Fall 12-18-2015

# Process Mining of Medication Revisions in Electronic Health Records

Deevakar Rogith
*University of Texas Health Science Center at Houston, School of Biomedicao Informatics*,
deevakar.rogith@uth.tmc.edu

# Process Mining of Medication Revisions in Electronic Health Records

A
Dissertation

Presented to the Faculty of
The University of Texas
Health Science Center at Houston (UTHealth)
School of Biomedical Informatics
in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

By

Deevakar Rogith, M.B.B.S.

The University of Texas Health Science Center at Houston (UTHealth)

2015

Dissertation Committee:

Jiajie Zhang, PhD[1], Advisor
Amy Franklin, PhD[1]
Elmer V. Bernstam, MD, MSE[1]
M. Sriram Iyengar, PhD[2]
Muhammad F. Walji, PhD[3]

[1] The University of Texas Health Science Center at Houston (UTHealth) School of Biomedical Informatics, Houston, TX
[2]Texas A&M Health Science Center, Houston, TX
[3] The University of Texas Health Science Center at Houston (UTHealth) School of Dentistry, Houston, TX

**Acknowledgements**

I am deeply indebted to my mentors, collaborators, friends and family who provided me the encouragement and support critical to completing this research.

I am grateful to my primary mentor Jiajie Zhang for his tremendous support and guidance through my doctoral research. I would also like to thank my committee members Amy Franklin, Elmer Bernstam, Sriram Iyengar and Muhammad Walji for their guidance and assistance. Special thanks to Amy Franklin for all the help, support and encouragement. I am fortunate to have a committee of this calibre.

I am also grateful to all the faculty, staff and students at the UTHealth School of Biomedical Informatics. In particular, I would like to thank Craig Johnson, Trevor Cohen, and Hua Xu for their expert guidance on various aspects of the research design.

I am also grateful to the UTH BIG team for maintaining the research data warehouse used in this research. They have been very helpful and patient in solving technical difficulties I faced in working with the dataset. Special thanks to Susan Guerrero and Charles Bearden for their help.

Finally, I am grateful to my family and friends for their love that helped me through the rigorous phase of completing the dissertation.

## Abstract

**Objective:** The objective of this work is to develop process mining techniques for analysing Electronic Health Record (EHR) events in order to uncover factors contributing to the event, and understanding deviations in the process. We have outlined a method for combining data mining with expert review to model the EHR process and develop automated algorithms that can be used to detect potential deviations for a defined process. **Introduction:** To analyse EHR events meaningfully, process mining can be applied to distil structured process description from a set of real executions. Process mining can be applied for 1) Discovery, 2) Conformance, and 3) Enhancement of processes. This can be used for improving efficiency and safety in the process. Extending process mining to EHR system use, user activity can be analysed to model EHR use behaviour and detect deviations from expected use. Understanding these behaviours could be used to optimize systems through redesign. Here, we explore the application of process mining of medication revisions in an EHR. **Methods:** We first apply exploratory data analysis (EDA) of medication revisions (i.e. instances of altering a previous medication order) in EHR data to understand the occurrence of revision in the data. Data was retrieved from 6 U.S. ambulatory clinics, and 35,833 medication revision events were analysed. To add domain knowledge to the EDA, physicians manually reviewed a subset of events (n=100) to identify probable cause for these revisions. From the resulting causes, a categorization scheme was developed and fault trees were constructed to model the medication revision process. Additionally, from access pattern of EHR elements used in the expert review, an algorithm for automated detection of revisions was developed. Sensitivity and specificity

were calculated for the algorithm used to categorize an order as a revision event. **Results:** Revisions were classified into 5 categories - Cancel, Discontinue, Duplicate, Update, and Wrong Medication. 55% of the revisions were used as system workarounds to discontinue/update medications. The process model indicated that system issues were most prevalent, including problems in data entry and item selection. An automated algorithm was developed to categorize a medication order as a revision event. Given prevalence of 1.1%, the algorithm performed with 66% sensitivity, 85% accuracy and PPV of 4.8%. **Discussion:** EHR medication events were process mined by applying both data mining and domain knowledge. For the majority of cases, medication revisions are used as system workarounds. The fault tree analysis also suggests a common cause of these alterations is system issues. Although our automated methods showed lower sensitivity because of these workarounds, they were able to classify successfully those medication revisions that did reflect errors such as Wrong Medication or Duplicate Medications. **Conclusion:** Process mining was applied to medication revisions and was shown to detect revision events in the data. The process model uncovers factors responsible for revision events and can be used for improving EHR use. The detection algorithm can be useful in real-time monitoring. The vision is to develop monitoring tool for EHR similar to flight recorders and antivirus software.

## Vita

2010.....M.B.B.S., Medicine

Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER)
Puducherry, India

2010-2011 .....Residency (Non-Academic), Psychiatry

Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER)Puducherry, India

2012-2015 .....Graduate Research Assistant

The University of Texas Health Science Center at Houston (UTHealth) School of Biomedical Informatics

Houston, TX

## Field of Study

Health Informatics

**Table of Contents**

## List of Tables

## List of Figures

# Chapter 1: Introduction

In any system with human involvement, there is a risk of harm. Whether introduced by the system or the user, harm can occur regardless of the care given in design and use (Leape, 1994; Reason, 2000). Harm can increase with the increased complexity of an environment, as is the case with Electronic Health Records (EHR). EHR actions are influenced by many socio-technical factors (Campbell, Sittig, Ash, Guappone, & Dykstra, 2006), that make modeling of all potential interactions difficult. This in turn makes it difficult to evaluate the effect of individual actions on overall patient outcomes. While EHR actions have been shown to be associated with patient harm (Koppel et al., 2005; Nanji et al., 2011; Reckmann, Westbrook, Koh, Lo, & Day, 2009), in the majority of cases, it has been difficult to quantify or precisely locate the problem or root cause within the system. There are also often challenges to studying such activities in live clinical environments (Edwards, Moloney, Jacko, & Sainfort, 2008).

EHRs are being increasingly used in the clinical workflow. In 2014, over 400,000 providers and 4,500 hospitals used EHRs (The Office of the National Coordinator for Health Information Technology (ONC), 2014b). As an estimate of use, approximately 10 million electronic prescriptions were made every month (The Office of the National Coordinator for Health Information Technology (ONC), 2014a). With this sheer volume, EHRs have moved into the domain of big data. The big data consists of not only clinical information, but also EHR activity logs that give hints of the underlying EHR activity. EHR big data offers the opportunity to study EHR actions and their relation to the entire clinical scenario. The full complexity of EHR interactions and their effect on patient outcome may now be examined.

Big data provides opportunity to utilize the data to uncover many models and designs. To meaningfully analyze big data events, process mining can be used (W. Van Der Aalst, 2011). Process mining can be applied to EHR actions, to uncover factors contributing to the action and model the process. Such knowledge would help in strategies to improve system use, detect deviations in system use and identify potential risks from system use.

In this dissertation, we demonstrate the use of process mining of EHR data to understand factors contributing to an EHR action – in particular, to medication revisions. Based on the description of the medication revision event provided by manual review, we developed an algorithm to predict such events. The prediction of user activity from the example of revisions may be applied to other EHR events that have impact on patient safety.

**Process Mining:**

Process mining helps in understanding the events in big data by uncovering knowledge about use. Process mining uncover workflows. It can be used to complete conformance testing, and can aid in detecting deviations from expected behaviors. All of which can be used to enhance or optimize the performance of a given process (W. Van Der Aalst, 2011). Process mining provides a promise of improving systems by learning from previous use. For example, online vendors have analyzed their customer activity logs to understand why purchases are not completed that is why shopping carts are abandoned mid-purchase (Kohavi, Rothleder, & Simoudis, 2002; Montgomery, Li, Srinivasan, & Liechty, 2004). With this knowledge, the vendors have modified their interface to support shopping to completion. Similarly, a software designer might use services like Google Analytics to monitor user activity within the app and uncover workflow and user interests (Hasan, Morris, & Probets, 2009). Process mining in health care includes assessing

business processes of patient registration (Guo, Wagner, & West, 2004; Jun, Jacobson, & Swisher, 1999), tracking laboratory samples (Tarkan, Plaisant, Shneiderman, & Hettinger, 2011), evaluating care protocols (Poelmans et al., 2010), and evaluating emergency wait times (Kolker, 2008).

As with other examples of process mining for system improvement, here we consider EHR activity to improve the use of this system as part of healthcare processes. Our focus is on recognized EHR activity deviations and events that may have an impact on patient safety. We propose that the knowledge derived from process mining of EHR activities can be used to develop a detection algorithm for detecting potentially risky events. Here, we explore instances of medication revisions as an opportunity to understand the factors that contribute to a physician recognizing and revising a medication order. We use this data for developing a detector of similar events.

**Research Strategy:**

In this study, we applied process mining on medication revision events--medication orders labeled as 'Entered in Error' (EIE)--in an ambulatory clinic database to understand the human and system factors contributing to the event. Medication revisions represent conscious decisions by providers to make a change in a previous order. This change may be driven by patient preferences, alterations to treatment decisions, as a response to an adverse event, or result from system problems, including workarounds. We first began with exploratory data analysis of raw data from the EHR, exploring the circumstances in which medication revisions occur. To add to the knowledge about the processes, we then manually reviewed a subset of cases to uncover reasons for medication revisions and the factors responsible for the events. Next, we applied fault

tree analysis (FTA) to understand the relationship between the factors contributing to medication revisions. Learning from the review process, we developed a rule-based algorithm to predict the medication revisions events. We show that by applying process mining on human recognized medication revisions, an automated method can be developed to detect the medication revisions. From this we demonstrate a proof of concept of using process mining to understand EHR events and how it can be used for predicting user activity.

**Contributions and Innovation:**

The practical contributions of this work are the application of process mining to EHR to understand the human and system factors leading to medication revisions. As we identify problems in the EHR system, including issues with use and workflow, we may ultimately improve the efficiency and safety of care processes. Using the process mining results, we developed a rule-based algorithm to predict when a medication order might be revised. Application of this algorithm can be used to monitor user interaction in EHR and warn of potential deviations. We demonstrate a novel use of fault trees for EHR activity and provide new descriptions of medication revisions. We discuss how EHR data can be used as an audit trail of user activity, its limitation and recommendations for improving audit log. The methods described in this dissertation can be extended to other EHR events that may have impact on patient safety.

