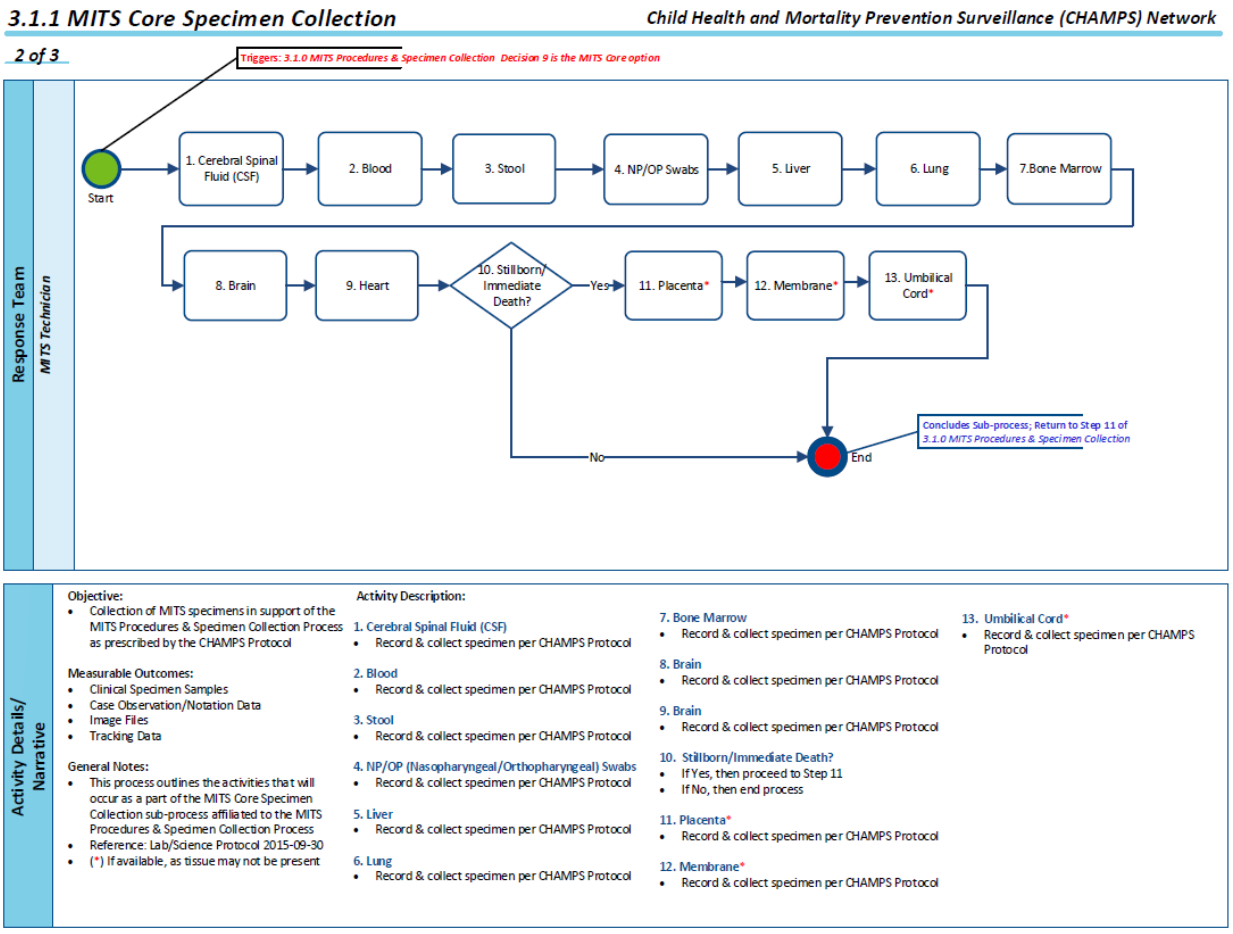


Figure 4 Example of a CHAMPS Workflow diagram [10]

1.2.2.4 System/Functional Requirements

A set of system/functional requirements are a form of documentation whereby a listing of enumerated functional requirements is recorded and organized. These functional requirements may be one or more descriptive statements that help define the functional requirement. Figure 5 and Figure 4 are simple examples of system (2.5) and functional (2.7) requirements are provided below.

2.5	PROVIDE CROSS-CHECKING OF SCHEDULES.
System should be able to provide constraints such that staff hours are not over-allocated. The system should be capable of cross-checking availability on all staff covering multiple coded departments.	

2.7	PROVIDE A USER-FRIENDLY METHOD TO ADJUST CUSTOMIZED FIELDS.
The system should be a common interface that can be customized to make it more appropriate to a program or department. The scheduling parameters should be flexible. View of schedule information should be configured for appropriate departmental staff. Each individual should be able to configure the interface, within limits, to import, view, or update information to their needs. The system should, where applicable, assist the users in updating scheduling information by presenting the users with suitable prompts. The system should, where applicable, permit established limits on certain fields.	

Figure 5 Example of a System and Functional Requirements [12]

1.2.3 Role of CRDM in Public Health Informatics

The CRDM has been an influential methodology within the field of public health informatics since its development and applied practice in 2002. A number of publications within the field cite the application of the methodology [18,19,20,21,22,23]. Moreover, the Public Health Informatics Institute (PHII) continues the application of this approach across several projects, initiatives, and programs currently. Conventional wisdom in public health informatics is, "If you have seen one public health system, then you have seen one public health system." The CRDM sought to evaluate and share findings regarding workflow and system requirements as a way to show that while public health agencies are different, they share common needs, deliverables, high-level workflows, and therefore, have similar system requirements [12]. Over time, a catalogue of industry white-papers, guidance documentation, facilitated collaborative sessions/conferences, blogs, and publications have been developed in association with PHII, contributing to the broader knowledge base of public health informatics.

While these contributions have been informative, offering general guidance and insight into a number of the facets of public health informatics, the information provided is not specific enough for actual technical development of software systems or surveillance networks; they provided high-level context, general workflows, and generic system/functional requirements. Technical architecture and system developers require different contextual perspectives and data-focused detail to produce actual functioning systems to support the public health system/surveillance needs. In other words, the documentation resulting from the CRDM provides an initial context of operations and

some understanding of business activities, but is incomplete relative to system design, technical development requirements, system implementation guidance, and fails to provide specifications regarding data elements, a significant deficiency for data-driven surveillance systems.

1.2.4 Literature specific to critiques of the CRDM and technical diagrams of public health surveillance networks.

Looking to build off the work the CRDM, it was reasonable to explore for scholarly articles critical of the methodology and its prescribed artifacts so that these critiques could be evaluated and considered. A thorough literature review did not yield scholarly articles or reputable publications that were specifically critical of the CRDM nor were there materials expressing weakness or challenges with CRDM as a method of public health system requirements development.

1.2.5 Literature Regarding Alternative forms of Informatics Diagrams

A review of alternative approaches for informatic technical diagrams was investigated. A thorough literature review did not yield scholarly articles or reputable publications that were specifically regarding prescribed methodologies for technical diagrams of public health surveillance systems. It was feasible to find some reference to the need for and the benefit of data flow, workflow, or information flow diagrams. Krishnamurthy and St. Louis pointed out that, “By process of abstraction, flow diagrams should provide a simplified yet rich view of the critical transactions that occur from the point of original capture of data all the way through the final dissemination of findings for public health

action”[7]. While this source did offer small-scale diagram examples, it did not expound on the methods or techniques for development.

It was clear from efforts to find referential articles or potentially alternative methods to direct the development of “a simplified yet rich view” [7] of the CHAMPS surveillance project that the requirements of this thesis project and the resulting method/artifacts were born out of the need for something truly novel.

1.3 Problem Statement

1.3.1 Framing the problem in the context of CHAMPS

In the fall of 2015, as the CHAMPS program was starting up and hiring staff, the funder required, as a grant payment contingent milestone of CHAMPS program, that a formal and comprehensive information technology plan to support the grant surveillance proposal be authored and submitted to the funder for review and evaluation by December of 2015 [10]. In addition to the expected CRDM workflows, the foundation recommended development of supplemental user personas and use cases that correlated to some of the workflows and add some potential user perspectives to the documentation. Examples below highlight the related user persona (Figure 6), use case (Figure 7), and workflow diagram (Figure8) for the CHAMPS *Expert Panel to Determine the Cause of Death*.

With consent of families, the CHAMPS project curated demographic information, autopsy procedure data and specimens, laboratory results data, histopathological findings, clinical data, verbal autopsy data, and other evidence regarding the deaths of

children under five years of age into a comprehensive case file [14]. These case files were submitted for review to a panel of public health experts (the *CHAMPS Expert Panel to Determine the Cause of Death*) to derive the immediate and underlying cause(s) of death for each case [3,14]. The informatics team developed three artifacts within the technical documentation [10] related the *CHAMPS Expert Panel to Determine the Cause of Death*. First, a user persona regarding one of the panel members was created to articulate details regarding the role and anticipated user needs (Figure 6). Then, the panel member persona was incorporated into a use case diagram to provide user context amongst the technical system(s) and introduce some data/information flow (Figure 7). Finally, the persona and use case were linked to the business process workflow diagram to illustrate the specific procedures and decision points that an expert panel member was expected to encounter (Figure 8).

1.2.10 DeCoDe Expert Panel Member

Persona	Dr. Clinician	**Persona ID:**	KP_10
Role	Cause of Death Panel Member	**Date:**	2015-10-12
Role Variation(s)	n/a		
User Group	Final Cause of Death Panel		
CHAMPS Data Profile	Data Entry: X ● ○ ○ ○ ○	Field Responsibilities: X ○ ○ ○ ○ ○	
	Data QA/QC: X ● ○ ○ ○ ○	Operational Responsibilities: X ● ○ ○ ○ ○	
	Data Consumption: X ● ● ● ● ●	Data Decision Making: X ● ● ● ● ●	
Data Family (3 of 6)	Surveillance Data, Specimen Collection Process Data, Clinical Results and Diagnostic Report Data		
Narrative	Dr. Clinician is an accomplished physician and respected global health researcher who specializes in neglected tropical diseases. She serves the CHAMPS Program as an expert panel member on the Final Cause of Death Panel. Dr. Clinician is expected to independently review detailed case information regarding the death of individual children enrolled in the CHAMPS program to diagnose and grade primary and contributory determinants that lead each child's death. She is also expected to convene with other panel member to review and compare outcomes. Dr. Clinician is eager to serve the CHAMPS mission, but has some concerns about the time commitment.		
Use Case Diagram(s)	UCD_10 Expert Panel to Determine Cause of Death (DeCoDe)		
Taskflow(s)	7.1 MITS Final Cause of Death Panel Review Process		
Objectives	Review CHAMPS case data, both independently and as panel member, to assign final cause of death and grade the primary and contributory determinants		
Actions	• Independently review comprehensive case data and artifacts to diagnose a final cause of death (asynchronously) • Grade/Score primary and contributory determinants (asynchronously and then with panel) • Participate in panel sessions to compare, review, and discuss aggregate panel findings in order to settle on a final cause of death for each case		
Needs	• Simple and convenient virtual method to schedule, access, and review all available case data from the CHAMPS Network (asynchronously) • Access to pertinent epidemiology information, regional surveillance data, and the ability to consult the results of the CHAMPS predictive model(s) and data reliability rankings • Record her observations, questions, and graded findings for future reference during panel meetings, where she may have to adjudicate more than one case file with the panel		
Technology Needs	• Ability to access comprehensive case data (clinical, images, laboratory, pathology, surveillance) (web portal) • Method to select and grade determinants (web form/application) • Individual Assessment - Method to record her observations, questions, and graded findings for future reference during panel meetings, where she may have to adjudicate more than one case file with the panel. (web form/application) • Panel Assessment - Access to a convenient platform to seamlessly communicate with other panel members and simultaneously view (physical/ virtual) all the same relevant case data and CHAMPS resources along with aggregate grading data from each of the panel members. (web application, WebEx, Conference Line) • Method to document and communicate the panel findings to the CHAMPS Program Office (web form/application)		

Figure 6 Example CHAMPS user persona for a member of the *Expert Panel to Determine Cause of Death* (later referred to as the "DeCoDe" panel) [10]

