CANCER REPORTING: TIMELINESS ANALYSIS AND PROCESS

REENGINEERING

Abdulrahman M. Jabour

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in the School of Informatics and Computing

Indiana University

June 2016

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Josette Jones, PhD, RN, Chair

Doctoral Committee

_____

Brian Dixon, PhD, MPA

November 9, 2015

_____

David Haggstrom, MD, MAS

_____

Bolchini, Davide, PhD

DEDICATION

I dedicate this dissertation to my parents, Mohamed and Maryam, for their inspiration and encouragement. I also dedicate this dissertation to my wife, Alanoud, for her unwavering support and to my wonderful kids, Yazeed and Dana, for being a source of joy and hope for the future.

ACKNOWLEDGMENT

Abdulrahman M Jabour

CANCER REPORTING: TIMELINESS ANALYSIS AND PROCESS

REENGINEERING

Introduction: Cancer registries collect tumor-related data to monitor incident rates and support population-based research. A common concern with using population-based registry data for research is reporting timeliness. Data timeliness have been recognized as an important data characteristic by both the Centers for Disease Control and Prevention (CDC) and the Institute of Medicine (IOM). Yet, few recent studies in the United States (U.S.) have systemically measured timeliness.

The goal of this research is to evaluate the quality of cancer data and examine methods by which the reporting process can be improved. The study aims are: 1- evaluate the timeliness of cancer cases at the Indiana State Department of Health (ISDH) Cancer Registry, 2- identify the perceived barriers and facilitators to timely reporting, and 3-reengineer the current reporting process to improve turnaround time.

Method: For Aim 1: Using the ISDH dataset from 2000 to 2009, we evaluated the reporting timeliness and subtask within the process cycle. For Aim 2: Certified cancer registrars reporting for ISDH were invited to a semi-structured interview. The interviews were recorded and qualitatively analyzed. For Aim 3: We designed a reengineered workflow to minimize the reporting timeliness and tested it using simulation.

Result: The results show variation in the mean reporting time, which ranged from 426 days in 2003 to 252 days in 2009. The barriers identified were categorized into six themes and the most common barrier was accessing medical records at external facilities.

We also found that cases reside for a few months in the local hospital database while waiting for treatment data to become available. The recommended workflow focused on leveraging a health information exchange for data access and adding a notification system to inform registrars when new treatments are available.

Josette Jones, PhD, RN, Chair

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ACoS	The American College of Surgeons
ACS	American Cancer Society
AHRQ	Agency for Healthcare Research and Quality
BPR	Business Process Reengineering
CDA	Clinical Document Architecture
CDC	Centers for Disease Control and Prevention
CDC/NPCR	Centers for Disease Control and Prevention/ National Program of Cancer Registries
CoC	Commission on Cancer
DES	Discreet Event Simulation
Disease Index or Indices Codes	Diseases and conditions identified in medical records coded as ICD-9-CM or CPT
DQ	Data Quality
EHR	Electronic Health Record
EMR	Electronic Medical Record
E-Path	Electronic Pathology Laboratory Reporting System
FORDS	Facility Oncology Registry Data Standards
FTP	File Transfer Protocol
HIE	Health Information Exchange
HIM	Health Information Management
HIMSS	Health Information Management Systems Society

HL7	Health Level Seven
MRN	Medical Records Number
NLC	National Learning Consortium
SD	System Dynamic Models
SEER	Surveillance, Epidemiology, and End Results Program
ICD	International Classification of Diseases
ICRA	Indiana Cancer Registrars Association
INPC	Indiana Network For Patient Care
IOM	Institute of Medicine
ISDH	Indiana State Department of Health
NAACCR	North American Association of Central Cancer Registries
NCI	National Cancer Institute
NPCR	National Prostate Cancer Register
VPN	Virtual Private Network

CHAPTER 1: INTRODUCTION

1.1 Introduction.

Data are generated, captured, stored, shared, and managed by people, organizations, and systems that are in pursuit of information, knowledge, and wisdom. Data are considered to be of high quality if they are a fit for their intended use by data consumers (Wang & Strong, 1996). Data quality (DQ), a sub-discipline within the broader fields of computer science and informatics, is concerned with the study and evaluation of the quality of data. The discipline defines the concept of DQ as a complex construct that is composed of multiple dimensions. Some of the dimensions that are frequently discussed with respect to health data include accessibility, accuracy, completeness, and timeliness (Weiskopf & Weng, 2013).

DQ is essential for the efficiency and effectiveness of healthcare organizations. The quality of data that are used for evaluating and monitoring population health underlies the quality of decision-making processes that rely upon those data. For example, cancer registry data are used to guide planning and evaluation programs for cancer prevention. This includes the assessment of prevention, screening, and treatment programs that guide decisions for health resource allocation (ACS, 2014).

A dimension of DQ that is commonly examined and is important for the evaluation and assessment of cancer registries and programs is timeliness—the quality of

arriving or being ready on time (e.g., when consumers desire or need data). The importance of collecting timely data is well documented in the Institute of Medicine (IOM) and the Centers for Disease Control and Prevention (CDC) reports (Centers for Disease Control and Prevention, 1999; Simone & Hewitt, 2000).

Timeliness standards used to maintain and monitor reporting are applied at the state and national levels. For example, at the state level, the Indiana State Department of Health (ISDH) requires health facilities to report cancer cases regularly at different frequencies based on the average number of cases diagnosed annually (Table 1-1) (ISDH). Currently, hospitals with an average of 1 to 59 cases annually are required to report their cases once each year, while hospitals with an average of 300 or more cases annually are required to report them on a monthly basis (ISDH). As is the practice with most cancer registries, records in the ISDH Cancer Registry are reported from different sources using different intervals.

Annual caseload	Frequency of reporting
1-59	Once per year
60-149	Quarterly
150-299	Every other month
300 and more	Every month

Table 1-1: Reporting requirement based on facility caseload.

Adopted from: *Policy and Procedure Manuals for Reporting Facilities*. Indiana State Department of Health Cancer Registry. May 2015.

As with most evaluation programs, establishing an objective standard for timeliness in measurement is fundamental. At the national level, timeliness requirements vary. As seen in (Table 1-2), while the Surveillance, Epidemiology, and End Results (SEER) program requires 98% of cases to be reported within 22 months, the Centers for Disease Control and Prevention, National Program of Cancer Registries (CDC/NPCR) require 95% of cases be reported within 24 months. Other registries, such as the North American Association of Central Cancer Registries (NAACCR), require 95% of cases be reported within 23 months from the date of diagnosis for Gold Certification and 90% for Silver Certification. Therefore, to meet the SEER requirements, both reporting hospitals and state registries should take fewer than 23 months to report. Because they are part of a larger system, state registries need to meet the international standards of registries, such as SEER, CDC/NPCR, and NAACCR. Currently, the ISDH Cancer Registry participates in the CDC/NPCR and the NAACCR programs. Since the ISDH Cancer Registry is not part of the SEER program, meeting its standard is not required.

Standard	Percentage of cases required within:			
	12 months	22 months	23 months	24 months
CDC/NPCR	90%		90%	95%
NAACCR Silver			90%	
NAACCR Gold			95%	
SEER		98%		

Table 1-2: National requirements for cancer registries.

Adopted from: NPCR Education and Training Series (NETS) Module 3: Quality Control for Central Registries. By Scientific Applications International Corporation (SAIC), CDC/NPCR.

1.2 Problem Statement

Timely reporting of cancer data is an important aspect of cancer surveillance, care delivery, and research (Simone & Hewitt, 2000). Unfortunately, studies have shown that reporting delays can, in some instances, take up to 5 years (Boscoe, 2014; Clegg, Feuer, Midthune, Fay, & Hankey, 2002). Avoiding delays in reporting plays a critical role in the accuracy of cancer incidence rates and the availability of data for cancer research. Longer delays in reporting make monitoring trends of cancer cases more difficult (Midthune, Fay, Clegg, & Feuer, 2005). Delays in reporting can lead to underestimations of cancer incidence rates by causing inaccurate decline signals (INDH; Smith-Gagen, Cress, Drake, Felter, & Beaumont, 2005). False decline signals can be observed with incomplete data

for recent years compared with complete data from prior years (Midthune et al., 2005). With cancer care delivery, reporting delays also hinder timely intervention, as most intervention programs are data-driven. Likewise, reporting delays can prevent access to recent data that is necessary for research. Assessments of prevention, screening, and treatment programs are more valuable if they are presented in a timely manner (Meyer, Carpenter, Abernethy, Stürmer, & Kosorok, 2012).

Local hospitals and treating facilities often collect and store data in local registries, and subsequently report them to state registries. State registries report cancer cases to larger registries, including SEER, NAACCR, and the CDC/NPCR. The process of reporting from one registry to another may lead to delays in final data reporting. As a result, timeliness remains a central, prevalent issue in accessing and using cancer registry data.

1.3 Gap in Knowledge

Several IOM and CDC reports have called for improving the timeliness of cancer case reporting; however, the issue has rarely been investigated (Centers for Disease Control and Prevention, 1999; Simone & Hewitt, 2000). While a handful of studies in Europe have evaluated the timeliness of cancer registry data, their findings are not relevant, as the U.S. has established its own cancer registry standards. Within the U.S., a study by Gagen and Cress (2005) investigated the factors associated with delays in reporting by examining the association between reporting delays and gender, race, type of

facility, cancer site, and/or stage at diagnosis (Smith-Gagen et al., 2005). The study reported that timeliness varies between reporting facilities and that geographical region and facility type (e.g., hospitals, laboratories, or the physician's office) have the greatest impact on timeliness.

To our knowledge, all previous studies investigated timeliness from a DQ perspective. None of these studies has investigated the reporting process or the complex series of steps that are involved in the reporting process from the time when the cancer case is first identified until the state cancer registry certifies it is complete. The purpose of this study is to evaluate timeliness as well as the detailed processes that are involved in cancer case reporting, explore the barriers and facilitators to timely reporting, and re-engineer the current reporting process for cancer registries.

1.4 Reporting Process

Cancer reporting involves a series of chronological tasks. Completion of each task depends on completion of the previous one; delays in completion of one task will result in delays in subsequent tasks and will, thus, delay the overall reporting time. Whereas previous studies have evaluated the timeliness of cancer reporting by measuring the reporting time as a single process from beginning to end, availability of the date and time stamp data that can support examination of different stages in the reporting process is unknown. Separating the reporting process into multiple steps will provide more detailed information about reporting at the state level. A review by Jajosky and Groseclose (2004)

6

showed that surveillance timeliness studies lacked detailed descriptions of reporting stages, such as processing and analyzing (Jajosky & Groseclose, 2004). The study also emphasized that the collection and assessment of time intervals is a very important part of surveillance systems and timeliness assessment (Jajosky & Groseclose, 2004).

Process improvement requires the right information to identify challenges and devise ways to overcome them. Understanding the details of the reporting process will aid in guiding the future development of cancer reporting technologies by identifying gaps and prioritizing needs. In addition to analyzing time intervals of the subtasks involved in the reporting process, this study also includes a field interview for a deeper understanding of cancer reporting workflows.

Searching for patients' records can be a complex process. Data is often stored within different systems that have varying standards. The timely collection of a comprehensive record can be very challenging. Workflow analysis and process reengineering techniques are commonly used for optimizing workflow and reducing its complexity. Workflow analysis and process reengineering techniques have been applied in many healthcare settings, including pharmacies and emergency departments (Bertolini, Bevilacqua, Ciarapica, & Giacchetta, 2011; Chou et al., 2012; Leu & Huang, 2011).

1.5 Aims

The goal of this research is to evaluate the quality of cancer data with respect to timeliness and examine methods by which the reporting process can be improved. The study has the following aims:

Aim 1: Evaluate the timeliness of cancer reporting at the Indiana State Department of Health (ISDH) Cancer Registry.

Several studies have demonstrated the lengthy time period facilities have taken to report cancer cases to state registries (ISDH, 2015; Izquierdo & Schoenbach, 2000; Midthune et al., 2005). However, states vary in their reporting performance. Currently, the timeliness of cancer reporting in Indiana is unknown. Evaluating the timeliness of ISDH Cancer Registry data will inform its users regarding the quality of data and how this data compares to other national standards. The comparison may include the following programs: SEER, NAACCR, and CDC/NPCR.

Aim 2: Identify the perceived barriers and facilitators to timely reporting.

Despite the magnitude of the problem, barriers and facilitators to timely reporting have rarely undergone systematic investigation. This study will explore the tasks that are involved in the cancer reporting workflow, as well as barriers and facilitators that are

involved with each task. Knowing the barriers and facilitators is important for analyzing and identifying potential solutions.

Aim 3: Reengineer the current reporting process to improve timeliness.

The information from Aim 2 will be employed to inform process reengineering in Aim 3. The goal of reengineering the reporting process is to design an alternative reporting method that minimizes delay. Accomplishing Aim 3 will include the following:

- Model the current workflow and identify tasks that generate the most delays in the reporting process.

- Identify tasks within the reporting process that can be modified to optimize reporting time.

- Propose alternative reporting processes and compare them with the current process using workflow simulation.

1.6 Study Design Overview

This study was conducted using two sources of data: date and timestamp data from cancer cases reported to the ISDH Cancer Registry and interviews with cancer registrars. Using triangulation techniques provides more insight into the reporting process and overcomes some of the limitations that might be found in one of the sources. Using the date and time stamps recorded to report cases provides an overview of the reporting process at the state level for a long period of time (9 years). The interviews inform the

details of the reporting processes and shed light on practices that could not be answered through analysis of temporal data, alone. Having a comprehensive understanding enables redesigning of the reporting process to optimize timeliness. Following the construction of an alternative workflow, we performed a simulation study to compare the alternative process with the existing one.

1.6.1 Timeliness analysis

Examination of the delay in reporting is conducted by analyzing the timestamps of cancer cases reported in the ISDH Cancer Registry. Overall timeliness is calculated by measuring the number of days from the date of diagnosis to the date when the data becomes available for reporting by the registry. In addition to the overall timeliness assessment, the study analyzes the dates and timestamps that are associated with different phases within the reporting process. This includes comparing the time taken by the reporting facility to abstract the case report and send it to the registry, to the time taken by the cancer registry to process the report and upload it to their database (Figure 1-1). Tasks within the facility reporting process are also examined. These tasks include the time taken to initiate the abstract, complete the case report, and export and send reports to the registry (Figure 1-1).
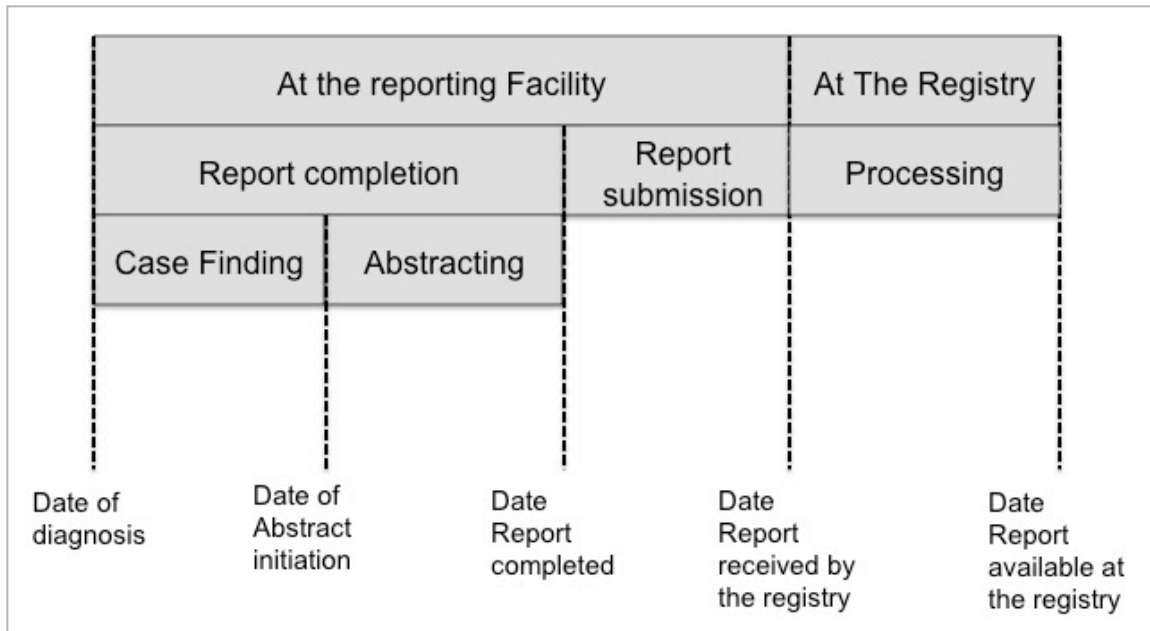
Figure 1-1: Detailed view of the reporting process with key dates associated with different stages within the reporting process.

1.6.2 Interview and field observation

To explore the reporting workflow, cancer registrars from major hospitals within Indiana were interviewed. The goal of the interviews was to understand the workflow and barriers encountered during reporting. The interviews were in-depth, semi-structured, and task-oriented (Appendix 6.1). The interviews were recorded and later transcribed. Throughout each interview, the researcher collected the information needed to develop the process model, analyze the workflow, and simulate the process.

The information collected at the interview was analyzed at two stages: thematic analysis and workflow modeling. The thematic analysis was conducted based upon

grounded theory, wherein empirical observations were derived directly from data in order to explain barriers and facilitators to timeliness. The key ideas were coded and classified.

The workflow model was then developed to describe the stage of the process delays. Re-engineering principles and methods were then applied to identify non-value-added activities and develop a new way to construct the process (Davenport, 2013; Hammer, 1990).

1.6.3 Simulation

This study also simulated the current process to offer insights into its capacity and performance. After development of the new process, simulation was employed to compare the generated process with the existing process (Curry & Prodan, 2013). The simulation modeled the current reporting process as a baseline and compared its performance with the suggested process.

CHAPTER 2: LITERTURE REVIEW

2.1 Introduction

The goal of this research was to evaluate the quality of cancer data and examine methods for improving the reporting process. In order to accomplish this, we followed three steps: measure the timeliness of the cancer registry data, identify barriers within the reporting process, and provide recommendations related to process redesign and simulation. This chapter will review the previous studies and applications related to these main phases. Descriptions of each phase can be seen in the following table:

Phase	Goal	Theme
Phase 1: Data Timeliness in Cancer Registries	Evaluate the current status	Examine studies related to DQ in cancer registries. In this section, a particular focus will be on timeliness and methodological techniques.
Phase 2: Workflow in Healthcare	Process understanding	Examine studies related to workflow and process redesign in healthcare.
Phase 3: Simulation in Healthcare	Redesign and testing	Examine studies that use simulation in healthcare.

Table 2-1: Overview of the literature review sections.

2.2 Data Timeliness in Cancer Registries

2.2.1 Data quality

Data quality (DQ) has been defined as "the totality of features and characteristics of an entity that bears on its ability to satisfy stated and implied needs" (Arts, De Keizer, & Scheffer, 2002). In healthcare and public health, the quality of data is important to many cases. Incomplete data can, for example, impact health outcomes if patients and clinicians do not have all relevant data when making decisions. In public health, epidemiologists and medical officers need complete, accurate data when planning community-level interventions for cancer prevention. Finally, DQ impacts research, as incomplete or inaccurate data will affect the findings of a study.

Cancer registries are important sources of data in cancer-related studies. Essentially, registries provide statistics about populations and assist researchers monitor trends in cancer cases (Clegg et al., 2002). With respect to cancer, DQ generally needs to meet certain standards and expectations for reporting. For instance, the NAACCR sets standards and quality attributes for data in cancer registries (2014).

2.2.2 Timeliness

An important factor for ensuring the quality of data is timeliness in reporting. Timely reporting supports assessment and quality improvement programs, as well as

epidemiological studies, by minimizing turnaround time. Delays in reporting undermine the quality of results by underestimating the incidence rate, which results in an inaccurate decline signal (Clegg et al., 2002). Longer delays in reporting make monitoring trends in cancer cases more difficult (Midthune et al., 2005). Similarly, such delays limit the accessibility of timely data required for research. According to the IOM, it is important to establish ways of improving the timeliness of cancer data (Simone & Hewitt, 2000). To explore ways that timeliness could be examined, we reviewed published studies related to timeliness evaluation in cancer registries. The review explored the type of studies, evaluation method, and outcome.

2.2.3 Previous studies

Norway conducted a study to evaluate DQ of the Norway Cancer Registry. The study evaluated data comparability, completeness, accuracy, and timeliness from 1953–2005. The timeliness analysis was limited to 2001 and 2005 (Larsen et al., 2009) and described timeliness as the "time from diagnosis to registration" and "the time from registration to the reporting of incidence via the annual report" (Larsen et al., 2009). The study assessed the timeliness of cases diagnosed in 2005 by comparing the number of cases reported (published) during November 2006 (one year after diagnosis) and during November 2007 (two years after diagnosis); it reported a difference of about 2.2%, indicating an underreporting in the one year after diagnosis publication. The study reported that the average time taken from diagnosis to registration was 525 days during 2001 and 261 days in 2005. Given the significant improvement in timeliness reported

during the study period, the study predicted greater improvement in the future due to use of electronic data reporting (Larsen et al., 2009).

Another DQ study was conducted to evaluate the Icelandic Cancer Registry data. The study evaluated data comparability, validity, timeliness, and completeness in 41,994 cases between 1955–2009. In terms of timeliness, the study examined cases from 2007 to 2011 (Sigurdardottir et al., 2012). It examined the time taken from the date of diagnosis to the time when the data was reported to the registry, and the time taken by the registry to process the data and make it available for reporting. During 2007, the facility took an average of 238 days to report cases to the registry. However, this time ranged widely from 49 to 1,445 days. The study reported that 85% of the cases diagnosed were reported to the registry within one year. The study also reported that it would take two years for 96.9% of the cases to be available in the registry (Sigurdardottir et al., 2012).

Using data from 9 states' cancer registries within the U.S., Limin and Clegg (2002) examined the impact of reporting delays and errors on incidence rates and trends. The goal of this study was to develop a statistical model to adjust for reporting delays in cancer rates and to predict future corrections, including both additions and deletions (Clegg et al., 2002). In this context, the reporting delay was defined as elapsed before a diagnosed cancer case was reported to the registry. The standard delay time selected was 2 years. The study included the following cancer types: female breast, colorectal, lung/bronchus, prostate, and melanoma. The study analyzed the reporting of adjusted cases from 9 North American states' cancer registries that participated in the

Surveillance, Epidemiology, and End Results (SEER) program from 1981 to 1998. The participating states included Connecticut, Hawaii, Iowa, New Mexico, and Utah, and the metropolitan areas of Atlanta, Detroit, Seattle-Puget Sound, and San Francisco-Oakland. The study measured the delay by calculating the difference in completeness rates during each diagnosis year. The delay distribution was considered completed after 19 years (Clegg et al., 2002). The study reported that 88% to 97% of the diagnosed cases were reported within the first 2 years. The study also reported that it would take 4 to 17 years for 99% of the cases to be reported (Clegg et al., 2002). Based on the cancer site, the rate of reporting delay could vary widely. Since this study revealed that reporting can take up to 17 years, an important factor to consider is the period during which the data is evaluated. The study evaluated data from 1981 to 1998, and thus, many changes were introduced to the field since that time period, such as increasing the adoption of Electronic Health Records (EHR) and the use of Electronic Pathology Reporting Systems (E-path).

Gagen and Cress conducted a more recent study in the U.S. in 2005 to examine the timeliness of the California Cancer Registry's data. The goal of the study was to explore the factors associated with reporting delays at the population level (Smith-Gagen et al., 2005). The study included cases diagnosed between January 1 and December 31, 2000. The data was retrieved 4 years after diagnosis (August 2004) and included the following cancer types: breast, prostate, colorectal, lung, non-Hodgkin's lymphoma, leukemia, melanoma, corpus, bladder, and pancreas (Smith-Gagen et al., 2005). The study defined timeliness as the interval between the date of diagnosis and the date the

17

case was available for research at the regional registry. The analysis split timeliness into the following 2 parts: the first part was the time from diagnosis to when the case was received by the registry, and the second was the time taken at the registry to make the data available for research. The analysis examined the variation in timeliness based on the following variables: type of reporting source, type of cancer, stage, gender, and age. The study applied univariate and multivariable analysis to examine the effects of these variables on timeliness (Smith-Gagen et al., 2005).

The study showed that 45% of the diagnosed cases were available at the registry within 12 months from the date of diagnosis and 96% were available within 24 months. The study showed that the average timeliness was 382 days. This included both the time taken by the facility to send the reports and the time taken by the registry to process the reports and make data available for research. The average time taken by facilities to send the reports was 268 days. This time was slightly shorter among hospitals than non-hospital facilities. The median number of days taken by the registry to process the cases and make them available for research was 69 days; however, this time varied widely, ranging from 32 to 208 days. The study also indicated that timeliness varied by geographical region, but no significant variation was related to gender, age, or stage of cancer (Smith-Gagen et al., 2005).

Tomic and Sandin, et al. (2015) conducted a study in Sweden to assess completeness, timeliness, validity, and comparability of the National Prostate Cancer Register (NPCR). Both completeness and timeliness were evaluated by cross-linking the

2 data sources: Swedish Cancer Register (SCR) and NPCR. The 2 data sources collect data differently from the diagnostic pathology department and from the treating clinical department. By linking both of these sources, the study compared the date of registration and registration status at the NPCR of Sweden with data in the SCR.

The study reported noticeable improvement in timeliness after 2007, which might be attributed to the introduction of electronic reporting. In 2008, for instance, 34% of the cases were reported within 3 months, 51% within 6 months, and 77% within 12 months. In 2012, 45% of the cases were reported within the first 2 months, 76% within 6 months, and 95% within 12 months. The study also reported a shorted timespan in NPCR compared to the other cancer registries, such as the Nordic National Cancer Registers, Icelandic Cancer Register, or Norwegian Cancer Register (Tomic et al., 2015).

2.2.4 Measuring timeliness

When evaluating timeliness, some studies report timeliness as the average number of days from the date of diagnosis to the date when data is available for research. This includes the sum of both the time taken by the facility to abstract and send the report and the time the registry needs to process the report and make it available for publication. However, other studies separate these 2 periods to distinguish between the time taken by the reporting facilities and the time taken by the registries.

Hunt (2004) outlined 2 approaches for timeliness measurement (National Cancer Registrars, 2004). The first approach calculates the proportion of cases abstracted in a given time of the year, as compared with the number of cases anticipated in that year. This percentage is then compared with the amount of time that has elapsed to date in the current accession year, minus the allowable reporting timeframe (National Cancer Registrars, 2004). The second approach calculates the difference between the date of diagnosis and the date of data entry. Some of the common data items employed in this approach are date of diagnosis, date of admission or first contact, and date when the records are transferred to the registry (National Cancer Registrars, 2004).

2.2.5 Standards, technology, and timeliness

Another factor that could potentially influence the timeliness of states' registries reporting is the standards and regulations of the national registries in which the states participate. Most states' registries are required to report their cases to larger national registries, because national registries have certain standards and regulations regarding the timeliness of the data received. For instance, SEER requires all cases be reported within 23 months of the date of diagnosis. While, on the other hand, the CDC/NPCR requires 90% of cases be reported within 12 months of the date of diagnosis and 95% of the cases be reported within 24 months (Larsen et al., 2009).