**Organization of Dissertation:**

The organization of this dissertation is as follows. We first describe process mining in Chapter 2. In Chapter 3, we present the overview of the methods and research design employed in this dissertation. In Chapter 4, we review the medication revisions in the data and apply exploratory

data analysis. In Chapter 5, we show how medication revision events are defined and categorized by a human review process. Based on the human review descriptions in Chapter 6, we describe fault tree analysis (FTA) as a method for analysis of the medication revision events. We manually constructed the FTA and analyzed the factors that could lead to medication revisions events. Learning from the human review process, we developed an algorithm for detection of the medication revisions, described in Chapter 7. In Chapter 8 there is a discussion of the dissertation, its limitation, recommendations and future directions. Finally, Chapter 9 offers concluding remarks on the contributions of the dissertation in this domain.

## Chapter 2: Process Mining in Electronic Health Records

With the opportunities provided by big data in EHRs, we review how to meaningfully analyze EHR data to understand processes in EHR activity. In this chapter, we describe process mining, the steps involved therein, and its application to EHR data.

**Motivation:**

Health information technology, including e-prescribing and CPOE, has been proven to improve patient safety (Ammenwerth, Schnell-Inderst, Machan, & Siebert, 2008; Bates et al., 2001; Fiumara et al., 2007). However, as the EHR is a complex system introduced into an already complex workflow of clinical care processes, the probability for error is high if the EHR is not properly implemented. In a study on unintended consequences of use of EHR, unexpected increased mortality was noted (Sittig, Ash, Zhang, Osheroff, & Shabot, 2006). It revealed new factors that were introduced into an existing workflow when an EHR system was introduced. Koppel et al's study of (Koppel et al., 2005) case scenarios demonstrate how EHRs actually facilitate errors by fragmenting data (e.g., multiple sections, delay in updating data, lack of notification of system failures), hindering performance with usability flaws (e.g., selection of medication, lack of visibility and feedback of automated process), and negatively impacting workflow issues (e.g., inability of modify ordering process, problem in sending prescription). Research by Campbell (Campbell et al., 2006) and others point out the number of socio-technical contributors to these unanticipated events.

While system methods have been applied in health care to understand how EHR actions can impact patient safety (Barach & Small, 2000; Carroll, Rudolph, & Hatakenaka, 2002; Rex,

Turnbull, Allen, Voorde, & Luther, 2000), not all problems within the EHR system lead to visible events. Many EHR systems only log clinical task data making it difficult to identify all potentials causes of unintended events. Given the absence of more granular data, the full potential of EHR actions on patient safety is unknown. Further, much of this work has been completed using observational studies and interviews. As the volume of EHR data increases, it will be difficult to scale exploration in this area to the volume of data produced. The challenges that are faced include: how to detect and investigate an error event in EHR, how to identify system issues leading to error event, and how to continuously monitor error events. With formal methods for process mining in EHR, we can meaningfully analyze EHR events to improve our understanding of EHR activity.

**What is Process Mining?**

Process mining can be defined as "distilling structured process description from a set of real executions" (Maruster, Weijters, Van der Aalst, & van den Bosch, 2002) Process mining is an abstract method that using an array of tools and techniques, specified to a particular context for domain usage. Process mining can be applied at various levels. At the highest level, it can be applied to workflows, such as the registering and control of processing public works contracts, which involves people and tasks (W. M. van der Aalst et al., 2007). Process mining can also be applied at the task level for specific tasks within a workflow, such as steps taken by a user to complete online purchase (Cho & Kim, 2004). Depending on the context, the methods vary.

## Methods for Process Mining

Process mining derives from process modeling methods and data mining methods. Methods include defining and analyzing the process in a workflow or in the tasks needed to develop a model. They are applied in developmental stages, and are used for simulation of the process. These models can be done by expert review and observations. They are used for discussions, documentation, performance analysis, and specification and configuration of the system (W. van der Aalst, 2013).

Data mining methods are particularly useful with big data to both explore and analyze data. Useful methods include exploratory data analysis, clustering algorithms, decision trees, rule-extraction systems, and other machine learning algorithms that aide in classification and regression studies.

## Applications of Process Mining

Three major areas of applications include 1) discovery 2) conformance and 3) enhancement of process. The three components are illustrated in Figure 1.

Discovery: Process mining can be used to uncover previously unknown workflow or user behavior. An example would be exploring workflows in an emergency room (García, Alfonso, & Armenteros, 2015).

Conformance: Process mining can be used to test if user behavior or a process conforms to the modeled or expected behavior. Deviations can be detected and analyzed to uncover factors

responsible. Such conformed models can be used to monitor and predict deviations. An example would be analysis of patient data to assess quality of care in heart failure (Baker et al., 2007).

Enhancement: Process enhancement can be done by assessing the overall outcome of the process. An example would be assessing time spent in an ordering process in an EHR (Zheng, Padman, Johnson, & Diamond, 2009). While such studies are common in usability testing, process mining can also provide analysis from real time execution.

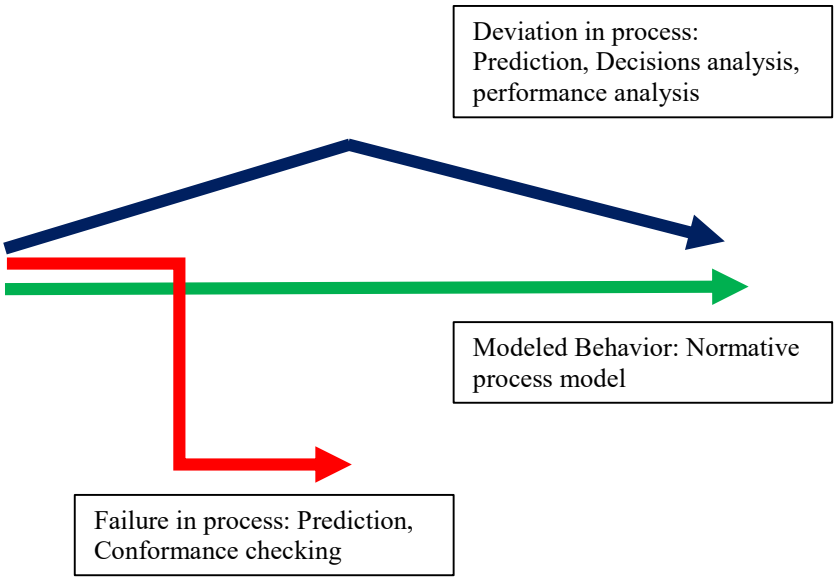



Figure 1 Process mining applications

Note: Adapted from (W. van der Aalst, 2013)

In the event of an error from an EHR user, it is difficult to the identify cause of the event. We can uncover the adverse outcome of EHR activity, but not the specific action by the user that led to error. Applying process mining in EHRs, we can uncover EHR processes that impact patient safety and monitor such events.

## Conclusion

With widespread adoption of EHRs, increasing volumes of clinical data will be available in electronic format. However, even in the era of "big data", without process mining or similar automated methods – we are limited to a human review of EHR events and their influence. While EHRs log clinical activity and some amount of administrative data, auditing of activity log have been limited in EHRs. In this dissertation, we define methods for process mining in EHR for medication revisions events with the aim of developing formal methods to aide understanding of EHR events and their impact on patient safety.

## Chapter 3: Research Design for Process Mining

Using EHR data, we seek to mine the process of medication revisions to uncover knowledge about the process, such as what system factors may lead to changes during medication ordering task. Process mining can be done to uncover factors resulting in medication revision that can be detected to improve patient safety. In this chapter, we present an overview of the methods used in the dissertation.

**Objective:**

Our objectives are to build upon human-recognized medication revision events in EHR, to apply process mining to these events to understand the factors contributing, and to build an automated system to detect medication revisions in electronic health records. Our long term goal is to develop a system that can monitor user activity and predict EHR activity deviation, thus potentially improving patient safety.

**Dataset:**

In this study, we used EHR data from six ambulatory clinics in the United States. Our unit of study is the process of making a medication revision to a previous medication order in a patient record. Medication orders represent an initiating order or entry in a patient's chart. Each medication order can have one or more refills, updates, or status changes. The EHR data covers patient data from April 2004 to May 2014, yielding 306,345 patient records. From this data, we chose a sample dataset consisting of medication orders in adult patients (18-60 years), and the medications ordered by Internal Medicine physicians.

**Medication Revisions:**

Medication revisions can be considered as a process within medication ordering, a higher level clinical task. Within that task, a number of EHR processes contribute to the final outcome. For example, selecting a medication item from a list, completing the medication instructions, marking the order for additional authorization, temporary deferral of the order, saving the order for later review and sign off. One such process is medication revision. The EHR system allows providers to revise medication orders by labelling them as 'Entered in Error' (EIE). EIE is a status of prescription that can be used to correct entries in the system. Once a medication is marked as EIE, the medication is removed from the patient's past and current medication history. However, it remains accessible when specifically choosing to view all medications. The EIE like feature is available in major EHR systems in the United States, though the labelling (i.e., delete rather than EIE) of this function varies.

Medication revisions are good candidates for process mining, because they can be considered both as deviations and as "good catches". They are deviations because the medication ordering task ends in altered state when revised using EIE. They are instances of conscious choice regarding a previous action. Given the nomenclature of 'Entered in Error', the use of this EHR feature was intended to allow remediation of previous action or a "good catch". However, the use of this function for this purpose is an assumption. Here we explore how and when EIE is actually used.

Examples of medication revisions:

a) A physician wants to order 400 mg of Ibuprofen. He searches for the medication and selects it from the drop-down menu. But after selection, he finds out that he has chosen

600 mg of Ibuprofen instead. So he removes the wrong selection by marking it as 'Entered in Error' and makes a new order.

b) A physician wants to order Albuterol, but instead selects Atenolol because of the proximity and similarity of the two drug names. He removes the Atenolol prescription by marking it as 'Entered in Error'.