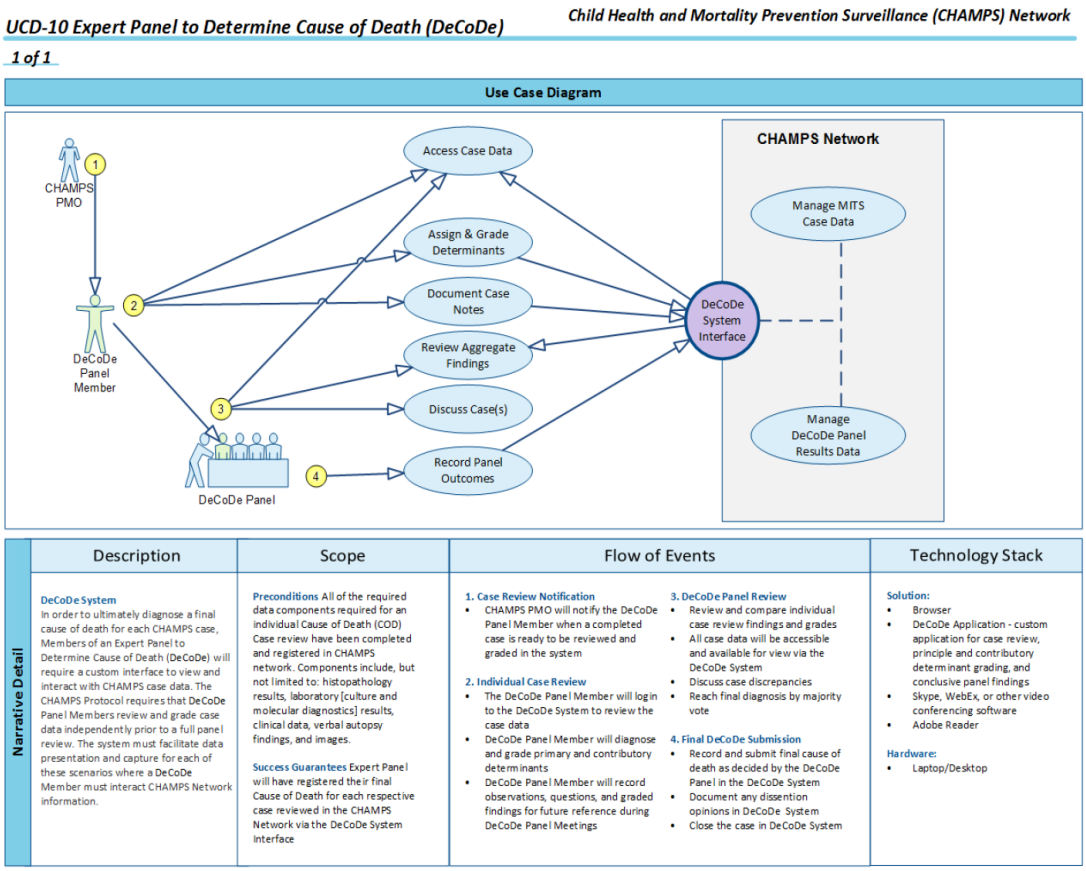


Figure 7 Example CHAMPS use case diagram for the *Expert Panel to Determine Cause of Death Panel* member [10]

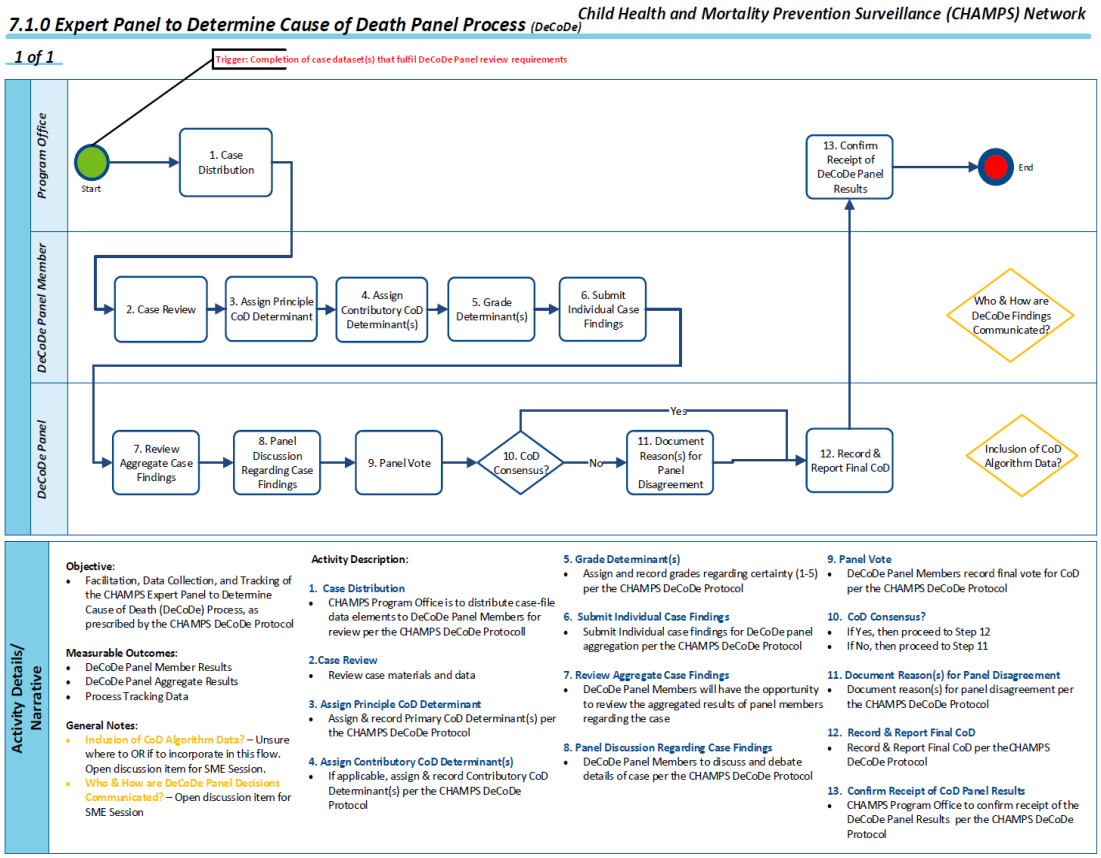


Figure 8 Example CHAMPS Workflow diagram for the *Expert Panel to Determine Cause of Death Panel* Process [10]

The subsequent CRDM evaluation of the CHAMPS program technical needs yielded eleven key personas with corresponding use cases and revealed eight (8) core business processes comprised of twenty-five (25) detailed workflows. Each of the artifacts

were intelligible and informative studies of hyper-focused aspects of CHAMPS. However, consumption of the individual artifacts failed to articulate the larger scale of the surveillance network, including actual surveillance data procurement and exchange. An enumeration scheme was devised to link and correlate related concepts. However, the issue remained that the volume and granularity of the detailed workflows obfuscated the comprehensive picture. Figure 9 captures printouts of the workflow diagram pasted to a large whiteboard with hand-drawn connectors and notations. This was how the workflows were "connected" in early attempts to piece together the whole picture.

Figure 9 Complex Surveillance Program/Network Required Linkage of Numerous Workflows

Inspired by subway maps, a diagram was developed to show the relationships of the enumerated materials. While this illustrated the relationships of the documented business process components and presented a flow of information, this format lacked sufficient detail to inform data collection or system implementation; it resulted in an over-consolidation of information flow.

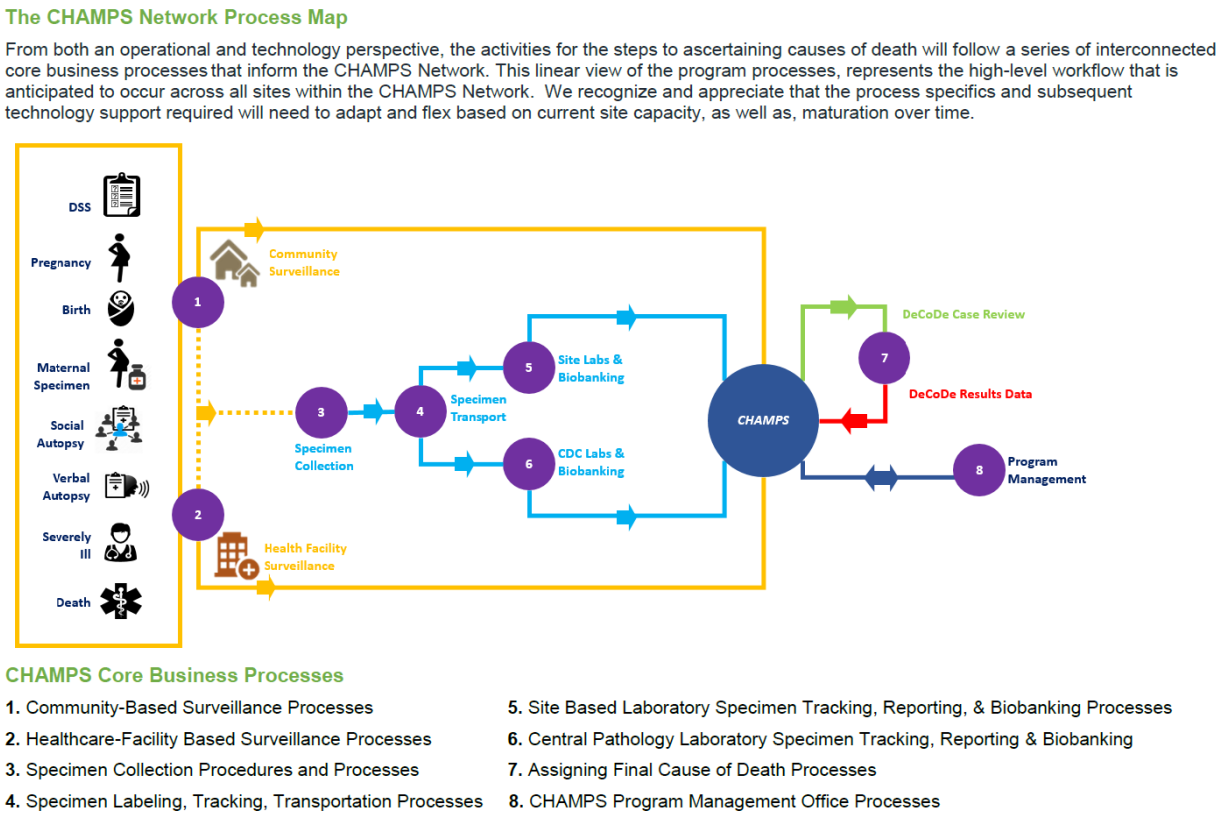


Figure 10CHAMPS Network Process Map from IT Plan [10]

Ultimately, the team needed to develop a better way to visualize the surveillance network with the right level of detail and information.

1.3.2 Formal problem statement

A public health informatics team was charged with the requirements gathering, development, and implementation of a technical systems architecture for the Child Health and Mortality Prevention Surveillance (CHAMPS) Network. CHAMPS stakeholders lacked a broadly understandable and comprehensive programmatic and/or technical documentation to describe and communicate the complexity, interdependency, timing, and orchestration of the surveillance case-level deliverables to a multitude of international stakeholders and data consumers.

1.4 Purpose statement

The purpose of this thesis is to describe how the creation of unique form of technical diagramming was able to holistically depict the complex series of workflows, decision points, operational hand-offs/information exchanges and dependencies required to collect and consolidate the near-real-time mortality surveillance data necessary for an in-country panel of experts to reach an informed and definitive cause of death determination for deceased under-five children. The documentation, in concert with intentionally facilitated discussions about the diagram, proved to be a significant strategic resource for the originating project stakeholders, helping them formulate the overarching surveillance design. This resulting in a consensus blueprint which served as the keystone document for orientation to the project, discovery of local processes, training, and implementation at each CHAMPS location: Soweto, South Africa, Bamako, Mali, Manhica, Mozambique, Kisumu, Kenya, Harar/Kersa, Ethiopia, Baliakandi, Bangladesh and Makeni, Sierra Leone.

1.5 Key Aims

Aim 1: Methodology for the development and refinement of novel data flow diagram for the CHAMPS case-based public health surveillance program

Aim 2: New type of technical documentation artifact for the CHAMPS case-based public health surveillance program that could be generalizable for other case-based public health surveillance programs

1.6 Significance statement

While the purpose and objective of a public health surveillance system may be clear, challenges to design, implement, operate and evolve technical systems to support the increasing data demands to support surveillance objectives is ever more complex and must keep pace with the speed of technological advances.

This project-based thesis presents an innovative technique of technical documentation developed in service of a complex global public health surveillance network that may offer a more comprehensive, versatile, and intelligible illustration technique to capture the flow of surveillance data, potentially disrupting conventional techniques within the industry.

Chapter 2 Methodology

2.1 Introduction

The charge of this thesis project was to develop new form of informatics diagram that would capture the broader surveillance logistics of CHAMPS by incorporating the confluence of processes, study artifacts, standard operating procedures, and data flow. Moreover, this diagram would be needed as a visual communication tool and baseline guide for system implementation at each site. This chapter will describe, in detail, how the data flow diagram was created, how it catalyzed the development of stakeholder consensus regarding the development of the surveillance protocol, and how it was then leveraged as the foundation for informatics planning and implementation at each surveillance location.

2.2 Description

Programmatic reference documentation provided written descriptions of the newly conceived surveillance scope, goals, desired case data, specimen collection methods, and anticipated laboratory and histopathology diagnostic results. Subject matter expert interviews and facilitated discussions added context, clarification, and opportunities to elicit new insights or rethink preconceived concepts as the program evolved. Both the documentation and interviews were used as inputs to develop and hone iterations of the data flow diagram artifact.

2.2.1 CHAMPS Bill & Melinda Gates Grant Proposal (2015)

The CHAMPS grant proposal to the Bill and Melinda Gates Foundation provided strategic-level description of the overall scope of work, the planned contributions of program partnerships, key data and specimen collection methods, definitions of surveillance site selection criteria, and set general expectations for the timeliness of data curation and data dissemination.

2.2.2 CHAMPS Laboratory and Diagnostics Plan (2015)

The CHAMPS Laboratory and Diagnostics Plan was a comprehensive and detailed plan describing the specimen types (both tissue and non-tissue) to be collected from surveillance cases, the procedures for specimen collection, sample preparation, sample testing, and long-term storage among local and central biorepositories.

2.2.3 Information Technology Milestone Deliverable (2015)

The Information Technology Milestone Deliverable [10] was an overarching information technology plan devised to identify the potential classes of systems, applications, and hardware anticipated to facilitate the data collection and management across the CHAMPS surveillance sites and the central CHAMPS Program Office. This plan incorporated a description of the anticipated categories data to be collected (i.e., data families), twelve detailed user personas and corresponding use case diagrams, a suite of 24 individual business process flow diagrams detailing eight core business processes, a network reference architecture, and budget allocation projections.