Other standards, such as those outlined by NAACCR, have different levels of certification for state registries. To meet the NAACCR golden certification, states need to

report 95% of the cases within 23 months of the date of diagnosis. On the other hand, to meet NAACCR silver certification, states need to report 90% of their cases within 23 months of the date of diagnosis (NAACCR, 2014).

Another factor that could impact reporting timeliness is the use of health information technologies. Larsen and Smastuen (2009) reported a 50% improvement in the Norway Cancer Registry's reporting time from 2001 to 2005. The reported improvement was attributed to the introduction of electronic data reporting (Larsen et al., 2009). Tomic and Sandin (2015) reported a similar finding, indicating that an improvement in timeliness coincided with the introduction of electronic reporting in 2007.

2.2.6 Reporting process and timeliness

Cancer reporting involves a series of interdependent tasks. The completion of each task depends on completion of the previous one; delays in the completion of one task will result in a delay in subsequent tasks and, thus, a delay in the overall reporting time.

Often, the data collected by local hospitals and other treating facilities is stored in the hospital's local registry, which is then reported to state registries. After that, state registries report to larger registries, including the NAACCR and CDC/NPCR. The process of reporting from one registry to another may lead to a delay in final data

reporting. Moreover, errors that may exist with certain tasks will be reflected in an overall report. Most of the reviewed studies evaluated the timeliness in cancer reporting by measuring the reporting time as a single process from the date of diagnosis to the date when the data is made available at the registry. Some separated the facility reporting time from the registry processing time. However, none of the studies examined the various stages of reporting within hospitals. The ability to examine the different stages of the process remains unknown. Separating this process into multiple steps will provide more detailed information about reporting at the state level. A review conducted by Jajosky and Groseclose (2004) reported that surveillance timeliness studies lack detailed descriptions of reporting stages, such as processing and analyzing. The study also emphasized that collection and assessment time intervals are an important part of surveillance systems and timeliness assessment (Jajosky & Groseclose, 2004).

Improving timeliness requires understanding the reporting process. Understanding the reporting process can help to identify gaps, prioritize needs, and support the future development of cancer reporting technologies. Prior studies lacked details of the reporting process and factors that may contribute to lengthy reporting; therefore, the issue remains under-investigated.

2.3 Workflow in Healthcare

The Agency for Healthcare Research and Quality (AHRQ) defines workflow as a "series of steps performed by different staff members, and often dependent on related

workflows, that accomplishes a particular task" (Aggarwal, Backman, Sager, & Sannyasi, 2010). Given the complexity of each organization, the efficiency of workflow is determined by the system as a whole. Like most fields, the use of information technology in healthcare is rapidly increasing, and the changes induced by them are improving how healthcare organizations function. Adapting to these changes often requires a consistent evaluation of process and reengineering to ensure integration of the organization's component parts (Champy & Greenspun, 2014).

Healthcare organizations are complex systems and the integration of organizational, human, and technical elements plays a critical role in their success and/or failure. The increasing demand for better service and the recent introduction of health information technology makes it more important than ever to systemically evaluate and analyze an organization's workflow and redesign it to its full potential. Knowledge from different domains, such as design, human-computer interaction, and business management, are increasingly applied in healthcare to cope with the rapid change induced by health information technology (HealthIT, 2015).

2.3.1 Workflow and process redesign theories

Many theories support workflow and process redesign. Theories like Six-Sigma, LEAN, Deming's Cycle of Continuous Improvement, Total Quality Management (TQM), and Business Process Reengineering (BPR), are well known for increasing performance and productivity (Aggarwal et al., 2010). Most of these theories were designed for the

purpose of enhancing performance; therefore, some put more emphasis on identifying inefficiencies within the process, where others are more focused on providing a solution. Other workflow and process redesign theories have some differences and similarities in how they function and perform.

Overall, most workflow and process redesign theories will include techniques for understanding the system being studied, analyze the system, provide an alternative, and implement the alternative (Table 2-2). Six-Sigma, for instance, contains the following 5 steps: define, measure, analyze, improve, and control (Aggarwal et al., 2010). Similar approaches can be seen with Deming's Cycle of Continuous Improvement, which contains these 4 steps: plan, do, check, and act (Stankard, 2002). Another example is the LEAN system, which involves the following steps: identify the key activities that create value, determine the current state, modify the current state by eliminating non-value activities, implement changes, and count improvement (Aggarwal et al., 2010).

	Six Sigma	Deming's Cycle of Continuous Improvement (Standard, 2002)	LEAN System
Step 1	Define: Define the problem, project goal, stakeholders' requirement.	Plan: Study the current situation. Develop theory about improvement. Design a test of theory.	Identify the key activities that create value.
Step 2	Measure: Develop and implement a data collection plan.	Do: Run a test of theory on the pilot.	Determine the current state.
Step 3	Analyze: Analyze data. Identify the issues within the process.	Check: Study how the actual result compare with the anticipated results and learn.	Modify the current state by eliminating the non-value activities.
Step 4	Improve: Address the issues, and eliminate the causes of dysfunction.	Act: Expand the use of what worked and eliminate what didn't.	Implement changes.
Step 5	Control: Control the new plan and create a future plan.		Count improvement.

Table 2-2: Overview process and redesign theories.

As seen in these examples, there are common, fundamental steps that seem essential for most workflow improvement activities. Those steps are: 1) information gathering, 2) analyzing the information and understanding the existing status, 3) redesigning the workflow, and 4) implementing the new workflow. The following section will explore these fundamental steps and review some of the applications to perform each of them.

2.3.2 Information gathering

A study's purpose and context often determines the design of a workflow study. As stated previously, healthcare organizations vary widely from one department to another; given this diversity, data collection techniques may also differ. In general, assessment studies can generate data by a variety of techniques, such as system-generated data, surveys, interviews, and/or observation (Unertl, Novak, Johnson, & Lorenzi, 2010). As there is no rule for which approach is ideal, some studies use a combination of techniques. Commonly, they take a qualitative empirical approach that seeks understanding, rather that confirming a hypothesis. For example, a systematic review conducted by Unertl and Novak (2010) found that the majority of workflow studies were qualitative (65 citations), as compared to quantitative (13 citations) and the mixed approach (35 citations). Their review also concluded that the most frequent technique was ethnographic observation (65 citations), followed by interview (58 citations), and artifact collection (29 citations) (Unertl et al., 2010). Lastly, their study found that within the healthcare setting, the most frequently studied subjects were nurses (51 citations) and physicians (45 citations), followed by other healthcare staff, such as administrative staff, pharmacists, and laboratory and radiology technicians (25 citations). Where some studies used saturation to decide the number of subjects, no further information was given regarding the number of subjects or sites studied (Unertl et al., 2010). Most were conducted over a period of weeks or months.

2.3.3 Workflow analysis

Workflow analysis is key to understanding a process and potential problems. Workflow analysis is performed before redesigning or proposing changes. This allows for the understanding of workflow to identify the tasks that need to be redesigned. The goal of this stage is often to identify what parts need to be redesigned and how critical they are. One of the most commonly applied methods of workflow analysis is workflow mapping.

Workflow mapping is a diagraming technique performed by creating a graphical illustration of the workflow process to determine the big picture and how the different elements of the system interact with one another. The goal of this step is to make the invisible visible. It clarifies the interaction between tasks and enables designers to communicate their ideas with stakeholders. With the introduction of health information technologies, workflow mapping has become more popular in health organizations. It facilitates an understanding of the changes induced by the system and allows decision makers to find ways to adjust the system to the work environment. Organizations, such as Healthcare Information and Management Systems Society (HIMSS) and AHRQ, are now providing toolkits for EHR workflow analysis and mapping (AHRQ, 2015; HIMSS, 2015).

The National Learning Consortium (NLC) lists 3 levels of workflow mapping based on the level of detail included in the mapping work. Those levels are macro (higher

level), mini, and micro (lowest) (National Learning Consortium, 2012). Depending on the level of detail and the amount of information collected during the information-gathering phase, this step can be very specific, in terms of reflecting every detail of the system process, or it can be very abstract, pointing out only the basic elements. Given the informal nature of workflow studies, such details are often determined by the study purpose and context.

2.3.4 Process redesign

Workflow redesign is often conducted in an informal manner, where the designer analyzes the workflow and brainstorms ways to improve the current process. The efficacy of this approach depends on the designer's creativity and experience, as well as the domain and field of study. Different backgrounds may emphasize different areas of concern and/or they may apply different techniques for redesign. Designers with a psychology background, for instance, will have a different perspective from designers with an engineering background. A review study by Unertl, el al. (2010) further exemplifies this idea. The review included 127 workflow and redesign studies. By categorizing studies based on the research perspective, the review showed 14 different areas of research, ranging from management, computer science, and health services, to anthropology and sociology (Unertl et al., 2010). One of the most advanced fields in workflow and process redesign is business and supply chain management.

One of the most popular workflow redesign techniques was established by Micheal Hammer in 1990; it was known as Business Process Reengineering (BPR) (Hammer, 1990). The principles of BPR aim to identify the underlying causes of process deficiencies and dysfunctions (Table 2-3).

- Organize around outcome, not tasks.

- Have those who use the output of the process perform the process.

- Subsume information processing work into the real work that produces the information.

- Treat geographically dispersed resources as though they were centralized.

- Link parallel activities instead of integrating their results.

- Put the decision point where the work is performed.

- Capture information once and at the source.

Table 2-3: BPR principles.

These principles are designed to fundamentally change and recreate how organizations work, instead of offering minor modifications to fix individual tasks. A key advantage of such an approach is that it deals with the system as a whole and considers all tasks involved when proposing changes. This allows for more integration between system parts and minimizes pitfalls induced by changing one of the system's elements. However, such a big change can be hard to implement. Introducing fundamental change can create people resistance and require much effort when dealing with human and organizational factors.

The LEAN system offers a more subtle approach. With the LEAN system, the redesign phase focuses on eliminating waste and identifying non-value-added activities. This is performed by evaluating efficiency of the system parts and involves searching for unnecessary waste within the process. Such waste could be time or resources (such as people, capital, etc.). Since the LEAN technique is more subtle and incremental, the changes produced are expected to be closer to the current system status and produce less resistance compared to BPR. On the other hand, due to the iterative nature of this approach, it may be least prepared to address systemic issues/problems.

2.3.5 Implementation and testing

Introducing a new system can be complicated and difficult. It requires changing the way people operate. Managing such change is often costly and time-consuming (Fleurant et al., 2012), requiring training staff and collaboration between stockholders and IT management (Goroll, Simon, Tripathi, Ascenzo, & Bates, 2009). Dealing with lost productivity and unexpected errors are common concerns in system implementation. Some of the most-discussed topics include people resistance, system integration, and inadequate infrastructure.

Many workflow improvement and process redesign theories suggest a continuing cycle of redesign after implementation. The reassessment phase enables refinement and optimization of the work process (Unertl et al., 2010). In some cases, simulation software is employed to test the proposed process before the implementation phase. This can

mimic real-life performance and predict potential issues in a risk-free environment. The next section will review some of the simulation applications in healthcare.

2.4 Simulation in Healthcare

Simulations in healthcare are sometimes employed to minimize the shortcomings of implementation and detect some potential issues without risking damage to the actual process (Borycki, Kushniruk, Kuwata, & Kannry, 2006). Simulation aids in identifying disparity between the expected outcome and the observed outcome. Such a disparity can be attributed to a variety of reasons, including confounding variables that are rarely captured during the information collection process, individual resistance, or unpredictable system behavior. This is especially valuable in organizations, including healthcare, where intervention can be complex. When considering the probability of compromising safety, productivity, or financial gain associated with implementing a new system, a risk-free alternative, such as simulation, can be a superior option (Baldwin, Eldabi, & Paul, 2004). Although many simulation studies use simulation software for processes testing, some utilize it to understand existing ones.

A slightly different approach to applying simulation was suggested by Blandwin et al. (2004). The study highlighted the advantages of employing simulation as a problem-solving tool, as it can promote a better understanding of the system, thus enhancing decisions (Baldwin et al., 2004). Most engineering applications adapted simulation for system testing. Given the advancement of simulation use in the

engineering field, many healthcare studies have adopted a similar approach by testing the potential impact of health information technology in a corresponding environment. However, healthcare organizations are structured differently. They often include more stakeholders and decision makers, with a high degree of interaction between them. With such complexity, making a decision that satisfies all stakeholders can be more complicated. Utilizing simulation to illustrate workflow and simplify intricacy can highlight understanding the process and identify dysfunctions (Baldwin et al., 2004).

Shared and carefully planned decisions are especially important in healthcare, where the outcome of change is critical. A review of the literature where simulation was applied in surgical care reported that only 50% of the reviewed content used simulation to address the needs of policymakers, whereas 26% included stakeholders, such as health system managers and policymakers, in the study (Sobolev, Sanchez, & Vasilakis, 2011). An example of this approach is a case study conducted by Bertolini et al. (2011). The study applied simulation to improve the process in the surgical department. The study applied mapping and simulation to analyze all existing processes. Through simulation, the study identified some of the current system's dysfunctions. The study also emphasized the need to overcome the traditional approach in most simulation tests, where designers are disconnected from stakeholders. Instead, the stakeholders were described as an interactive group that contributed to the process development. The study also reported that using simulation contributed to understanding the most efficient management choices (Bertolini et al., 2011).

2.4.1 Simulation in different healthcare settings

Simulation techniques were used in diverse healthcare settings. A review by Gunal and Pidd (2010) reported that the most common healthcare area to apply simulation was Accidents and Emergencies (A&E), followed by inpatients facilities. Other less common areas included Intensive Care Units (ICU), laboratories, surgical units, pharmacies, and screening units (Günal & Pidd, 2010). Another literature review focusing on simulation studies for surgical units reported that very few studies have used simulation to redesign workflow in this particular setting. The study found only 34 related publications from 1957 to 2007 (Baldwin et al., 2004).

While the focus of this research is the reporting process of cancer cases, none of the conducted studies included the use of simulation to analyze the process of public health reporting, such as reporting cancer, infectious disease, or sexually-transmitted disease.

2.4.2 Simulation models validation

Simulation validation refers to the process of ensuring the model representative. In a situation where simulation is intended to reflect the existing workflow status, observing its output can help to ensure the accuracy of the system's elements. This is known as internal validity. On the other hand, external validity ensures that the system's variables represent the real-world process. The variables include subject, setting, case,

task, and/or scenario. A system that captures accurate, real-life events and produces representative results can describe the behavior of the process being studied. Having a system that describes the real-world process can aid in solving workflow problems and identify dysfunctions (Borycki et al., 2006).

In a study that used simulation to redesign a pharmacy's workflow, similar validation techniques were applied. After modeling the existing status, the study used the system yield to compare it with real-world output as a validation method (Wong, Geiger, Derman, Busby, & Carter, 2003). A similar approach was also applied in a simulation study of surgery units. The study compared the simulation system output with real-world data, which was collected during a 1-year period. The study referred to the observed data as the real value of operation and the system-generated value as the predicted value of operation (Bertolini et al., 2011).

Another method for validating simulation is using experts or subjects who are knowledgeable about the workflow (Unertl et al., 2010). In this technique, experts validate that the models' assumptions are correct. Working with subjects who are familiar with workflow will help in identifying issues in the designed conceptual model (Banks, 1998).

2.4.3 Simulation techniques in healthcare

As for the simulation approach, studies reported that Discreet Event Simulation (DES) is the most commonly used technique in healthcare studies. DES presents the workflow process as a series of chronological activities where each activity is described as an event. Each event is associated with characteristics, such as time, to describe performance. This allows the model to capture detailed information to represent the real-life process behavior. These features promote the widespread use of DES in healthcare studies. A review by Sobolev, et al. (2011) indicated that around 75% of simulation studies identified applied DES. Other approaches included System Dynamic (SD) models (9%), the Monte Carlo model (6%), and Makov (3%) (Sobolev et al., 2011).

When comparing DES with other simulation techniques, such as SD, Brailsford and Hilton reported that SD lacked the detailed description needed for healthcare simulation studies (Brailsford & Hilton, 2001). Studies favor DES for its ability to express more details, as well as provide animated and dynamic illustrations to communicate findings with clients (Brailsford & Hilton, 2001).

The findings stressed that SD was appropriate for strategic or conceptual level uses, but it also reported that SD models were not always simulated. An example would be the influence diagrams, which are a useful part in the modeling process. DES referred to this as "the technique of choice for modeling healthcare systems, which are characterized by variability, uncertainty, and complexity" (Brailsford & Hilton, 2001).

2.4.4 Simulation software

In terms of simulation software, review studies showed that Arena simulation software was the most commonly used software in healthcare (20%); followed by the programming language Borland Delphi (17%); Simul8 (10%); followed lastly by others, such as Pascal, AutoMod, and SIGMA (7%) (Sobolev et al., 2011).

CHAPTER 3: METHODS

This study used a mixed-method approach of objective and subjective data. The two sources of data were time intervals data associated with previously reported cases and interviews. This method consisted of the following parts: data tracking, interview, and simulation. Data tracking was used to evaluate the time taken to report previously reported cases. The interviews, however, were used to identify challenges to timely reporting, while informing workflow and simulation development. Lastly, simulation was developed to examine the proposed workflow and compare it to the existing workflow.

3.1 Data Tracking: Context and Data Retrieval

The study was approved by the Institutional Review Board at Indiana University. The dataset was retrieved from the ISDH Cancer Registry. The ISDH Cancer Registry collects all malignant cancer cases required for reporting by federal regulation or the National Program of Cancer Registries (INDH, 2015). The registry contains information on cancer cases needed for performing epidemiological, preventive, and control studies. It also contains demographic data, tumor-related data, and some treatment information (INDH, 2015) from state hospitals, physician clinics, and radiology centers. The cohort of this data was patients' diagnosed with breast, colorectal, or lung cancer from 2001 to 2010.

In total, there were 76,259 de-identified cases chosen from among highly prevalent types of cancer: 28,782 breast cancer cases, 19,530 colorectal cancer cases, and 27,947 lung cancer cases. The variables retrieved for this part of the study were cancer type, reporting sources, date of diagnosis, and the date associated with each phase of the reporting process.

3.2 Phase 1: Data Tracking - Timeliness Overview

Past studies described timeliness as "the rapidity at which a registry can collect, process and report sufficiently reliable and complete cancer data" (Bray & Parkin, 2009). When measuring cancer registry data timeliness, prior studies define it as the as "The interval between date of diagnosis" and "the date the case was available in the registry for research" (Smith-Gagen et al., 2005). Phase 1 focuses on overall timeliness from the time of diagnosis to the time when data was made available at the state registry. This includes the sum of both the time taken by the facility to abstract and send the report to the state registry, and the time taken by the state registry to process the report and make it available for external use by researchers and other stakeholders.

Hunt (2004) outlines 2 approaches for timeliness measurement (National Cancer Registrars, 2004). The first is to calculate the percentage of the number of cases abstracted in a given time of year compared with the expected number of cases in that year. This percentage is then compared with the amount of time that has elapsed to date in the current accession year, minus the allowable reporting timeframe (National Cancer

Registrars, 2004). The second approach is to calculate the difference between the date of diagnosis and the date of data entry; some common data items employed in this approach are date of diagnosis, date of admission or first contact, and date when the records were transferred to the registry (National Cancer Registrars, 2004). The date of first contact refers to the date the patient was first seen or admitted to the facility for malignancy diagnosis or treatment, whichever is first (North American Association of Central Cancer Registries , 2012). In quality control, this date is used with the "date tumor records are made available" to measure overall timeliness (North American Association of Central Cancer Registries , 2012). "The date of first contact" is also used by the Commission on Cancer (CoC) as the starting point to measure abstracting time (Surgeons, 2014). In this study, we applied the second approach to the timeliness evaluation.

3.2.1 Analysis

Measuring timeliness was conducted by analyzing the timestamps of previously reported cases using the Statistical Package for Social Sciences (SPSS) software version 21. The analysis measured the number of days from the date of diagnosis to the date when data became available for reporting by the registry. The "date of first contact" was the starting point and the "date when the report was made available for reporting" was the end point.

To calculate the difference between the date of diagnosis and the date when data was made available, we needed to have both dates. Out of the 76,259 cases retrieved,

8,175 were excluded, because of missing data. The dates of the time intervals were recorded in day/month/year format. Cases where any of these values were missing were also excluded from the analysis. Cases diagnosed during the last year of the study period (2010) were also excluded, due to the low number of cases retrieved (1,066 cases compared to an average of 7,377 in previous years).

The data were later checked for data validity. Given the logical meaning of the time intervals, the date of diagnosis was expected to be earlier than the date the data was made available at the state registry. Cases that did not follow this order were identified as invalid data. Invalid data also included outliers at the 0.99% end of the distribution. Reporting timeliness was then stratified by type of cancer and year of diagnosis.

Finally, we calculated the mean, median, and percentage of records that met the predefined measures for reporting performance. The measures were drawn from the timeliness requirements of national registry standards, such as SEER, NAACCR, and CDC/NPCR. NAACCR, for instance, requires hospitals to report 95% of cases within 23 months of the date of diagnosis. In contrast, CDC/PNR allows 12 months for 90% of cases to be reported and 24 months for 95% of cases. The SEER registry standard, however, requires the total count to be reported to NCI within 22 months (NCI, 2014). The length of time used to report performances was 6 months, 12 months, 22 months, 23 months, and 24 months.

To assess consistency in reporting speed, we calculated historical data for the 9-year period from 2001 to 2010. We calculated the overall annual mean of reporting time, as well as the mean for each of the 3 cancer types. The number of cases within each year was determined based on the date of diagnosis.

3.3 Phase 2: Data Tracking - Reporting Stages Timeliness

In Phase 1, we evaluated the time needed to report cancer cases to the state registry. In this section, we will explore details of the various stages of reporting, provide an overview of the fundamental phases of the reporting process, and offer a description of the method used to evaluate the various stages of cancer reporting.

3.3.1 Basic Steps of Cancer Reporting

Analyzing time intervals at each phase will require an understanding of the reporting process. In this part of the analysis, we will outline the reporting steps measured, as well as the time intervals associated with each phase. The cancer reporting process involves the following steps.

Case Finding

Case finding is the process of identifying new cancer patients for a given time period. At this step, the list includes cancer records associated with any diagnosis,

treatment term, or code, indicating a reportable cancer condition. These lists are often defined by registry-approved programs (SEER, 2015). Once a case is reviewed and identified as a reportable case, it is added to a suspense file in preparation for abstracting. The time from diagnosis to the beginning of abstracting can represent the time from the "date of diagnosis" to the "date of abstract initiation" (Figure 1-1).

Abstracting

Abstracting is the process of collecting and compiling information about each reportable patient in preparation for sending the report to the state registry. Abstracts are often comprehensive reports that include all the cancer-related information required by the receiving registry. The abstract may include data related to demographics, tumor information, staging, diagnostic studies, and treatment. Once the abstract is completed, it is saved in preparation for reporting.

Reporting facilities, in general, save completed abstracts for purposes of reporting them to registries on a fixed schedule. The abstracting time is described as the time needed to find the required information, collect it, and then compile it into an abstract. Once all the required information is collected and the abstract meets the receiving registry's requirements, the abstract is considered completed and ready for reporting. In this study, the abstracting time is represented by the time from "date of abstract initiation" to "date report completed by the facility" (Figure 1-1).

Report Submission

Completed abstracts are sent to the state registry when they are due for reporting, which is often done at fixed intervals. The state registries require facilities with higher caseloads to report at higher frequencies. The time from abstract completion to when the state registry receives the report can be calculated by measuring the time from "date report completed by the facility" to "date report received by the registry" (Figure 1-1).

Editing and Processing

State registries receive reports from multiple facilities that may need to be checked and edited before they are made accessible to researchers. The registry processing time is calculated by measuring the time from "date report received by the registry" to "date report available at the registry" (Figure 1-1).

3.3.2 Analysis

As shown in Figure 1-1, the time taken for each phase within the reporting process is represented by the difference between the corresponding time interval dates. These time intervals indicate the start or end of some stages of reporting. The first part of the analysis separates the time taken by the reporting facility from the time taken by the registry. Time was measured by number of days using box plots to illustrate distribution of the reporting times.

Time taken by the facility was then examined by separating the report completion time (the time taken for case finding and abstracting) from report submission time (time the completed reports remain in the facility database before they are received by the state registry).

The "abstract initiation date" (see Figure 1-1) can be used to distinguish case finding from abstracting time. However, because of the high rate of missing data in this variable, distinguishing these steps was not possible. Instead, we simply measured the report completion time.

Timeliness within Facilities

To examine the variation in reporting time among facilities, we calculated the annual average time taken for each individual facility at each stage of reporting. This includes report completion time and report submission time by the reporting facilities.

Since the reporting frequency requirements vary based on facilities' annual caseloads, we focused on larger facilities with higher numbers of cases, due to their greater impact on the mean. In this section, we included facilities with a total of 1,000 cases or more for the entire study period, which equals an average of 100 cases per year for the 3 cancer types included.

Facilities with an average of 60 to 149 cases a year are expected to report their cases every quarter. Facilities with higher numbers are expected to report at a higher frequency. Facilities included in this part of the analysis are thereby either expected to report every quarter, every month, or every other month.

3.3.3 Reporting Frequency

For facilities with a significantly longer report submission time, we performed further analysis to evaluate the frequency of reporting, which was assessed by examining the distribution of cases reported each quarter. For instance, if a facility reported many more cases in one quarter than in others, it is an indication of irregular reporting.

Due to the chronological nature of the cancer reporting process, frequency in reporting can be influenced by many factors. A common example is the rate of patient visits. When physicians' offices receive higher numbers of patients in some quarters because of the holiday season, more cases will be diagnosed in that season; thus, the cases that need to be reported for that season will be higher. Similar effects in the rate of reporting might be experienced if patients come for treatment at irregular rates, as some treatments are required in order for the abstract to be complete and become ready for reporting. Another possibility might be the staffing of registrars who do the reporting. If the reporting registrars are part-time or freelance and happen to be overworked during particular months of the year, more cases will be reported when they have more free time.

To narrow down possibilities, we tracked down the stage when irregularity of case distribution started. As seen in the table below, we examined distribution during the following 3 phases: time of diagnosis, time of report completion, and time the report was received by the state registry (Table 3-1).

Time when irregular distribution of cases started	Potential reasons for irregular distribution of cases
Time of diagnosis	More patients visiting the facility and are diagnosed during certain quarters.
Time of report completion	More part-time or freelance registrars working during particular quarters. More patients receive treatments, tests, or procedures needed to complete reports during certain quarters.
Submission time or time the report is received by the state registry	More completed reports are delivered to the registry during certain quarters.

Table 3-1: Potential reasons for irregular distribution of cases at each phase of reporting.

3.4 Phase 3: Interviews, Barriers, and Workflow

3.4.1 Study Context and Data Collection

This study was approved by the Institutional Review Board (IRB) at Indiana University. Cancer registrars reporting to the State of Indiana were invited to participate in the interviews, which were conducted either in person or by telephone.