**Revision Dataset**

The distribution of medication orders, orders labelled as 'Entered in Error' by the providers, is presented in

Table 1. Given in this table are details regarding the overall occurrence of medication orders, the sample data set (limited to only those orders within Internal Medicine meeting our inclusion criteria), the training set used as part of the algorithm, and the test set (Table 1). 100 patient cases were selected from the sample set for the manual review described in Chapter 5.

Table 1 *Description of EHR data*

| | **EHR Data** | **Sample Data** | **Training Data** | **Test Data** |
|---|---|---|---|---|
| Total medication orders | 2,110,385 | 270,774 | 3,056 | 267,718 |
| Medication orders with revisions | 35,833 (1.70%) | 3200 (1.18%) | 213 | 2,987 |
| Total patients | 306,345 | 24253 | 100 | 24,153 |
| Patients with at least one revision | 18089 | 2040 | 100 | 1,940 |

**Aims:**

Process mining in EHR would explore/describe:

1. Human and system factors leading to the event
2. Differences across factors revealed by the process model
3. Means to proactively detect, support or mitigate the impact of system factors

**Research Design:**

For process mining of medication revisions events, we followed the steps:

1. Through data mining, we first explored the prevalence of revisions in a large dataset
2. To uncover the process, we used expert review to explore the tasks underlying revisions
3. To synthesize the process model in process mining, we modeled revisions using fault trees

Exploratory Data Analysis of medication revisions

The first step in process mining was to apply exploratory data analysis (EDA) to the data to understand the medication revisions in the dataset. We applied EDA from the levels of the organization, the patient and the event. We generated visualizations and descriptive statistics. As an exploratory study, and due to limited data, we only performed a descriptive analysis. No statistical model and confirmatory analysis was conducted. Using EDA, we sought to gain an understanding of how medication revisions occur in the system, and searched for any trends or outliers. EDA would not reveal why the events occurred, i.e. factors contributing to the event, but EDA did help us conceptualize our ideas and direct our further explorations.

Expert Review of medication revisions

To uncover the factors responsible for the revision process, we conducted expert review of selected cases. Physicians were recruited to review the medication revisions. We considered it important for the local dataset of local events, to be reviewed locally to understand the factors specific to this implantation and the use in these sites. We recruited physicians to review 100 cases of medication revisions and to provide possible reason or cause for the revision. From the review, categories were defined. The review process also provided data for modelling the process using fault trees and for developing the algorithm for detection.

Development of process model using fault tree analysis

The final step in process mining was synthesizing a process model using fault trees and Petri nets. We created a model using fault trees for individual categories described by expert review. We constructed fault trees for the error categories, based on the physician review of the 100 cases reviewed. From the fault trees we then synthesized a Petri net model of the process. The fault trees showed interaction between the factors, and the process model showed different factors contributing to the medication revisions.

Development of algorithms for automated detection of medication revisions.

With the process mined medication revisions, we sought to show application to clinical scenarios. We developed an algorithm for detection of medication revisions based on previous user activity in the EHR. This knowledge was obtained from the process mining. We modelled a rule-based classifier algorithm to detect the medication revisions in the data. Training on the 100 patients that were expert reviewed, we evaluated the sensitivity, specificity and accuracy of the algorithm on the test data set (Table 1).

In the coming chapters, we present the study design, findings and discussion for each step in detail. We conclude with a discussion of future direction, vision and contributions.

## Chapter 4: Exploratory Data Analysis of Medication Revisions

In this Chapter, we explore the medication revision events in the data. We begin our process mining by applying data mining to understand the raw data. We used exploratory data analysis (EDA) to obtain descriptive analyses and visualizations towards understanding the occurrence of medication revision events in the EHR data.

### Introduction:

Exploratory data analysis is a quantitative approach for understanding data when little or no statistical hypothesis exists (Behrens & Yu, 2003). Exploratory data analysis does not require probability, significance or confidence. (Tukey, 1977). It provides clues to the data and discovery of unexpected events. EDA focuses on fit and residual, helping in understanding outliers. It assists with understanding what happened along with an event, without needing to consider conclusions of significance or confidence. It isolates features of the data and directs design of ideas and further analysis.

### Methods

EDA is more focused on the goal of understanding data than modeling the data. We explored the data from three perspectives (W. Van Der Aalst, 2011):

1) Organization perspective: Provides understanding of resources such human and material, behind the processes. They provide insight into the ‘actors’ and ‘props’ involved in the process. They are useful in understanding people, roles, authority, policies, location influencing the processes.

2) Patient perspective: Provides understanding of case (patient) factors in the path of the process. These factors are external to process and useful in characterizing the environments in which the process occur.

3) Event perspective: These are factors that are components of the process – mostly timing and frequency of events. They are useful in understanding bottlenecks, service loads, resource utilization, execution time.

Organization perspective: We calculated the overall prevalence, the raw counts of occurrences by clinics, the proportion of patients with revisions, and the proportion of providers who have ever used the revision feature. We calculated the overall average medication revisions per patient.

Patient perspective: We mined the top 10 diagnoses, and top 10 medications by counts of occurrence. We explored the relationship between patient age and the number of medication revisions, and the number of visits made by patient till the time of occurrence of medication revisions event by using a bar graph. For both the graphs, we also plotted the average number of revision per patient in each category.

Event perspective: The following were calculated and graphed for medication revisions:

Number of medication at time of revision: total number of medication in the patient record at the time a medication was revised. We used a bar graph, with the average number of revision per patient as a secondary line graph.

Number of medication in an encounter at time of revision: in a single encounter at the time the medication revision was done, the number of medications that were signed off in that encounter. We used a bar graph, with the average number of revisions per patient as a secondary line graph.

Time of concurrence: at what time of the day the revision event occurred. We used a bar graph, with average number of revision per patient as a secondary line graph.

Number of instances ordered: the number of instances the medication that was revised had been edited or refilled. We used a pie-chart to show proportion of medication revisions in each category.

Time relationship between first occurrence and revisions: time in days between the first time the medication was ordered and the time it was revised. We used a pie-chart to show proportion of medication revisions in each category.

Time between last update and revisions: time in days between the last time the medication was updated and the time it was revised. An update suggests the last time the medication was reviewed by provider. We used a pie-chart to show the proportion of medication revisions in each category.

Who made the revision:  whether the revision made by the person who first ordered the medication, or by someone else. We used a pie-chart to show proportion of medication revisions in each category.

We conducted the EDA to understand the occurrence of medication revisions in the data set. As an exploratory work, no statistical analysis was done nor models created. From EDA, we expect to understand how, when and why medication revisions occur. EDA will also help us identify

limitations in the dataset. Based on the knowledge, we design our questions and experiments for detailed analysis, to uncover specific factors from the data.

**Results:**

Organization perspective: We want to see which provider and clinic uses medication revisions. Is there any particular provider group, say residents, or nurses, or specific clinic using the process more than the others? This can help in focusing strategies to improve process on specific actors handling the process. Table 2 describes the prevalence of medication revisions by patients and providers. A majority of providers (55%) have used the process, but we could not further characterize the providers due to study restrictions.

Table 2 *Organization perspective of medication revisions*

| | Total count | With medication revisions | Prevalence |
|---|---|---|---|
| Medication orders | 2,110,385 | 35,833 | 1.7% |
| Patient | 306,345 | 18,089 | 5.9% |
| Providers | 2688 | 1481 | 55% |

On average, 1.9 (SD: 2.4) medication revisions occur per patient.

Figure 2 shows the distribution of medication revision events by clinics. These are raw counts. Given the number of clinics that fall into internal medicine, it is not unexpected that this type of clinic, in an ambulatory setting, has the highest number of revisions.

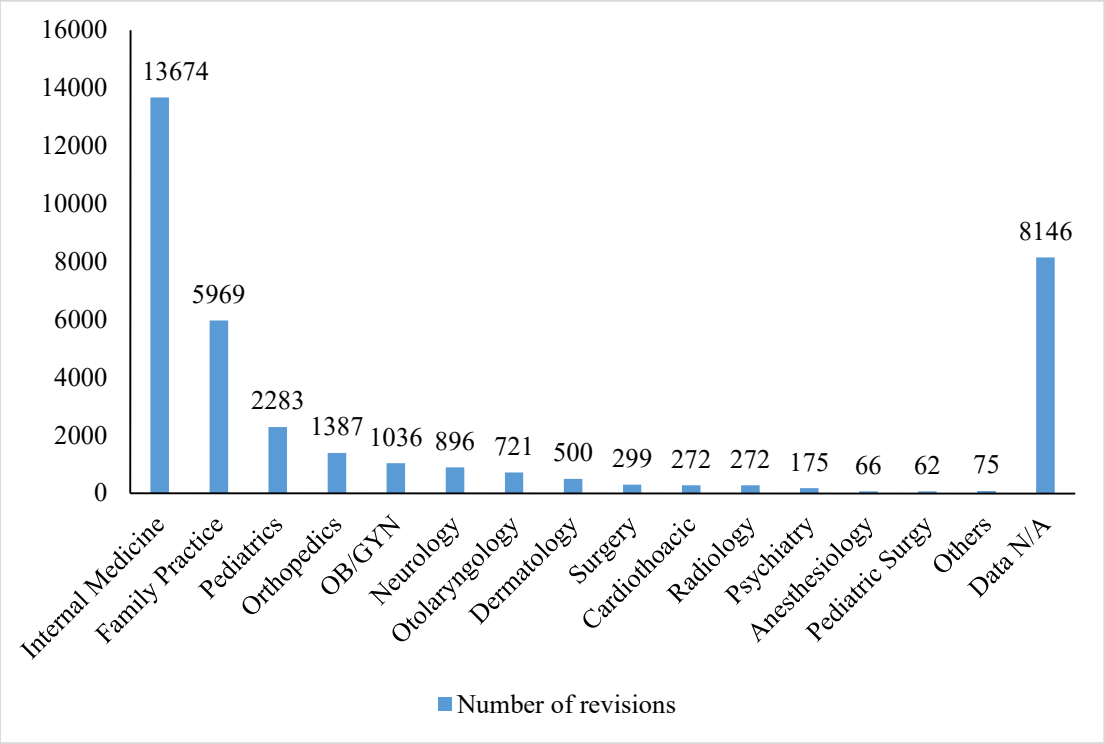



Figure 2 Medication revisions by clinic

Once we learn organizational perspective, we next want to see in a given case (i.e. patient) what occurs during the process of medication revision.