2.2.4 Business Process Workflow Documentation

A suite of twenty-four individual business process flow diagrams which documented the underlying workflows of eight core business processes: (a)community-based surveillance processes, (b) healthcare-facility based surveillance processes, (c) specimen collection procedures, (d) specimen labeling, tracking, and transportation processes, (e) site-based laboratory specimen tracking, reporting, and biobanking processes, (f) central pathology laboratory specimen tracking, reporting, and biobanking (g) assigning the final cause of death processes, and (h) CHAMPS program management office processes.

2.2.5 Subject Matter Expert Interviews and Facilitated Discussions

Subject matter expert (SME) interviews and facilitated discussions were utilized in two primary contexts/formats: CHAMPS Program Office SME discussions (2.2.5.1) and CHAMPS Site/Field SME discussions (2.2.5.2)

2.2.5.1 CHAMPS Program Office SME Discussions

Individual interviews and facilitated group discussions were held with stakeholders, partners, consultants, and staff representing the Emory Global Health Institute (EGHI), United States Centers for Disease Control and Prevention (CDC), International Association of National Public Health Institutes (IANPHI), Task Force for Global Health, Deloitte, and the World Health Organization (WHO). These discussions covered both broad and highly specific topics in the fields of Social Behavior Sciences, Laboratory Sciences, Histopathology, Demography, Epidemiology, Verbal Autopsy, and Mortality surveillance (e.g., causal chain). The primary intention of gathering information in these

sessions was to develop consensus agreement among respective stakeholders regarding methods (both case-level and network-level), processes, data collection, data exchanges, and outputs of each surveillance component required to compile mortality case files necessary for formal cause of death determinations.

2.2.5.2 CHAMPS Site/Field SME Discussions

Individual interviews and facilitated group discussions were held with CHAMPS site personnel responsible for implementation and operation of each surveillance component (of data stream) prescribed by the CHAMPS mortality surveillance protocol. Site personnel included physicians and other health providers, social behavioral experts, demographers, laboratorians, pathologists, community health workers, public health policy makers, and local information technology professionals. The primary goal of these discussions was to provide a comprehensive overview of the CHAMPS surveillance requirements to local site staff, with subsequent discussions focused on the current state of resources, methods, processes, and systems that may be locally utilized to facilitate the CHAMPS surveillance requirements (e.g. does the site utilize a laboratory information management system? Paper or electronic health records?).

2.3 Project design

2.3.1 Development of a prototype diagrams

In the context of this applied project, the original diagrams were effectively draft sketches captured in MS PowerPoint to help think through what was needed.

Figure 11 shows an example of an attempt to expand on the subway map concept originally developed for the Information Technology Milestone plan. The iteration in Figure 11 served more as a mind-map of the business processes that would need to be included in the more formal version.

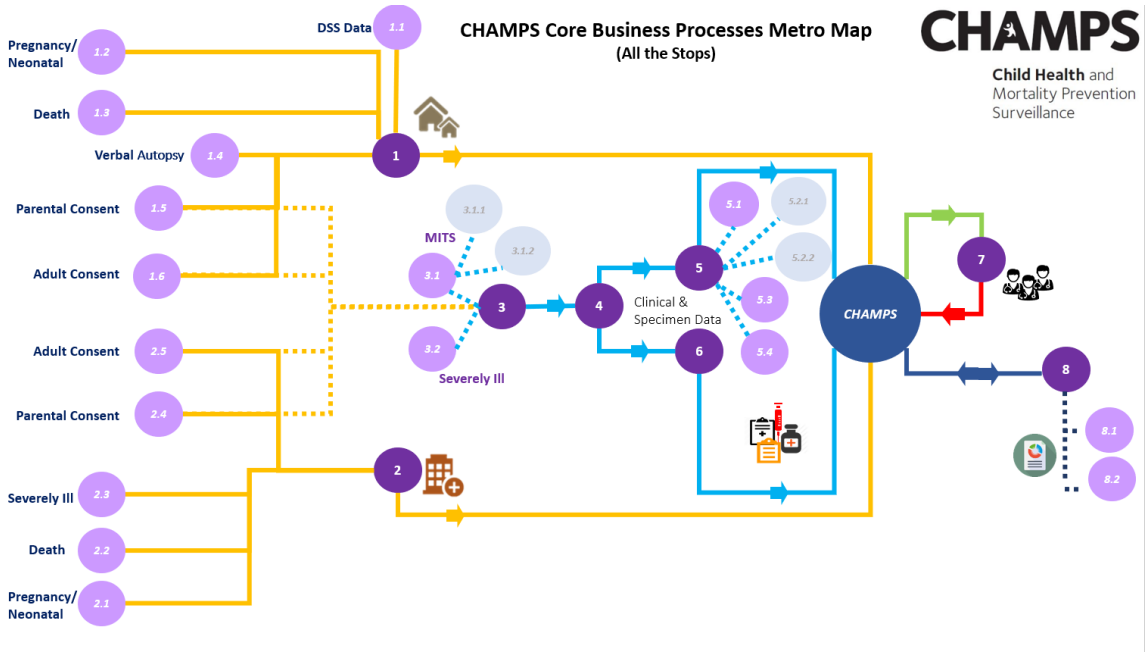


Figure 11 Data flow diagram prototype (version 1.0)

Additional informal drafts tried to indicate where consents, data capture/case report forms, specimen collection, and other case artifacts may have been required. These drafts began to incorporate more narrative detail and notes for consideration in a more formal model (Figure 12).

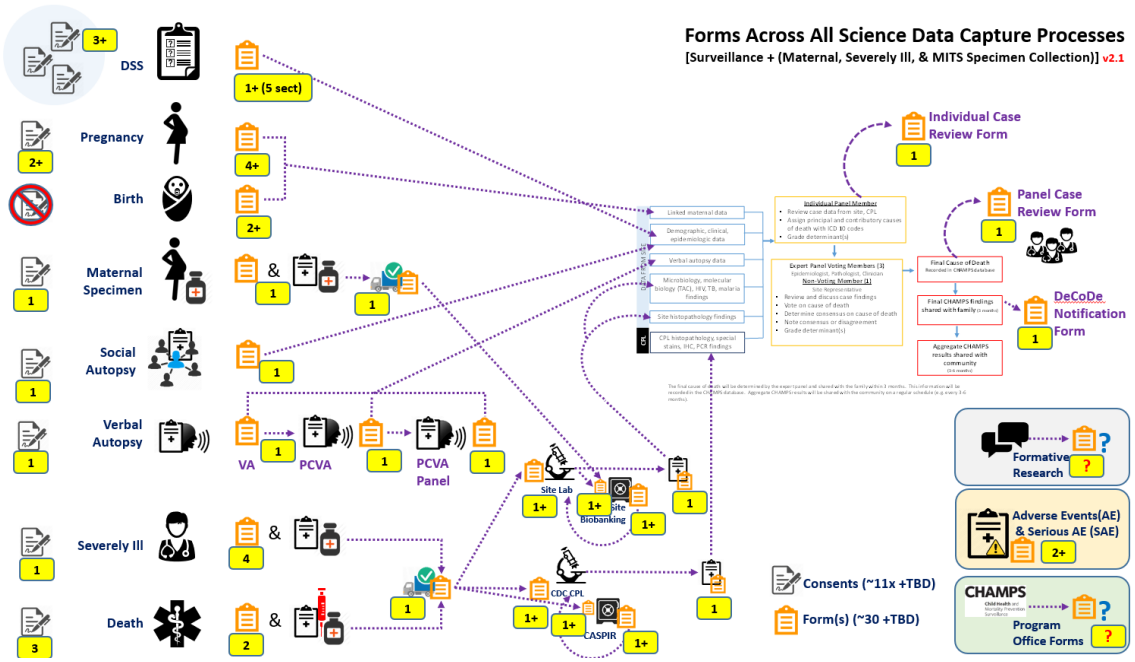


Figure 12 Data flow prototype version 2.1.

The third significant iteration of the diagram development was the creation of the more formalized model developed in MS Visio (Figure 13). This model was purposely cleaner and more uniform with professional and polished appearance. This model brought together the key components of the previous sketches. Furthermore, these items where organized in a logical sequence of events with indications of significant decision points and flow of information thorough and among the business processes. Lastly, a person model and diagram legends were added. The formal prototype diagram was reviewed and modified iteratively with the field advisor and the informatics team until there was consensus that the prototype fulfilled the initial requirements and was ready for stakeholder review and engagement.

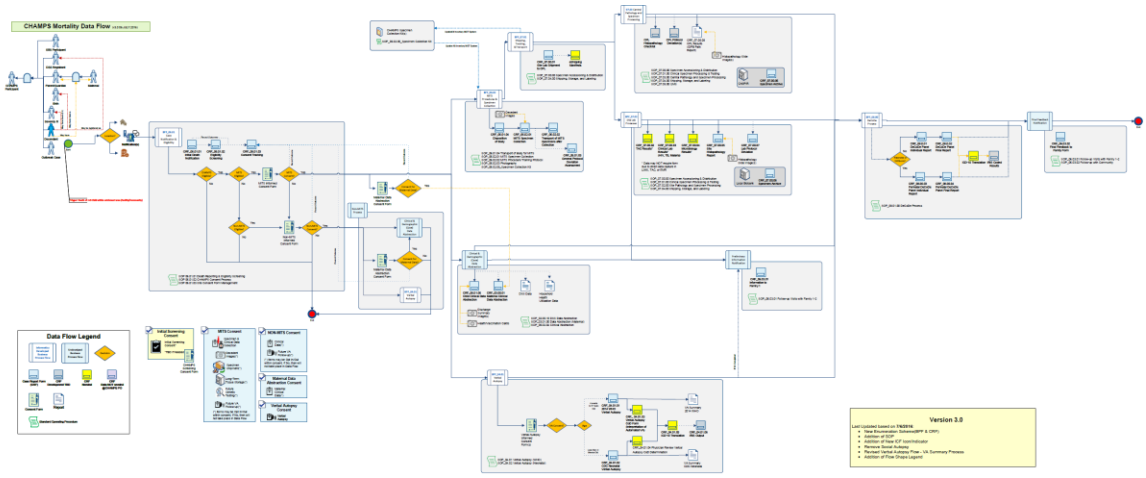


Figure 13 Final working prototype the formal model, version 3.0

2.3.2 Internal programmatic baseline version

The formal prototype data flow diagram, that evolved from model sketches and informatic team review, was then intentionally engaged in iterative design and refinement sessions. The diagram was subjected to a series of collaborative reviews and design sessions whereby SMEs, partners, and other stakeholders could comment, question, and ultimately validate incremental changes until a satisfactory generic version of the document was derived. The outcome of these sessions was the inaugural baseline version of the comprehensive and novel data flow diagram. This version of the data flow diagram (Figure 14) represented the expected or ideal state of how a participating site would implement CHAMPS; it would serve as a blueprint for the model house prior to the discovery of site-specific capabilities, gaps, and constraints.

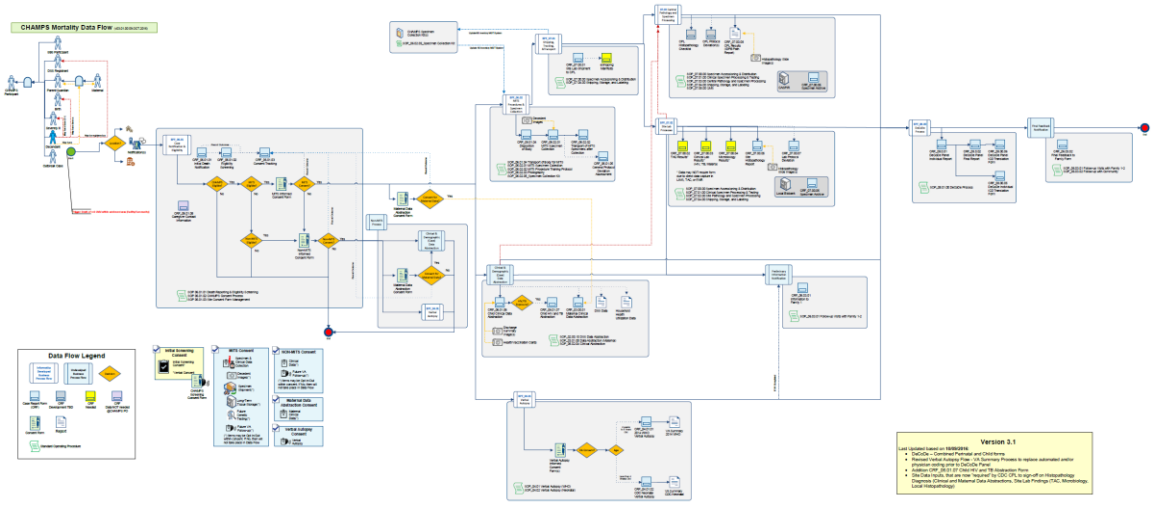


Figure 14 Inaugural baseline version of the data flow diagram, version 3.1

2.3.3 External in-country overview and site-specific capabilities and gaps

The generic baseline version of the diagram was then presented to CHAMPS site teams in-country where it was initially used to provide an overview of how the mortality protocol functioned among and across the varied surveillance modalities (or work streams). Then the diagram was used in small group settings with the specific work stream groups (e.g. laboratory technicians, pathologists) to facilitate discussions around data collection methods and local capacities. For example, laboratory data collection could occur within a pre-existing laboratory information system (LIMS) or the local group may require a custom paper/electronic case report form. During these sessions, CHAMPS informaticians could articulate the specific needs of the surveillance network to the site staff while also discovering the site-specific capabilities, gaps, and constraints that would be needed to plan for future system implementation and operation.