Participant selection was based on purposeful sampling that focused on potential participants from larger hospitals. The researcher identified participants during hospital visits and through the Indiana Cancer Registrars Association (ICRA) directory. Snowball sampling—wherein initial contacts identify other individuals who may have insight into the topics of interest, e.g., barriers and facilitators to timely cancer reporting—was also used.

Potential participants were contacted either by telephone or email. Participants were approached until thematic saturation was achieved for a grounded theory analysis. Recruitment took place over 5 months, March 22 to August 26, 2015. The average interview duration was 28 minutes. Interviews were recorded and later transcribed for analysis. At the end of the interview, each participant received a $20 gift card. The final interview sample comprised 14 registrars.

Participants were asked to describe the reporting process routine, the steps involved, and the estimated time spent on each task. Participants were also asked to identify obstacles that affect timeliness of reporting during the process. The interview format was semi-structured and task-oriented (Appendix 6.1).

Information from the interviews was used to develop simulations around variables such as the average time taken per task and the capacity (both human and capital resource) needed for each task. In this context, capacity is described as the number of cases that can be performed daily and the time required for each by a registrar. In

47

addition to describing the current state of the cancer registry reporting system, interviewees were asked to use a "blue sky" approach to consider optimal cancer reporting mechanisms. "Blue sky" thinking transcends concerns about current resources or structural limitations.

3.4.2 Analysis

Data collected during the interviews was analyzed in two stages: 1) thematic analysis and 2) workflow modeling. For thematic analysis, the transcribed text was imported into NVivo 10.2.0 and analyzed. The analysis was based on grounded theory and focused on identifying factors that may affect timely reporting and thematic analysis. The thematic analysis was derived directly from the data and was used to explain barriers to timeliness. The key ideas were identified, then grouped into similar themes.

The main focus of workflow modeling was to develop an alternative reporting process to reduce reporting time through process reengineering. The first step was to understand and analyze the existing process through a review of current training materials that were both related to cancer reporting and directed at cancer registrars, followed by interviewing a sample of registrars. In this way, the researcher was familiar with the overall process and, therefore, able to explore process details and flaws through the interview process. Thus, the researcher was able to notice gaps between policy and practice, assess the efficiency of various steps, and document challenges that were encountered. During the interviews, registrars were asked to estimate the time spent on

each stage of the reporting process. This data was used to model the existing workflow and develop a simulation model of the process. Data flow charts were then applied to illustrate the complete reporting process.

3.5 Phase 4: Simulation Development and Validation

We simulated the current workflow to demonstrate the time spent per task. The simulation was performed using Anylogic 7.1. The simulation provided an indication of the time spent at each phase of reporting (e.g., processing time, waiting time, and the time cases spent in queue before being processed). The simulation inputs were obtained from interviews. During the interviews, participants were asked to estimate the time needed for each step in the reporting process.

We then redesigned the reporting process to reduce reporting time and minimize some of the barriers identified during the thematic analysis. The redesigned process was also simulated to estimate the difference in reporting time compared with the current workflow.

The simulation was also validated to ensure an accurate representation of real-world workflow. The simulation model went through an iterative process of calibration and comparison with the existing workflow. Model development occurred concurrently with the interviews to investigate model assumptions and enable the iteration process.

Model input and output was also compared with reporting time in the real world to ensure an accurate representation. After simulating the current workflow, the proposed workflow model was simulated and the outputs compared with respect to time.

CHAPTER 4: RESULTS

The purpose of this research was to evaluate the quality of cancer data with respect to timeliness and examine methods by which the reporting process can be improved. The timeliness evaluation was conducted in 2 phases: overall reporting time evaluation and reporting step timeliness evaluation. Timeliness improvement started by identifying barriers to timely reporting through interviews. The interview also informed the subsequent phases of our study: workflow redesign and simulation development.

4.1 Phase 1: Data Tracking-Timeliness Overview

The initial number of cases retrieved was 76,259; there were 28,782 breast cancer cases, 19,530 colorectal cancer cases, and 27,947 lung cancer cases. To calculate the difference between the date of diagnosis and the date when the data was made available, we needed to have both dates available. Out of the 76,259 cases retrieved, 8,175 were excluded because of missing data.

Cases diagnosed during the last year of the study period, 2010, were also excluded, due to the low number of cases retrieved (1,066 cases compared to an average of 7,377 for previous years). Data cleaning also included removing outliers and invalid variables.

After cleaning the dataset and excluding cases from 2010, 66,395 cases remained. The number of breast, colorectal, and lung cancer cases was 25,013 (37.6%), 17,074 (25.7%), and 24,308 (36.6%), respectively. The annual number of cases ranged from 7,083 to 7,726, except for 2009, which had 6,693 cases.

Table 4-1 shows the yearly reporting delay for the 3 cancer types combined. The mean for the entire 9-year period was 323.4 days; however, there was a large variation in reporting time mean—from 252 to 426 days—across the years. The lowest reporting time was 252 days (2009) and the highest 426 days (2003). The reporting time for the rest of the data shows less variation: from 292 to 368 days. In some years, the results also show a noticeable difference between the mean and median, indicating a skewed distribution of data. The difference between the mean and median was highest in 2007 (299 for the mean and 220 for the median) and 2008 (313 for the mean and 219 for the median).

Year	Mean (days)				Median (days)	Median-mean difference	N
	Breast	Colorectal	Lung	Total			
2001	317.1	315.5	332.4	322.08	292	30	7566
2002	296.24	280.11	296.32	292	279	13	7702
2003	427.2	427.2	424	426.05	424	2	7571
2004	364.59	364.86	374.14	368.19	335	33	7208
2005	338.68	323.28	337.14	334.08	273	61	7083
2006	300.5	295.87	309.79	302.79	237	11	7553
2007	295.84	290.26	310.3	299.83	220	79	7293
2008	298.28	303.4	334.93	313.35	219	94	7726
2009	237.72	246.22	273.79	252.88	203	49	6693

Table 4-1: The average timeliness for each cancer type.

As for stability in timeliness, the results show inconsistency across the 10-year period. The average timeliness declined from 2001 to 2002 and then increased by about 46% (from 292 days in 2002 to 426 days in 2003; Figure 4-1). This was followed by a gradual decline until 2009. The graph also shows a similar trend for the 3 types of cancer: breast, colorectal, and lung.
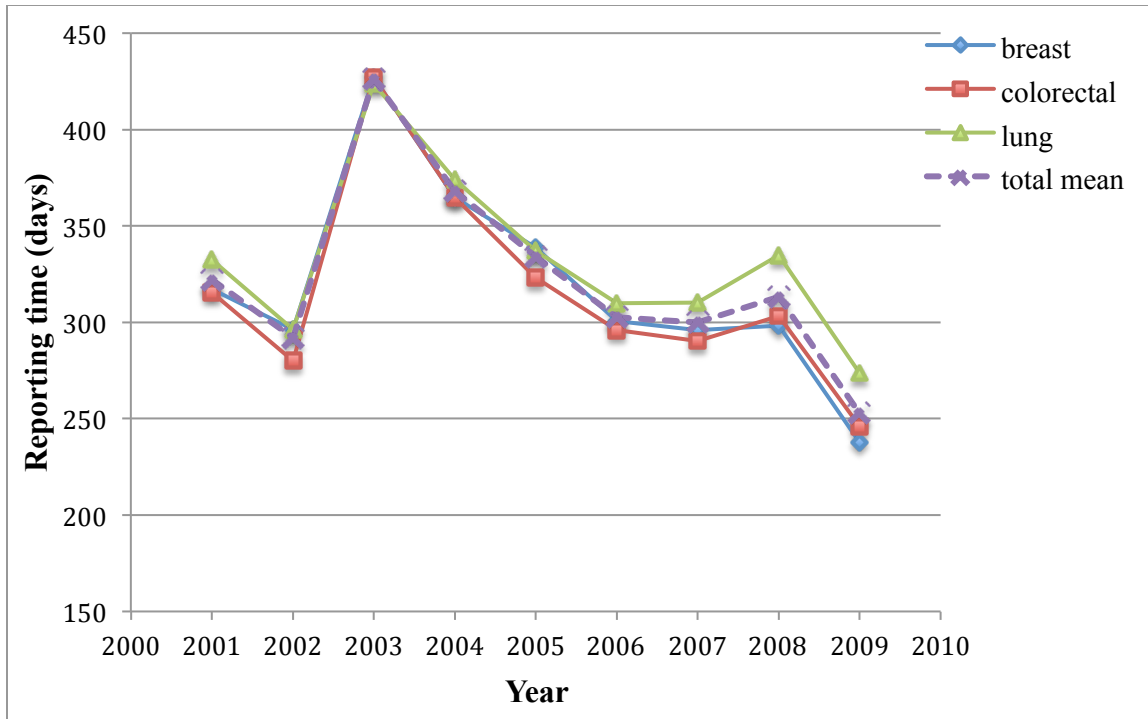
Figure 4-1: Timeliness for different cancer types from 2001 to 2009.
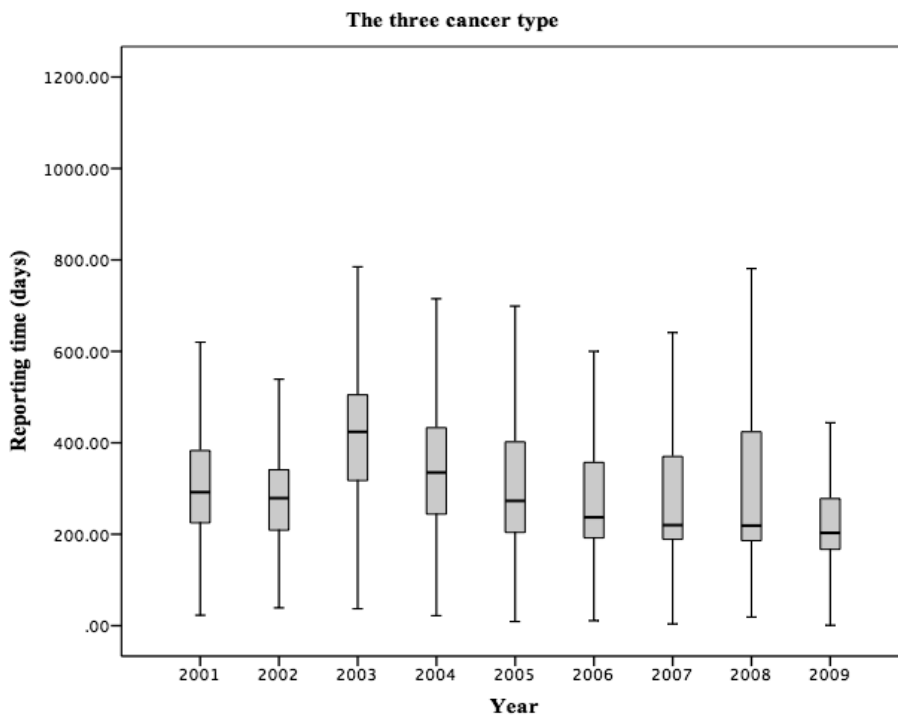

Figure 4-2: Average overall timeliness from 2001 to 2009.

	2001	2002	2003	2004	2005	2006	2007	2008	2009	Mean
6 month	11.9 %	16.5 %	03.9 %	10.8 %	15.1 %	19.6 %	21.1 %	23.0 %	37.2 %	17.68 %
12 month	71.4 %	82.1 %	34.4 %	59.2 %	68.8 %	76.6 %	74.7 %	70.6 %	82.6 %	68.93 %
18 month	92.3 %	95.7 %	82.6 %	87.2 %	88.7 %	92.0 %	88.2 %	84.6 %	93.8 %	89.45 %
22 month	97.1 %	98.3 %	93.7 %	93.3 %	94.3 %	94.2 %	95.3 %	92.9 %	97.5 %	95.18 %
23 month	97.6 %	98.5 %	95.0 %	94.1 %	95.3 %	94.7 %	96.0 %	94.5 %	97.8 %	95.94 %
24 month	98.1 %	98.8 %	95.8 %	94.8 %	95.7 %	95.3 %	96.6 %	95.9 %	98.2 %	96.58 %

Table 4- 2: Percentage of cases reported within each time period.

For the entire 9-year period, the results show that an average of 17.6 cases was reported within the first 6 months, 68.9% were reported within the first 12 months, and 89.4% were reported within the first 18 months (Table 4-2). The highest percentage of cases reported within the first 6 months was 37.2 % (2009) and the lowest was 3.9 % (2003). For the remaining years, the percentage of cases reported within the first 6 months ranged from 10.8% to 23% (Table 4-2).

The percentage of cases reported within the first 12 months was also highest in 2009, with 82.6%, and lowest in 2003 with 34.4%. For the remaining years, this percentage ranged from approximately 60% to 75%. Cases reported within the first 18 months were also lower during 2003, with 82.6 %, but the highest was during 2002, with 95.7% (Table 4-2). Over the whole dataset, the average percentage of cases reported within 22, 23, and 24 months ranged from 95.1% to 96.5%. Cases reported within the first 22, 23, and 24 months were highest in 2002, with 98.3%, 98.5%, and 98.8%,

respectively. As for the percentage of cases reported within the first 22 and 23 months, both were lowest in 2010, with 92.2% and 93.6%. However, cases reported within the first 24 months were lowest in 2004, with 94.8% (Table 4-2).

4.2 Phase 2: Data Tracking-Reporting Stages Evaluation

4.2.1 Reporting Stages

As previously discussed, the reporting process comprises 4 steps: case finding, abstracting, report submission, and report processing (Figure 1-1). The first 2 steps, case finding and abstracting, were performed at the reporting facility. The completed cases then resided at the facility database before being submitted to the central registry for processing.

The results distinguished between the time taken by facilities to complete and send reports, and the time taken by the state registry to process the cases received. The reporting time taken by the facility to complete and send the report for each year is shown in Figure 4-3. A higher variation in median reporting time was observed for the period 2001 to 2005, followed by a more consistent trend for the remaining period of 2006 to 2010.

For most years, there was a large variability in reporting time between cases. Variability was higher in 2008 and 2009, especially at the upper whisker. The annual number of cases for the observed time period ranged from 7,270 to 7,812.
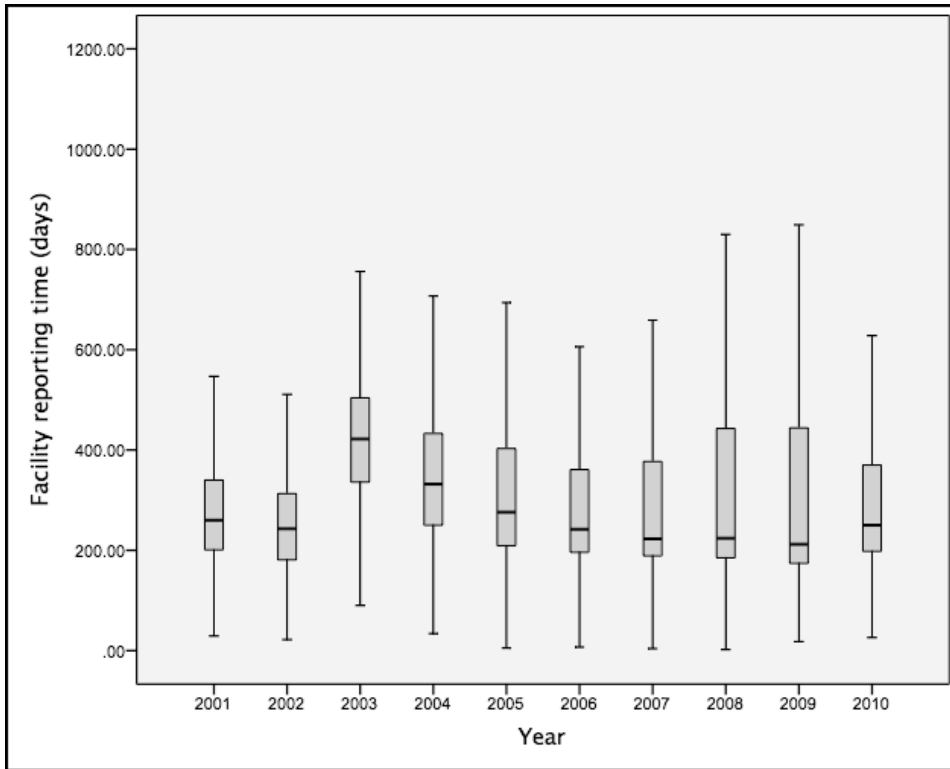

Figure 4-3: Box plot showing the distribution of time taken by facilities to complete and submit reports.

The time taken by the state registry to process and edit cases was much lower compared to the time taken by the reporting facilities (see Figures 4-3 and 4-4). The time taken by the state registry was higher during the first 2 years, 2002 and 2003, with medians of around 30 and 20 days, respectively. A noticeable reduction in processing time was observed in 2004 and beyond. The processing time was much lower during the last 6 years, 2006 to 2010.
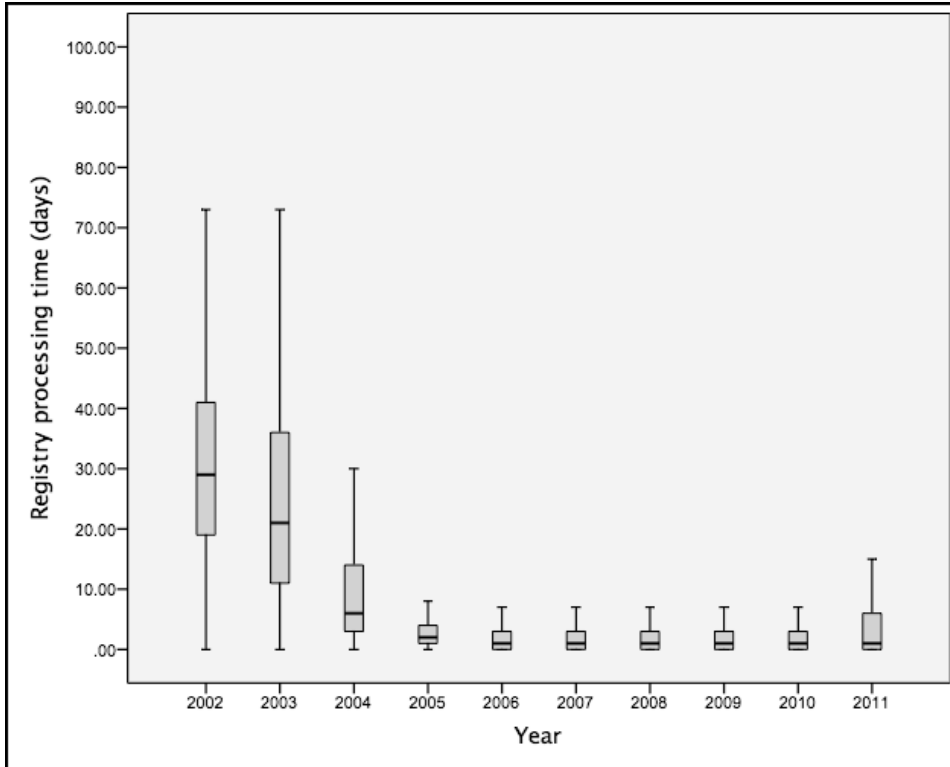
Figure 4-4: Box plot showing the distribution of time taken by the central registry to process received reports.

Time taken by the reporting facilities was further investigated, which revealed the difference between the time needed to complete the reports and the time taken to send the reports after completion. The median report completion time was highest during 2003 and 2004, as shown in Figure 4-5. The median reporting time was also more consistent during the last 5 years, from 2006 to 2010. Moreover, a noticeable reduction in variability in case reporting time was observed starting in 2006.
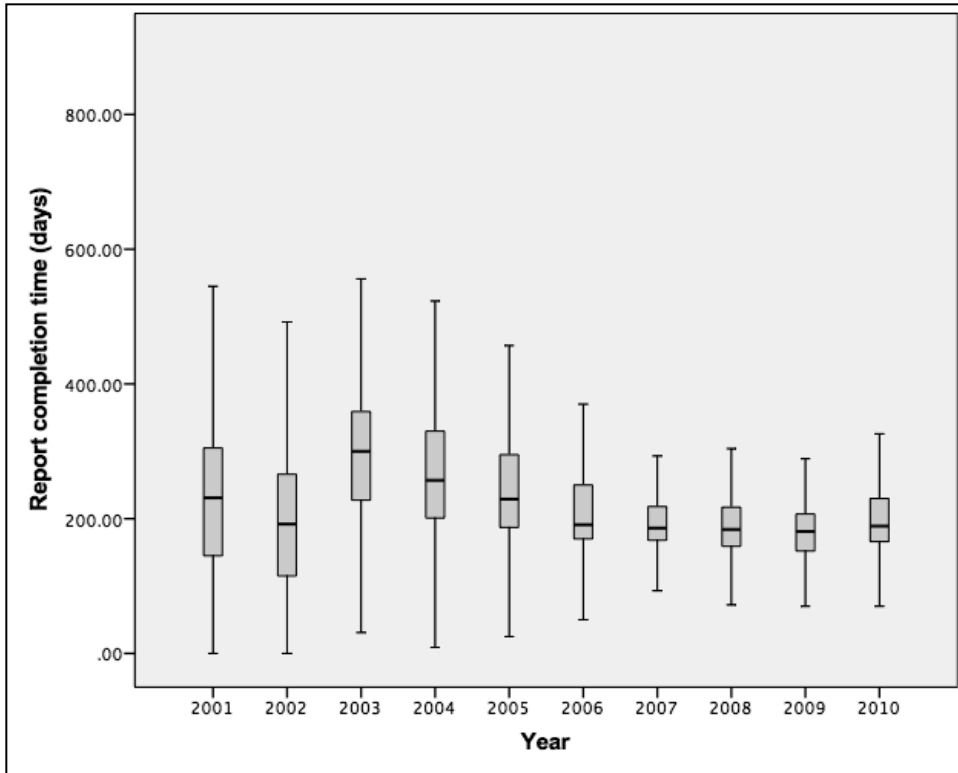
Figure 4-5: Box plot showing the distribution of time taken by facilities to complete reports.

Report submission time represents the time taken from report completion to receipt by the state registry. With the exception of 2003, the median time was around 50 days or less for the entire 10 years (see Figure 4-6). The box plots show that the data was skewed toward the lower limit with longer upper whiskers. Though most cases (75th percentile) took fewer than 120 days, some cases in the upper whiskers took up to a year.
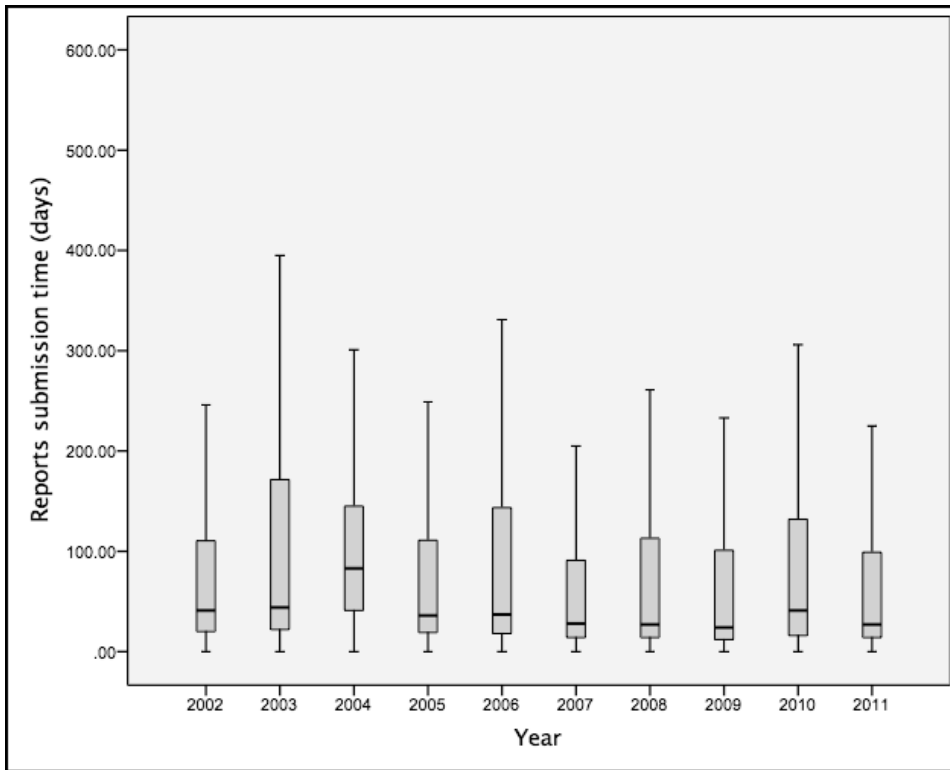
Figure 4-6: Box plot showing the distribution of time taken by facilities to submit completed reports.

4.2.2 Variation between Facilities

The reports retrieved for the entire 10-year study period were reported from 49 unique facilities. Of the 49 facilities, 22 were classified as higher-load facilities. The 22 facilities selected accounted for 90.5% of the cases. The remaining 9.5% came from the other 27 smaller facilities.

Figure 4-7 shows the average time taken to complete and submit reports at higher-load facilities. There was a considerable difference among facilities in average reporting time (Figure 4-7); whereas, most facilities took around 200 to 300 days to report their cases, others exceeded 400 days.
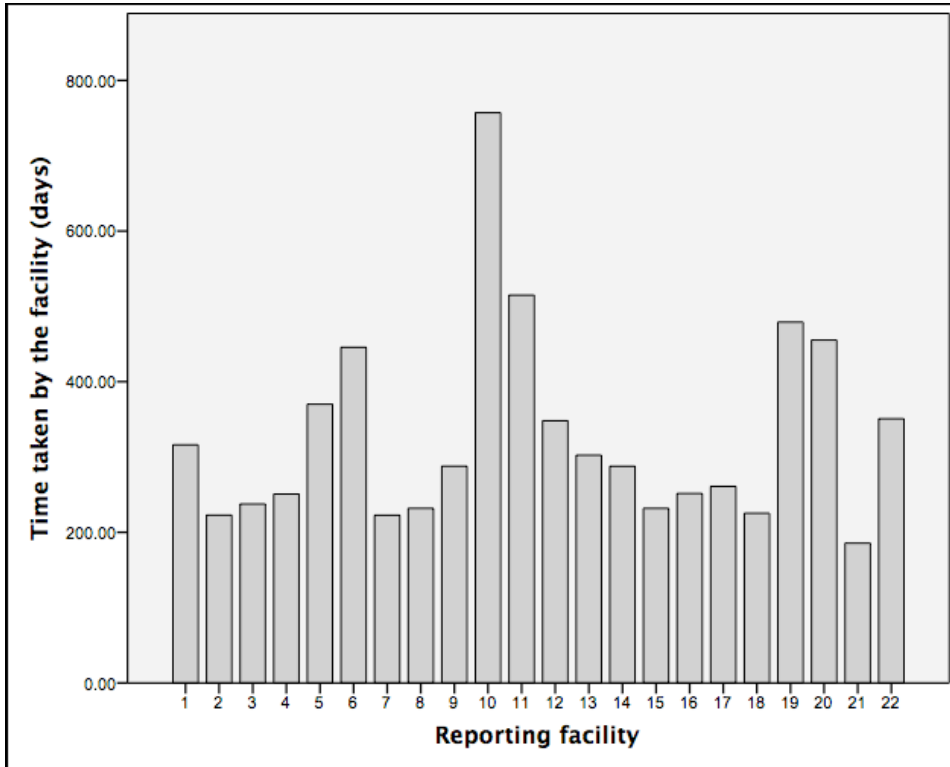

Figure 4-7: Average time taken to complete and submit reports at higher-load facilities.