Patient perspective: Patient perspective helps us to understand the data elements relating to the processes. We wanted to see which patients have medication revisions. So we explore medication revisions from patient demographics - age, patient clinical condition (diagnosis, type of medication, clinical visit frequency)

In order to determine whether or not certain diseases or medications might be more at risk for revisions, we considered the most frequent diseases or medications with revisions. Table 3 and Table 4 show the top 10 diagnosis and medications that occur with medication revisions, sorted from highest to lowest. They do not show any pattern specific to diagnosis or to medications.

Here we see that there are no clusters or groupings such as cardiac medications or common conditions.

Table 3 *Top 10 diagnosis occurring with medication revision*

| |
|---|
| Hypertension |
| Hypercholesterolemia |
| Ischemic Heart Disease |
| Atherosclerosis |
| Sinusitis |
| Diabetes |
| Respiratory difficulty |
| Asthma |
| Thyroid Disorders |
| Infectious Diseases |

Table 4 *Top 10 medication entries associated with revisions*

| |
|---|
| Metoprolol (2%) |
| Hydrocodone-Acetaminophen (1.9%) |
| “No Medications” (1.8%) |
| Lisinopril (1.7%) |
| Levothyroxine (1.6%) |
| “None” (1.4%) |
| Atorvastatin (1.3%) |
| Aspirin (1.3%) |
| Amlodipine (1.2%) |
| Simvastatin (1.1%) |

Note: The proportion of medication revisions with the entry is shown in parenthesis

Patient age relates to risk of certain illnesses, type of medications, and number of medications. We wanted to see if there are specific age groups at higher risk for medication revisions. Figure 3 shows the distribution of patient age with medication revisions. The occurrence is higher in older patients. The average age of patients with medication revision is 51.6 years (SD 30.3).

As a proxy for chronicity of patient condition, we studied number of visits to clinic. Figure 4 shows the distribution of the number of visits made by patients with medication revisions. A slight increase in average revision rate per patient is present when medication entry is made prior to patient visit (prior to a patient visit is possibly a telephone order or by pre-registration documents). No statistical inference of the association is made.

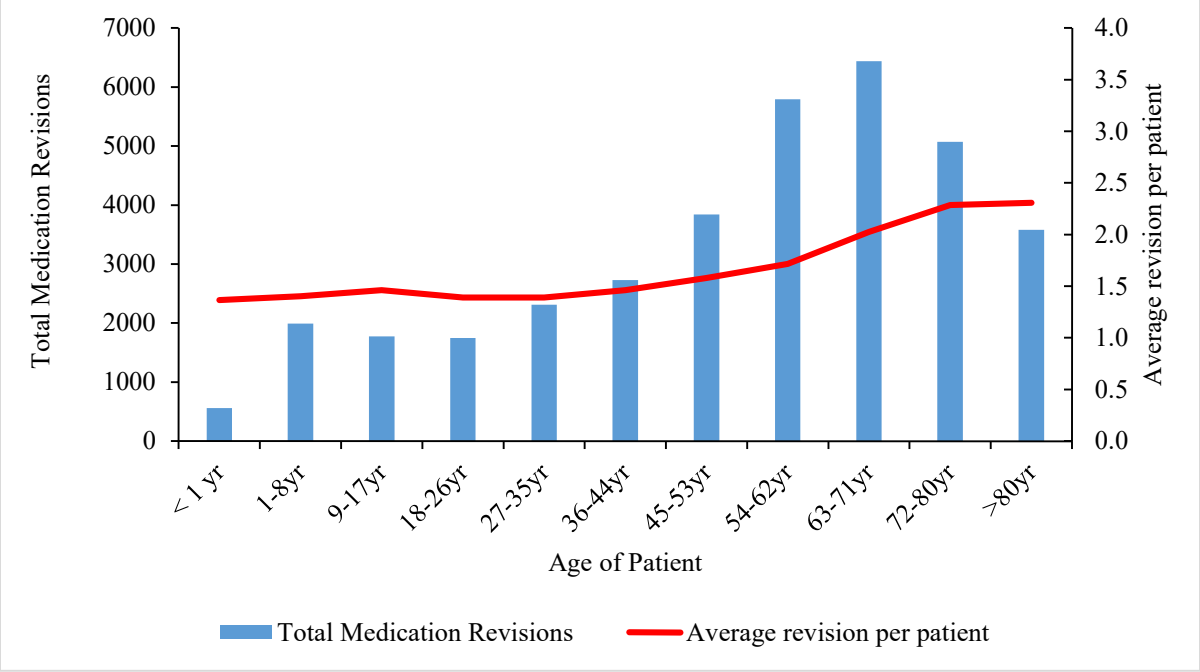



Figure 3 Age of patients with medication revisions

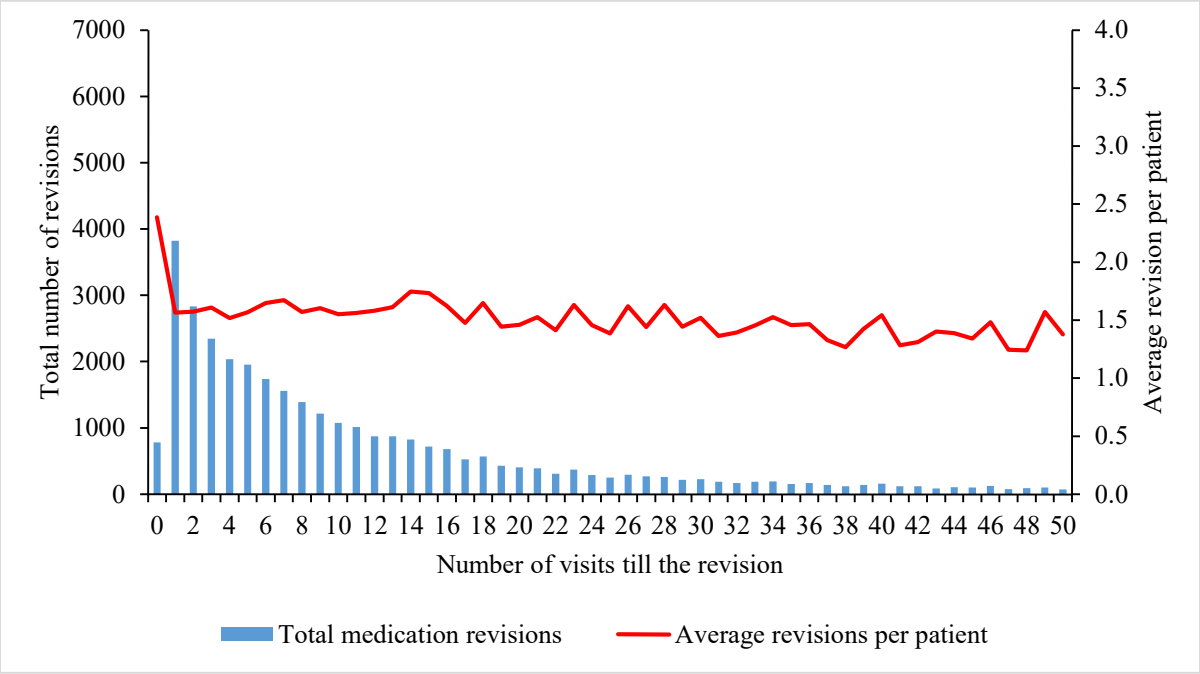



Figure 4 Number of visits till medication revision

The Patient perspective did not reveal any specific cluster or groups of patient with higher risk for medication revisions. But it did reveal some outliers in form of increased incidence in circumstances such as medication orders made prior to any visit. Now moving from organization to patient, our next level is at the event – medication revision – itself.

Event perspective:

Event perspective helps un understand the timing and frequency of events. We wanted to see when does medication revision occurs? How long does it take for a medication to be revised? How medications are revised in an encounter? Such questions help us understand the factors predisposing to the process or resulting from the process. In Figure 5, patients with more medications have a higher occurrence of medication revisions. Furthermore, those with more medications signed-off in one encounter had a higher occurrence of medication revisions per

patient (Figure 6). More revisions are made if medications are entered prior to patient visit or entered outside of working hours (8am-6pm) (Figure 7). The majority of revisions (67%) are made within the first 2 instance of the medication (Figure 8). Additionally, the majority of revisions (52%) are made to medications that have been in the system for a longer duration (over one month) (Figure 9). Most of revisions (41%) are made on the same of day of the last review (Figure 10). The majority (57%) of revisions are made by a different physician than the physician who initially ordered the medication (Figure 11).