2.4 Procedures

2.4.1 Document Components and Features

2.4.1.1 Data Flow Icons or Symbols

The data flow icons or symbols are images used in the diagram to represent artifacts, processes, decision points, and other applicable concepts. Whenever feasible, commonly recognized business process notation symbols were used or adapted. New symbols or icons were created as needed. Common data flow symbols include icons for case report forms, standard operating procedures (SOP), data files, reports, decision points, and business processes to name a few. Figure 16 is a legend with data flow icons. Other symbols and icon are found among additional figures herein.

It is important to highlight the indication of SOP and case report forms within this diagraming method, as it is novel to this method. Inclusion of these items expands audience and utility of the diagram as these components are important linkages to programmatic/operational activities and artifacts that are beyond traditional information technology diagrams. Their presence in the diagram serves as a roadmap for where SOP/data collection occurs or is needed and is use for staff training.

2.4.1.2 Person/Case Entity Relationship Model

Unlike a specific disease surveillance system, where a case-type is defined by the presence or absence of a specific condition, the original CHAMPS surveillance scope and planned duration allowed for the possibility that a person (or case) could be engaged in CHAMPS surveillance in a number of contexts over time.

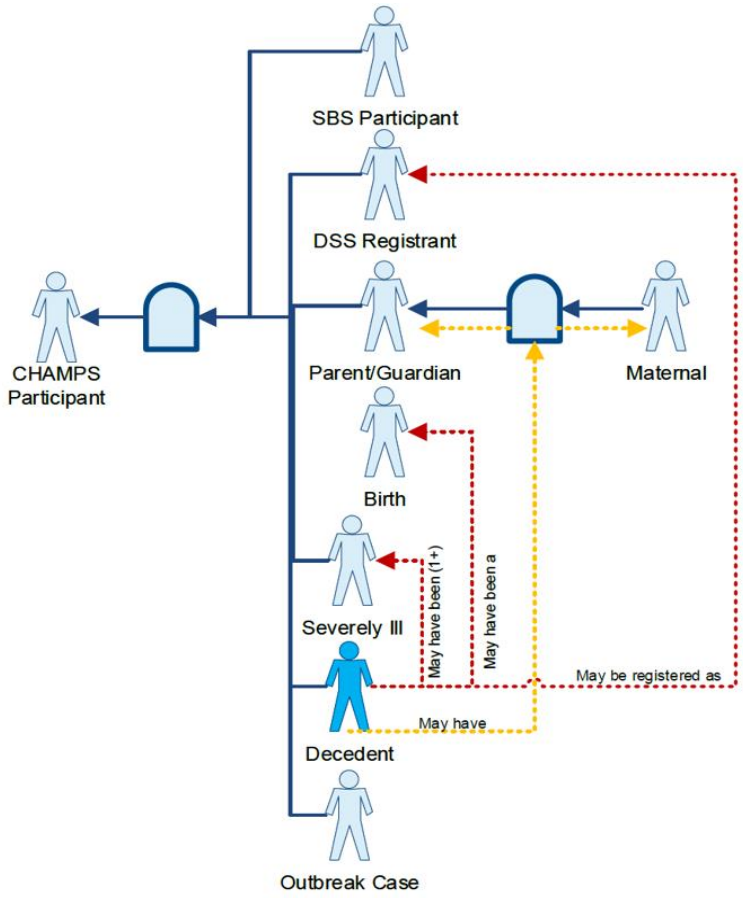


Figure 15 Person/Case Entity Relationship Model for CHAMPS Morality Surveillance

2.4.1.3 Data Flow Legend

The data flow legend is included to provide an explanation of common symbols or icons used within the diagram to facilitate a better understanding or interpretation. Key symbols were included in the legend with generic labeling to describe the shape/symbol/ purpose within the diagram (Figure 16).

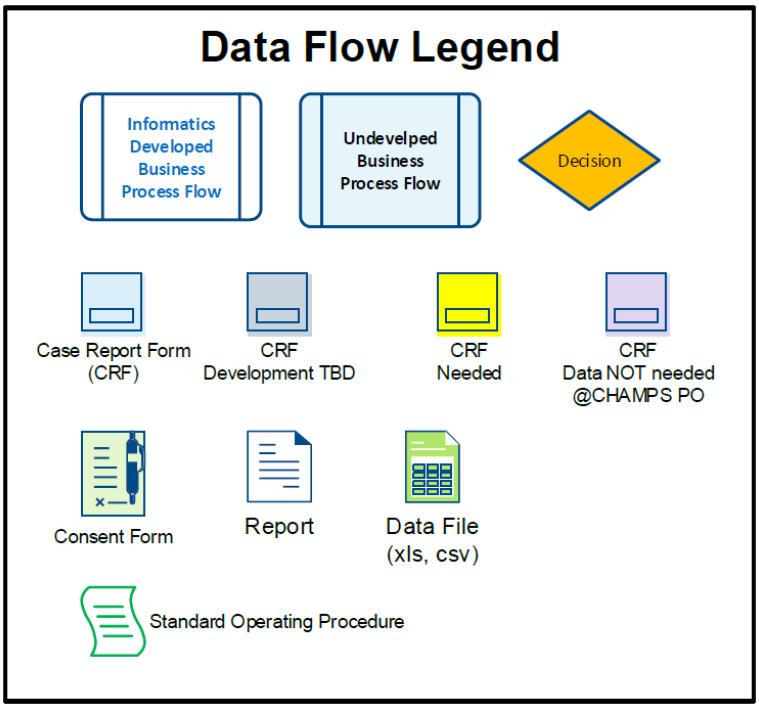


Figure 16 Depiction Data Flow Legend

2.4.1.4 Data Flow Connectors

Data flow connections are directionally indicative solid (event/exchange shall occur)or dashed lines (event/exchange is optional) that express the flow of information, data, and/or artifacts between and among business process groupings, decision points, and other icons/symbols. Colors were applied to emphasize specific data flow connections, relationships, or specific types of information (e.g. orange for image files, red for data/artifacts shared with both the CHAMPS program office and the central pathology laboratory).Figure 17 provides examples of data flow connectors within a diagram.

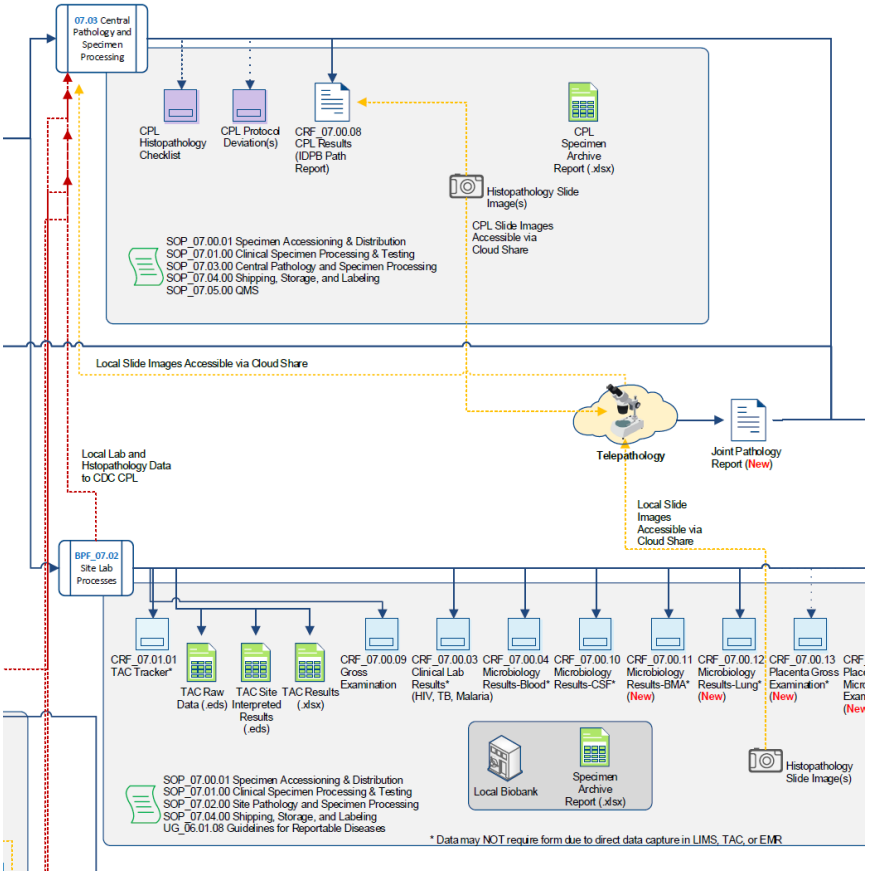


Figure 17 Image depicting large number and types of connectors (solid, dashed, and colors)

2.4.1.5 Data Flow Decision Points

Standard diamond decision icons were used to indicate significant data flow decisions points that impact case data collection pathways or hand-offs. Data flow connectors leaving a decision icon were labeled with the applicable decision outcome associated to the flow (e.g., “yes” or “no”). Figure 18 provides examples standard decision diamonds.

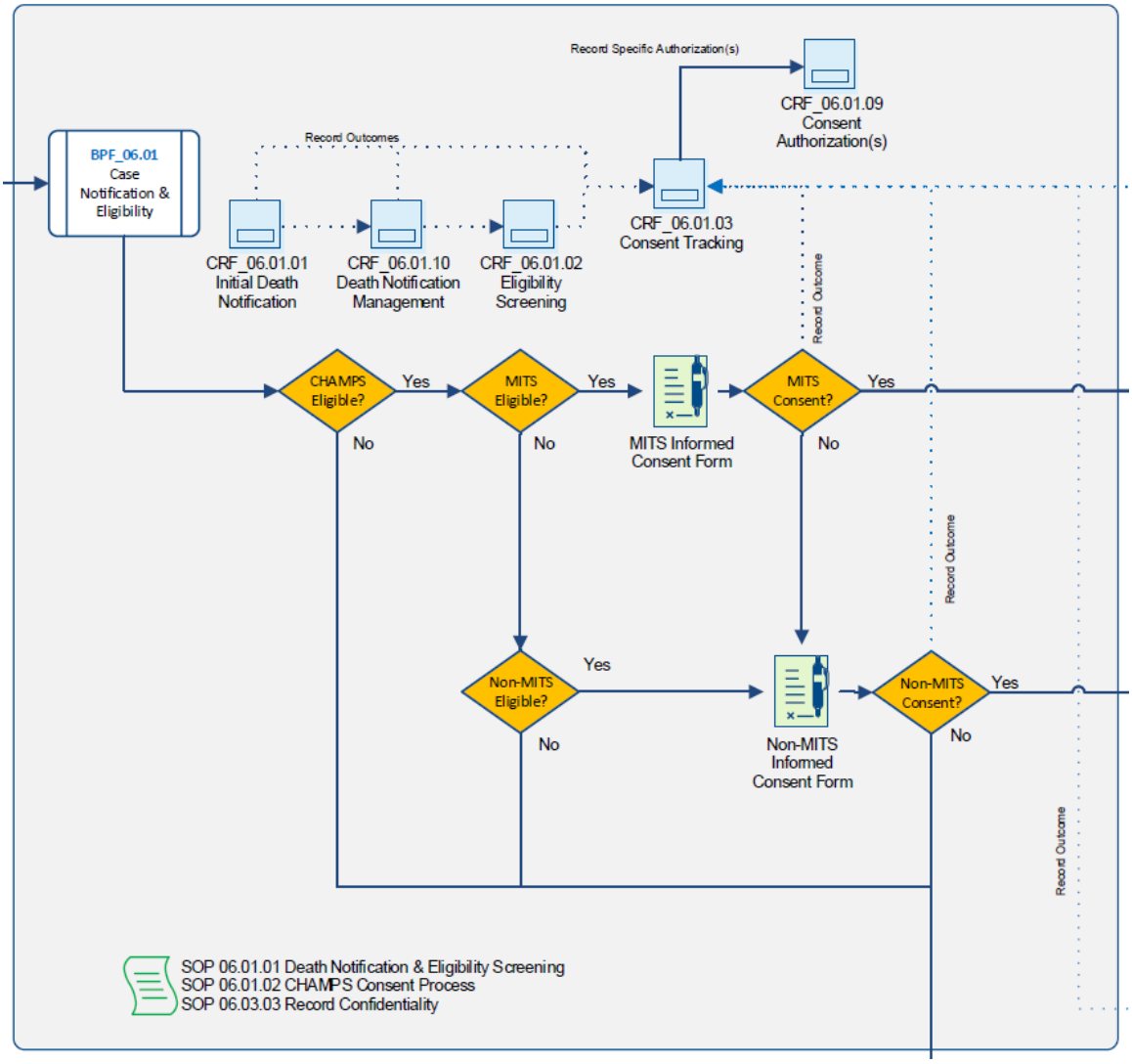


Figure 18 Section of the data flow diagram with decision points (orange diamonds)

2.4.1.6 Business Process Grouping

A business process grouping is a collection of symbols, icons, data flow connectors and decision points that share a common core business process and are encapsulated within a background shape. The core business process may represent one or many sub-business process workflow diagrams associated with it, representing the larger or overarching process (Figure 19).