Facility reporting time was further investigated by distinguishing report completion time from report submission time. The time taken to complete reports varied among facilities (see Figure 4-8); whereas, most facilities' reported completion time ranged from around 175 to 225 days, a few exceeded 300 days.
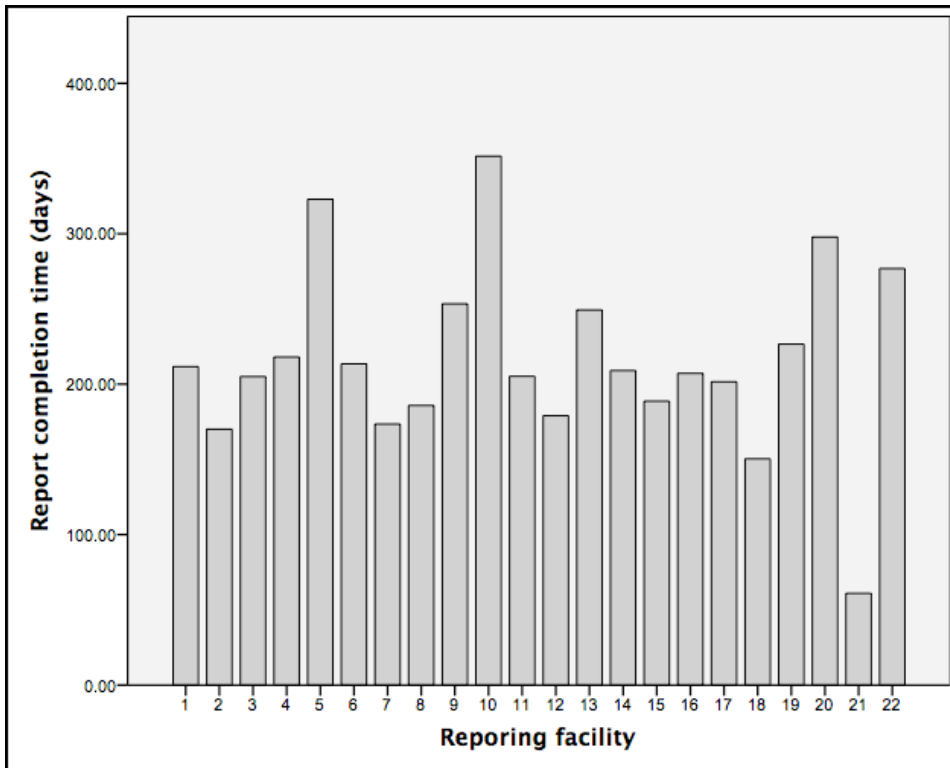
61

Figure 4-8: Reported completion times at higher load facilities.

Report submission time is presented in Figure 4-9. A significant difference among facilities was found in report submission time; whereas most facilities took less than 50 days on average to send completed reports to the state registry, 7 of the 22 facilities took more than twice that time (around 100 to 300 days). Facility numbers 1, 6, 11, 12, 19, 20, and 21 took 106, 274, 325, 172, 256, 155, and 123 days, respectively.
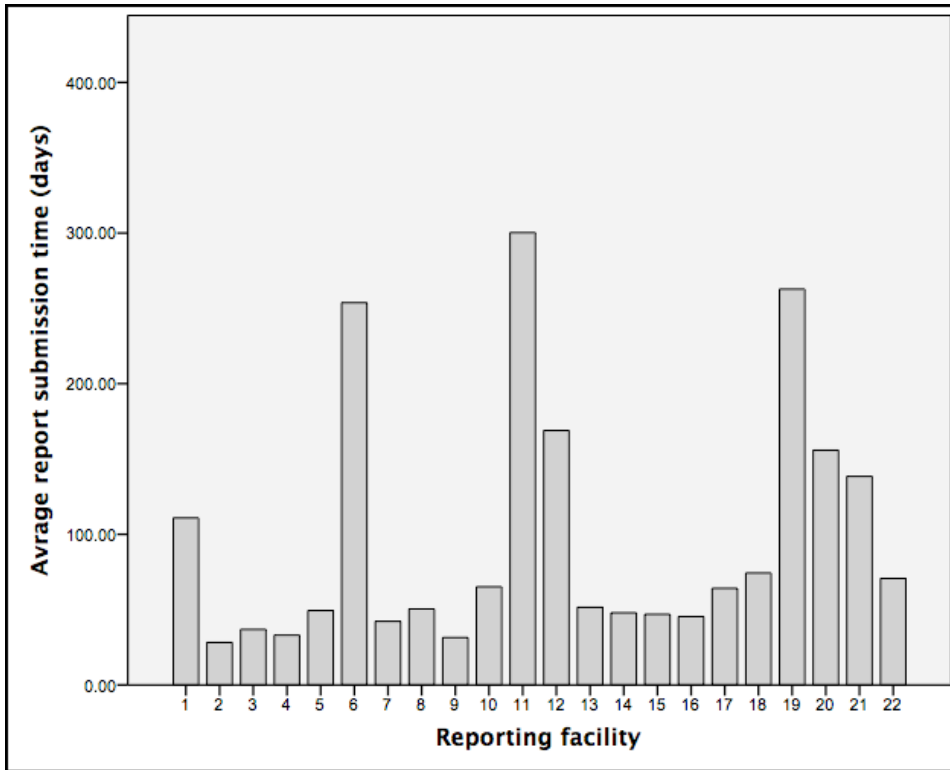
Figure 4-9: Report submission times at higher load facilities.

To investigate report submission times for these facilities, we examined the frequency of reporting. Figures A-1 to A-7 in Appendix 6-2 illustrate the annual number and quarterly distribution of cases received by the state registry from each facility. All 7 facilities showed an inconsistency in the annual number of reports received by the state registry. Higher inconsistencies were observed in facility numbers 1, 12, and 19.

As for quarterly distribution, most facilities showed significant irregularities in the number of cases received each quarter in a given year. This irregularity was higher in facility numbers 6, 11, and 20. In many years, facilities were submitting all of their reports during only 1 or 2 quarters.

Since distribution of reports at the report submission phase might have been inherited from one or more of the preceding phases, we also investigated the distribution of reports in terms of when they were completed and when the cases were diagnosed.

The distribution of cases diagnosed and reports completed is illustrated in Figures B-1 to C-7 in Appendix 6-2. The results show that some facilities experienced inconsistency in the number of cases diagnosed annually (see Figures C-1 to C-7 in Appendix 6-2). The greatest inconsistency was observed in facilities 1, 6, and 12. The quarterly distribution of cases diagnosed within the same year was nearly equally distributed among all 6 facilities. In facility 1, however, slightly fewer cases were diagnosed in the third quarter of 2005 and 2007, respectively.

At the report completion phase, results show that the distribution of reports completed each quarter within the same year was nearly equally distributed. Some irregularity was observed in a few years, when the number of reports completed in a given quarter was relatively smaller than the rest. However, compared to the quarterly distribution during the report submission phase, the reports appeared to be completed at a more regular frequency.

When comparing the quarterly distribution of the 3 phases (date of diagnosis, date of report completion, and date the reports were received by the registry), we find that most of the irregularity originated in the last phase, the time between report completion and receipt at the registry. This suggests that the increased submission time was not a

result of patient visit time or staffing issues, but rather, a workflow issue regarding when the completed reports were sent to the state registry.

4.3 Phase 3: Reporting Workflow and Barriers to Timely Reporting

4.3.1 Workflow Description

The developed data flow chart comprises the 3 major steps: case finding, abstracting, and report submission (Figure 4-10). The details of each step are described below.
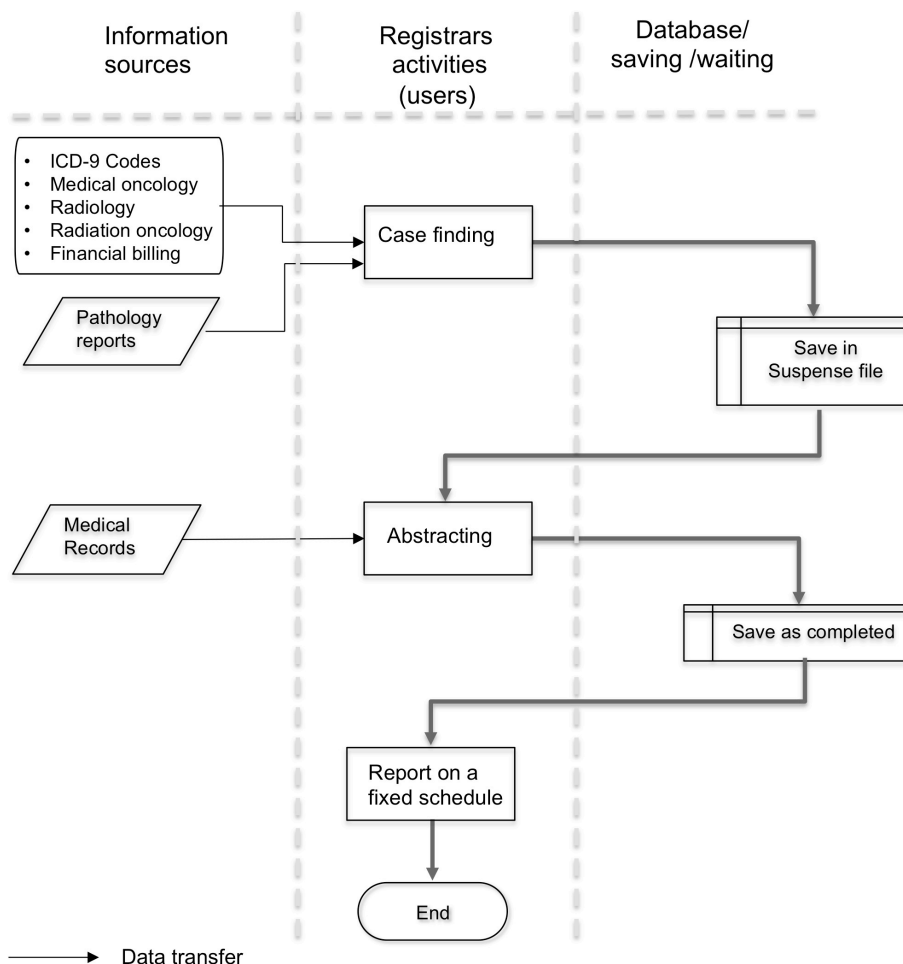
Figure 4-10: Cancer reporting flow chart (existing status).

Case Finding

Case finding is the process of finding new cancer cases diagnosed within a given time period; it is the first step in the cancer reporting system. It applies to all patients—inpatients and outpatients—as long as they are diagnosed and/or treated with a reportable tumor.

Registrars use several data sources for case findings. Most cases, 90–95%, are identified through pathology reports. Pathology reports are also useful, because they contain detailed information about the cancer (such as diagnosis, histology, and behavior). Some facilities use additional sources for case finding, such as hospital admission and discharge records, surgery schedules, cytology reports, oncology reports (nuclear and medical), radiology reports, and financial billing records. These sources, however, are less informative than pathology reports, and registrars often use multiple sources or refer to medical records to find the information they need.

Once a case is confirmed as reportable, it is added to a suspense file to await abstraction. In most facilities, case finding is performed weekly or monthly, but cases may reside in the suspense file for up to 6 months before abstraction. The rationale for this is to wait for tests and treatments to be performed and added to the EHR.

Abstracting

While case finding generally provides an overview of the case, itself, abstracting is more comprehensive and detailed. Abstracting uses different parts of medical records to collect demographic information, tumor-related information, and information about staging, diagnostic studies, and treatment. In general, abstracting is less structured than case finding, because registrars use different forms and notes to create a summary. Depending on the cancer type and stage, as well as the abstracting registrar's knowledge

and experience, registrars often have an understanding of what treatment paths patients are likely to follow and, therefore, they will search records accordingly.

Some registrars indicate that physician notes or discharge summaries can serve as a starting point to help them form an overview of the patient journey and guide patients as they navigate through sources within medical records. Unfortunately, notes and discharge summaries may not become available until long after a patient's diagnosis.

Interviews show that a patient's data is not always available in local medical records. This happens when patients receive care at different hospitals/facilities. The interviews also show that registrars use different methods to access records at external facilities, and that the methods vary among registrars. In some cases, registrars contact health professionals at the external facility to ask for patients' records. This was found to be easier when most patients are referred to a particular facility, thereby, enabling registrars to explain the need for collecting patients' data to health professionals at the external facility and establish a relationship with them. An alternative approach followed by some registrars is to contact their counterparts working at hospitals where care is provided. After all the required information is collected and the abstract is considered complete, it is saved in the suspense file preparation for submission. The suspense file acts as a local temporary database, where completed abstracts are saved before sending them to the central registry.

Submission

Completed abstracts are sent to the state registry at fixed intervals. Facilities with a higher number of cases are required to report to the state registry at a higher frequency. For example, facilities with an annual caseload of 300 or more are required to report monthly, while facilities with an annual caseload of 150-299 are required to report every other month.

4.3.2 Time per Task

When examining the process cycle time, we see that it contains both activity and waiting times. Activity time includes the time registrars spend accessing and retrieving data, reviewing the records, and entering information into the system. Meanwhile, waiting time refers to the time cases or records reside in the system while no activity is being performed. This includes the time cases reside in the suspense file before abstracting and the time completed reports reside in the local system before being sent to the state registry. A list of the estimated time spent at each phase is available in Table 4-3.

Activity Time		Non-Activity Time	
Task	**Time**	**Phase**	**Time**
Case finding from pathology reports.	Daily: 1 hour	Suspense file.	3–6 months. Varies among hospitals.
Case finding from the ICD-9 list.	Monthly: 1 day	Completed cases reside at the local registry before submission.	An average of 15 days for hospitals with higher caseloads.*
Abstracting	Daily: 45 minutes to 1½ hours per case.		

*Hospitals are considered high load if they have 300 or more reportable cases a year.

Table 4-3: Estimated time for each reporting phase.

4.3.3 Barriers

Identified barriers were categorized into 6 themes. The most critical and common themes identified were access to data at other facilities and case findings from index codes. Details of identified barriers are discussed in the following paragraphs.

Access to Information-Related Barriers

In this context, access to medical records can result from a technical inhibitor or a non-technical inhibitor. The technical inhibitor can be seen as the inability to exchange data between systems as a result of using legacy systems, a paper system, or data interoperability, among others.

However, non-technical barriers are often the result of the hospital network effect. The network effect can be seen when facilities are parts of different hospitals that do not exchange data. Exchanging patients' records in such cases often requires additional steps and additional paperwork to ensure security and confidentiality of the data. Generally, cancer registrars work as hospital employees and have access provided by the recruiting hospital. Registrars are often able to access data available within the hospital they work for, but when patients receive care in a facility outside the hospital network, access becomes more challenging. The following section will discuss access-related barriers in more details.

Access to External Records

During abstracting, registrars are required to complete a predefined data file. Once the case is identified as reportable at the case-finding stage, registrars will search for the required data fields in hospital medical records. In some cases, data is not available in local medical records, which occurs when patients receive care at an external facility. This is common when dealing with treatment data, as many of the facilities providing oncology treatment are external or independent.

Interviews show that registrars were applying different methods to access records at external facilities. In some cases, registrars can have electronic access to medical records at external facilities or know staff who can provide them with the needed information. Examples of electronic access include remote desktop connection, virtual

private network (VPN), and Health Information Exchanges (HIE), such as Indiana Network for Patient Care (INPC). Remote desktop connections and VPNs are employed when many patients visit some specific outside facility so that registrars can ask for access or establish a relationship with the external facility.

When registrars depend on staff at the external facility to request information (such as nurses or physicians), often they are constrained by staff time and availability. Since searching for and sending information to registrars is not part of their core job requirement or is a direct patient service, this is done on a voluntary basis.

Interviews show that the percentage of cases that require contacting external facilities varies widely, from 10% to 40%. One registrar stated that it can be difficult and time-consuming to find out what physicians are doing in their offices; for example, a physician might be contacted to find out if the patient received chemotherapy, hormone therapy, or refused treatment. While describing the barriers encountered, one interviewee stated:

> "Getting the information from them 'physicians and nurses' and letting them know they are not breaking HIPAA if they give us this information. Telling them even if we are not face-to-face with the patients, we are still doing patients' care."

The interviewee also stated that physicians who are active on cancer boards are more likely to help, because they understand both the process and the needs of cancer registries.

Staff at external facilities are contacted through phone calls, mail, or faxes. The method of contact and the access granted varies based on each registrar's ability to establish a relation with the external facility. An interviewee stated that, being a non-contracted registrar makes it more difficult to collect information from external facilities. The interviewee stated:

"Because I am not a contracted employee, I tend to go onsite and meet people. I don't call a lot, because some places are not happy giving that information. They want to know who I am and where I am from, so the contact I do have, I build a rapport with and I get the information from them."

Since the majority of interviewees experienced some level of barrier to access, the extent of the issue varies between registrars. Registrars who experienced fewer obstacles with data access also stressed the need to build a rapport and know staff at other departments and nearby facilities.

When patients receive treatment at an external facility, the abstracting registrar sometimes contacts the registrar working for that facility, instead of contacting the physician or nurses. This is often expressed as a preferable alternative, because they are familiar with the reporting process and requirements. The state registrars' association provides the registrars' directory, but our results show that few registrars have used it or mention it during the interview.

Another barrier is knowing where the patient has been transferred to. One interviewee stated that the real obstacle comes with finding out where patients receive treatment, as this is not always indicated in their medical records. When this is the case, registrars need to guess, using the patient's address and nearby facilities.

Technical Barriers

Access barriers related to technical issues were less common. In this context, system compatibility was indicated as being an issue. Although it did not prevent registrars from accessing the needed information, it prevented them from using the oncology management system to its fullest potential.

One registrar stated that the hospital system doesn't support some of the technology they wish to use. Because the hospital system is not compatible with the oncology management reporting system, registrars were not able to use all of the features that require data sharing.

System Integration and Electronic Data Exchange

System integration can support the case-finding process. In most cases, data needed for case finding is retrieved electronically from case-finding sources, such as pathology reports and disease indices. Disease indices, for example, are retrieved electronically by health information management departments. Screenings for disease

codes function as a wide net that form an inclusive search that screens for cancer-related codes. However, this itself is not enough for case ascertainment. Once the case is identified as potentially reportable, registrars will search other data sources, such as pathology reports, surgery reports, and physicians' notes, in order to confirm case repeatability.

This process can be time-consuming, because registrars have to search and review several data sources before they can decide whether or not the case is eligible for reporting. The list of International Classification of Diseases (ICD) codes used for screening is inclusive, which often results in a large number of non-reportable cases. In many cases, patients with a history of cancer are flagged in the system and added to the disease indices list. To distinguish the history of cancer from newly diagnosed cases, registrars will need to check the local registry to figure out if the case has been previously reported.

With large numbers of cases, relying on human skills to distinguish the new from previously reported cases can be difficult and time-consuming. Some oncology management systems provide tools to help with records matching. Using such tools often requires system integration and data sharing. Currently, EMR screening is performed by the health information department on a weekly or monthly basis, and the list is then sent to the registrars in Excel format. Using a hospital EMR and oncology management system that is not comparable will prevent data sharing and thus prevent using some of the oncology management system tools that require data sharing. This also happens when

the hospital system relies on a paper format for data sharing. Interviews show that departments that use a paper-based system will send their records in paper format; whereas, others sometimes fax to share their data.

System Login and Session Timeout

Abstracting requires the collection of a significant amount of data from different sources using different systems. Due to HIPAA regulations, access to medical records is protected in a variety of ways. One measure is the automatic logging-out of users who are inactive for a set period of time. Due to the large number of data needed from diverse sources, registrars sometimes need to go back and forth between these data sources, which include accessing different systems, paper records, or phone calls. However, being busy with one source will result in inactivity in the previous one, and, in turn, the system will log the user out automatically. One registrar commented:

> "My most time consuming thing for me lately is getting the medical records to work, logging in to the system, staying logged in, and dealing with connection."

This is more problematic when hospitals use segregated systems. For example, when patients are diagnosed in a hospital that refers patients to the Radiology Department that uses a different system, registrars will need a different system in order to access the radiology records. In some cases, where registrars are working remotely or from home, they will often require a VPN in addition to their login credentials.

Cognitive Load

One commonly time-consuming part of abstracting is the mental load. A mental load can arise when registrars are trying to predict the sequence of events that patients went through and when registrars are trying to interpret information provided in the physician's notes.

During case abstracting, registrars search for the diagnosis, procedures, and treatments the patients received. They collect information from multiple sources and arrange it in chronological order, thereby, building a series of events. Building this series of events often requires vast knowledge and experience in cancer. Moreover, it requires a familiarity with the protocols of the facility, as well as the resources available. One interviewee described the process as putting pieces of a puzzle together, where they try to find the answer to what they are looking into. One registrar commented:

> "When you do that abstracting for patients, you are writing their story, you are the author. You want to make sure you have all the facts, the dates, the treatment collection, date of birth, name, so when you write, your comments, you have to be clear as to what happened to that patient."

The sequence of events has to follow a logical treatment path using the available data. This process can become complicated when some of the expected events (such as treatment or procedures) are missing. Registrars will then try to find out which data is missing or which procedures were not performed.

The second factor that contributes to mental load is interpretation of the physician's notes. Some of the information that needs to be reported is usually found in the physician's notes. Although physicians' notes can have a lot of the information that registrars need, they are not designed for that purpose, but rather they are unstructured and written as text. Moreover, the terminology used may differ from what is required by central registries. After finding the required information, registrars will have to interpret the notes and make sure they answered the questions posed by the central registry. Interpreting this information requires not only a solid understanding of the domain, but also an understanding of the patient's individual situation and contexts.

Many False Positive Cases Identified from ICD-9 Codes

As described previously, case finding is the process of identifying newly diagnosed cases that need to be reported. The second commonly used source for case findings after pathology reports is the ICD-9 list. During the case finding, registrars search for codes relevant to cancer. Searching a hospital's database for a predetermined set of codes may have the added advantage of leveraging technology for fast retrieval. However, this may result in many non-reportable cases also being retrieved by the system. The range of the estimated cases identified through ICD-9 lists reported by interviewees was only 2.5% to 11%. To filter them, registrars manually review the results and verify their eligibility for reporting.

One reason that a retrieved result may not be reportable is that a patient has a history of cancer. During the manual review, registrars will need to determine whether it is a new case that needs to be reported, there is a history of cancer, or this is a second primary or recurrent incident of cancer. For instance, in the pathology reports in which 90% to 95% of case findings are identified, classification can be determined by reading the report. However, with the disease indices, classification will be much harder to determine, because registrars receive a list of codes, often in ICD-O, with insufficient details to classify.

When the case is identified through the indices code and there is no pathology report available, the physician's notes might be used. However, registrars have reported some issues associated with physicians' notes, such as lack of information and ambiguous terminology. When this occurs, registrars will search additional sources to determine case repeatability. Such a review process can be difficult and time-consuming, especially because a large number of cases are often found to be non-reportable.

Case Related Barriers

Registrars seem to agree that abstraction time can vary widely among cases. That variation was not only found among cancer types, but also among different cases within the same cancer type. The results show that some cancer types are known to be more time-consuming than others. Most interviewees have indicated that head and neck cancer is often harder and requires more time to abstract. Registrars have often indicated that

cases referred to as complicated, such as head and neck, generally take longer to abstract than simpler cases, such as breast or lung. With head and neck cancer, patients could receive multiple treatments and procedures, where the treatment path can be more complicated than other cancer types. Moreover, due to the nature of the close overlapping anatomy in the head and neck, it may take longer to determine the amount of tissue involved and the extent of tumor spread. Unlike head and neck cancer, kidney cancer is known for requiring less treatment and is often easier and faster to abstract. The results also show that even within the same cancer type, some cases may take longer than others.

Additional Activities

Some registrars have indicated that administrative tasks, such as reviewing compliancy and serving on a tumor board or cancer committees, could be time-consuming. For instance, the frequency of committee board meetings can range from 1-to-6 per month, depending on hospital size and activities. Meetings require preparation of reports, and PowerPoint presentations, among other documents. Additionally, cancer committees can be conducted 4 times a year, at which times registrars have to review all standards (around 25), make goals, and vote. It is important to note that our interviews were focused on the time needed to report to the state registry, whereas most hospitals are also required to report to the Commission on Cancer (CoC) and the state registry. One registrar commented: "If I just abstract for the state, I would be part-time."

In addition to reporting, the CoC requires participating hospitals to perform continuous follow-ups, which involve updating patient status, cancer status, any recurrence, new cancer, or new treatments. To perform follow-ups, registrars continue searching and updating patient information for life.

Other barriers were software-related. Some registrars indicated that open-source software, like Rocky Mountain, only provides basic features and does not provide any of the additional functionalities that can promote an efficient workflow, especially for matching cases and case follow-up.

Although registrars were using different systems, most of them tend to agree that abstracting time ranges from 45 minutes to 1.5 hours. Variation in the usage of oncology management systems seems to have been more significant at the case-finding and follow-up stages, with activities such as matching records and importing data, than it did at the abstracting phase. Only one registrar reported dissatisfaction with the Rocky Mountain system at the abstracting stage, and that dissatisfaction was related to checking for errors after report completion.

Physician's Notes and Information Contradiction

A less common barrier identified during abstracting is contradiction in information found in the records. As described previously, registrars collect information from various sources to describe patients' cases and the treatment path that was followed.

In some rare cases, registrars find contradictions between these sources, such as physicians' notes and pathology reports, or even within physicians' notes. Contradictory information is not common, but it can be found more often when multiple physicians treat the same patient. The results also indicated that, when this happens, it is often a complicated case.

Another barrier is cancer confirmation. Most cancer diagnoses are confirmed through biopsy, but when pathology reports are not available, other information sources, such as physicians' notes or diagnostic imaging reports, are used. The difficulty arises when uncertain language is used. Terms like "probable, suspected, likely, questionable, or possible" will lead registrars to seek more data sources for diagnosis confirmation.