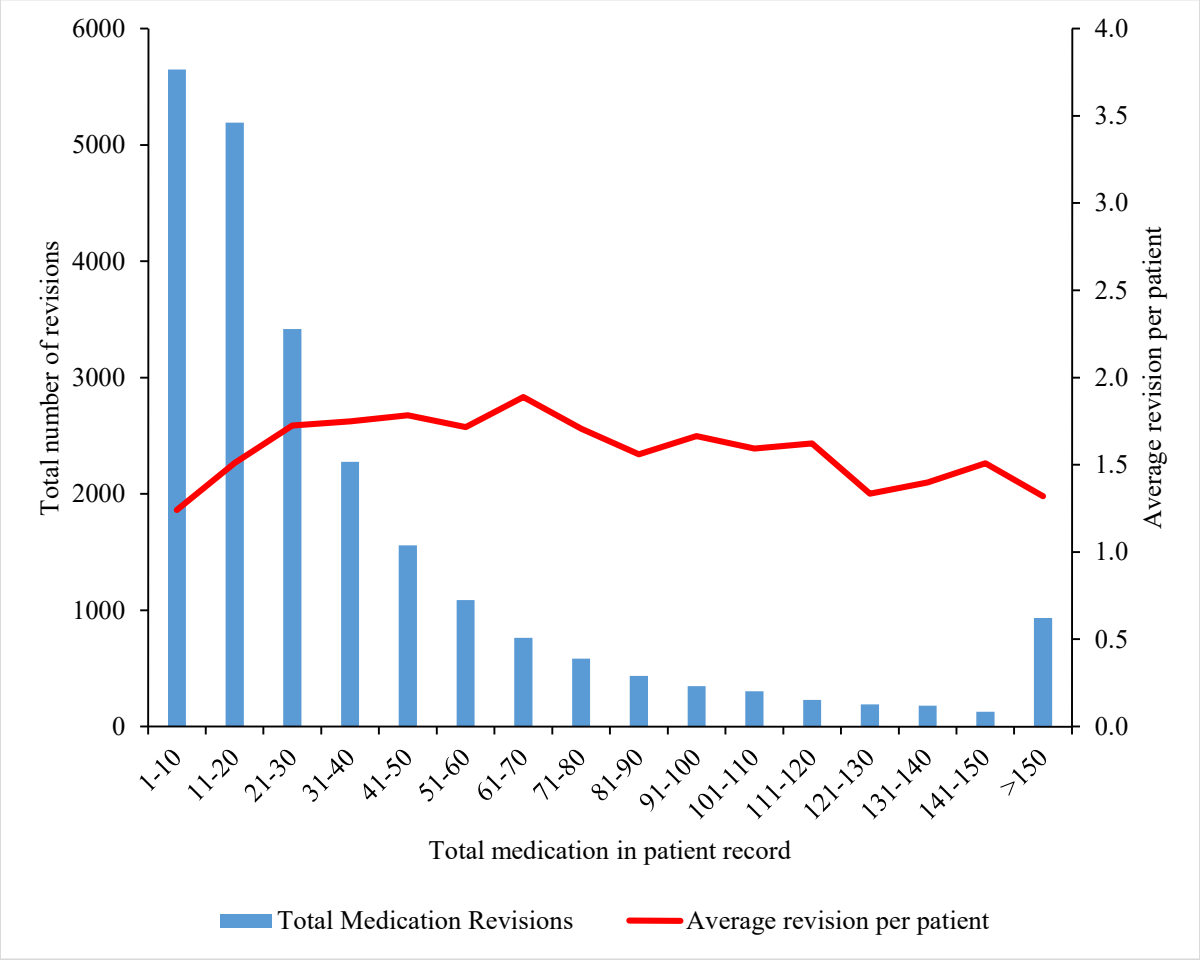



Figure 5 Medication revisions by number of medications in patient record

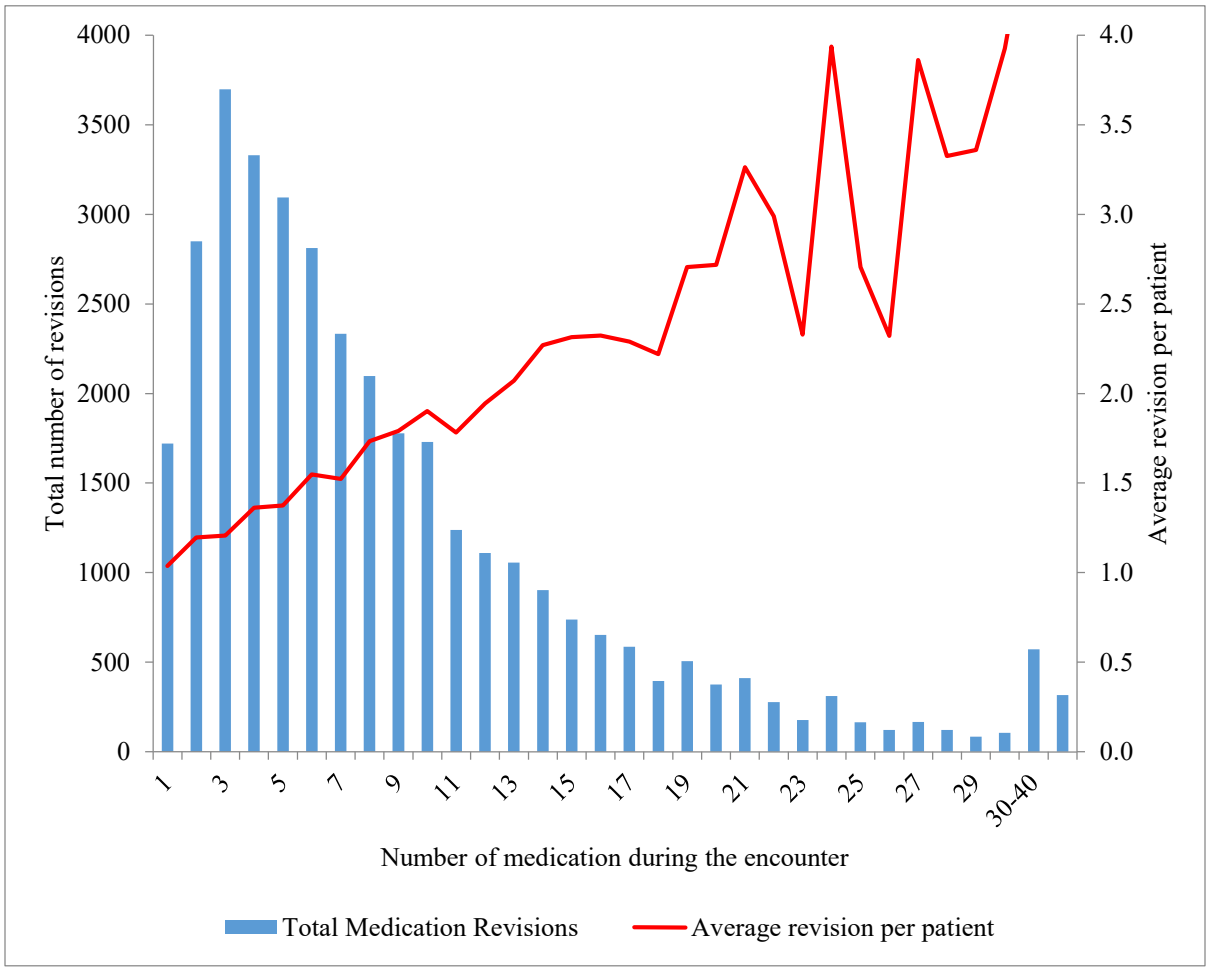



Figure 6 Medication revisions by total medication during an encounter

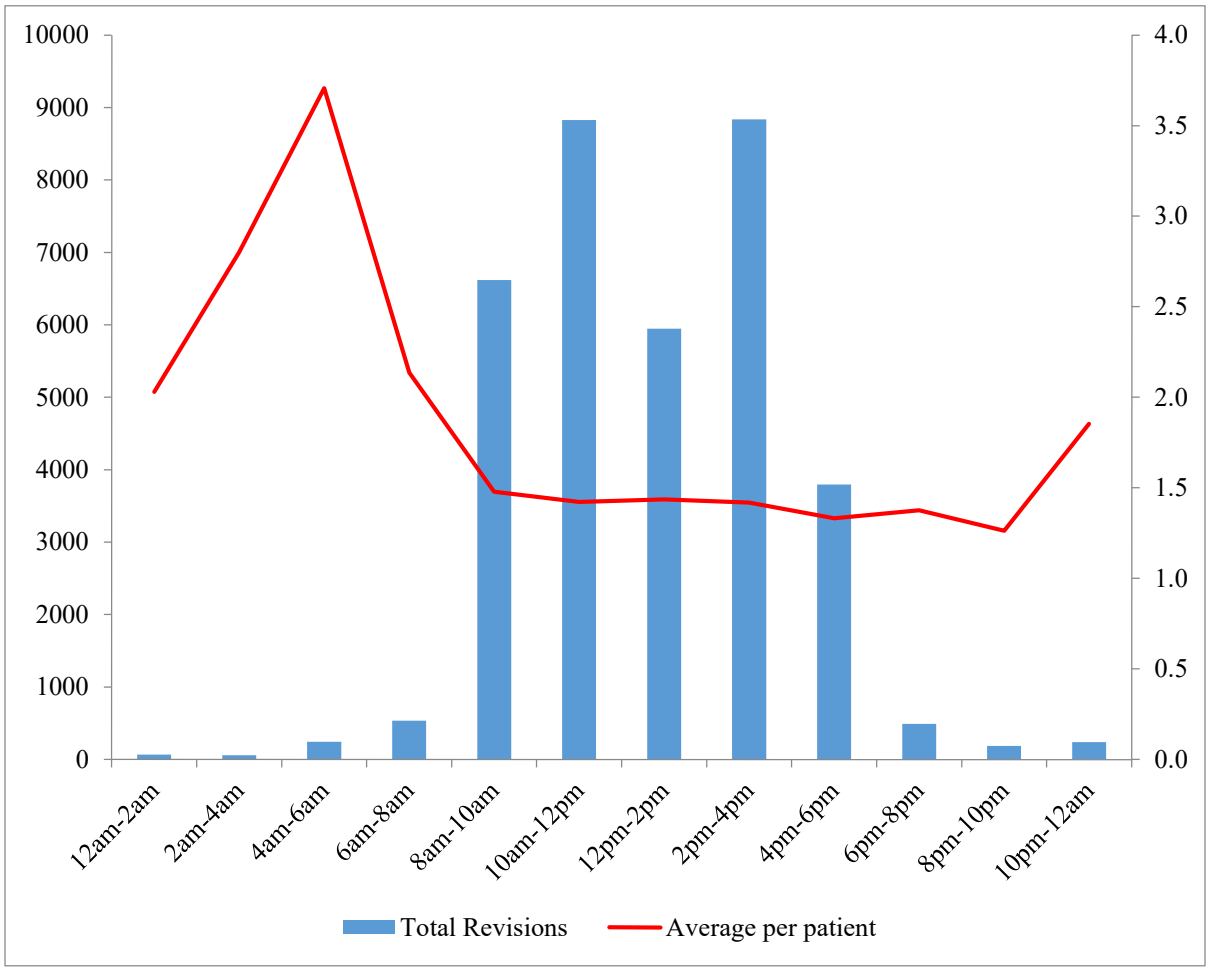



Figure 7 Medication revisions by time of the day

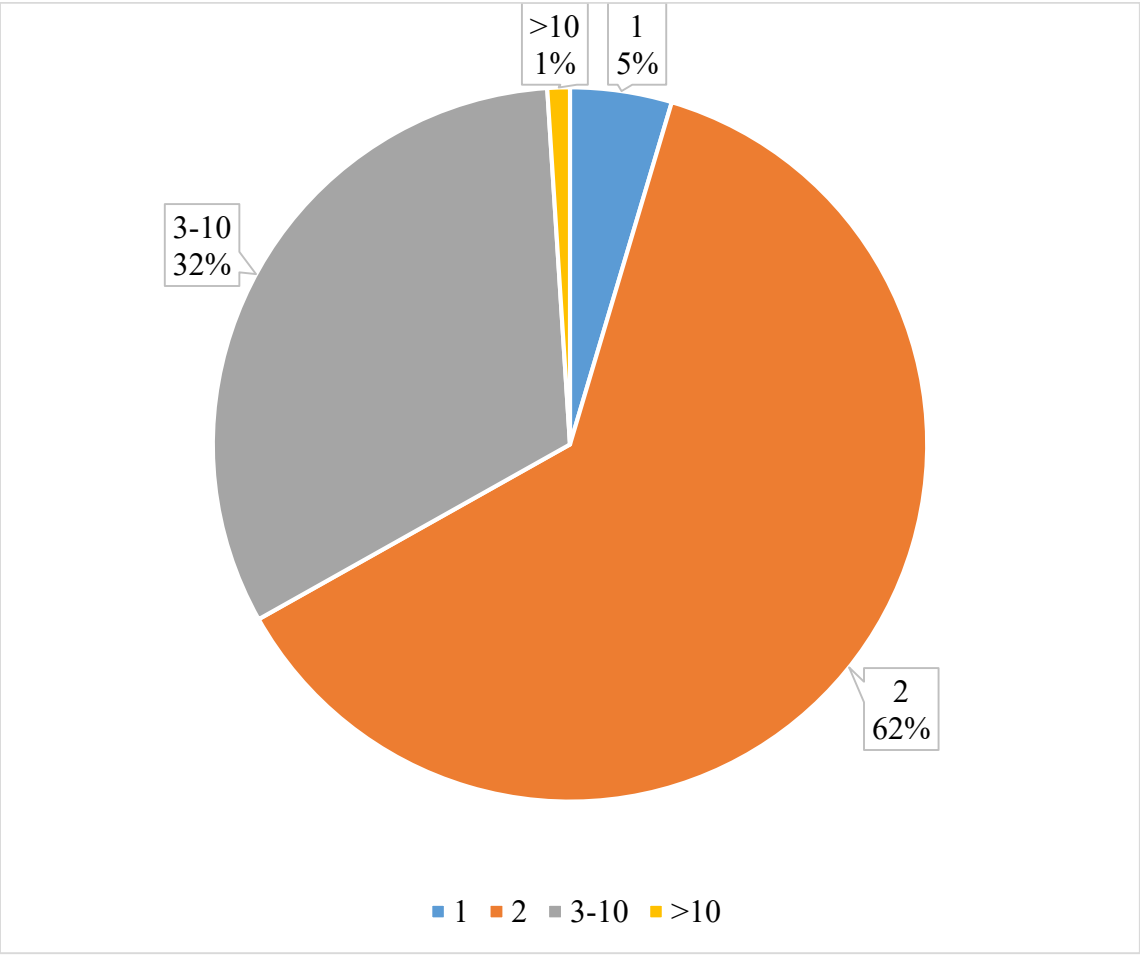



Figure 8 Number of instances prior to revision

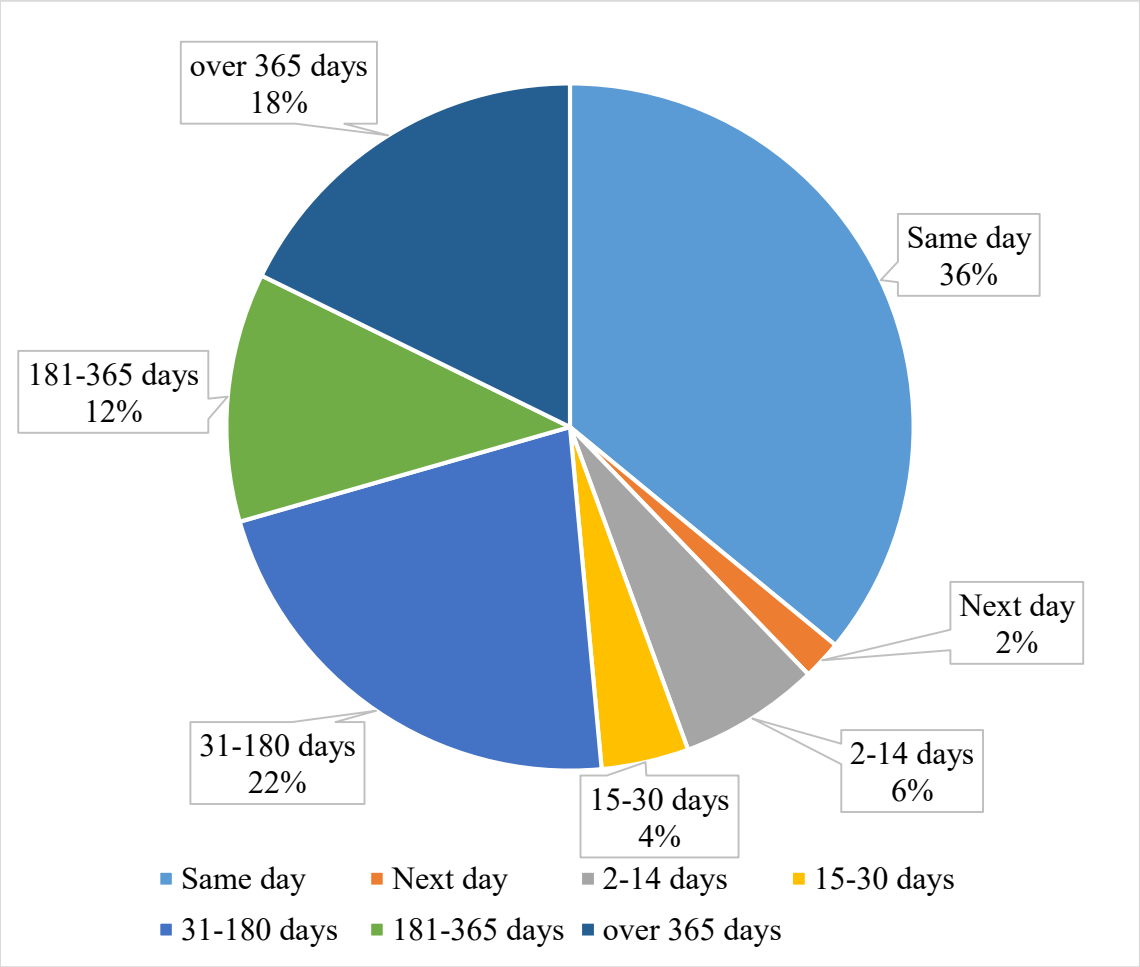



Figure 9 Time between first order and revision

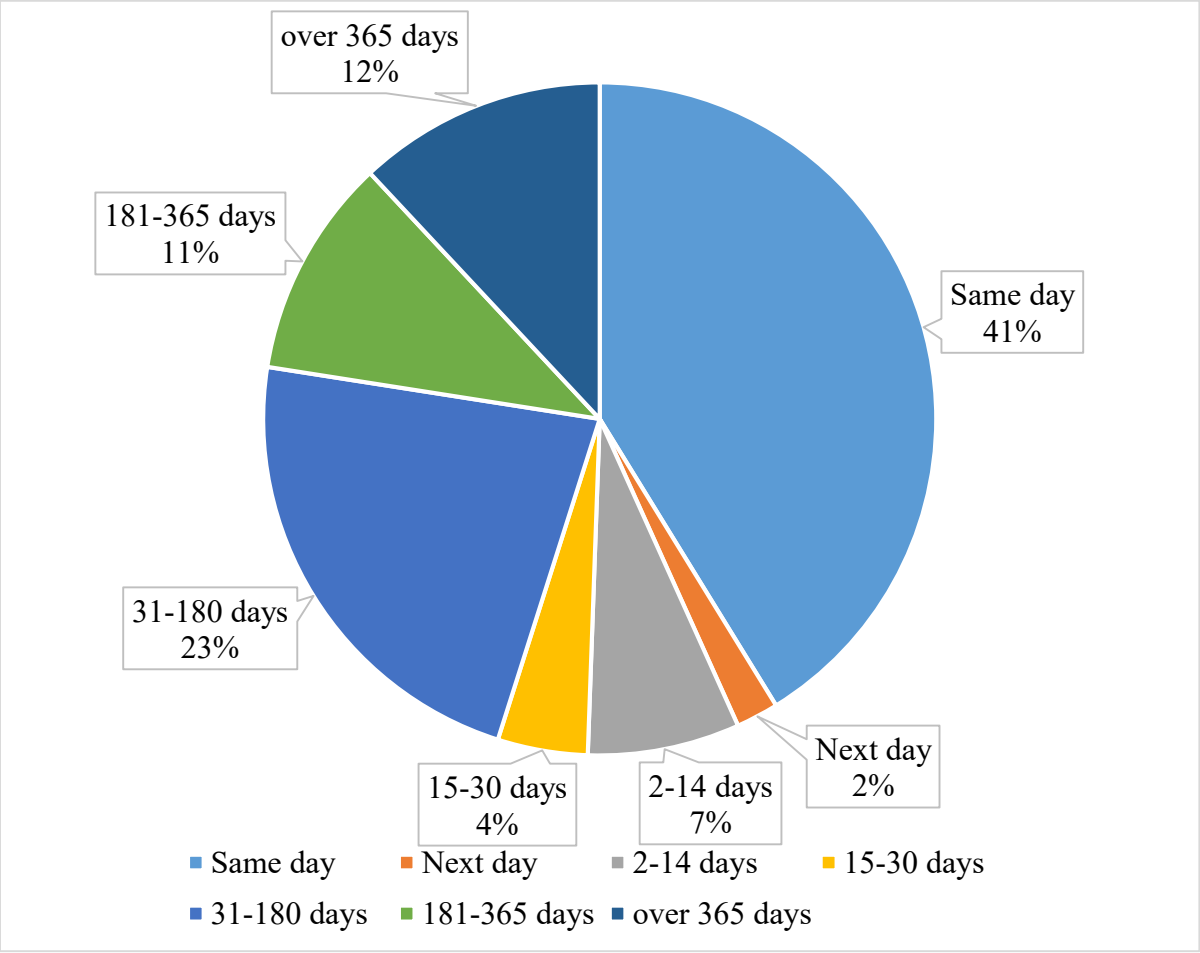



Figure 10 Time between last update and revision

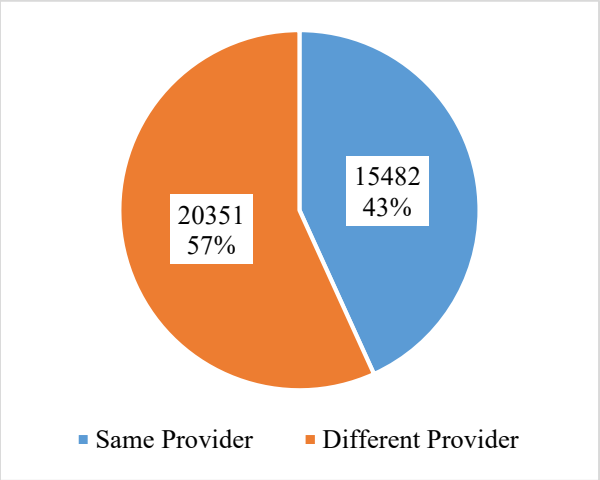



Figure 11 Provider who made the revision

**Discussion**

We performed EDA to understand the medication revisions in the data. Though we did not find specific patterns, we identified some interesting occurrences in the data, such as average revisions per patient increased when medication details are entered without a patient visit, when medications ordered outside of regular clinic hours, when there are more medications signed off in a given encounter. These occurrences suggest possible cognitive factors and workflow factors contributing to the revisions. However, using EDA alone, we could not identify the factors from the data. The EDA also revealed limitations in the data logs and relationships between data elements. We found that event logs were relating more to user access activity, such as when and who made a change, than to event state, such as what was the previous state, and why the state changed.

The patient perspective revealed that medication revisions occur in common illnesses cared for in ambulatory clinics, and the top medications match the diagnoses. Older patients tend to have

more revisions, which may be due to presence of more medications, poly pharmacy and chronic diseases care. In the majority of visits, we see a slight increase in average revisions per patient when medication is entered prior to the visit. This could be from telephone orders, or documents sent by patients (from other clinics) prior to first visit. This could be due to correcting information on patients visit. The number of revisions tends to decrease by number of visits. This is due to earlier visits being the time when newer medications are added and when changes are made to therapy.

A higher occurrence of medication revisions was seen when patients had more medications in their chart at the time of revision, and also when more medications were signed-off in one encounter. This could be due to the increase information load and clutter in the interface leading to poor review by the providers. This noise in the data could occur due to a number of reasons, such as inability to remove or reconcile older medications, and system issues that facilitate providers to add new medication rather than review and update older medication. The increased average revision per patient on medications made outside of working hours could be due to cognitive factors (fatigue, distress) and system factors (non-availability of data). The majority of revisions were made within the first 2 instances of the medication, which is a likely indicator that they were revised immediately. However, upon review of the time data, we found this not to be the case, as the majority were entered in system over a month prior to revision. This can be due to the tendency to order new medication rather than review and update older orders, and thus leaving behind an active order to be revised later. It is also to be noted that the majority of revisions were made to medications without reviewing the medication list. The difference in the provider could be due to increased turnover of providers in a teaching hospital setup, longer

duration of medications in the system, or difficulties in managing medication orders made by other physicians.