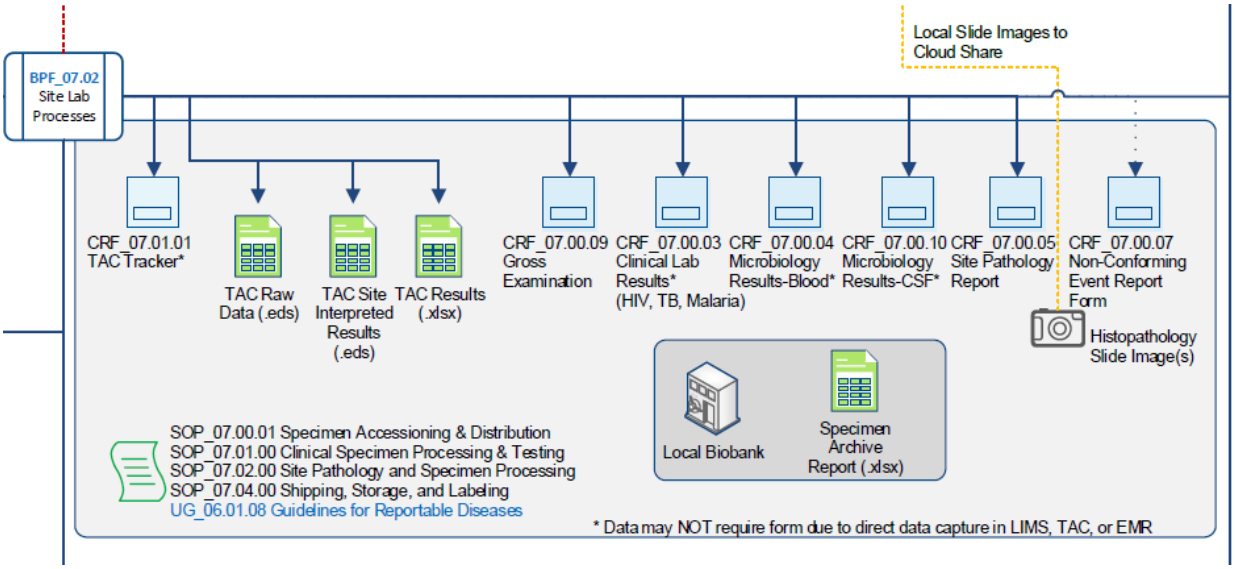


Figure 19 depicts a business process group for the site laboratory process. It includes icons for applicable case report forms, data files, images, and the governing standard operating procedures.

2.4.1.7 Consent Attributes

Consent attributes were depicted as distinct symbols within small tables or listings that detail the specific attributes a subject is consenting to when agreeing to participate in the surveillance program (Figure 20).

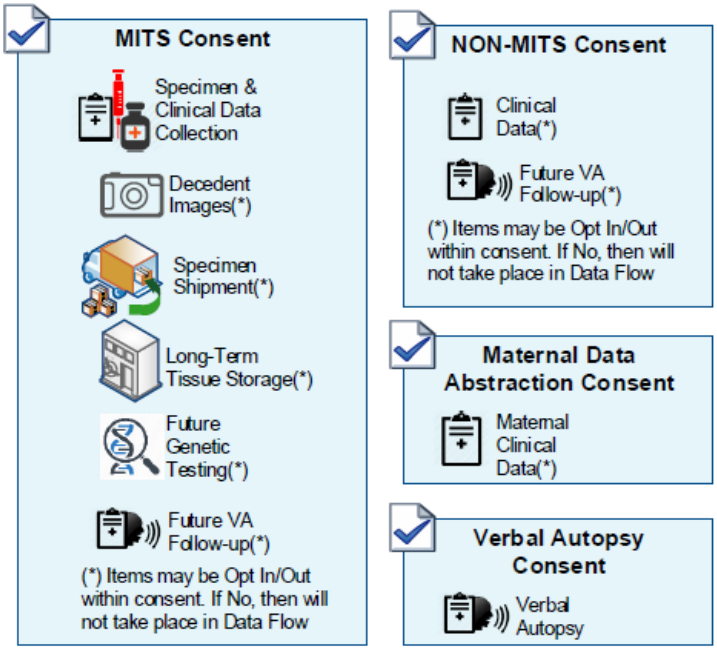
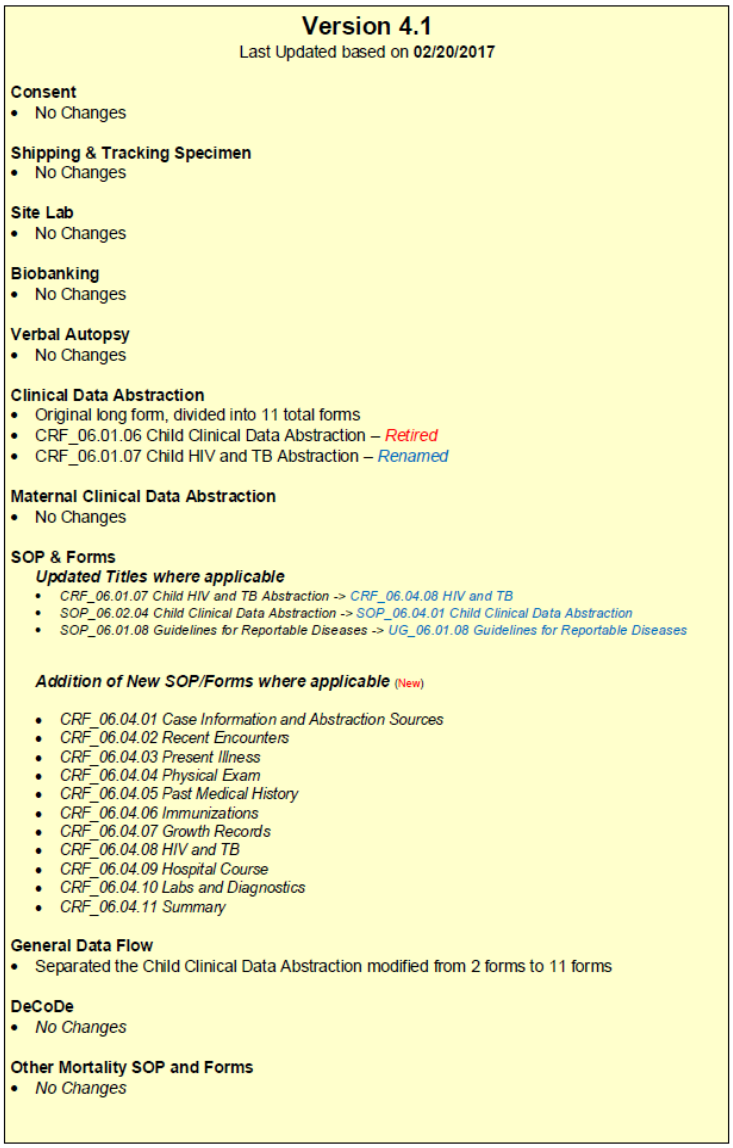


Figure 20 Depiction of the consent attribute information found on the CHAMPS mortality data flow diagram

2.4.1.8 Diagram Version Notes

The diagram version notes is a table or notes section that recorded the current diagram version and documents any changes to the diagram since the previous version. Here, data flow elements that are changed, removed, or added are documented (Figure 21).



Version 4.1

Last Updated based on **02/20/2017**

Consent
- No Changes

Shipping & Tracking Specimen
- No Changes

Site Lab
- No Changes

Biobanking
- No Changes

Verbal Autopsy
- No Changes

Clinical Data Abstraction
- Original long form, divided into 11 total forms
- CRF_06.01.06 Child Clinical Data Abstraction – *Retired*
- CRF_06.01.07 Child HIV and TB Abstraction – *Renamed*

Maternal Clinical Data Abstraction
- No Changes

SOP & Forms

Updated Titles where applicable
- *CRF_06.01.07 Child HIV and TB Abstraction -> CRF_06.04.08 HIV and TB*
- *SOP_06.02.04 Child Clinical Data Abstraction -> SOP_06.04.01 Child Clinical Data Abstraction*
- *SOP_06.01.08 Guidelines for Reportable Diseases -> UG_06.01.08 Guidelines for Reportable Diseases*

Addition of New SOP/Forms where applicable (New)
- *CRF_06.04.01 Case Information and Abstraction Sources*
- *CRF_06.04.02 Recent Encounters*
- *CRF_06.04.03 Present Illness*
- *CRF_06.04.04 Physical Exam*
- *CRF_06.04.05 Past Medical History*
- *CRF_06.04.06 Immunizations*
- *CRF_06.04.07 Growth Records*
- *CRF_06.04.08 HIV and TB*
- *CRF_06.04.09 Hospital Course*
- *CRF_06.04.10 Labs and Diagnostics*
- *CRF_06.04.11 Summary*

General Data Flow
- Separated the Child Clinical Data Abstraction modified from 2 forms to 11 forms

DeCoDe
- *No Changes*

Other Mortality SOP and Forms
- *No Changes*

Figure 21 Sample Diagram Version Notes section

2.4.1.9 Retired Diagram Components

This is a section of the diagram where icons of case report forms, SOP, decision points or other key elements were placed if they were retired from the current version and were present in a previous version of the document. This served as a reminder that the items were present at one point in the past (Figure 22).

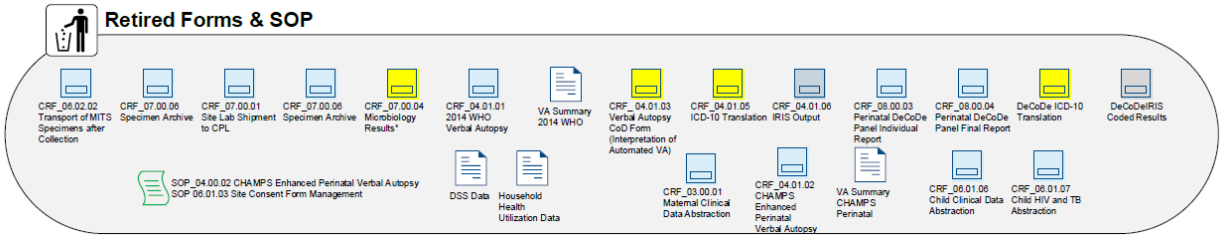


Figure 22 Depiction of Retired Diagram Components

2.4.2 Internal Programmatic Design

The purpose of internal programmatic design sessions were to depict the perceived initial or current state of a single or series of business process groupings in order to engage stakeholders and subject matter experts in reaching a consensus version.

A review of applicable reference materials was required for the informatician to account for the key diagram components needed to depict an initial/proposed current state model prior to discussion. The initial model was presented to the applicable SME. The informatician stepped through the process steps, flow of information and related

artifacts, reports, and data capture needs. Any changes, decision points, new information, and undecided aspects of the diagram were noted during the discussion. A new version of the document was generated that represents the feedback from the previous discussion. Then, the diagram version notes, the version number, and date were updated. In addition, the changes applied to diagram are outlined: any new icons added to diagram are indicated as "new", changed, or relabeled icons were noted, and any icons removed were copied to the retired diagram component section. The updated version document (i.e., revised state) was shared with the SME along with the previous version. A follow-up facilitated discussion where the previous and revised versions are presented and compared was organized. If the revised version receives consensus approval from the SME, then it is adopted. However, if consensus approval from the SME is not achieved, the process of iterative revisions and reviews were to be continued until a consensus can be reached. The ultimate outcome expected is a clear depiction of the most current consensus thinking of a specific business process grouping or series of interconnected process groupings.

2.4.3 In-Country Overviews

The purpose of the in-country overview sessions were to engage and orient the disparate local working groups to the full scope of the CHAMPS surveillance requirements such that the local working groups had a sense of the big picture and their respective place in contributing to the overall surveillance goals. Secondly, the in-country overview sessions allowed for the facilitation of small groups sessions whereby local working groups and program informaticians could communicate the location-specific capacities and workflows relative to the generic standardized dataflow (e.g., local laboratory results case

report forms may not be applicable if the site has a laboratory information system in place).

For large group presentations, a soft copy of the most current version of the data flow diagrams were incorporated into a slide presentation whereby a general overview of the whole diagram and drilldowns of specific sections were projected.

Several copies of the most current version of the data flow diagrams were printed on large paper prints that were three feet by four feet (3'x4') in dimension. More intimate small group meetings were arranged with local SMEs to perform a review of the generic flow and they were engaged in a detailed discussion regarding how the generic flow would be implemented locally. Key details regarding the staff involved, facilities, systems, equipment, timing, strengths, exceptions, gaps, and risks were discussed. Handwritten notes and modifications were captured on the paper data flow in real-time. Following the small group sessions, the notes regarding the local context were recorded

(Figure23).