4.3.4 Redesigned Workflow

The redesigned workflow focused on minimizing the time cases reside in the suspense file, as well as facilitating communication between registrars and their access to information. To achieve this, we proposed the following:

- An electronic pathology (E-path) reporting system

- A notification system for treatments

- A secure messaging system

- Access to HIEs, such as INPC

Incorporation of the above elements into the workflow process is illustrated in Figure 4-11.
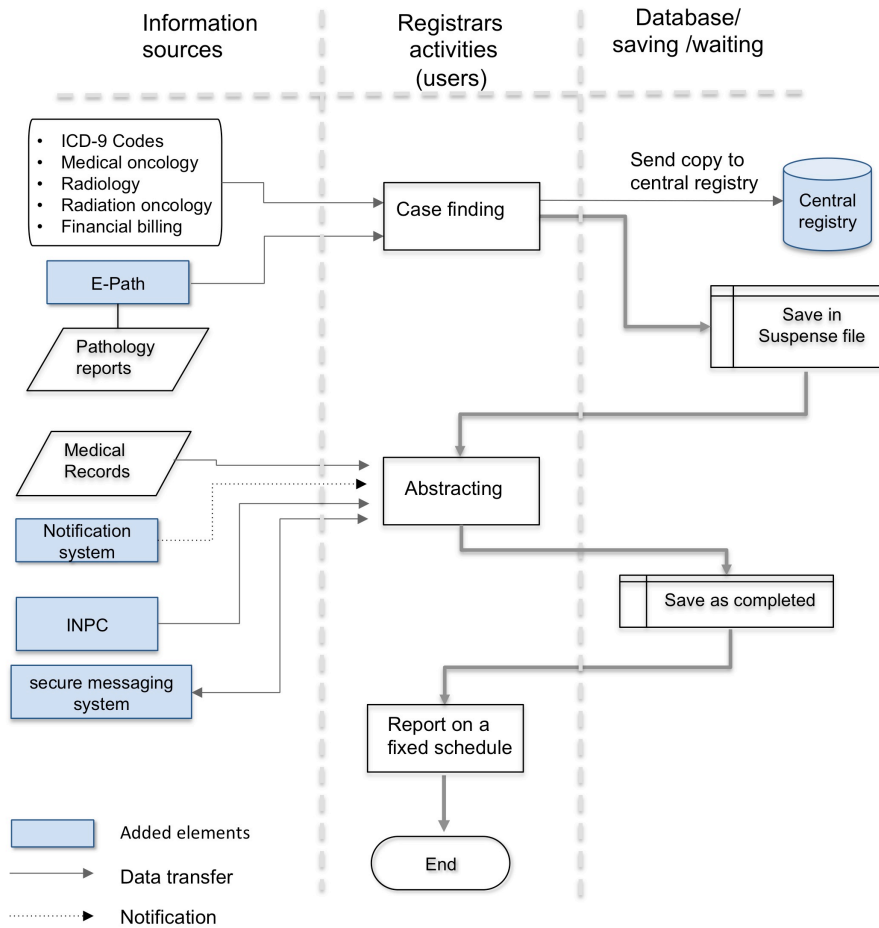
Figure 4-11: Cancer reporting flow chart (systems added in the redesigned workflow).

Workflow Steps

Figure 4-12 shows the flowchart for the proposed workflow. The steps for the proposed workflow will be as follows:

1. Cancer cases are identified through pathology reports using an automated tool (E-path).

2. Registrars visually check cases identified by E-path for reporting eligibility.

3. Case findings from other sources are done manually by registrars and matched with pathology reports (if available).

4. Cases identified as reportable are saved into the suspense file for abstracting. A copy of the identified cases is sent to the state registry and marked as incomplete.

5. EHR sends a notification to the oncology management system of any new cancer-related treatment. If the notification matches any of the cases in the suspense file, then the case will be flagged.

6. The registrar will check flagged cases and start abstracting. If no new treatment is received within 6 months of the case finding date, then the registrar will start abstracting and check the physician's notes and discharge notes regarding whether treatment was provided elsewhere.

7. If an abstract is completed, then it is reported to the central registry. If treatment is received at an external facility, then the registrar will search the HIE—INPC in this case—for additional data and use the secure messaging system to contact registrars at the external facility.
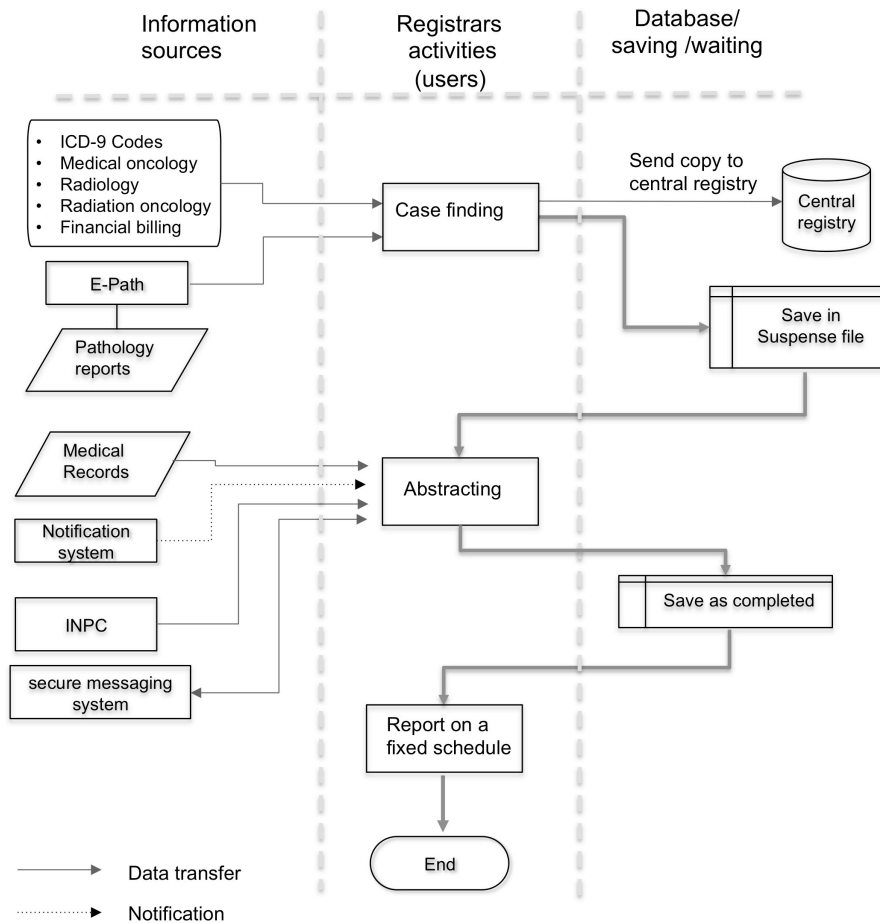
Figure 4-12: Cancer reporting flow chart (redesigned workflow).

4.4 Phase 4:Simulation

4.4.1 Current Process Simulation

The simulation was constructed using Anylogic software version 7.1. The current workflow was simulated using Discreet Event Simulation (DES). The flow of data and reporting steps were analyzed and the time spent was measured. The simulation model was then validated to reflect the actual processing time. Details of the validation process

are described in the methods section. The simulation diagram elements and model are described below (Table 4-4, Figure 4-13).

Entity	Location	Service	Action
Tumor case	Case finding (pathology and indices codes)	Add the new incoming pathology reports and indices codes.	Source
Tumor case	Case finding (pathology and indices codes)	Stack incoming pathology reports to be reviewed.	Queuing
Tumor case	Case finding (pathology and indices codes)	Review pathology reports by registrars.	Services
Tumor case	Case finding	Supply registrars to perform the case finding.	Resource pool
Tumor case	Filtering	Filter non-reportable cases.	Select output
Tumor case	Case finding	Remove non-reportable cases.	Sink
Tumor case	Suspense file	Store cases while waiting for treatment and other procedures to be performed.	Delay
Tumor case	Abstracting	Split cases requiring data retrieval from external sources from those available in local records.	Select output
Tumor case	Abstracting	Abstract patients' information.	Delay
Tumor case	Save completed reports	Stack cases before reporting to the central registry.	Delay
Tumor case	Reporting	Report to the central registry.	Sink

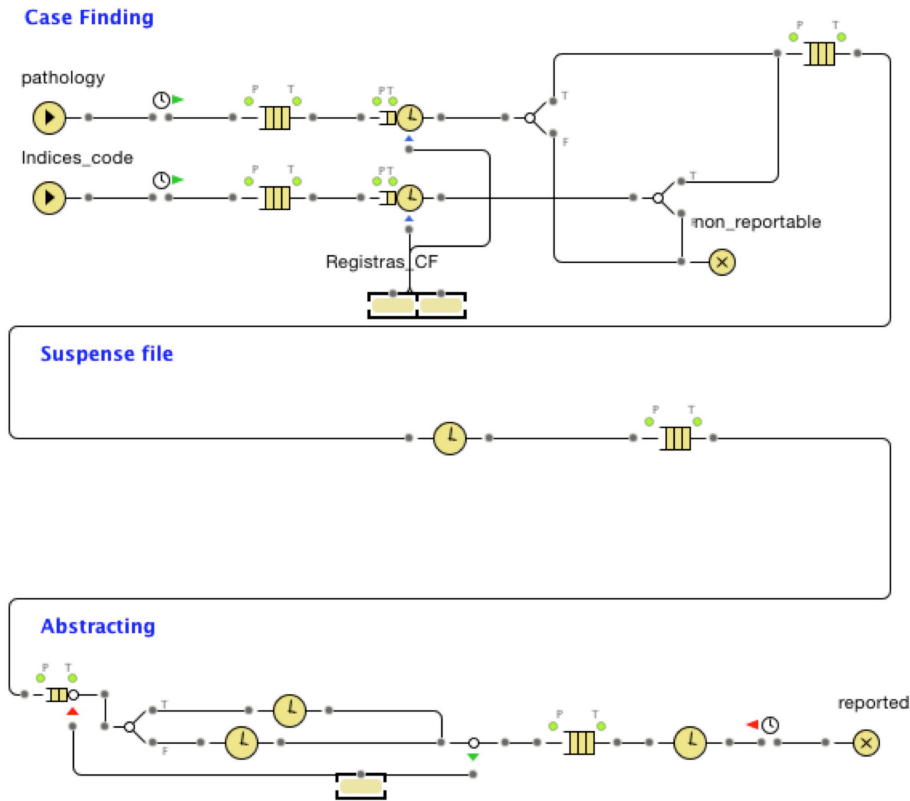Table 4-4: Description of simulation model parameters.

Figure 4-13: Cancer reporting (current status) discrete events model.

The simulation shows the lower, average, and upper times needed for reporting. Running the simulation for the existing workflow resulted in the following:

Running Time		Reporting Time (min, avg., max)	
Business Days	Calendar Days	Business Days	Calendar Days
260	One year	(73, 99, 127)	(102.2, 138.6, 177.8)
520	Two years	(73, 99,129)	(102.2, 138.6, 180.6)

Table 4-5: Simulation output (current status).

4.4.2 Proposed Workflow Simulation

This section describes simulation of the proposed workflow. For the proposed process, the time that cases should reside in the suspense file could not be estimated by the registrars who were interviewed. However, the registrars seemed to agree that this time can vary, largely based on factors like cancer site, stage, and hospital resources. To estimate this time, we used previous studies that calculated the time between diagnosis and treatment (Bilimoria et al., 2011). The study analyzed 1,228,071 patient records from 1995 to 2005. Data retrieved from the National Cancer Database represented around 1,443 hospitals in the U.S. The results included the median days and interquartile range waiting times for breast, colon, esophageal, gastric, liver, lung, pancreatic, and rectal cancers. Our study focused on breast, lung, and colorectal cancer. The waiting time for breast, lung, and colorectal cancer was (14; 24–40), (20; 37–63), and (13; 26–46), respectively.

The proposed notification system was simulated for the 3 cancer types studied (breast, colorectal, and lung). Based on the Indiana "Cancer Facts and Figures 2012" report, the distribution among these 3 cancer types was as follows: 33% breast, 26% colorectal, and 41% lung (ISDH, 2012). This distribution was used to simulate the probability of cases added to the suspense file.

The simulation for the proposed model can be seen in Figure 4-14. The results of the proposed model simulation can be seen in Table 4-6. The redesigned process simulation shows an average reporting time of 51 days, compared to 138 days with the current process simulation (Table 4-7).

Running Time		Reporting Time (min, avg., max)	
Business Days	Calendar Days	Business Days	Calendar Days
260	One year	(14, 37, 68)	(19.6, 51.8, 95.2)
520	Two years	(14, 37, 72)	(19.6, 51.8,100.8)

Table 4-6: Simulation output (redesigned status).

Simulation Time	Existing Reporting Time (min, avg., max)	Redesigned Reporting Time (min, avg., max)
One year	(102.2, 138.6, 177.8)	(19.6, 51.8, 95.2)
Two years	(102.2, 138.6, 180.6)	(19.6, 51.8,100.8)

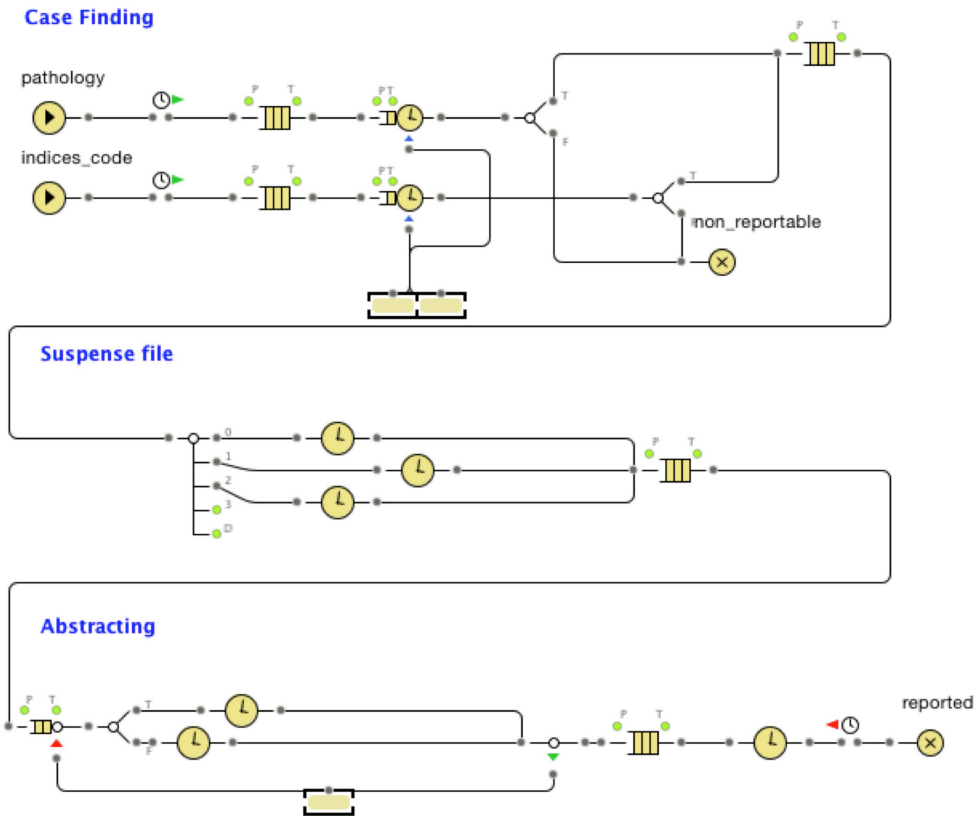Table 4-7: Comparing simulation output (current and redesigned status).

Figure 4-14: Cancer reporting (redesigned workflow) discrete events model.

CHAPTER 5: DISCUSSION

5.1 Introduction

The goal of this study was to evaluate the timeliness of cancer registry reporting while exploring ways that the reporting process might be improved. This study was motivated by the increasing demand for timely cancer data to support research, quality care, and public health surveillance. Phases 1 and 2 of the results evaluate timeliness of data reported to the state registry and steps within cancer reporting processes, respectively. Phase 3 explores the barriers encountered during reporting and offers a redesigned workflow for timelier, more efficient reporting.

As described in previous sections, cancer reporting involves a series of steps that include case finding, abstracting, and submission, all of which are fundamental to the reporting process. Reporting steps involve the information search, retrieval, and comprehension, in addition to data entry. To situate the evaluation, it is necessary to understand the steps involved, the time needed, and the challenges experienced during each stage.

Previous studies have evaluated timeliness in cancer reporting by measuring reporting time as a single process from start to finish. Although this provides a clear indication of the overall reporting time, the exact duration of each stage within the

process remains unknown. By separating this process into multiple steps, we gain more detailed information about reporting efficiency at both the organizational and state level.

In a review by Jajosky and Groseclose (2004), the authors concluded that surveillance and timeliness studies in public health lack detailed descriptions of reporting stages, such as processing and analyzing. Their review also emphasized that both collection and assessment of time interval dates are an important part of any surveillance system or timeliness assessment (Jajosky & Groseclose, 2004).

The novelty of our study involves leveraging the time interval dates to analyze the reporting process historically and at the state level. Doing so enables us to calculate the time each step took within the reporting process, as well as examine changes in timeliness for each step over the years.

Examples of the knowledge acquired using this approach include: the time taken by the reporting facility to abstract the case report and send it to the registry, the time taken by the facility to complete the abstracts, the amount of time completed reports remain at the facility database before being sent to the state registry, and the time taken by the cancer registry to process the report and upload it to their database.

Longitudinal data (over the course of 9 years) enabled us to assess changes in reporting timeliness for each step. As seen in Figure 4-4, the time taken by the state registry, for instance, was reduced significantly by 2004.

A key benefit of the applied method is the granularity of results, because the outcomes were stratified by different facilities. The stratification allowed us to determine whether observed timeliness for a given step is universal or unique to a particular facility or step. One example is the variation among facilities in the time taken to send completed reports to the state registry.

Time intervals at each step were utilized to identify potential causes for observed delays at some facilities, and some facilities were found to keep completed records at the local database longer than expected. The interviews offer one possible reason for the delay: an uneven flow of patients visiting the facility. Patient flow, in turn, is affected by holidays and the rush at the end of the year to meet health insurance deductibles. Using time intervals allowed us to check the distribution of cases, the time of diagnosis, the time of report completion, and the time of report submission.

Whereas prior studies examined timeliness as part of a data quality evaluation, they generally lacked description of the reporting process. While this emphasis is helpful with understanding the quality of data, it does not provide sufficient knowledge of the efficiency of the process or how the reporting process could be improved.

In this study, we investigated the process of data production by collecting objective and subjective data. Objective data was obtained from dates and timestamps to evaluate the process at the state level over an extended period of time. Subjective data were collected during registrar interviews, and it provided deeper insight into the details

of the process. To explore this understudied area, we conducted semi-structured interviews to support the grounded-theory approach of deeper understanding and theme discovery. During the interviews, registrars were also encouraged to think beyond technical limitations through "blue sky" thinking.

5.2 Timeliness Overview

The results showed a sharp increase in reporting time during 2003, which was concurrent with changes in the policy and procedure requirements for reporting facilities. Starting in 2003, the ISDH Cancer Registry implemented the Facility Oncology Registry Data Standards (FORDS) coding standard. FORDS was developed by the Commission on Cancer (CoC) and is now required by all CoC-approved cancer programs to define data collection and coding requirements. Another 2 major changes occurred in 2004. By the beginning of that year, reporting facilities were required to report all benign and borderline brain and central nervous system (CNS) tumors. Additionally, the ISDH Cancer Registry also required all reporting facilities to start coding using the Collaborative Staging System (ISDH, 2004).

Recent studies indicate the importance of the secondary use of data and the role of health information systems in supporting this use. Special attention has been given to cancer and the importance of data in understanding and improving treatment and care. Accessibility of recent data is key for timely intervention and quality improvement (Abernethy et al., 2010). Currently, our results show an average timeliness of 323 days,

which suggests that data users will have to wait that amount of time after diagnosis before they can access this data via the state cancer registry. Our results also show that researchers are able to access an average of 17.6% of cases within the first 6 months of diagnosis, 68.7% after the first year, and 96.5% after 2 years. The percentage of cases available within these time periods improved as of 2009, reaching 37.2%, 82.6%, and 97.5% for 6 months, 1 year, and 2 years, respectively. Although this seems like a long time from a research perspective, it is tacitly accepted by most national registry standards.

National cancer registries provide funding to participating hospitals and states to collect cancer data. Participation in cancer registry programs requires meeting their standard for data quality. The following paragraphs will discuss the timeliness of ISDH Cancer Registry data compared to some of the most common cancer registry standards.

The ISDH Cancer Registry was certified gold by NAACCR in 2001, 2002, 2004, 2006, 2007, 2008, and 2009. Gold certification requires 95% of cases to be reported within 23 months from the date of diagnosis. In 2003, 2005, and 2010, the ISDH was silver certified, indicating that 90% or more of the cases were reported within 23 months (NAACCR, 2015).

Compared to our results, the percentage of cases reported within the first 23 months was 95% or higher, except in 2004, 2006, and 2008 (Table 5-1). The discrepancy between the NAACCR certification level and our results can be found in Table 5-2, but,

in general, the difference was less than 1%. Such a small amount illustrates that our data, although limited to a subset of cancer types, was representative of all data reported to ISDH. Furthermore, it underscores that ISDH is meeting the timeliness requirements for NAACCR. However, there is still room for improvement in meeting the timeliness standard for SEER, or for enhancing efficiency of the process.

Months (N)	2001	2002	2003	2004	2005	2006	2007	2008	2009
12	71.4%	82.1%	34.4%	59.2%	68.8%	76.6%	74.7%	70.6%	82.6%
22	97.1%	98.3%	93.7%	93.3%	94.3%	94.2%	95.3%	92.9%	97.5%
23	97.6%	98.5%	95.0%	94.1%	95.3%	94.7%	96.0%	94.5%	97.8%
24	98.1%	98.8%	95.8%	94.8%	95.7%	95.3%	96.6%	95.9%	98.2%

Table 5-1: Comparing national registry timeliness standards to our results.

SEER 98% within 22 months.

NACCAR: Golden 95% within 23 months.

NACCAR: Silver 90% within 23 months.

CDC/NPCR 12 months for reporting 90% of cases after the date of diagnosis or 24 months for 95%.

Year	Certification level	Target	Our results	Discrepancy %
2001	Gold	95%	97.6%	0
2002	Gold	95%	98.5%	0
2003	Silver	90%	95.0%	0
2004	Gold	95%	94.1%	0.9%
2005	Silver	90%	95.3%	0
2006	Gold	95%	94.7%	0.3%
2007	Gold	95%	96.0%	0
2008	Gold	95%	94.5%	0.5%
2009	Gold	95%	97.8%	0

Table 5-2: Comparing NAACCR certification level to our results.

Another known cancer registry standard is SEER. Currently, the ISDH registry is not part of SEER. To meet the timeliness standard for SEER, state registries are required to report at least 98% of their cases within 22 months from the date of diagnosis (Table 5-1) (NCI, 2014). Our results show that SEER standards were only satisfied in 2002, with 98.3% of the cases reported within 22 months. For the remaining 9 years, the highest percentage of cases reported within 22 months was 97.1% in 2001 and 97.5% in 2009, respectively.

The third cancer registry standard is CDC/NPCR. The ISDH is currently participating in CDC/NPCR, which allows 12 months for reporting 90% of cases after the date of diagnosis and 24 months for 95% of cases (CDC, 2013). Our results show that the percentage of cases reported within 12 months was less than 90% for the entire 9 years. However, the percentage was higher than 95% within 24 months for all 9 years except in 2004, when it was 94.8% with a 0.02% difference (Table 5-1).

Our study only included 3 of the most common types of cancer, while national registries consider all cancer types. The low variation in reporting timeliness among the different cancer types may have led to a more accurate estimation. Moreover, registry standards take several DQ measures into account, including measures like the missing data rate and the overall data validity. The NAACCR gold certification, for instance, requires the following:

- less than 3% of cases use a death certificate for identification,

- less than 1% are duplicate files,

- 100% error-free for some variables,

- less than a 2% missing rate on age, sex, and county, and less than a 3% missing

  rate on race (NAACCR, 2015).

The calculation method used to arrive at these results should also be considered. Where some studies use a date stamp to calculate the difference between the start and end date, others use the observed and expected number of cases. In this study, we calculated the difference between the date of diagnosis and the date when the data was made available on the state registry by using data items that conform to NAACCR standards. However, the CDC/NPCR uses the observed-to-expected ratio to determine the rate of completeness within a specified period (CDC, 2013). This method uses the registry's previous reporting experience to determine an expected number of cases. Because this method may have the advantage of estimating rates of completeness for recent years, historical data can leverage the date stamp for a more accurate measurement. This is an especially important factor if the number of annual incidents is inconsistent. The annual number of cases included in this study range from 6,693 to 7,726, with the lowest being 6,693 observed in 2009. Due to the lengthy process of cancer reporting, cases diagnosed within recent years often take more time to be reflected.

Prior work reported that timeliness varied by cancer types and the variation was explained by the fact that different types of cancer are often diagnosed and reported by different sources, such as hospitals, laboratories, or physicians' clinics (Smith-Gagen et al., 2005). Gagen et al. (2005), for example, reported that breast and colorectal cancer are often diagnosed and reported by hospitals and, therefore, tend to be reported quicker than melanoma or prostate cancers, which are often diagnosed and reported by non-hospitals (Smith-Gagen et al., 2005). In our study, the 3 cancer types examined showed similar reporting times. The impact of the reporting source on timeliness could not be examined in our study, because around 98% of the cases used were reported by a hospital. A recommendation for future studies is to examine the variation in timeliness by reporting source.

Our study shows similar reporting times for the 3 cancer types. The difference in average reporting time was insignificant, with a high of around 35 days during 2008. The similarity in reporting times among the different cancer types could be explained by understanding the reporting workflow. As seen in Part 3 of the results, different cancer types are grouped together and processed similarly (more details in Phase 3 of the results).

5.3 Reporting Stages Timeliness

A key step for process improvement is identifying the steps that consume most of the reporting time. Reporting time can be affected by many factors, including caseload,

access to information, disease type, and staffing. Such factors are often assessed through subjective tools, including surveys and interviews.

To enhance our understanding of the impact of these factors and to understand the extent to which each affects timeliness, we also used objective data. Each factor can affect different stages of the reporting process, and in Phase 2, we examined the time taken at each stage. The main stages examined were registry processing time, report submission time, and report completion times.

Phase 2 also examined the variation in timeliness among reporting facilities. Phase 1 shows that the overall timeliness of cancer report submissions to the State's Department of Health did not differ based on the type of cancer. However, vast differences in reporting time between cases was demonstrated by the widely distributed box plots. In this section, we examine the differences in reporting time among facilities.

5.3.1 Registry Processing Time

By separating the time taken for case finding, abstracting, and report submission at each facility from receipt of the reports, we found that the registry requires a very short amount of time for processing. A remarkable reduction in registry processing time was found after 2004. By 2004, the time taken by the registry to process and edit the cases received started to decrease from 20 to 30 days to only a few days (less than 10 days). Additionally, the variation in processing time among cases substantially decreased after

2004; whereas in 2002 and 2003, processing time could extend up to 75 days, and in most of the remaining years from 2005 afterward, all cases were processed in less than 10 days. The change in processing time was concurrent with implementation of a registry system using the File Transfer Protocol (FTP) (ISDH, 2004). FTP is an Internet protocol standard for exchanging files online. It was implemented to substitute the traditional means of physically sending discs. Submitting reports online can enable the real time transfer of data compared to post mail.

5.3.2 Report Completion Time

To further examine timeliness at the level of reporting facilities, we separated the time taken to complete the reports from the time taken for submission. The results show clear variability in reporting speed among facilities. Whereas most facilities took 175–200 days to abstract and complete reports, some took almost 300 days. Many factors, such as staff numbers, hospital recourses, or registrars' access to medical records, could have influenced variability in report completion time.