**Conclusion:**

EDA is subjective and not comprehensive. One of limitations is that very focused exploration was done, keeping in mind only the most evident factors. This was because of our limited knowledge of the system and the data relationships prior to expert review. The EDA only provided insight, and based on our knowledge and expectation of the system, we formed ideas to further explore. In our case, our goal in process mining was identifying factors contributing to the medication revision event. From EDA, we received an idea of what could have happened. Based on this, we proceeded to seek definitive factors responsible for medication revisions.

Limitations

We could not precisely calculate the proportion of medication revision by total orders by clinic. This is due to the billing data used in determining the origin of the medication order. Further, all the medical specialty clinics, such as Neurology, Cardiology, etc., were grouped under Internal Medicine clinics. Therefore, we treat a number of clinics as falling under a single clinic label. Due to constraints within the study, we could not show provider characteristics in relation to the medication revisions.

## Chapter 5: Describing Medication Revisions by Manual Review

The data mining approach presented an overview of medication revisions in the system. In this chapter, we describe the expert review of a sample of cases to describe the events and uncover factors contributing to medication revisions. Using manual review, physicians explored 100 medication revision events and categorized them in terms of trigger for the revision. From this review, we obtained the workflow used to review each chart as we well as a proposed reason for each revision. These results are used in the development of a process model using fault trees and a standard dataset (training dataset) for generation of an algorithm for detection.

### Introduction

From the exploratory data analysis, we found no specific pattern for medication revisions in the system. This led us to anticipate that multiple types of medication revisions could be present. Thus, we directed our goal to classifying the medication revision events based on possible reason for the revisions, and in the process, describing the factors contributing to the revisions. To have a background on possible reasons for revision, we learnt from medication error literature that describes possible reasons for errors/deviation in ordered medications. As medication revisions can be considered as "good catches", we further explored the error literature.

From studies on medication errors in prescription process (Barker, Flynn, Pepper, Bates, & Mikeal, 2002; Lesar et al., 1990; Tesh & Beeley, 1975), we derived three major error domains: decision, ordering and specification. Decision errors include wrong drug for the condition, inappropriate drug because of allergy, drug interactions, and adverse events. Ordering errors include illegible writing, wrong information, missing information, and wrong patient.

Specification errors include errors in dosage, dosage forms, and quantity to dispense. For a more specific description of such causes, we reviewed studies on medication errors in EHRs (Bobb et al., 2004; Shamliyan, Duval, Du, & Kane, 2008; Shulman, Singer, Goldstone, & Bellingan, 2005; Walsh et al., 2006; Westbrook et al., 2012; Zhan, Hicks, Blanchette, Keyes, & Cousins, 2006). From these studies, we synthesized a list of error types, shown in Table 5.

Table 5 *List of error types complied from articles on EHR medication errors*

| | **Error Category** |
|---|---|
| Decision errors | Drug not indicated |
| | Drug-drug interaction |
| | Drug-laboratory interaction |
| | Drug allergy |
| Ordering errors | Wrong patient |