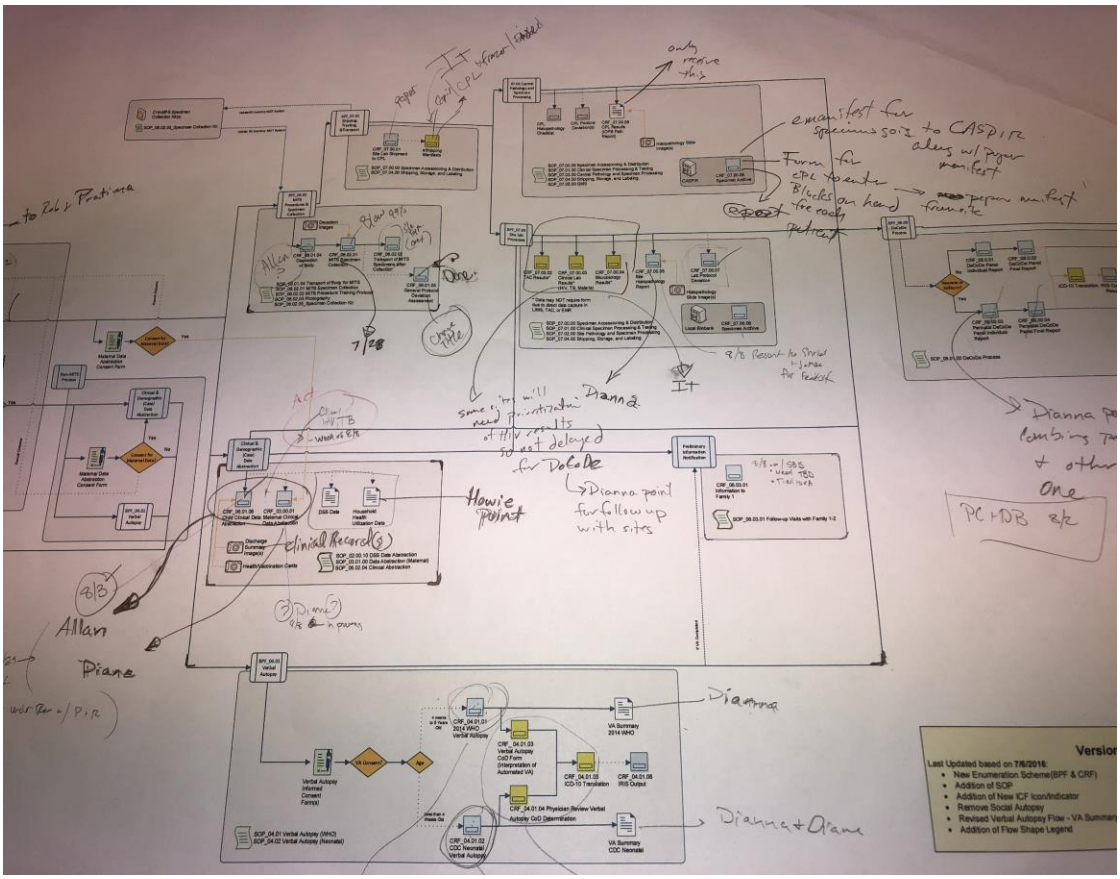


Figure 23 Depiction of over-sized, high-resolution paper printout of the data flow diagram with hand-written notations collected during facilitated discussion(s)

The information gathered during the in-country sessions were incorporated into the enterprise technical architecture considerations and site implementation planning, informing budgets, staffing needs, timelines, equipment procurement, and technical assistance planning.

2.5 Instruments

2.5.1 Microsoft Visio Professional 2016

Microsoft Visio Professional 2016 is a diagramming and vector graphics software application produced and distributed by the Microsoft Corporation. This is the primary application used for the development of the data flow diagram. The application provides a library custom and industry standard stencils (or icons) from a variety of industries (e.g., electrical engineering, network engineering). Many of the standard business process notation icons were employed for use in the data flow diagram. Other custom icons were developed using the native drawing and formatting tools within the application (Figure 24).

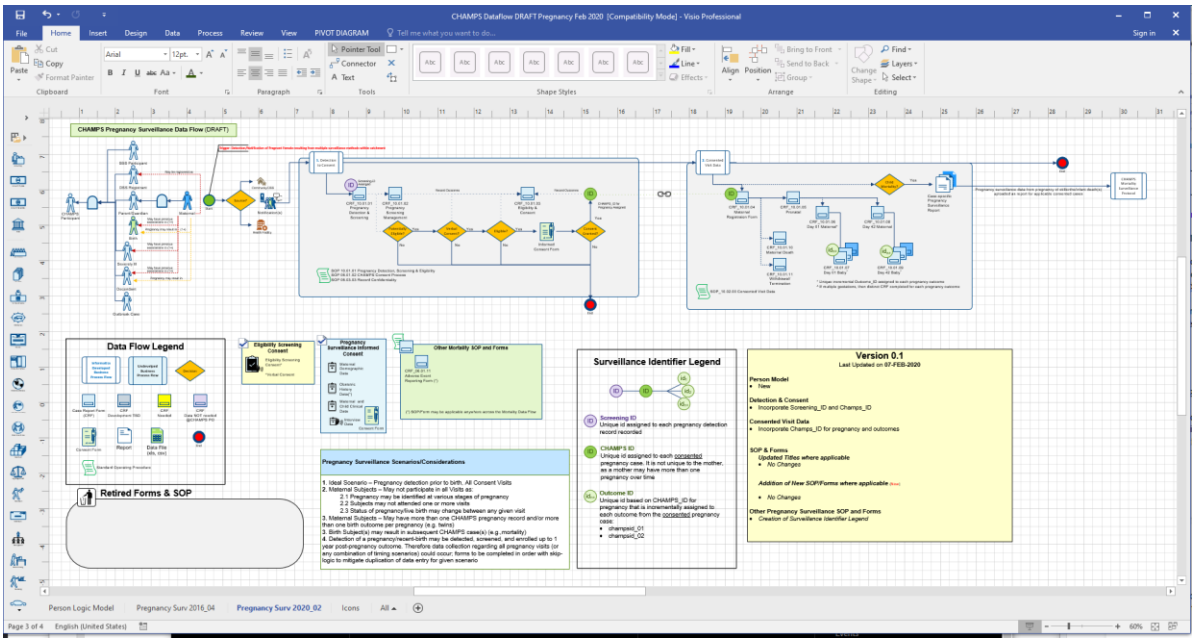


Figure 24 Depiction of data flow diagram development within Microsoft Visio Professional 2016 software

2.5.2 Microsoft PowerPoint 2016

Microsoft PowerPoint 2016 is a presentation development software application, produced and distributed by the Microsoft Corporation, that uses electronic slides to convey information at text, images or other forms of multimedia. This application was used for two primary purposes: (a) development of custom icons or images for incorporation into data flow diagrams and (b) facilitation of interviews and facilitated discussions with subject matter experts (Figure 25).

Figure 25 Depiction data flow diagram discussion facilitated with aid of Microsoft PowerPoint software [16]

2.5.3 High Resolution PDF (portable document format) of Data Flow Diagram

PDF document files allow for the data flow diagram (which is developed in MS Visio and then exported as an image in PDF format) to be shared and displayed as a high-resolution image file. Adobe provides a software product, Adobe Acrobat Reader, to open, view, and print PDF files that is easily accessible at no cost (Figure 26).

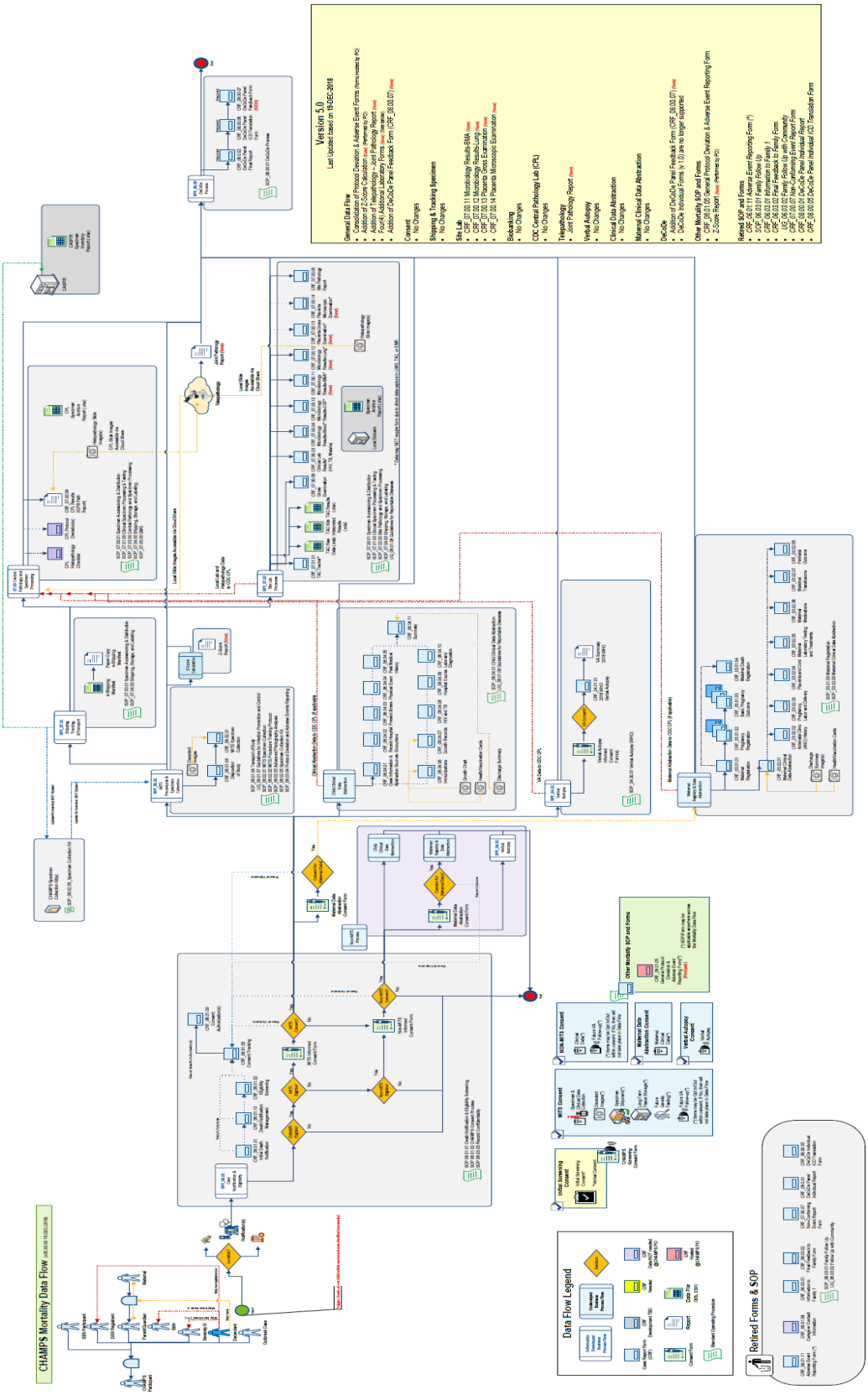


Figure 26 Depiction data flow diagram as high-resolution pdf format, version 5.0

2.5.4 High Resolution Oversized Paper Printouts

Data flow diagrams were printed on large paper prints that were three feet by four feet (3'x4') in dimension and used or local context notes and additional copies were left for local reference. Please see Section 2.4.3 and Figure 23 for procedural details regarding use of oversized paper printouts of the data flow diagram

2.6 Analytic Methodology

The analytic methodology for the development the data flow diagram for this thesis project was an iterative trial and error method based on immediate qualitative feedback from subject matter experts that were engaged in facilitated discussions using the diagram as a requirements facilitation tool (i.e., discovery) and/or those SME that independently reviewed the diagram content for accuracy. This feedback was incorporated in an updated version of the diagram that was then used in the next iterative round of facilitated discussion or review. Versions of the document were saved over time as changes were made. Changes and innovations to the diagram that achieve consensus acceptance were retained (i.e., validation) and maintained going forward in future versions.

Chapter 3 Final Product

3.1 Introduction

Through multiple iterations and rounds of changes, the data flow diagram achieved the overall objective of providing a comprehensive depiction of the surveillance data flow that is informative, instructive, and universally engaging, striking the appropriate balance of communicating both technical and non-technical system workflow without overwhelming in the minutia task-level activities. This balance is similar to a theme park kiosk that clearly informs the viewer "you are here" and engages the viewer at the right perspective level, such that the viewer is promptly acclimated to his/her immediate position within the park ecosystem and ultimately enabled to discern how or where to go next.

3.2 Key Findings

3.2.1 Universal Engagement Tool

The data flow diagram served the informatics objective of providing an engaging visual blueprint of the surveillance network that facilitated both technical and operational sessions among and across a variety of stakeholders. The diagram at its core is a picture transcending language barriers and technical/professional jargon with its use of non-verbal and written communication of process and flow.

3.2.2 Method Accelerated Findings and Implementation

The data flow diagram and facilitated requirements sessions improved stakeholder engagement by accelerating overall comprehensive requirement gathering through discussions and reviews at the appropriate level and context of review. These data flow informed sessions improved general requirements collection by delaying deep-technical task flow sessions until higher-level decisions were finalized. The diagram shifted the focus of the enterprise technical architecture from collection of completed case results to facilitating an ongoing transactional procurement and redistribution of data to complete a case file. Finally, the resulting comprehensive view of all the case data depicted in the data flow served as an implementation roadmap for the implementation of the underlying systems required to collect and transmit case data.

3.2.3 Requirements

Early in the development of the CHAMPS program, the challenge of gathering technical requirements was confounded by the translation of proposed concepts (e.g., collection of verbal autopsy was required) to actual tactical procedures and operational decisions (e.g., which specific verbal autopsy instrument would be used and if/how/when would it be analyzed). A task flow diagram is ideal when a task/method is known and therefore can be documented and evaluated. However, when the task is complete, how that task contributes to a larger operation, and what is to become of the task output is out of scope for the task flow documentation.

The data flow method allowed for the facilitation of requirements gathering at a higher level of focus. Again, the verbal autopsy is a good example within the CHAMPS context.

Review of the serial versions of the data flow diagrams with a focus on the verbal autopsy (VA) section reveals a constant change in both approach and expected outputs.

First, Version 3.0 requires a standard WHO VA to be implemented for cases between the ages of 4 weeks and 5 years old with a second custom CDC VA to be implemented for neonates and stillbirths. Version 3.0 then called for both VAs to be summarized, interpreted, assigned an International Classification of Diseases Tenth Revision (ICD 10) code and run through automated review software (Figure 27).