Abstracting cancer reports is a labor-intensive and time-consuming process. Cancer registrars are required to go through long training sessions and pass certification requirements. Registrars need to meet demand, and shortages of qualified registrars appears to be common. In 2004, Tangka and Subramanian et al. conducted an economic evaluation study on the 4 central registries in operation. The study collected information with the aim of quantifying the cost effectiveness of registry operations. The study

101

reported that staffing shortages were the most frequently cited problem. Staffing shortages were reported by more than 50% of the registries and over 70% of the reporting facilities (ACS; Tangka, Subramanian, Beebe, Trebino, & Michaud, 2010).

The number of registrars needed often depends on the facility's caseload. Though the number of diagnosed cases was nearly equally distributed across different quarters in the same year, the annual caseload at the same facility was inconsistent across years. This means that patients come to facilities at a constant rate within the same year, but the number of patients coming from one year to the next can change significantly. In facility 1 (Appendix 6-2), for instance, the number of reportable cases for the 3 cancer types diagnosed in 2007 was around 300, and it nearly doubled in the subsequent 2 years, reaching around 600 cases in 2009. Similarly, with facility 6 (Appendix 6-2), the number of reportable cases diagnosed during 2005 was less than 100 cases, but this number increased to nearly 175 cases in 2008. To enable facilities to complete reports at the same speed, they need to maintain a consistent registrar-to-caseload ratio. Having a highly inconsistent number of cases diagnosed each year creates staffing challenges.

Hospital structures and registrars' access to medical records also can affect report completion times. Hospitals with segregated systems require more time and effort to collect the information needed to complete the reports. The use of segregated systems can be attributed to technical or organizational reasons. One technical reason involves having a paper-based or legacy system that is hard to access. However, organizational reasons can include having outside facilities that are not part of the same hospital network. If

patients are diagnosed in an office, for example, and receive treatment or procedures at an independent facility with a separate medical records system operated by another organization, registrars will have difficulty accessing these patient records. Such a situation can prolong the time taken to complete the report, and it is more common when patients tend to receive care at multiple locations.

5.3.3 Report Submission Time

Similar to reporting completion times, variations among facilities reporting submission times was examined by calculating the average submission time for facilities with higher caseloads. Both report completion time and report submission time showed variability among facilities; however, a much higher variability was found in report submission times. In this context, submission time refers to the time between report completion and receipt of the report at the state registry.

In this study, 22 out of 49 unique facility IDs were classified as higher-load facilities. These 22 facilities accounted for 90.5% of the total number of cases, leaving only 9.5% for the remaining 27 facilities. By examining report submission times at the 22 facilities, we identified 7 facilities that took significantly longer than others (at least twice as long). Although a longer submission time was found only in 7 of the 22 higher-load sources, those 7 accounted for around 28.4% of the total number of reported cases. Such a high percentage can heavily affect registry standards. With most cancer registry

standards, the completion rate is weighted heavily. The SEER standard, for instance, requires 98% of the cases to be reported within 22 months of the date of diagnosis.

Nowadays, transferring reports from facilities to state registries is mostly handled electronically. Unlike case finding and abstracting, submitting reports is easier and less time consuming. One way to speed up the reporting process is to encourage facilities to submit completed reports more often. Submitting reports more promptly can potentially improve overall registry performance, especially if facilities with longer submission times are targeted. The observed difference in how quickly reports are submitted among facilities could also indicate a variation in their reporting practice. This highlights the need for examining variations in reporting practices among facilities.

5.3.4 Reporting Frequency

Submission time was further examined by calculating the frequency of reporting. In this step, we focused on facilities with longer submission times. Those reporting facilities classified as high-caseload facilities in this part of the analysis were expected to report their cases every 90 days or less. The results show that the distribution of cases submitted every quarter differed from facility to facility. In some years, facility 6 and 20, for instance, submitted all their annual cases in one quarter.

We examined the frequency of reporting in previous steps to decide whether the frequency of irregular reporting originated during the submission time or was inherited

from prior steps. The results also show that both the rate of cases diagnosed and the rate of reports completed within a given year were relatively equally distributed most of the time. This suggests that the irregular reporting frequency observed during the report submission phase occurred at this step and was not inherited from previous steps. It also suggests that barriers to regular reporting are unlikely to be encountered during the diagnosis or report completion phases.

One of the barriers observed during the diagnosis phase was the irregular flow of patients in a given year, which means that patients visited the facility and were diagnosed more during certain seasons. At the report completion phase, the frequency of irregular reporting can occur when more registrars (e.g., part-time or contract employees) work during particular quarters. Another potential cause for irregular reporting involves having more patients receiving treatment, tests, or procedures that are needed to complete the report during certain quarters.

5.4 Barriers to Timely Reporting

The identified barriers encountered during the reporting process were categorized into 6 themes. The most critical and common themes were access to data at other facilities and case finding using index codes. The table below lists the identified themes in their order of importance.

Identified Themes	Key Barriers Identified
1) Access to information-related barriers.	• Difficulty accessing information within facilities that are outside the hospital network. • Data exchange between electronic systems. • System session timeout.
2) Many false-positive cases identified from the ICD-9 codes.	• A large number of non-reportable cases that are flagged for review. • ICD-9 codes are not sufficient for confirming the repeatability of flagged cases. • Some systems do not have the ability to distinguish previously reported cases from new cases.
3) Cognitive load-related barriers.	• Combining different events into a single coherent abstract. • Interpreting some of the information in the medical records and translating it to fit registry requirements.
4) Case-related barriers.	• Cases with overlapping anatomy. • Complicated cases with many procedures.
5) Additional activities required.	• Administrative tasks, such as reviewing compliance, as well as serving on the tumor board and cancer committees. • Reporting for other institutions, such as the CoC, with different requirements.
6) Physicians' notes and information.	• Different treating physicians sometimes report contradictory information. • Text reports using uncertain language, such as "probable," "suspected," "likely," "questionable," and "possible."

Table 5-3: Summary of barriers to timely reporting.

During the interview, registrars were encouraged to describe optimal cancer reporting mechanisms through "blue sky" thinking. By asking them to describe optimal status, the responses often focused on eliminating the barriers they encountered. For most registrars, optimal status was associated with better access to information or easier case

findings. Access to information was expressed as more access to medical records or envisioning a central database for all information.

However, easier case finding was often associated with system integration and the ability of a system to exchange information. Registrars wanted system integration to enable record matching during the case-finding process. As described above, matching potentially reportable cases with previously reported ones can help registrars identify duplicates and focus on the new cases. This function is available in some commercial oncology management systems, but it requires a level of compatibility between the EHR and oncology management system used.

Only 3 registrars associated optimal status with automating the reporting process. One registrar referred to E-path as a successful example of automation in cancer reporting.

5.5 Workflow Recommendations

The Results Chapter presented several barriers to timely reporting. Many of the identified barriers can potentially be minimized by redesigning the workflow at the hospital level. Another alternative approach is to leverage the HIE for centralized workflow. In this section, we will discuss 2 approaches to redesigning reporting workflow: hospital-based and national-based (centralized) workflow.

5.5.1 Hospital Based Workflow

The Results section described a recommended workflow at the hospital level. To assist registrars with case finding from pathology reports, we recommend using the E-Path system. The redesigned workflow that we also recommend minimizes the time cases reside in the suspense file by using a notification system for treatment. Finally, we recommend utilizing HIE and the secure messaging system during the casefinding and abstracting phases.

Electronic Pathology Reporting System (E-path)

As indicated previously, 90–95% of cases identified at the casefinding stage are identified through pathology reports, followed by the ICD-9 list and others (Appendix 6-3). The number of cases that need to be reviewed during casefinding can be reduced using the E-path system. E-path utilizes a text mining to automate the identification and coding of pathology reports. Using E-path for casefinding has been shown to improve reporting timeliness and increase reporting efficiency (Dale, Golabek, & Chong, 2002). Many states adopted E-path in their workflow to save registrars time and reduce casefinding times. Some of the challenges include the cost associated with the software implementation and infrastructure needed to send HL7 messages (NAACCR, 2011).

The second casefinding source is the ICD-9 list (disease indices). Unlike pathology reports, index codes have much less information, because they display code

without further information on how or why the case was coded in such a way. To confirm the diagnosis, registrars often need to cross-reference the disease indices with additional sources, including pathology reports, surgery reports, and/or oncology records. In this proposed workflow, cases that are identified as reportable by pathology reports and have codes associated with them will distinguish them from others. If the information collected from the pathology report is sufficient to confirm the case as reportable, that case will be ready for abstracting and further review, since index codes will not be needed. Index codes can also be matched with previously reported cases to help distinguish new cases from recurrent ones. Such functions can be made available in some commercial oncology management systems.

Notification Systems

After the case is identified at the casefinding stage, it is stored in a suspense file. Generally, cases reside in the suspense file for a few months, as treatments and procedures are performed and entered into the EHR system. However, procedures and treatments can be done at different speeds, depending on factors, such as the cancer type, stage, and facility resources, among others. Using a standard waiting time for all cases can create an unnecessary delay if treatments are made available earlier than expected.

In the proposed workflow, we recommend a communication system between the local oncology management system and the EHR to determine when to start abstracting. A notification system will alert the reporting registrar when new treatments are available

in the EHR for any cases in the suspense file. Using a notification system, registrars will be able to abstract the case as soon as treatment is available. If a patient receives no treatment within 6 months from the date of the casefinding, then the registrar can start abstracting. The unavailability of treatment in medical records may be due to possibilities, including receiving treatment at an external facility, refusing treatment, or death.

Leveraging the health information infrastructure for cancer reporting can positively affect the workflow. The use of notification systems for workflow optimization has been noted in other healthcare areas, such as care coordination (Moore et al., 2012). System notification can be implemented using Health Level Seven (HL7) Clinical Document Architecture (CDA) notification messages. Once a new treatment is added to the EHR, the notification system will match it with patient lists in the suspense file. If a match is found, registrars will be notified of the new treatment (see Figure 4-36). Some CDA specifications can be found at the *CDC Implementation Guide for Ambulatory Healthcare Provider Reporting to Central Cancer Registries, Release 1.0* (CDC, 2014).

Access to Health Information Exchanges (HIEs)

Access to external records was one of the biggest barriers that registrars encountered. Using some of the existing health information exchange systems to facilitate access to information can minimize many obstacles. Studies have shown the benefits of

HIEs for improving access to clinical data (Fontaine, Ross, Zink, & Schilling, 2010). Several states already have the HIE infrastructure to address this gap.

Especially in Indiana, the ISDH should explore the benefits of the Indiana Network for Patient Care (INPC) to address some cancer registry challenges. INPC includes a large portion of the medical records needed for cancer reporting, such as laboratory, radiology, and dictation. It covers 25,000 physicians, 106 hospitals, and 110 clinics and surgery centers within the State of Indiana (Indiana Health Information Exchange).

Moreover, the INPC presents the opportunity for automating the identification and detection of cancer, which is especially useful for monitoring and surveillance. Most cancer cases are diagnosed through laboratory histology. Many studies document the feasibility of automatic cancer identification and detection through pathology reports (Hanauer, Miela, Chinnaiyan, Chang, & Blayney, 2007; Kasthurirathne et al., 2016; Kavuluru, Hands, Durbin, & Witt, 2013). As the study by Kasthurirathne et al. (2016) showed, the automatic information extraction of pathology reports is a feasible and practical approach. The study also suggested that this approach could be generalized to other clinical data (Kasthurirathne et al., 2016).

In addition to cancer monitoring and surveillance, INPC can support a wide range of cancer research activities. Within the State of Indiana, the Regenstrief Medical Record System (RMRS) and INPC have previously provided data needed in diverse research

activities and domains (Dixon, Whipple, Lajiness, & Murray, 2015). Research activities include observational studies, health services research, and comparative effectiveness research (Dixon, Whipple, et al., 2015).

Secure Messaging Systems

Results show that some registrars encounter difficulties when asking clinicians for patient information at external facilities; because of this, they then contact registrars at external facilities in order to access patient information. This was perceived as more efficient in terms of accessing the information needed, given that registrars' understand each other's job roles and reporting requirements. In this workflow model, we propose usage of a secure messaging system to facilitate communication among registrars and to minimize the access barrier. Studies have shown that use of secure messaging systems in other clinical settings improves communication effectiveness among health professionals (Joos, Chen, Jirjis, & Johnson, 2006).

Potential for Improvement

Simulation was applied to estimate improvements in reporting speed. The redesigned workflow is expected to reduce reporting time from an average of 138 days to 51 days (Tables 4-7). This is based on the simulated assumption of time that cases reside in the suspense file (Figure 4-14). Despite barriers to the current reporting process, analyses show that cases spend the most time in the suspense file as registrars wait for

treatment data to become available. Even though registrars tend to agree that the time patients wait for treatment can vary among cancer types or patients with the same types of cancer, abstracting cases from the suspense file is performed in a chronological order that uses the FIFO (first-in, first-out) queuing system. By applying FIFO, the average wait time needed is more likely to be determined by the time needed for cases with higher waiting times so as to avoid early abstracting. We should note that early abstracting will increase the workload, as registrars may have to search, wait, and search again if treatment has not become available.

5.5.2 National Based (Centralized) Workflow

This section provides a brief workflow overview that describes a vision for a cancer reporting system. In this visionary workflow, some of the health information technologies will be incorporated into data transmission and retrieval.

As seen in the Results Chapter, the current process has many challenges that prevent timely reporting. Common challenges, such as access barriers and lengthy searching and retrieval times, cannot be simply eliminated by adding staff or increasing working hours. Reducing reporting time will require changing the current structure and workflow practice to facilitate data sharing and access to information. To reduce reporting time, we suggest shifting to centralized reporting. We also suggest leveraging some of the public health information exchange techniques for information transfer and case identification.

The most common barrier that was identified during the interview was access to external records. When patients transfer from one facility to another, registrars from both facilities are responsible for finding and retrieving the information from external records, which often entails many challenges, including identifying the facility to which the patient was transferred and then seeking access to the relevant records. When this occurs, registrars from both facilities must create abstracts and reports that result in an increased workload. Centralizing the workflow can eliminate many access barriers and duplicate reporting.

Meaningful Use and Public Health Reporting

The increased use of certified EHR and the requirements for Meaningful Use (MU) compliance present an opportunity for transforming the traditional cancer case reporting workflow into a centralized model. Recently, there has been significant emphasis on the implementation and use of certified EHRs, those which comply with regulations stipulated in the MU incentive program (HHS, 2015). Eligible hospitals and providers are implementing certified EHRs in record numbers (Adler-Milstein et al., 2015; CMS, 2015). The number of U.S. hospitals adopting the basic EHR systems grew from 59% in 2013 to 75% in 2015 (Adler-Milstein et al., 2015). The increase in EHRs' implementation is often accompanied with adherence to MU criteria. Studies shows that the number of hospitals meeting the core stage 2 MU criteria increased from 5.8% in 2013, to reaching 40.5% in 2015 (Adler-Milstein et al., 2015).

The increasing adoption of certified EHRs and MU compliance facilitated the utilization of public health reporting. Studies have documented the positive association between informatics capabilities and public health service provision, especially with services targeting population health (Mac McCullough & Goodin, 2014). More hospitals are participating in immunizations, syndromic surveillance, and reportable infectious disease electronic reporting every year (Health IT Dashboard, 2014). In addition to surveillance activities, studies within the U.S. are shifting toward population health application at the state and local levels (Dixon, Pina, Kharrazi, Gharghabi, & Richards, 2015). Most of the population health application and activities depend on the EHR capability of data collection and transmission. Since most systems cannot communicate effectively, HIE was introduced to function as the middle layer and facilitate the exchange of data (Finnell & Dixon, 2015).

One of the biggest concerns in HIE is data interoperability and standards. However, with the increasing implementation and use of EHRs, health information technology innovation and interoperability solutions are expected to advance (Sheikh, Sood, & Bates, 2015). The adoption of health exchanges and interoperability standards was also promoted by CMS incentives. MU stage 2 requires hospitals and healthcare providers to incorporate Clinical Document Architecture (CDA) for information exchange. CDA is a common HL7 standard for specifying the encoding, structure, and semantics of clinical documents for exchange. One of the most appropriate CDA documents for sharing detailed patient information with public health entities is the Continuity of Care Document (CCD) (Health IT, 2013). CCD suits this purpose, as it is

source-independent and enables the exchange of demographic data that is necessary for matching and clinical data (D'Amore, Sittig, & Ness, 2012). A simplified example of a CCD format can be seen in Figure 5-1 (D'Amore et al., 2012).


```
<section>
  <templateId root = "2.16.840.1.113883.10.20.1.11"/>
  <entry>
    <effectiveTime>
      <low value = "1982" >
      <high nullflavor="UNK">
    </effectiveTime>
    <observation>
      <value displayname = "Asthma" code ="195967001"
      codeSystem="2.16.840.1.113883.6.96">
    </observation>
  </entry>
</section>
```

Section tags divide data in the CCD.
Template Id identifies that this section will document problems.

EffectiveTime states problem timing:
The problem started in 1982.
The problem is not known to have ceased.

An **observation** of asthma is documented.
The tag of **codeSystem** informs that terminology used is SNOMED-CT.

While only one **entry** is shown in this example, entries can repeat within a **section** to document multiple problems.

Figure 5-1: A simplified example of a CCD structure. Source: D'Amore, Sittig, & Ness. "How the Continuity of Care Document Can Advance Medical Research and Public Health." *American Journal of Public Health* 102.5 (2012): E1–E4.

Workflow Description

The recommended workflow is based on a centralized approach in which most of the processing is performed at the central registry, rather than by the reporting hospitals. The workflow steps are described below.

Data Transfer

In this workflow, patients' records are transferred to the central registry for processing. The transfer of patient information can be accomplished using CCD. Triggering events can be created for automatic detection and transfer of cancer-related information. A key component of CCD transmission is the specification of identified data elements, vocabularies, and other requirements. One of the established efforts with regard to the cancer registry data elements is the *CDA Implementation Guide for Ambulatory Healthcare Provider Reporting to Central Cancer Registries* (HL7 International).

Records Matching

Records transferred from each facility will then be matched at the patient level at the central registry. This minimizes the need for duplicate abstracting, if patients receive care at more than one facility, as well as provides more complete and comprehensive records.

At this phase, we should mention some of the challenges associated with records matching, which includes the lack of unique patients' identifier. Currently, EHRs work solo, where each hospital creates a Medical Records Number (MRN) for its patients. Since the MRN differs in each hospital, it cannot be used for records matching. Currently, the U.S. does not have unique patient identifiers that could be employed for matching medical records (McFarlane, Dixon, & Grannis, 2016).

One of the most commonly used techniques to overcome this challenge is known as probabilistic matching. In probabilistic matching, a set of demographic attributes, including name, date of birth, address, and/or Social Security number are being used instead of a single records number. One of the advantages to probabilistic matching is flexibility in working with incomplete information and errors. Instead of determining whether or not the demographic attributes are an exact match, it assigns a score based on the degree of matching. A higher score indicates more confidence that the records being matched represent the same patient. With probabilistic matching, a cutoff score is often assigned to determine if the records being compared will be accepted as a true match. Achieving a higher score will require more complete and accurate data from all sources. Therefore, is important to encourage the collection of more complete, accurate, and up-to-date demographic data (McFarlane et al., 2016).

Case Finding

E-path will still be used as the main case finding tool, due to its high sensitivity and the large percentage of cancer cases that are diagnosed through pathology reports. The second data source for case finding will be the ICD codes (indices codes). In this workflow, the ICD codes will be matched with other patients' diagnostic reports that are retrieved through CCD. Unlike ICD codes, diagnostic reports can provide more detailed information about the cases, which will provide a clearer and more comprehensive view by combining patient information from multiple sources, thus making it easier to determine case eligibility for reporting and saving some of the time spent on information

search and retrieval. This will make the demographics, tumor type, and stage available sooner for rapid case ascertainment without the need to wait for treatment data.

Abstracting

Most of the data related to demographics, cancer type, and stage, can be collected at the casefinding stage. For the complete abstract, treatment and procedure data is often necessary. By utilizing the event triggers for treatment data, registrars will be notified about the availability of new treatments for any case identified at the casefinding stage. This will involve matching the received CCD with the cases that are identified at the case-finding stage. Once the abstract is completed, the records can be added to the central registry. By adding records as soon as they are completed, we can eliminate the waiting period between the record submission time and record completion time.

5.6 Study Contribution

System design and planning, data collection, and data management are essential aspects of health informatics (Savel & Foldy, 2012). The study of data production and sharing serve as fundamental components of health informatics' research, and are consistently applied to create solutions that support healthcare delivery and research (American Medical Informatics Association, 2015). Particularly in cancer, the timeliness of registry data is critical for the discovery and improvement in quality of care (Centers for Disease Control and Prevention, 1999; Simone & Hewitt, 2000).

The significance of timeliness in cancer registry data is documented in several IOM and CDC reports (Centers for Disease Control and Prevention, 1999; Simone & Hewitt, 2000). Few studies in Europe have evaluated the timeliness of cancer registry data. In the U.S., a study by Gagen and Cress (2005) investigated some of the factors associated with reporting delays by examining the association between reporting delays and gender, race, type of facility, cancer site, and/or stage at diagnosis (Smith-Gagen et al., 2005). However, the studies found evaluated timeliness as part of the DQ evaluation. To the best of our knowledge, no previous studies have specifically investigated the efficiency of the cancer reporting process or the steps involved. The method described in this paper provides a systematic evaluation of data timeliness and the reporting process that can identify challenges and therefore contributes to our understanding and improvement of cancer reporting systems. Some of the key findings are presented in the following sections.

The proposed method reveals variations among facilities in terms of both reporting time and the time completed cases reside at the facility database before they are reported to the state registry. This was found to be the case at 7 of the 22 facilities with higher caseloads. The 22 facilities classified as higher-caseload facilities account for 90.5% of the total number of cases. The 7 identified facilities were found to take twice the amount of time to submit the completed records, compared with the remaining 15 (around 100‑300 days). The potential impact for improving timeliness in cancer reporting can be demonstrated by understanding the timeliness requirements of cancer registry standards. As seen in the previous sections, variations in timeliness requirements among

cancer registry standards are small. For example, the difference in timeliness requirements between the gold NAACCR certification and the silver NAACCR certification is 5% (NAACCR gold certification requires 95% of cases reported within 23 months and silver certification requires 90% within 23 months). Moving from silver to gold certification, for instance, requires focusing on the 5% of cases that take more than 23 months. During the registrar interview, the process of submitting completed reports was often described as fast and easy. Therefore, encouraging more frequent report submission could be a practical step to improving timeliness without investing in additional technology or staff.

This study also shows that access to information can be the biggest barrier encountered by registrars. The interview results suggest that directing research and development into addressing these challenges could have a significant impact on the reporting process. Many states already have the HIE infrastructure that could be used to address this gap (Fontaine et al., 2010).

More importantly, the results show that time spent on cancer reporting comprises not only task time but also waiting time, which consumes most of the overall reporting time. Most of this waiting time occurs while patients await treatments and procedures. Our findings suggest that this time can be significantly reduced by changing the currently followed queing method. Generally, cases reside in the suspense file for a few months, during which time treatments and procedures are performed and are entered into the EHR system. Thus, the first case entered into the suspense file will be the first case abstracted.

However, procedures and treatments can be done at different speeds, depending on factors such as the cancer type, cancer stage, and facility resources. Using a standard waiting time for all cases creates an unnecessary delay if treatments are made available earlier than prior cases. Adding a notification system to inform registrars of when treatments are available will enable them to abstract the case as soon as the treatment is available, instead of using a fixed period of time for all cases.

5.7 Implications

This study has implications in the following areas: systems design and development, quality management, and decision-making. Examples of the study implication are listed below.

Phase 3 of the study has implications for systems engineering, design, and development. Informing systems engineers and designers of the barriers encountered by users and the time spent at each step can assist them in creating systems that meet information and work needs (McNulty & Ferlie, 2002). For example, in Phase 3, we found that casefinding from ICD is often time-consuming. Some systems enable registrars to match the ICD list with previously reported cases, a function that was found to be valuable for registrars. By understanding how the ICD list is interpreted by registrars, we could then recommend matching the ICD list with path reports and other sources used during the casefinding stage.

This study also informs system developers and hospital administrators about the privacy policy related to cancer reporting. Some examples are listed below.

System session timeout is set as a safeguard measure for data protection. In Phase 3, however, we see that session timeout is often perceived as a barrier by registrars. The time set for the session timeout is often estimated based on the average user's activities. It sets the session expiration period to automatically log users out when they are away from the computer. With cancer reporting, however, registrars are multi-tasking, logging into more than one system, or making phone calls in search of information. Understanding the needs of registrars and the nature of their workflow can assist system developers and administrators in setting the appropriate privacy and security guidelines.

Phase 3 also shows that many registrars experience some difficulty while accessing information within facilities outside the hospital network. When electronic access is not available, registrars use other methods—such as phone calls—to contact outside facilities. Some registrars have indicated that inquiring about patient information via the phone can be difficult when nurses and physicians are unwilling to share patient information due to their concerns about the privacy of patient information. Educating other health professionals, such as nurses and physicians, about the role of cancer registrars, can promote communication between them.

5.8 Limitations

One of the limitations of Phase 2 was the high rate of missing data in one of the time intervals. The report completion phase consisted of case finding and abstracting. Separating the casefinding from the abstracting time requires the availability of "abstract initiation" dates. However, due to the high rate of missing data in the "abstract initiation" date, making the distinction between casefinding and abstracting was not possible. Instead, we calculated the combined time for report completion time.

The second limitation is the time span from which the study data has been retrieved. This study retrieved cases diagnosed from 2000 to 2010. Considering the declining trend observed in reporting time and changes in heath technologies since 2010, our results may not reflect the current situation.

A third limitation was the absence of filed observations in Phase 3. Data was collected in this study through interviews, and one of the advantages of the interview is the ability to capture the entire cycle of the process. This approach was appropriate for this study, because the entire reporting cycle can take months or years. However, similar to other self-reported techniques, interviews are subjected to self-reporting bias.

A fourth limitation was estimating the simulation input for the redesigned workflow. To conduct the simulation for the redesigned workflow, we needed an estimation of the expected time cases reside in the suspense file. This is represented by

the time from casefinding to availability of treatment results. Registrars indicated that abstracting could commence at the beginning of the first course of treatment. To estimate the time, we used a national study that measures the time from diagnosis to treatment. As this could underestimate the simulation input by disregarding the time needed to write the treatment result and add it to the EHR, it may also overestimate it by disregarding the time from diagnosis to casefinding.

Moreover, we should also recognize the tradeoff between completeness, accuracy, and timeliness when interpreting the simulation results. As stated above, registrars can currently abstract by the beginning of the first course of treatment. As the results of our simulation were estimated based on this scenario, they would not reflect situations wherein registrars decided to wait to add further data. The need to wait for more data was either to provide more complete abstracts or to confirm case diagnoses. In situations where the available data was insufficient to confirm the diagnosis, registrars might want to wait for more information to reach a higher degree of certainty.