| | Wrong timing |
| | Wrong place in chart |
| | Wrong information |
| | Missing information |
| | Typo/Keystroke error |
| | Drop-down menu selection error |
| | Duplicated therapy |
| Specification errors | Wrong dose/volume |
| | Wrong rate/frequency |
| | Wrong route |
| | Wrong strength |
| | Wrong formulation |

These error categories informed us on the reasons for medication revisions we could expect in our study data. We used these categories to develop the form for expert review.

Depending the type of errors, the causes and reasons for the medication errors also varies. Errors in decision making could arise from lack of knowledge, or failure to review for tolerability and adverse effects (Leape, 1994; Schiff et al., 2011). Some of the other factors such as workload, physical environment, physical and mental well-being (fatigue and stress), have also been related to errors (Blendon et al., 2002; Dean, Schachter, Vincent, & Barber, 2002). In EHR prescriptions, the causes of error include lack of knowledge and poor training in use of the system, poorly deployed systems and usability issues in the system (Ash, Berg, & Coiera, 2004). Other causes include wrong timing of alerts, failed alerts, length list of menu items, proximate screen items, obscured order hierarchies, poorly designed icons, and lack of explanation for automated computations (Khajouei & Jaspers, 2008). We expected to uncover such factors responsible for revision in the dataset.

**Methods**

To describe and ultimately categorize medication revisions, 4 physicians manually reviewed a selection of 100 total cases. Their review process included a think-aloud session in which they described their thought process as the navigated through the chart and the potential revision process. Their actions were captured with audio-video and screen recordings. Our participants included 4 physicians (3 residents in post graduate year 3 (PGY3) and 1 faculty) from Internal Medicine. The inclusion criteria required an internal medicine specialty and access to the EHR system in use for this study. No participants were asked to review their own charts. All participants were compensated for their participation.

Dataset: 100 cases were reviewed out of 2040 records (Table 1). Each record was curated to ensure that (1) it captured a unique patient (i.e. no patients with more than one revisions) and (2) that it was selected based on the number of edits/refills. Twenty-seven had 1 refill, 24 had 2, and 55 had three or more refills. The number of refills was included to determine differences between new medications and medications with a longer prescribing history. The first reviewer was given 30 randomly selected cases. Based on the categorization of cases by the first reviewer, each remaining physician received a set of unique cases, as well as cases that overlapped another physician set. The overlapping cases were selected to cover the full range of categories. Physician-1 and Physician-2 had ten cases in common, Physician-2, -3 and -4 had five cases in common (Figure 14).

Data collection: Each physician was allotted 30 cases. In each session, the physicians were asked to walk through their selection of cases. Their actions were recorded in order to identify the workflow used to uncover potential triggers. Movements through the EHR system (e.g. access of tabs, etc.) were captured with keystroke and screen recordings using Turf-usability software (Zhu, Rogith, Franklin, Walji, & Zhang, 2014). Verbalizations of their thoughts were also recorded during this time. Participants had full access to the medical records in each case, and were not constrained in how they interacted with the system. At the end of session, they were briefly interviewed about possible scenarios that would have led to the medication revisions event.