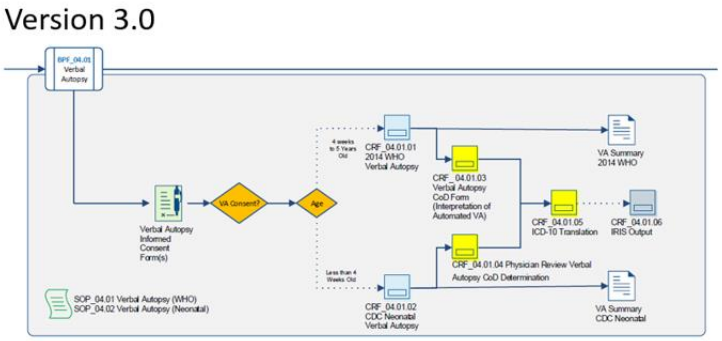


Figure 27 Version 3.0 of Verbal Autopsy section of data flow diagram

Then in Version 3.1 the interpretations, coding, and software analysis were abandoned (Figure 14).

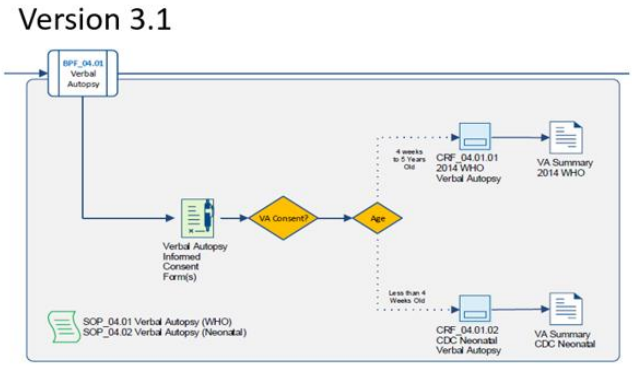


Figure 28 Version 3.1 of Verbal Autopsy section of data flow diagram

Lastly in version 4.0, a newer version of an integrated WHO VA was adopted for all ages with the requirement that both the raw data and a summary be provided as an output (Figure 29).

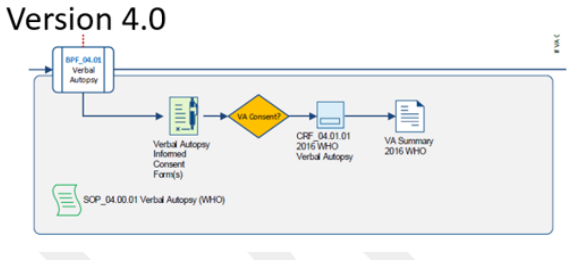


Figure 29 Version 4.0 of Verbal Autopsy section of data flow diagram

The requirements sessions at the level of the data flow of inputs, outputs, and how they contributed to the final case file were critical decisions that had to be made before it was prudent to engage in the technical facilitation of the final process. Engagement constrained to magnification of the task flow would not account for operational impact of timing, staffing, training, custom VA development, integrations into software tools, or the downstream contributions to the final case file. At the data flow level, we were able to raise important considerations regarding the technical facilitation of automated verbal

autopsy analysis, physician coding, impact of timing of these methods, and the complexity of potential conflicting or inconsistent results without expending analyst and SME time altering and reviewing the detailed task flow.

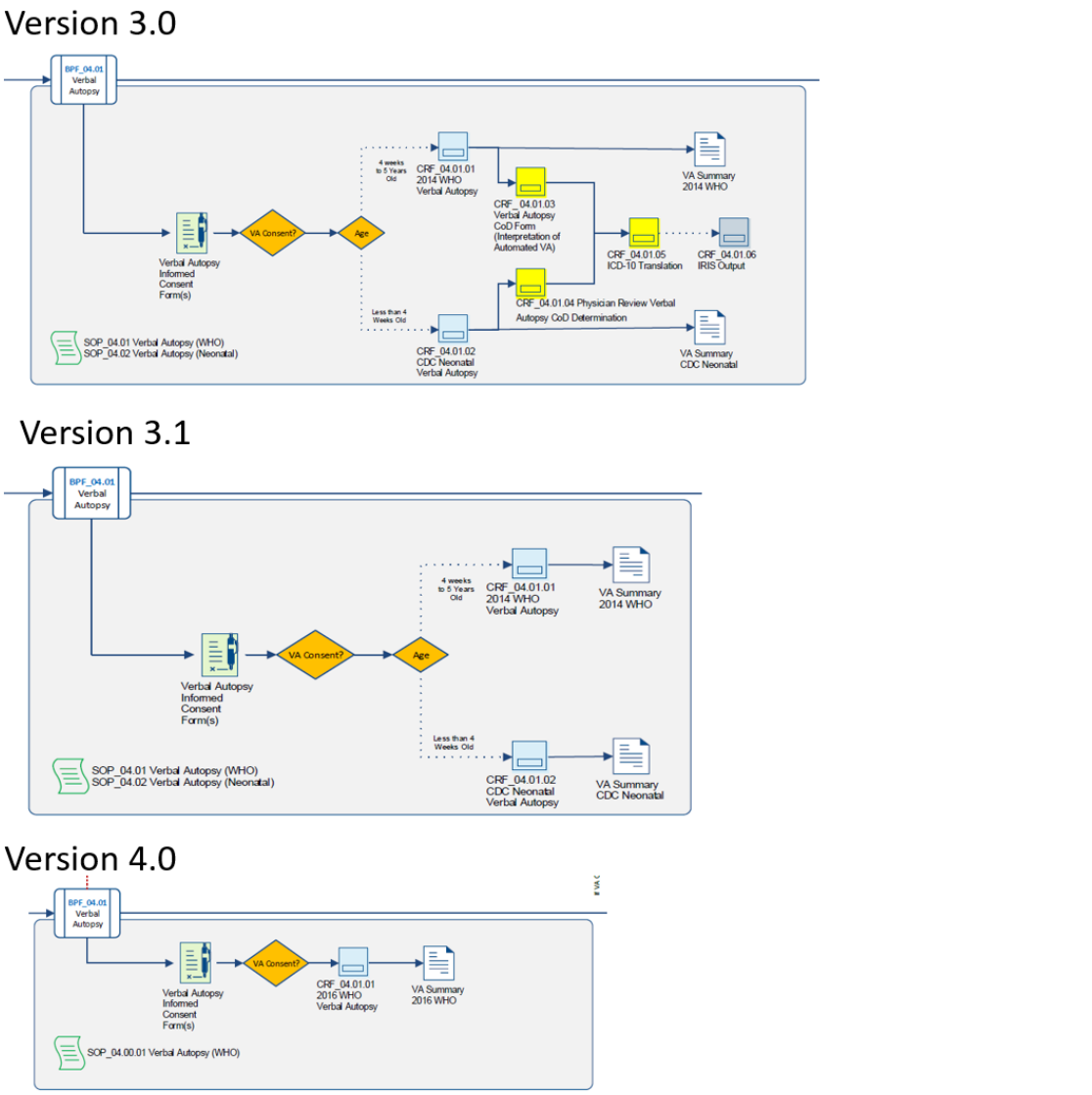


Figure 30 Depiction changes in sequence with the CHAMPS Verbal Autopsy high-level requirements over time and iterative versions of the data flow diagram

3.2.4 Technical Architecture

Early technical thinking on the CHAMPS project was that surveillance locations would collect, curate, and aggregate case information locally (or in-country); CHAMPS would collect the results when the case was completed. In fact, the first draft of the data flow diagram utilized in Mozambique was initially presented as an educational tool for the teams on what "they" needed to do. CHAMPS informaticians planned to document how the site teams were going to fulfill the requirements. As the CHAMPS informatics team engaged with the individual site team modalities (e.g., lab, clinicians, community contacts, demographers, etc.), it became apparent that local support and systems lacked the technical capacity to orchestrate and collate the disparate data in a manner that fulfilled the data requirements of individual case files that consolidated all the prescribed surveillance data into a single comprehensive report. This was a "technical" epiphany that changed the direction of the CHAMPS technical architecture planning, shifting to a model whereby a central CHAMPS system would need to receive transactional components of the case data, transform, and then redistribute the information throughout the case life cycle. This new technical paradigm required deeper consideration by the informatics team for more detailed and intentional data linkage strategies for all the surveillance data components to facilitate asynchronous data collection, asynchronous transactions of that data, and the eventual case-level assimilation of the most current case data available in support of operational hand-offs over time[1]. For example, surveillance site laboratory results were needed by CDC pathologists aid in their histopathology testing. The data flow diagram immediately found a broader purpose of documenting the central data flow blueprint for the honest brokering of data (i.e., collection, linkage, and consolidation of surveillance data on

behalf of the research sites) by a central system amongst stakeholders as the case developed, not just acceptance and validation of the final results.

3.2.5 Implementation Planning

Subsequent site implementation visits continued to leverage the data flow diagram to inform and educate the site teams on the broader process and procedures. However, the CHAMPS informaticians focus shifted to understanding which local systems were currently in-place to integrate with and where gaps existed whereby the CHAMPS informaticians needed to devise and implement solution offerings. The data flow diagram was then used to plan implementation steps and track progress of system implementation components. Both local and central stakeholders could see which components were complete and which remained outstanding.

3.2.6 Reference Document

This new form of diagraming serves as a keystone reference document for the project. When new surveillance sites are evaluated, the most current version of the data flow diagram is an essential artifact provided to represent the mortality surveillance operation. The data flow animates the protocol and procedures; visualizing the organized and sequential depiction of the significant decision points, consent milestones, case report forms, applicable standard operating procedures, and exchanges of data as it flows to the case-level summary report and final cause of death determination.

3.3 Other Findings

3.3.1 Cross-functional Organization

The CHAMPS surveillance project requires the curation and orchestration of case-level data across a number of modalities: community engagement, demographics, autopsy procedure, photos, and specimen collection, clinical laboratory results, child and maternal medical record abstraction, pathology images and diagnostics, biorepository data, verbal autopsy data, and cause of death determinations with ICD 10 coding. Surveillance of any one of these data streams from multiple countries is a daunting challenge. The data flow diagram resulted in a unifying depiction of the overall project allowing, without sacrificing, the representation and acknowledgement of the seemingly disparate parts.

This holistic view of the system facilitated awareness amongst the working groups regarding where their data elements or components fit within the project and how their contributions impacted the work of downstream consumers of the data, ultimately impacting the final case file. This appreciation for the bigger picture and broader goal aided in cross-functional organization and operation of the project, expanding perspectives beyond the seemingly siloed components/artifacts. Those working to respond to reported cases, administer consent, and perform the specimen collection procedures, understood that timely collection of high-quality specimens was essential to accurate and timely laboratory and pathology testing. Those working to procure comprehensive medical record abstractions and timely verbal autopsy data, appreciated how this data informed and influenced additional diagnostic testing for each case. Ultimately, all data streams converged to form comprehensive case files that provided

clinical decision makers with essential case evidence necessary for determining a cause of death.

3.3.2 Change Management

An unexpected outcome of the development cycles was the value the documentation played in process and system change management. The data flow diagram is now leveraged as a tool for the evaluation and/or impact analysis of potential workflow changes. Both technical and non-technical surveillance staff are able to consider the impact of additional forms, form version changes, new diagnostics, or new data streams by visualizing and considering where the change may impact the targeted surveillance component(s) and how that change may impact the broader network.

3.4. Summary

In summary, the call to action to improvise a novel product and method to better communicate and elicit the operational and technical requirements of a complex and undefined surveillance initiative succeeded in its goal. The data flow diagram captured the surveillance system/network at the appropriate level of magnification and incorporated essential operational and technical elements that served and affected the widest range and diversity of stakeholders without sacrificing the informatic-integrity of technical information needed to support design, planning, and implementation of the surveillance network.

In late 2019 and early 2020, the CHAMPS program began planning for a new line of surveillance - Pregnancy surveillance - set to rollout to three CHAMPS locations in the

late summer and early fall of 2020. One of the first requests of the CHAMPS program team was to commission a draft data flow diagram based on the proposed protocol. This request speaks to the value and the appreciation of the tool and its perceived benefits and potential advantages for surveillance "system" planning beyond just the informatics field.

Chapter 4 Discussion

4.1 Introduction

The dataflow methodology and resulting technical artifacts described in this review have the potential to advance the public health informatics field by building on the CRDM collaborative design approach to produce more robust and comprehensive technical documentation that captures an enterprise-level view of a surveillance system/network. This innovative style of documentation represents the network of workflows, coupling them with the data collection mechanisms, information flow, and operational procedures that comprise the holistic body of a surveillance system of systems. The data flow diagram is better suited to communicate the whole informatic story via a visual and tactile medium that informs a more diverse range of stakeholders.

4.2 Summary of Project

The informatics team tasked with architecture, development, and implementation of the CHAMPS mortality surveillance protocol began requirements analysis leveraging the Public Health Informatics Institute's Collaborative Requirements Development Methodology (CRDM). The CHAMPS mortality surveillance protocol was a complex

endeavor seeking to bring together case-level demographic information, verbal autopsy data, clinical record abstraction data, maternal clinical abstraction data, autopsy procedure and specimen collection data, laboratory testing results, images, documents, and both local and centralized histopathology image and diagnostic data for under-five childhood deaths from multiple low-resource setting with seven sites in Africa and Asia. Application of the CRDM yielded a library of intricate workflows that collectively obfuscated the broader technical objectives of the surveillance project and were an impediment to the alignment of stakeholders on the full scope of the project (i.e., "could not see the forest for the trees"). The project described herein was to create a new format of technical documentation that would honor and include the fundamental business process work while simultaneously depicting the full scope surveillance activities, interconnections, dependencies, and general flow of data/information for each case from discovery to final cause of death determination.