5.9 Future Work

The results of this study point to the value of examining the time taken from diagnosis to treatment to set realistic expectations in terms of timeliness. Although different cancer standards require hospitals to report their cancer cases within a specific time period after diagnosis, registrars need to wait for treatments and procedures to be performed and are entered into the EHR. Knowing the expected time for the required

information being ready in the EHR can help to determine the reasonable timeliness expectation for each cancer type.

Another area to investigate is the relative importance of different information with respect to time. This study provided an indication of the time needed for casefinding and abstracting and informed us of the time needed to collect demographic information, cancer type, and cancer stage compared with the time needed to collect treatment information. Given the long period of time between casefinding and abstracting, we have found that information collected during casefinding (demographic information, cancer type, and cancer stage) can potentially be available much sooner, if reported separately. A similar technique was applied by the CoC for breast and colorectal cancer (American College of Surgeons, 2015). Knowing what the data will be used for and what information will be more valuable, if received sooner, can determine if a two-stage reporting process is needed to accelerate the timelines of this information.

Future work may also include development and testing of the notification system. The role of the recommended notification system is to inform the abstracting registrars when new cancer treatment is available. A fundamental requirement for developing such a system is identifying the treatments and procedures to be included, as well as the corresponding specification for HL7 messages. The same will apply for enabling automatic public health reporting through the CDA. An established effort was made by the CDC to guide implementation of automatic reporting in EHRs and public health registries through CDA (HL7 International).

CHAPTER 6: APPENDICES.

6.1 Interview script

1. Demographic

•   Are you a CTR or a non-CTR?

    o   **Describe your role in reporting?**

2.Workload questions

- **Can you estimate the number of cases processed (daily or weekly)?**

- How many people in your workplace are involved in reporting?

    o   How many of them are CTR and how many are not?

    o   Do they perform the same task at the same time?

    o   **Do they share some of the resources, such as computers or phones? Is it enough?**

2. Process workflow questions

- **In order, what are the main steps for reporting?**

- **How does reporting start (what triggers the event to start)?**

    o   **Potential Probes; Are there other ways to initiate the process?**

This part can be repeated for the main steps: Case finding, abstracting, reporting, follow-up.

- **Describe the _____ process?**

    - **Potential Probes: What is the goal of this step?**

    - **Probes (casefinding): Are there cases/cancer types that take longer than others to find/identify? If yes, what are they? Why?**

    - **Probes (case finding): Do you work on an electronic or paper format?**

    - **Probes (abstracting): Are there certain types of information that take longer than other to abstract? If yes, how often do you have them?**

    o **Can you estimate the time it takes you to perform this step (daily or weekly) (range and average)?**

    o **At which point do you spend most of your time, and why?**

    o **Can you estimate the time you spend on this part?**

Repeat for abstracting, reporting, follow-up.

- **How do you know when a process is complete?**

    o **Are there any other possible outcomes? If yes, please describe?**

4.Efficiency Questions

- **How often you encounter a delay or wait in the reporting process? If often:**

Potential probes:

  - o Where in the reporting cycle does the delay exist?

  - o What could happen?

  - o Why would it happen?

  - o How bad is it?

  - o And when it happens, what you do about it? Is it effective?

- Are there any parts of the reporting process you wish you could eliminate or fix?

  If yes,

  - o Which part is it**?**

- Why? Is there a time when you have to start over or repeat work? If yes,

  - o **How often? And, if often,**

    - ▪ **When (at what part), and why?**

5. Other

- Regardless of the current constraints and limitations, what would the ideal reporting system look like?

- Any other comments you would like to add?

6.2 Reporting Frequency

Frequency of Report Submissions
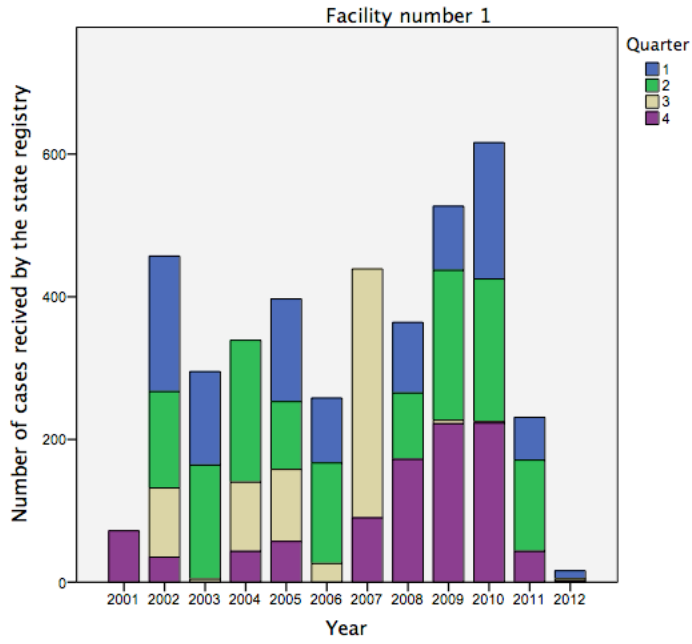
Figure A-1 (Facility #1): The frequency of report submissions.
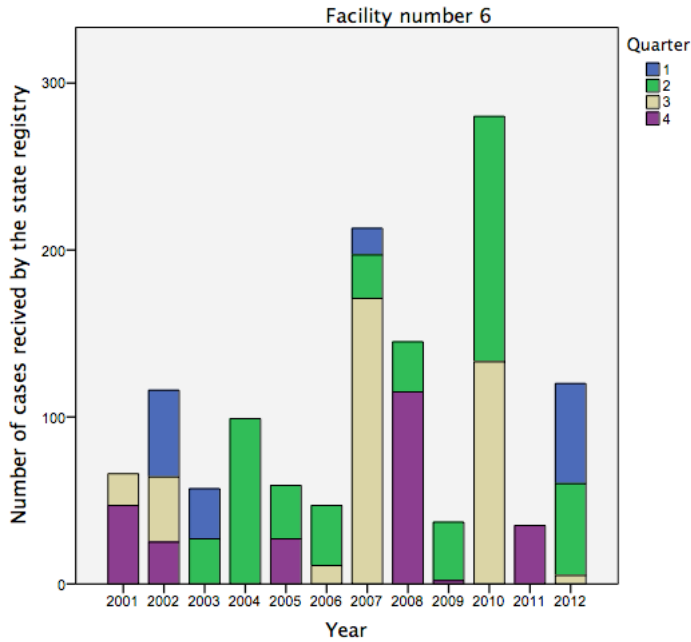
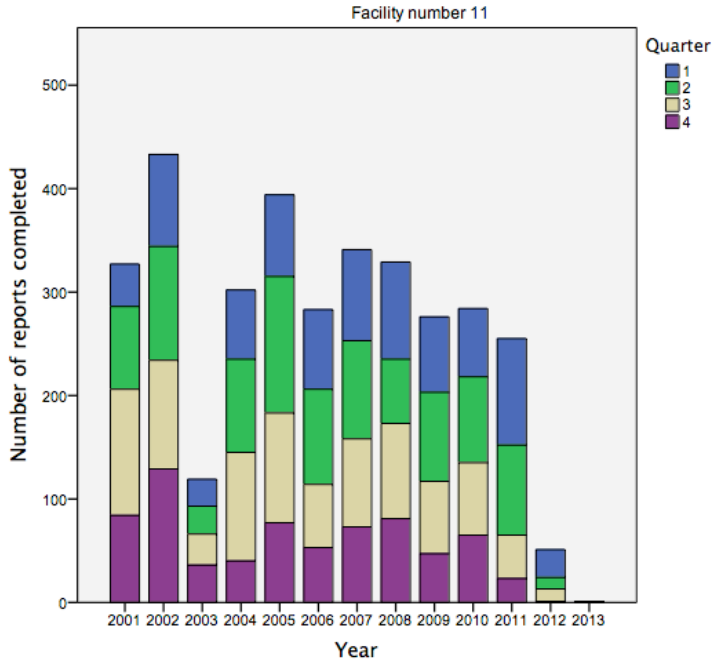Figure A-2 (Facility #6): The frequency of report submissions.
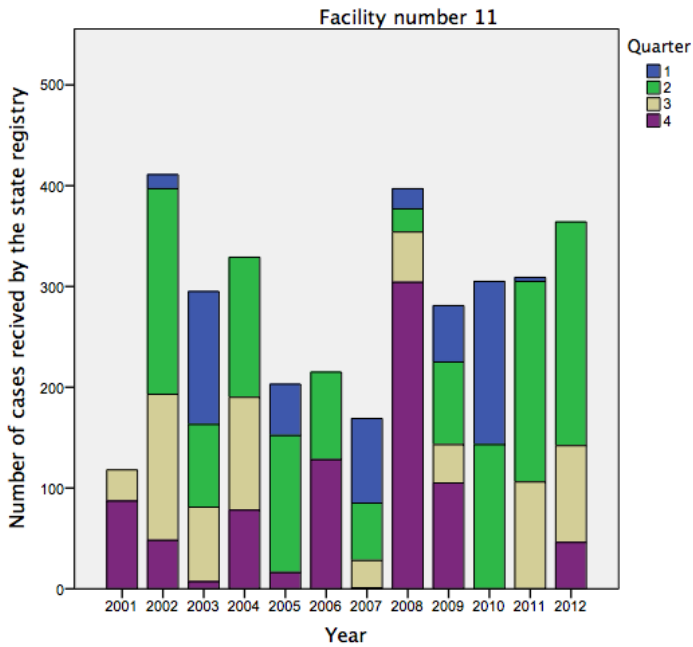
Figure A-3 (Facility #11): The frequency of report submissions.
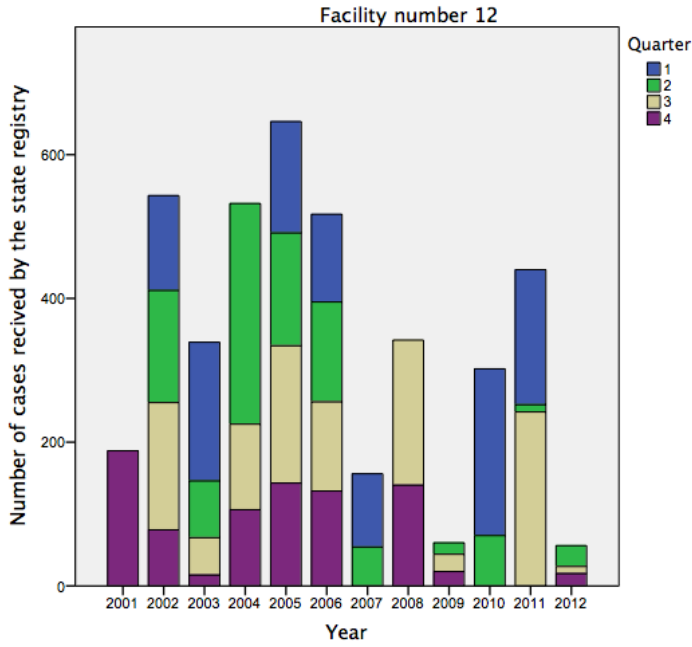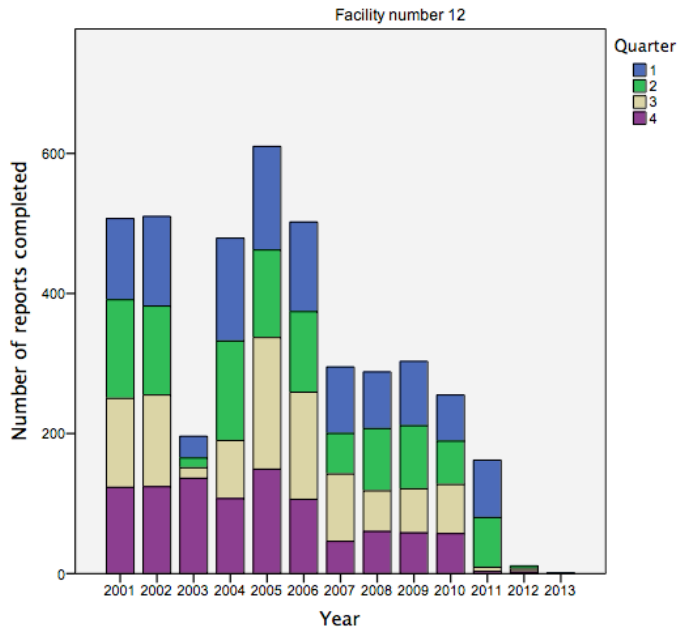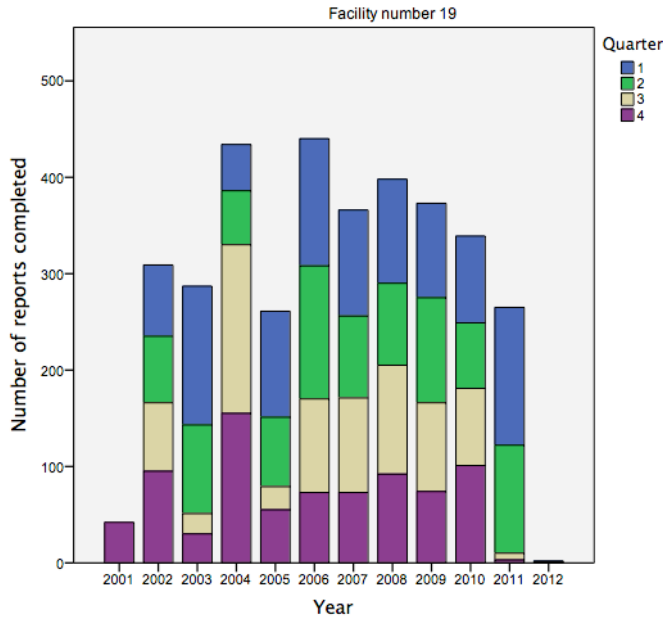
Figure A-4 (Facility #12): The frequency of report submissions.
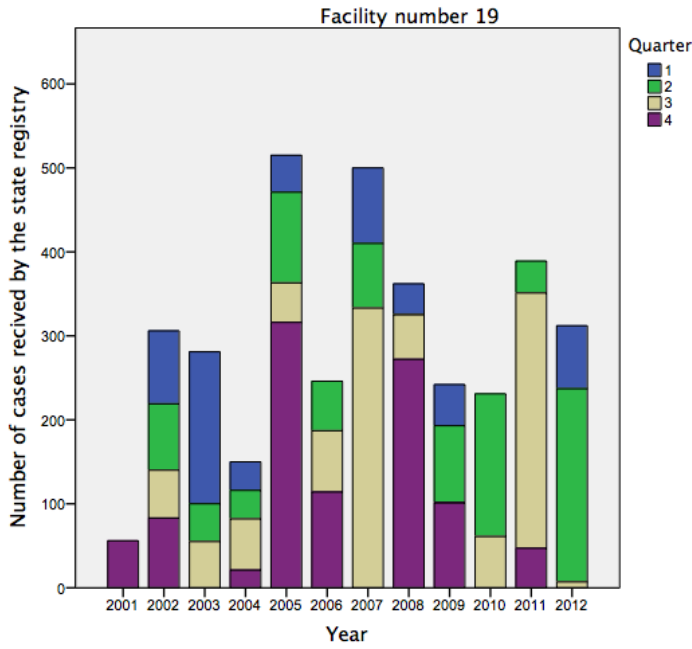
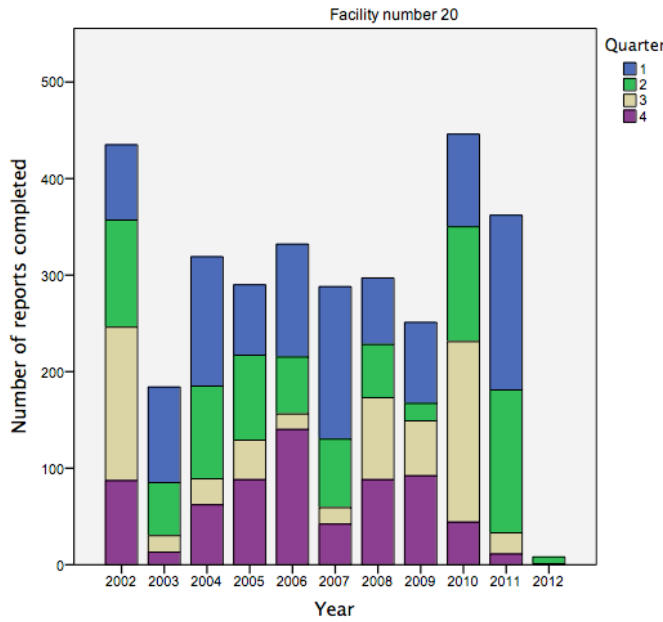Figure A-5 (Facility #19): The frequency of report submissions.
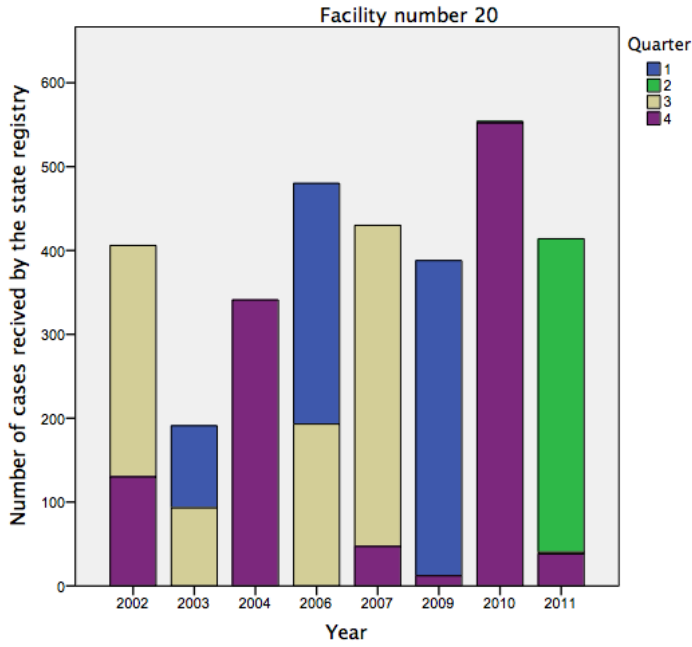
Figure A-6 (Facility #20): The frequency of report submissions.
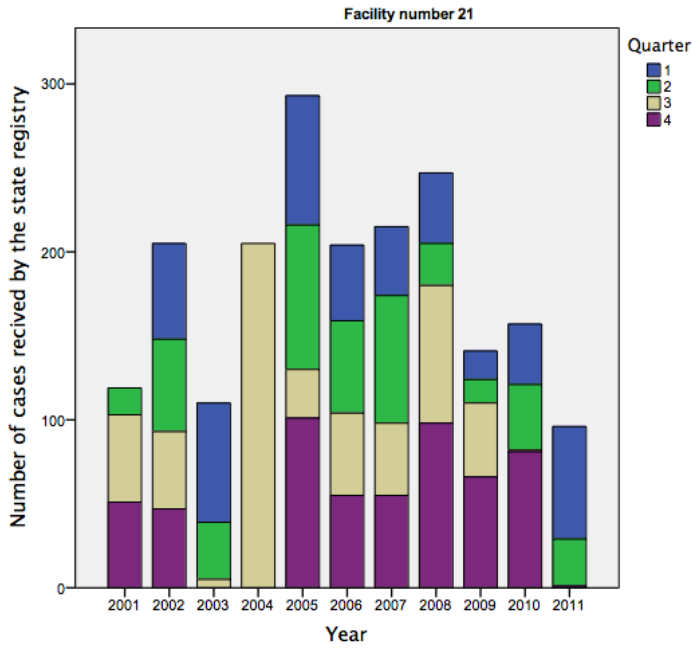
Figure A-7 (Facility #21): The frequency of report submissions.

Frequency of Reports Completions
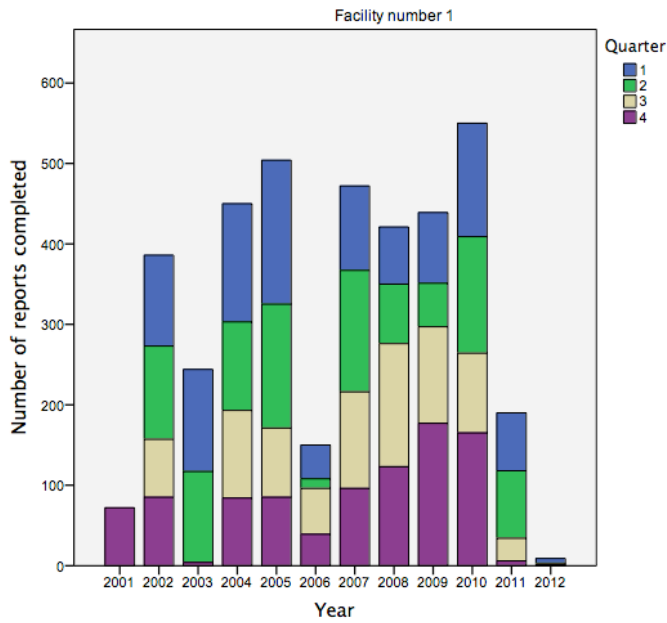
Figure B-1 (Facility #1): The frequency of report completions.
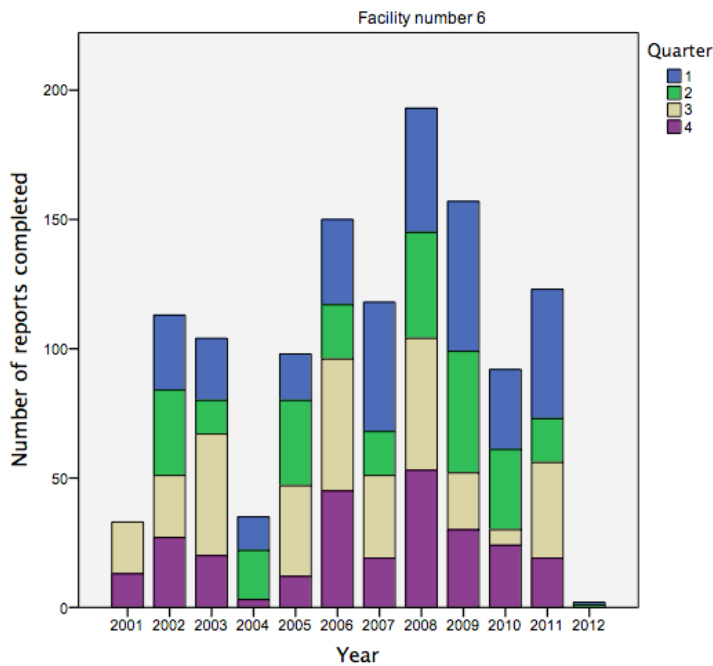
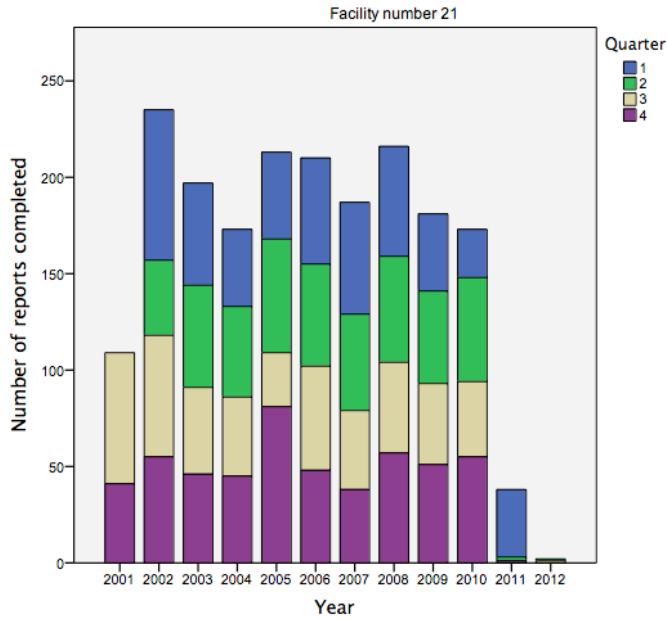Figure B-2 (Facility #6): The frequency of report completions.

Figure B-3 (Facility #11): The frequency of report completions.

Figure B-4 (Facility #12): The frequency of report completions.

Figure B-5 (Facility #19): The frequency of report completions.

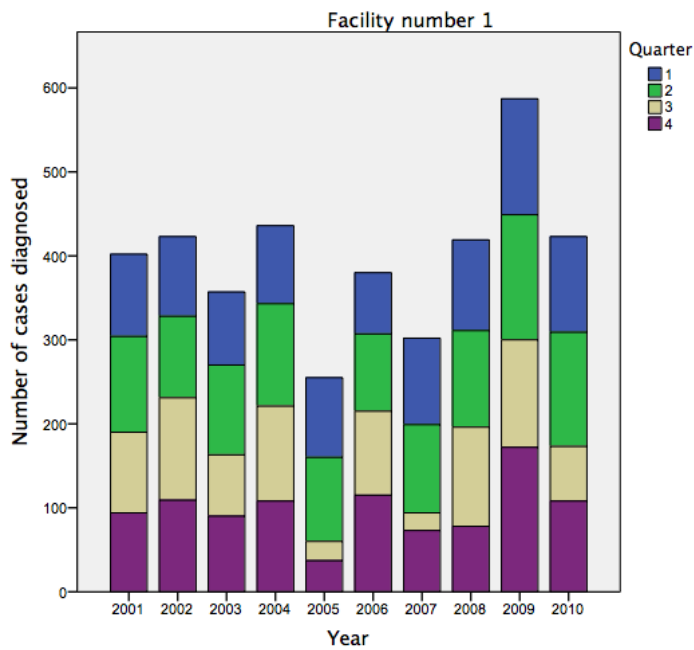Figure B-6 (Facility #20): The frequency of report completions.

136

Figure B-7 (Facility #21): The frequency of report completions.

Frequency of cases diagnosed

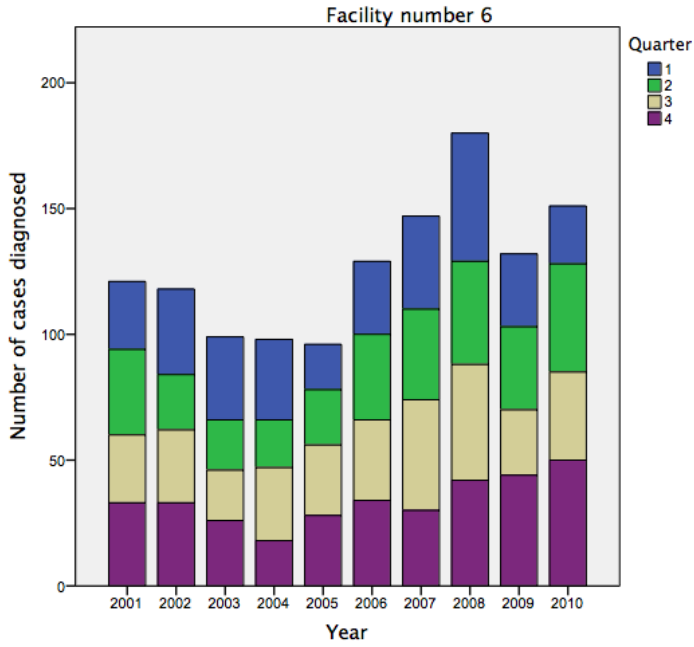Figure C-1 (Facility #1): The frequency of cases diagnosed.

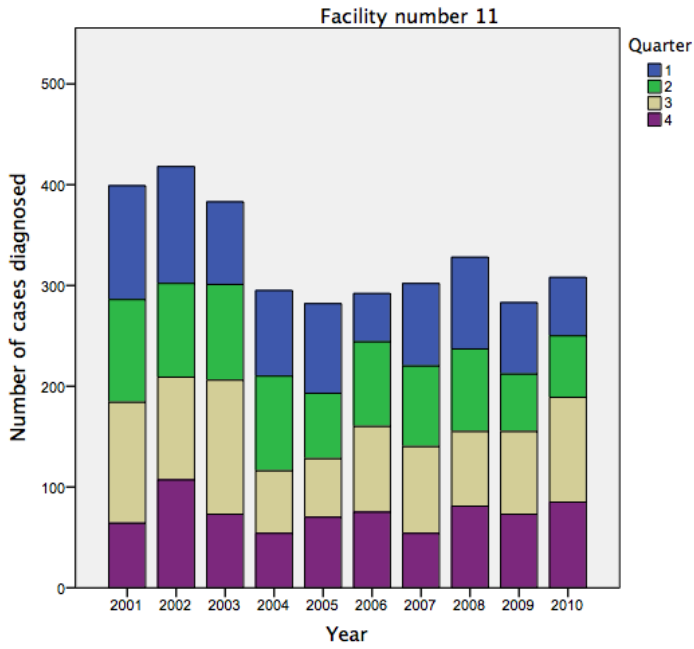Figure C-2 (Facility #6): The frequency of cases diagnosed.

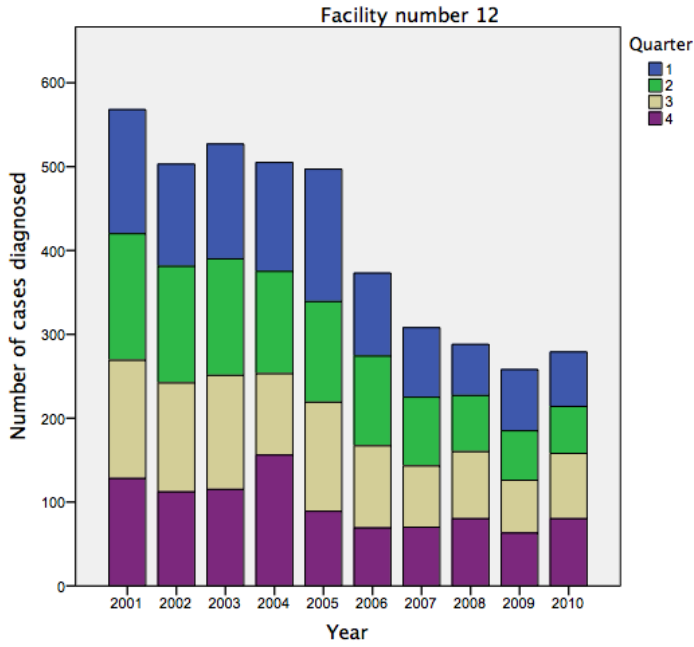Figure C-3 (Facility #11): The frequency of cases diagnosed.

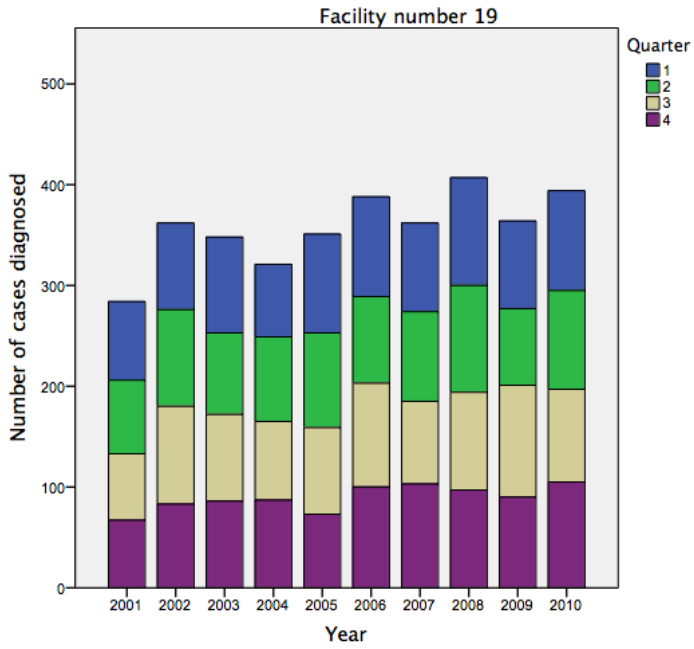Figure C-4 (Facility #12): The frequency of cases diagnosed.

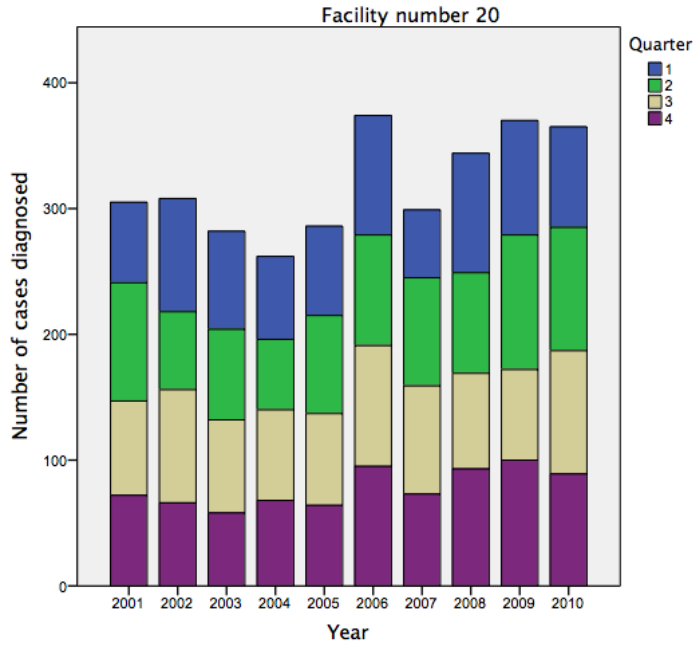Figure C-5 (Facility #19): The frequency of cases diagnosed.

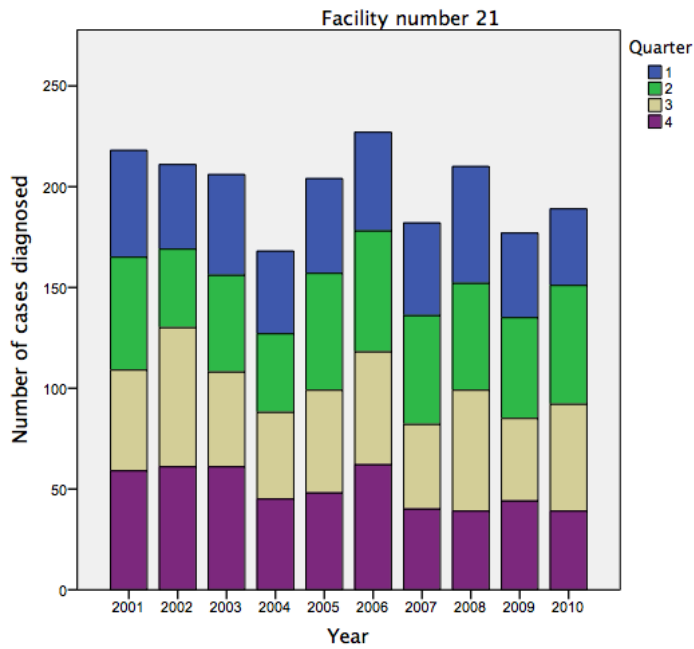Figure C-6 (Facility #20): The frequency of cases diagnosed.

Figure C-7 (Facility #21): The frequency of cases diagnosed.

6.3 Data sources used at the reporting process

Data sources for case finding:

- Pathology reports

- Indices code

- Admission and discharge reports

- Surgery schedules

- Cytology reports

- Radiation oncology logs

- Billing information

- X-ray reports

- Medical oncology

- Radiation oncology reports

- Chemotherapy

Data sources for abstracting:

- Pathology reports

- Radiation oncology

- Medical oncology

- Physician notes

- Oncology notes

- Chemotherapy reports

- Medical imaging

- Cytopathology reports

- Surgery and operative reports

- Treatments reports (such as hormones, biological response, or chemo).

- Discharge summaries

- Laboratory results

- Consultation reports

CHAPTER 7: REFERENCES

Abernethy, A.P., Etheredge, L.M., Ganz, P.A., Wallace, P., German, R.R., Neti, C. Murphy, S.B. (2010). Rapid-learning system for cancer care. Journal of Clinical Oncology, 28(27), 4268-4274.

ACS. (2014). Cancer Surveillance Programs in the United States. American Cancer Society Retrieved from http://www.cancer.org/cancer/cancerbasics/cancer-surveillance-programs-and-registries-in-the-united-states.

ACS. (2015). Registry Manuals. American College of Surgeons Retrieved from https://http://www.facs.org/qualityprograms/cancer/ncdb/registrymanuals/cocman uals.

Adler-Milstein, J., DesRoches, C.M., Kralovec, P., Foster, G., Worzala, C., Charles, D. Jha, A.K. (2015). Electronic health record adoption in US hospitals: Progress continues, but challenges persist. Health Affairs, 10.1377/hlthaff. 2015.0992.

Aggarwal, V., Backman, C., Sager, K., & Sannyasi, A. (2010). Workflow redesign in support of the use of information technology within healthcare.

AHRQ. (2015). Module 5. Mapping and Redesigning Workflow. 2015, from http://www.ahrq.gov/professionals/prevention-chronic-care/improve/system/pfhandbook/mod5.html

American College of Surgeons. (2015). Rapid Quality Reporting System (RQRS). from https://http://www.facs.org/quality-programs/cancer/ncdb/qualitytools/rqrs

American Medical Informatics Association. (2015). The Science of Informatics. from https://http://www.amia.org/about-amia/science-informatics

Arts, D.G.T., De Keizer, N.F., & Scheffer, G.-J. (2002). Defining and improving data quality in medical registries: A literature review, case study, and generic framework. Journal of the American Medical Informatics Association, 9(6), 600-611.

Baldwin, L.P., Eldabi, T., & Paul, R.J. (2004). Simulation in healthcare management: A soft approach (MAPIU). Simulation Modelling Practice and Theory, 12(7), 541-557.

Banks, J. (1998). Handbook of simulation: Wiley Online Library.

Bertolini, M., Bevilacqua, M., Ciarapica, F.E., & Giacchetta, G. (2011). Business process re-engineering in healthcare management: A case study. Business Process Management Journal, 17(1), 42-66.

Bilimoria, K.Y., Ko, C.Y., Tomlinson, J.S., Stewart, A.K., Talamonti, M.S., Hynes, D.L. Bentrem, D.J. (2011). Wait times for cancer surgery in the United States: Trends and predictors of delays. Annals of Surgery, 253(4), 779-785.

Borycki, E.M., Kushniruk, A.W., Kuwata, S., & Kannry, J. (2006). Use of Simulation Approaches in the Study of Clinician Workflow. Paper presented at the AMIA Annual Symposium Proceedings.

Boscoe, F.P. (2014). Trends in reporting delay in United States central cancer registries. Journal of Registry Management, 41(3), 135.

Brailsford, S.C., & Hilton, N.A. (2001). A comparison of discrete event simulation and system dynamics for modelling health care systems.

Bray, F., & Parkin, D.M. (2009). Evaluation of data quality in the cancer registry: Principles and methods. Part I: comparability, validity and timeliness. European Journal of Cancer, 45(5), 747-755.

CDC. (2013). National Program of Cancer Registries: Program standards. 2012-2017.

CDC. (2014). Implementation guide for ambulatory healthcare provider reporting to central cancer registries. HL7 Clinical Document Architecture (CDA) National Center for Chronic Disease Prevention and Health Promotion, Division of Cancer Prevention and Control.

Centers for Disease Control and Prevention. (1999). National program of cancer registries cancer surveillance system rationale and approach. Atlanta: CDC.

Champy, J., & Greenspun, H. (2014). Reengineering health care: A manifesto for radically rethinking health care delivery: Pearson Education.

Chou, Y.-C., Chen, B.-Y., Tang, Y.-Y., Qiu, Z.-J., Wu, M.-F., Wang, S.-C., Chuang, W.-C. (2012). Prescription-filling process engineering of an outpatient pharmacy. Journal of Medical Systems, 36(2), 893-902.

Clegg, L.X., Feuer, E.J., Midthune, D.N., Fay, M.P., & Hankey, B.F. (2002). Impact of reporting delay and reporting error on cancer incidence rates and trends. Journal of the National Cancer Institute, 94(20), 1537-1545.

CMS. (2015). The Official Web Site for the Medicare and Medicaid Electronic Health Records (EHR) Incentive Programs from http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html?redirect=/EHRIncentivePrograms.

Curry, J., & Prodan, A. (2013). Using Visual Tools to Improve Clinical Engagement and Group Understanding of Complex IT Concepts. Paper presented at the MedInfo.

D'Amore, J.D., Sittig, D.F., & Ness, R.B. (2012). How the continuity of care document can advance medical research and public health. American Journal of Public Health, 102(5), e1-e4.

Dale, D., Golabek, J.K., & Chong, N. (2002). The impact of E-path technology on Ontario Cancer Registry operations. Journal of Registry Management, 29(2), 52-56.

Davenport, T.H. (2013). Process innovation: Reengineering work through information technology: Harvard Business Press.

Dixon, B.E., Pina, J., Kharrazi, H., Gharghabi, F., & Richards, J. (2015). What's past is prologue: A scoping review of recent public health and global health informatics literature. Online Journal of Public Health Informatics, 7(2).

Dixon, B.E., Whipple, E.C., Lajiness, J.M., & Murray, M.D. (2015). Three decades of utilising an integrated information infrastructure for outcomes research: A systematic review of studies conducted at the Regenstrief Institute. Health Information & Libraries Journal.

Finnell, J.T., & Dixon, B.E. (2015). Clinical Informatics Study Guide: Text and Review. Public Health Informatics (1 ed., pp. p. 501-520). Zurich: Springer International Publishing.

Fleurant, M., Kell, R., Jenter, C., Volk, L.A., Zhang, F., Bates, D.W., & Simon, S.R. (2012). Factors associated with difficult electronic health record implementation in office practice. Journal of the American Medical Informatics Association, 19(4), 541-544.

144

Fontaine, P., Ross, S.E., Zink, T., & Schilling, L.M. (2010). Systematic review of health information exchange in primary care practices. The Journal of the American Board of Family Medicine, 23(5), 655-670.

Goroll, A.H., Simon, S.R., Tripathi, M., Ascenzo, C., & Bates, D.W. (2009). Community-wide implementation of health information technology: the Massachusetts eHealth Collaborative experience. Journal of the American Medical Informatics Association, 16(1), 132-139.

Günal, M.M., & Pidd, M. (2010). Discrete event simulation for performance modelling in health care: A review of the literature. Journal of Simulation, 4(1), 42-51.

Hammer, M. (1990). Reengineering work: Don't automate, obliterate. Harvard Business Review, 68(4), 104-112.

Hanauer, D.A., Miela, G., Chinnaiyan, A.M., Chang, A.E., & Blayney, D.W. (2007). The registry case finding engine: an automated tool to identify cancer cases from unstructured, free-text pathology reports and clinical notes. Journal of the American College of Surgeons, 205(5), 690-697.

Health IT. (2013). Implementing Consolidated-Clinical Document Architecture (C-CDA) for Meaningful Use Stage 2.

Health IT Dashboard. (2014). Hospital Selection of Meaningful Use Public Health Measures; Eligible Hospitals Reporting to the Medicare EHR Incentive Program for Fiscal Year 2014: U.S. Department of Health and Human Services.

HealthIT. (2015). What is workflow redesign? Why is it important? from http://www.healthit.gov/providers-professionals/faqs/ehr-workflow-redesign

HHS. (2015). HITECH Act Enforcement Interim Final Rule. Health Information Privacy. from http://www.hhs.gov/ocr/privacy/hipaa/administrative/enforcementrule/hitechenforcementifr.html

HIMSS. (2015). Process Management, Workflow & Mapping. 2015, from http://www.himss.org/resourcelibrary/TopicList.aspx?MetaDataID=739

HL7 International. Project Summary for CDA Implementation Guide for Ambulatory Healthcare Provider Reporting to Central Cancer Registries. Retrieved 08/22/2015, 2015, from: http://www.hl7.org/special/committees/projman/searchableprojectindex.cfm?action=edit&ProjectNumber=1069.

INDH. (2015). Policy and Procedure Manuals for Reporting Facilities. Indiana State Department of health Cancer Registry. Retrieved from http://www.in.gov/isdh/files/ISCR_Policy_Manual.pdf , p18.

Indiana Health Information Exchange. Indiana Network For Patient Care (INPC), Retrieved 12/14/2015, from http://www.ihie.org/indiana-network-for-patient-care

ISDH. (2004). Indiana State Cancer Registry News Briefs. Indiana State Department of Health.

ISDH. (2012). Indiana Cancer Facts & Figures 2012: Indiana State Department of Health.

Izquierdo, J.N., & Schoenbach, V.J. (2000). The potential and limitations of data from population-based state cancer registries. American Journal of Public Health, 90(5), 695.

Jajosky, R.A., & Groseclose, S.L. (2004). Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. BMC Public Health, 4(1), 29.

Joos, D., Chen, Q., Jirjis, J., & Johnson, K.B. (2006). An electronic medical record in primary care: Impact on satisfaction, work efficiency and clinic processes. Paper presented at the AMIA Annual Symposium Proceedings.

Kasthurirathne, S.N., Dixon, B.E., Gichoya, J., Xu, H., Xia, Y., Mamlin, B., & Grannis, S.J. (2016). Toward better public health reporting using existing off the shelf approaches: A comparison of alternative cancer detection approaches using plaintext medical data and non-dictionary based feature selection. Journal of Biomedical Informatics.

Kavuluru, R., Hands, I., Durbin, E., & Witt, L. (2013). Automatic extraction of ICD-O-3 primary sites from cancer pathology reports. Paper presented at the Clinical Research Informatics AMIA symposium (forthcoming).

Larsen, I.K., Småstuen, M., Johannesen, T.B., Langmark, F., Parkin, D.M., Bray, F., & Møller, B. (2009). Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. European Journal of Cancer, 45(7), 1218-1231.

Leu, J.-D., & Huang, Y.-T. (2011). An application of business process method to the clinical efficiency of hospital. Journal of Medical Systems, 35(3), 409-421.

Mac McCullough, J., & Goodin, K. (2014). Patterns and correlates of public health informatics capacity among local health separtments. Online Journal of Public Health Informatics, 6(3).

McFarlane., T.D., Dixon., B.E., & Grannis., S.J. (2016). Client registries: Identifying and linking patients. Health Information Exchange: Navigating and managing a network of health information systems (pp. 163-182): Elsevier Science.

McNulty, T., & Ferlie, E. (2002). Reengineering health care: The complexities of organizational transformation: The complexities of organizational transformation: Oxford University Press.

Meyer, A.M., Carpenter, W.R., Abernethy, A.P., Stürmer, T., & Kosorok, M. R. (2012). Data for cancer comparative effectiveness research. Cancer, 118(21), 5186-5197.

Midthune, D.N., Fay, M.P., Clegg, L.X., & Feuer, E.J. (2005). Modeling reporting delays and reporting corrections in cancer registry data. Journal of the American Statistical Association, 100(469), 61-70.

Moore, T., Shapiro, J.S., Doles, L., Calman, N., Camhi, E., Check, T., . . . Kuperman, G. (2012). Event detection: A clinical notification service on a health information exchange platform. Paper presented at the AMIA Annual Symposium Proceedings.

NAACCR. (2011). Electronic Pathology Reporting Guidelines. The North American Association of Central Cancer Registries.

NAACCR. (2014). Certification Levels. 2014 December from http://www.naaccr.org/certification/certificationlevels.aspx

NAACCR. (2015). certification level. from; http://www.naaccr.org/Certification/CertificationLevels.aspx

NAACCR. (2015). Who is Certified. from; http://www.naaccr.org/Certification/WhoisCertified.aspx

National Cancer Registrars, A. (2004). Cancer Registry Management: Principles & Practice: Kendall/Hunt Publishing Company.

National Learning Consortium. (2012). Workflow process mapping for Electronic Health Record (EHR) implementation.

NCI. (2014). Cancer Incidence Rates Adjusted for Reporting Delay. 03/01/2015, from http://surveillance.cancer.gov/delay/

North American Association of Central Cancer Registries (NAACCR). (2012). Standards for Cancer Registries Volume II, Data Standards and Data Dictionary. North American Association of Central Cancer Registries.

Savel, T.G., & Foldy, S. (2012). The role of public health informatics in enhancing public health surveillance. MMWR Surveill Summ, 61(Suppl), 20-24.

SEER. (2015). Comprehensive ICD-9-CM Case Finding Code List for Reportable Tumors. Casefinding Lists. 2015 from http://seer.cancer.gov/tools/casefinding/

Sheikh, A., Sood, H.S., & Bates, D.W. (2015). Leveraging health information technology to achieve the "Triple Aim" of healthcare reform. Journal of the American Medical Informatics Association, ocv022.

Sigurdardottir, L.G., Jonasson, J.G., Stefansdottir, S., Jonsdottir, A., Olafsdottir, G.H., Olafsdottir, E.J., & Tryggvadottir, L. (2012). Data quality at the Icelandic Cancer Registry: Comparability, validity, timeliness and completeness. Acta Oncologica, 51(7), 880-889.

Simone, J.V., & Hewitt, M. (2000). Enhancing data systems to improve the quality of cancer care: National Academies Press.

Smith-Gagen, J., Cress, R.D., Drake, C.M., Felter, M.C., & Beaumont, J.J. (2005). Factors associated with time to availability for cases reported to population-based cancer registries. Cancer Causes & Control, 16(4), 449-454.

Sobolev, B.G., Sanchez, V., & Vasilakis, C. (2011). Systematic review of the use of computer simulation modeling of patient flow in surgical care. Journal of Medical Systems, 35(1), 1-16.

Stankard, M.F. (2002). Management systems and organizational performance: The search for excellence beyond ISO9000: Greenwood Publishing Group.

Surgeons, A. C. o. (Producer). (2014). 5.2 Cancer Registrar Credentials. Cancer Program Standards 2012. Retrieved from https://http://www.facs.org/quality-programs/cancer/coc/standards/video/chap51/chap52

Tangka, F., Subramanian, S., Beebe, M.E.C., Trebino, D., & Michaud, F. (2010). Economic assessment of central cancer registry operations, Part III. Journal of Registry Management, 37(4), 152-155.

Tomic, K., Sandin, F., Wigertz, A., Robinson, D., Lambe, M., & Stattin, P. (2015). Evaluation of data quality in the National Prostate Cancer Register of Sweden. European Journal of Cancer, 51(1), 101-111.

Unertl, K.M., Novak, L.L., Johnson, K.B., & Lorenzi, N.M. (2010). Traversing the many paths of workflow research: Developing a conceptual framework of workflow terminology through a systematic literature review. Journal of the American Medical Informatics Association, 17(3), 265-273.

Wang, R.Y., & Strong, D.M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 5-33.

Weiskopf, N.G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. Journal of the American Medical Informatics Association, 20(1), 144-151.

Wong, C., Geiger, G., Derman, Y.D., Busby, C.R., & Carter, M.W. (2003). Healthcare process analysis: Redesigning the medication ordering, dispensing, and administration process in an acute care academic health sciences centre. Paper presented at the Proceedings of the 35th Conference on Winter Simulation: Driving Innovation.

CURRICULUM VITAE

Abdulrahman M Jabour

EDUCATION

Doctor of Philosophy in Health and Biomedical Informatics

School of Informatics and Computing.

Indiana University, Indianapolis, Indiana, United State.

Dissertation title: Cancer reporting: timeliness analysis and process reengineering.

Master of Health Service Management (with honors)

School of Medicine.

Griffith University, Brisbane, Australia.

Dissertation title: *Organizational culture and Electronic Health Record among GPs.*

Bachelor of Science (medical imaging)

School of Medical Imaging.

Curtin University of technology. Perth, Australia.

Graduation project: PACS has improved workflow efficiency of imaging departments in terms of decreasing turnaround time (TAT) by decreasing reporting time and decreasing technologist examination time.

Diploma in Diagnostic Radiological Imaging.

Riyadh and Alkharj Armed forces hospital, programs & training center , Riyadh, Saudi Arabia.

ACADEMIC EXPERIENCE

- Part time Research Assistant at Indiana University (2011 -2012).

- Lecturer; one year full time faculty member in Applied Medical Science Faculty, Jazan University, Jazan, Saudi Arabia (December 2009 to December 2010).

CLINICAL EXPERIENCE

- Radiographer; 2 years in CT Scan radiology in RKH hospital, Riyadh, Saudi Arabia From 2004 to 2006

- Radiographer; 3 years of general radiography in RKH hospital, Riyadh, Saudi Arabia From 2001 to 2004.

OTHER SERVICES

- Contribution to the establishment and development of Health Information Management HIM bachelor degree program at Jazan University, Jazan , Saudi Arabia.

- Participated with a group of professionals to add the Arabic translation to OpenEMR version 4.1.2.

PEER-REVIEWING ACTIVITIES

- American Medical Informatics Association Symposium, 2014.
- Computers in Biology and Medicine Journals 2015.

PUBLICATIONS

Jabour, A.M., & Jones, J. F. (2013). Facilitators and Barriers to Patients' Engagements with Personal Health Records: Systematic Review. In Universal Access in Human-Computer Interaction. Applications and Services for Quality of Life (pp. 472-481). Springer Berlin Heidelberg (Conference proceeding)

Dixon, B. E., **Jabour, A. M**., Phillips, E. O. K., & Marrero, D. G. (2013). An informatics approach to medication adherence assessment and improvement using clinical, billing, and patient-entered data. Journal of the American Medical Informatics Association, amiajnl-2013.

Dixon, B. E., **Jabour, A. M**., Phillips, E. O. K., & Marrero, D. G. (2014). Challenges When Engaging Diabetic Patients and Their Clinicians in Using E-Health Technologies

to Improve Clinical Outcomes. In Universal Access in Human-Computer Interaction. Aging and Assistive Environments (pp. 238-247). Springer International Publishing (Conference proceeding).

POSTERS

Jabour, A. M, Dixon, B. E. Jones, J. F, Haggstrom, D (2015) Examining Cancer Case Reporting Processes and Timeliness: Preliminary Results. AMIA 2015 Annual Symposium.

Jabour, A. M, Dixon, B. E, Phillips, E. O., & Marrero, D. G (2013). The Development and Implementation of a Novel Clinical Dashboard for Improving Medication Adherence. AMIA 2013 Annual Symposium. (poster)

Dixon, B. E., **Jabour, A. M**., Phillips, E. O., & Marrero, D. G. (2013). Improving Medication Adherence for Chronic Disease Using Integrated e-Technologies. Studies in health technology and informatics, 192, 929-929. http://www.ncbi.nlm.nih.gov/pubmed/23920703