4.3 Implications

The innovative technique of technical documentation developed and described within this paper has potential to change how both technical and nontechnical public health surveillance personnel view and consider surveillance systems documentation. The CRDM method recommends development of context diagrams that provide a high-level view of relationships and interactions amongst entities and are simple to follow. However, the context diagram lacks the level of specificity to capture the decision points or specific file exchanges. The CRDM workflow approach is helpful and appropriate at the task level. Unfortunately, in a large complex network, this level of detail is too granular and the resulting volume of workflows obscures the broader picture. Classic technical network architecture provides a wealth of technical infrastructure information

and detail, but lacks the operational context to resonate with non-technical audiences. The data flow diagram was intentionally designed and honed to communicate the appropriate ratio of operational/functional workflow to the amount of specificity/detail such that the document could inform the broadest audience of surveillance stakeholders, establishing a common denominator at the appropriate magnification. If refined and more broadly adopted, this data flow development methodology and the resulting artifacts could change how public health surveillance networks are designed, communicated, and comprehended. Moreover, this approach could benefit research or surveillance on a wide range of scales, from smaller projects (e.g., connections single reference lab to a local public health jurisdiction) to larger more complex multi-component international networks like CHAMPS.

While the actual implications of the data flow diagram on the broader public health informatics community remains to be seen, the CHAMPS program continues to utilize the data flow diagram as a principle reference document and planning tool. New staff are given the diagram to orient them to the surveillance operations. The diagram is used to train and orient technical informatics staff or technical contractors and is often referenced to think through potential system impacts or aid in troubleshooting issues. Updated versions of the diagram are requested by non-technical staff when standard operating procedures or case report forms are added, updated, or retired. A number of case report forms and procedures for the mortality protocol were upgraded or modified in 2019 and the data flow diagram has been updated to remain current. The diagram may be seen posted in staff offices across the network both stateside and abroad. The adoption of the data flow artifact into the fabric and operational culture of CHAMPS program by a variety of programmatic roles speaks to the success of this new instrument and serves as evidence that the key aims of this thesis project were met.

4.4 Limitations

Several limitations to this thesis project should be noted. This thesis project was charged with devising something new and innovative, as such, problem-specific informatic references for literature review was considerably limited. A broader more comprehensive and creative review of other industries (e.g., engineering, manufacturing, supply chain logistics) may yield more substantive guidance with potential application in the public health surveillance context. The thesis project was conducted on a single public health surveillance project which raises the issue of generalizability. The methods and products need to be applied on additional surveillance projects to further evaluate the merit and effectiveness of the approach. Subjective feedback gathered during the development iterations may have been affected by confirmation bias [13] whereby the desire and commitment to devise a better methodology and products in an applied rather than experimental setting fostered a bias in the perception of its effectiveness. Further research of the methods and products by a more objective party is required to ascertain the effectiveness without prejudice.

4.5 Recommendations and Next Steps

The dataflow methodology and resulting technical artifacts have potential to advance informatics practice by building on the Public Health Informatics Institute's Collaborative Requirements Development Methodology (CRDM), offering a more comprehensive next step for the operational and technical engineering of a surveillance network on a number of scales. Where the CRDM adds value by yielding public health workflows and defining functional requirements for information systems that support those workflows, the dataflow methodology contextualizes the workflows within the full

constellation of a surveillance system's network of systems, operational procedures, and work-streams of information to reveal a more comprehensive picture of the surveillance system at an enterprise perspective.

Potential next steps for advancing the data flow methodology within PHII include the formalization of the methods into more consumable and easily transferable instructional guidance followed by application of the method on additional informatics projects with the intent of further critical evaluation of both methods and end products. Direct benefits of these actions for PHII or other informaticians are an eventual refinement of the method from its originating form, such that the method is shareable to the broader public health informatics community and may be consistently repeatable. These actions set the stage for more objective study and evaluation by PHII or other researchers. Lastly, there is potential immediate value to the operators and stakeholders of the systems evaluated resulting from the system documentation curated within the data flow artifacts. This recommended approach aligns with the beginnings of PHII when in 2002 the institute developed and published the Collaborative Requirements Development Methodology™ (CRDM).

4.6 Conclusion

The purpose of this thesis project was to develop a new form of informatics diagram that captured the broader surveillance logistics of a complex global public health surveillance network by incorporating processes, study artifacts, standard operating procedures, and data exchange. This charge required the achievement of two key aims. The first aim was to devise and document a methodology for the development and refinement of novel technical diagram in support of a global case-based public health surveillance program.

The technique of iterative design sessions with stakeholders to solicit and apply feedback resulted in the creation of functional protypes. Continued application of this methodology on a broader and more diverse cohorts of stakeholders lead to the achievement of the second key aim of the project, an original type of technical documentation artifact for the global case-based public health surveillance program that could be generalizable for other case-based public health surveillance programs. The tangible manifestation of document itself coupled with the application of document in the field are evidentiary of the achievement of initial component of the second aim. The subsequent adoption of the data flow artifact into the fabric and operational culture of the CHAMPS program by a variety of programmatic roles speaks to the success of this new instrument within the CHAMPS program and serves as convincing circumstantial evidence for the potential generalizability of the diagraming technique for other surveillance initiatives.

Ultimately, a refined version of the novel data flow diagram artifact and a method of developed was established. The artifact had the appropriate balance of communicating both technical and non-technical system workflow consumable by a multitude of international stakeholders. The resulting technical document met the overall objective of providing a comprehensive depiction of the surveillance data flow that is informative, instructive, universally engaging, and could be applied to systems and programs of varied scale beyond the scope of this thesis project.

The general application and effectiveness of the innovative data flow methodology and technical diagram are apparent within the context of the CHAMPS surveillance program. The current data flow diagram and the development methods described herein shall serve as the foundation for the direct application and further research regarding the

efficacy of the tool for other public health surveillance initiatives. If refined and more broadly adopted, this data flow development methodology and the resulting artifacts could change how public health surveillance networks are designed, communicated, and comprehended.

5.0 References

1. Ahmed, K., Temate, Y., Tiagueu, J. A., Stover, D. L., Peters, P. J., Brooks, J. T., Sangareddy, S. P., & Dcruz, J. J. (2018). Integrating data from disparate data systems for improved HIV reporting: Lessons learned. *Online Journal of Public Health Informatics*, *10*(1).
2. Baker, E. L., Fond, M., Hale, P., & Cook, J. (2016). What Is "Informatics"? *Journal of Public Health Management and Practice*, *22*(4). https://journals.lww.com/jphmp/Fulltext/2016/07000/What_Is__Informatics__.15.aspx
3. Blau, D. M., Caneer, J. P., Philipsborn, R. P., Madhi, S. A., Bassat, Q., Varo, R., Mandomando, I., Igunza, K. A., Kotloff, K. L., Tapia, M. D., Johnstone, S., Chawana, R., Rahman, A., El Arifeen, S., Onyango, D., Kaiser, R., Seale, A. C., Assefa, N., Morris, T., … Breiman, R. F. (2019). Overview and Development of the Child Health and Mortality Prevention Surveillance Determination of Cause of Death (DeCoDe) Process and DeCoDe Diagnosis Standards. *Clinical Infectious Diseases*, *69*(Supplement_4), S333–S341. https://doi.org/10.1093/cid/ciz572
4. Chawana, R., Baillie, V., Izu, A., Solomon, F., Bassat, Q., Blau, D. M., Breiman, R. F., Hale, M., Houpt, E. R., Lala, S. G., Martines, R. B., Mathunjwa, A., Nzenze, S., Pathirana, J., Petersen, K. L., Raghunathan, P. L., Ritter, J. M., Wadula, J., Zaki, S. R., & Madhi, S. A. (2019). Potential of Minimally Invasive Tissue Sampling for Attributing Specific Causes of Childhood Deaths in South Africa: A Pilot, Epidemiological Study. *Clinical Infectious Diseases*, *69*(Supplement_4), S361–S373. https://doi.org/10.1093/cid/ciz550 *CitationList.ris*. (n.d.).
5. *Disease Surveillance Network in Africa and Asia*. (1 C.E., January 1). https://www.gatesfoundation.org/Media-Center/Press-Releases/2015/05/Child-Health-and-Mortality-Prevention-Surveillance-Network

6. Koo, D., O'Carroll, P., & LaVenture, M. (2001). Public Health 101 for Informaticians. *Journal of the American Medical Informatics Association*, *8*(6), 585–597. https://doi.org/10.1136/jamia.2001.0080585

7. Lee, L. M., Thacker, S. B., & Louis, M. E. S. (2010). *Principles and Practice of Public Health Surveillance*. Oxford University Press. https://books.google.com/books?id=FF78GCiUbwUC

8. Levine, M. M., Kotloff, K. L., Nataro, J. P., & Muhsen, K. (2012). The Global Enteric Multicenter Study (GEMS): Impetus, rationale, and genesis. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55 Suppl 4*(Suppl 4), S215–S224. PubMed. https://doi.org/10.1093/cid/cis761

9. Madhi, S. A., Pathirana, J., Baillie, V., Cutland, C., Adam, Y., Izu, A., Bassat, Q., Blau, D. M., Breiman, R. F., Hale, M., Johnstone, S., Martines, R. B., Mathunjwa, A., Nzenze, S., Ordi, J., Raghunathan, P. L., Ritter, J. M., Solomon, F., Wadula, J., … Chawana, R. (2019). An Observational Pilot Study Evaluating the Utility of Minimally Invasive Tissue Sampling to Determine the Cause of Stillbirths in South African Women. *Clinical Infectious Diseases*, *69*(Supplement_4), S342–S350. https://doi.org/10.1093/cid/ciz573

10. Morris, T., Caneer, J. P., Said, W., & Tchamako, J. (2016). *CHAMPS Information Technology Milestone Deliverable, A Description of the CHAMPS Network Solution Architecture* (Information Technology Plan Version 5.0; p. 140). Emory Global Health Institute, Emory University.

11. Public Health Informatics Institute. *Collaborative requirements development methodology (CRDM)*. 2016. Retrieved January 15, 2020 from, https//www.phii.org/crdm

12. Public Health Informatics Institute. *Taking Care of Business, A collaboration to define local health department business process*. 2006.Retrieved January 15, 2020 from, https://www.phii.org/resources/taking-care-business-second-edition

13. *Oswald ME, Grosjean S (2004). "Confirmation Bias". In Pohl RF (ed.).* Pohl, R., & Pohl, R. F. (2004). *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. Psychology Press.

14. Salzberg, N. T., Sivalogan, K., Bassat, Q., Taylor, A. W., Adedini, S., El Arifeen, S., Assefa, N., Blau, D. M., Chawana, R., Cain, C. J., Cain, K. P., Caneer, J. P., Garel, M., Gurley, E. S., Kaiser, R., Kotloff, K. L., Mandomando, I., Morris, T., Nyamthimba Onyango, P., … Child Health and Mortality Prevention Surveillance (CHAMPS) Methods Consortium. (2019). Mortality Surveillance Methods to Identify and Characterize Deaths in Child Health and Mortality Prevention Surveillance Network Sites. *Clinical Infectious Diseases*, *69*(Supplement_4), S262–S273. https://doi.org/10.1093/cid/ciz599

15. Yasnoff, W. A., O Carroll, P. W., Koo, D., Linkins, R. W., & Kilbourne, E. M. (2000). Public health informatics: Improving and transforming public health in the information age. *Journal of Public Health Management and Practice*, *6*(6), 67–75.

16. Informatics at CHAMPS – CHAMPS. (n.d.). Retrieved January 15, 2020, from https://champshealth.org/1901-2-2-2/

*17. Public Health Surveillance in the United States: Evolution and Challenges**. (n.d.). Retrieved January 15 2020, from https://www.cdc.gov/mmwr/preview/mmwrhtml/su6103a2.htm

18. Dixon, B. E., & Grannis, S. J. (2011). Why "What Data Are Necessary for This Project?" and Other Basic Questions are Important to Address in Public Health Informatics Practice and Research. *Online Journal of Public Health Informatics*, *3*(3).

19. Garcia, M. C., Garrett, N. Y., Singletary, V., Brown, S., Hennessy-Burt, T., Haney, G., Link, K., Tripp, J., Mac Kenzie, W. R., & Yoon, P. (2018). An assessment of information

exchange practices, challenges and opportunities to support US disease surveillance in three states. *Journal of Public Health Management and Practice: JPHMP*, *24*(6), 546.

20. Musabi, A. G., Thiga, M. M., & Karume, S. M. (n.d.). *Enabling Secure Maternal Health Information Exchange using Blockchain*.

21. Seymour, D., Werner, L., Mwansa, F. D., Bulula, N., Mwanyika, H., Dube, M., Taliesin, B., & Settle, D. (2019). Electronic immunization registries in Tanzania and Zambia: Shaping a minimum viable product for scaled solutions. *Frontiers in Public Health*, *7*, 218.

22. Vielot, N. (2014). Improving Public Health Preparedness: Strengthening Biosurveillance Systems for Enhanced Situational Awareness. *Online Journal of Public Health Informatics*, *6*(1).

23. Zaffran, M., Vandelaer, J., Kristensen, D., Melgaard, B., Yadav, P., Antwi-Agyei, K., & Lasher, H. (2013). The imperative for stronger vaccine supply and logistics systems. *Vaccine*, *31*, B73–B80.