Triggers or a possible rationale for the revisions were also provided by the participants in this study. The physicians were not asked to provide a reason in an open ended fashion, rather the first physician was provided a selection of reasons based our interpretation of the literature

(Figure 12). Following the first ten cases, this list was revised to include physician feedback. The amended list (Figure 13) was then provided to all participants. In addition to the list, from which a physician may select one or more potential causes for revisions, each physician was free to describe the probable cause for the revisions in free text. The potential causes or reason for the revisions were later transcribed to categories by the investigator.

**Reason for labelling as "Entered in Error"**

- Use for rectifying an error, because of
  - Duplicate medication
  - Typos
  - Wrong space/place/person
- Use for medication of treatment, because of
  - Dose
  - Drug name
  - Route / Form
  - Drug class
  - Drug formulation
  - Side effects
- Use for removing from medication list, because of
  - Cancelled medication
  - Discontinued medications
  - Remove and reorder same medication
- Insufficient data
  - Removal of medication that are indicated for diagnosis
  - No sufficient data

Figure 12 Preliminary form with list of reasons for medication revisions

**Reason for labelling as "Entered in Error"**

- Correct a keystroke / click error/ typo
  - Wrong person
  - Similar drug names
  - Drugs nearby in dropdown list
  - Other error in typing / selectin
- Medication list update
  - Patient was not taking medication
  - Patient self-discontinued
  - Provider discontinues a medication
  - Patient finished the course of medication
- Correct a medical decision
  - Medication not indicated
  - Medication allergy
  - Medication safety issues (other than allergy)
  - Update medication SIG - dose/dose form/ instructions
  - Medication drug name
  - Medication drug class
  - Cancel the medication
- Insufficient data

Figure 13 Final form with list of reasons for medication revisions

From the literature review (Table 5), the form with a list of reasons for medication revision was developed. Based on the physician response to the form and free-text description, five categories of triggers were identified (Table 6 ). Many of the responses were in free-text form (Table 7). Using grounded theory, the categories were obtained. The category terms were used by the physician reviewers - e.g. "physician wanted to discontinue", "this is a duplicate medication". Some of the responses were less declarative, but suggested possible reason, for example "drug is not appropriate for patient, so possible a wrong medication", "patient provided incorrect information, so physician updated it".

Category Definitions:

1. **Cancel medication:** When a drug was prescribed only once with no refills. The drug must be indicated for the patient.
2. **Discontinue medication:** When a drug is prescribed and had at least one refill.
3. **Duplicate medication:** When two similar drugs are present in the patient chart at the same time.
4. **Update medication:** When a prescription is updated by changing brand to generic, dosage form, dose, frequency, instructions, etc.
5. **Wrong medication:** When the drug is not indicated for the patient and should not be in the patient's chart.

Table 6 *Relation between error categories defined and form used by reviewers*

| **Category in the form provided to reviewers** | **Linked category** |
|---|---|
| Wrong person | Wrong medication |
| Patient was not taking medication | Cancel medication |
| Patient self-discontinued | Discontinue medication |
| Provider discontinues a medication | Discontinue medication |
| Patient finished the course of medication | Discontinue medication |
| Update medication SIG-dose/dose form/instruction | Update medication |
| Medication drug name | Update medication |
| Medication drug class | Update medication |
| Cancel the medication | Cancel medication |

The remaining categories in the form were used as an associated factor or cause, rather than a category. These 5 categories were grouped into two major subcategories: appropriate use of EIE and system workarounds.

From the physicians think aloud session recordings, access patterns in the EHRs were obtained. From the brief interviews, workflow and possible scenarios for the error categories were obtained. These were used in generation of the fault trees to be described in Chapter 6.

Analysis

Descriptive statistics were calculated for the categories of errors within the 100 cases. Reliability between the physicians was calculated for the overlapping cases using Fleiss Kappa.

**Results**

From the review process we obtained 120 reviews on 100 cases. 87% of the responses were determined from free text. The number of instance and categories selected from the form is shown in Table 7.

Inter-rater Reliability: As majority of responses were free-text, the inter-rater reliability was calculated based on the investigators transcription to defined categories. Due of the way the common cases were allotted, two sets of inter-rater reliability were calculated (Figure 14). Between Physician-1 and Physician-2, the reliability was 0.63. Among Physicians-2, -3 and -4, the reliability was 0.82. This is considered to be adequate to good reliability.

Table 7 *Counts of type of responses and its transcription to error categories*

| **Category selected in form** | **Count** | **Error Categories** |
|---|---|---|
| Wrong person | 3 | Wrong medication |
| Patient was not taking medication | 1 | Cancel medication |
| Patient self-discontinued | 1 | Discontinue medication |
| Provider discontinues a medication | 6 | Discontinue medication |
| Patient finished the course of medication | 5 | Discontinue medication |
| Update medication SIG - dose/dose form/ instructions | 11 | Update medication |
| Medication drug name | 1 | Update medication |
| Cancel the medication | 5 | Cancel medication |
| Multiple Categories | 8 | |
| Free Text | 79 | |
| **Total** | **120** | |

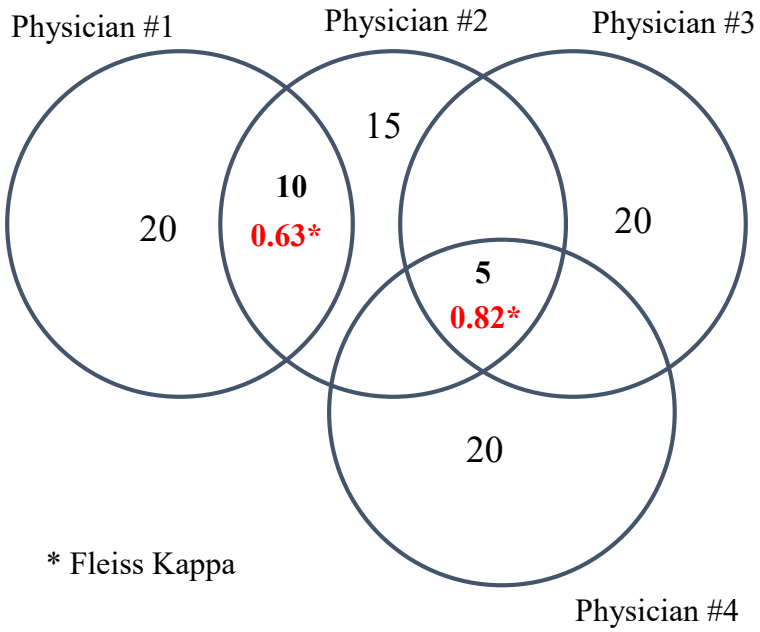



Figure 14 Case allotment and inter-rater reliability

Categories: From the review process, we obtained 5 categories. We found that the majority of EIE were not true medication errors, but workarounds. So we further grouped the categories into appropriate use of 'Entered in Error' function and System issues/workarounds (Table 8).

Table 8 *Error categories for medication revisions*

| **Categories** | **Count** |
|---|---|
| **Appropriate Use of "Entered in Error"** | **45** |
| Cancel medication | 9 |
| Duplicate medication | 27 |
| Wrong medication | 9 |
| **System Issues and Workarounds** | **54** |
| Discontinue medication | 27 |
| Update medication | 27 |
| **Insufficient Information** | **1** |
| **TOTAL** | **100** |

Additionally, we also present the event perspective data from exploratory data analysis across the categories in Table 9.

Table 9 *Exploratory data analysis applied to categories*

| | **Cancel** | **Discontinue** | **Duplicate** | **Update** | **Wrong** | **Overall** |
|---|---|---|---|---|---|---|
| N | 9 | 27 | 27 | 27 | 9 | 35833 |
| Number of visits (Mean) | 9 | 45 | 39 | 39 | 3 | 19 |
| Total medications in patient (Mean) | 30 | 103 | 110 | 79 | 19 | 25 |
| No. of medications in encounter (Mean) | 7 | 5 | 8 | 7 | 10 | 6 |
| Number of instances ordered (Mean) | 4 | 5 | 8 | 5 | 2 | 2 |
| Days between first order and revision (Mean) | 240 | 525 | 464 | 409 | 1 | 521 |
| Days between last update and revision (Mean) | 1 | 133 | 104 | 95 | 1 | 147 |
| Most occurring time | 2pm - 4pm | 10am - 12pm | 2pm - 4pm | 2pm - 4pm | 2pm - 4pm | 2pm - 4pm |
| Revision by different provider | 2 (22%) | 20 (74%) | 7 (25%) | 15 (55%) | 5 (55%) | 20351 (57%) |

Review Process: In addition to describing the types of triggers for medication revisions, the manual review process was conducted to determine the components within each record that may be required to determine a cause. This process was explored in order to: (1) inform the algorithm

for automated detection and (2) provide the structure for our later fault tree analysis (see Chapter 6). The overall flow of the review process by a physician using the EHR system is shown in Figure 15

During the review process, the physicians mostly accessed free text notes rather structured data elements to form their decision (Figure 16). In the frequency of accessing different elements, 10 cases reviewed Physician-1 as pilot cases were not included. From the review process, we discovered that in the case where a drug was not appropriate for the patient, the problem linked to the medication was not removed. So in very few cases, structured data was used. At times, there were no notes associated with the event, so to confirm the date time, vitas were used. Other reviewers solely relied on the clinical notes only.

The physician began with a medication details window for the given medication order. From there, the most commonly used element was to look for notes on the date of the event. If such note was not available, they looked for notes in the previous encounter or in a future encounter. Most of the times, previous encounter notes were reviewed to establish compliance of medication and appropriateness of medication. Future notes were reviewed at times when the event occurred during a telephone encounter or audit encounter. Future notes were also used to ascertain that the medication was never repeated after the event. At times, notes existed in the form of scanned patient history, and scanned notes from lab and other clinics. The least accessed elements were the structured data such as lab results and problem list. During the review, diagnosis was established mainly from the clinical notes.

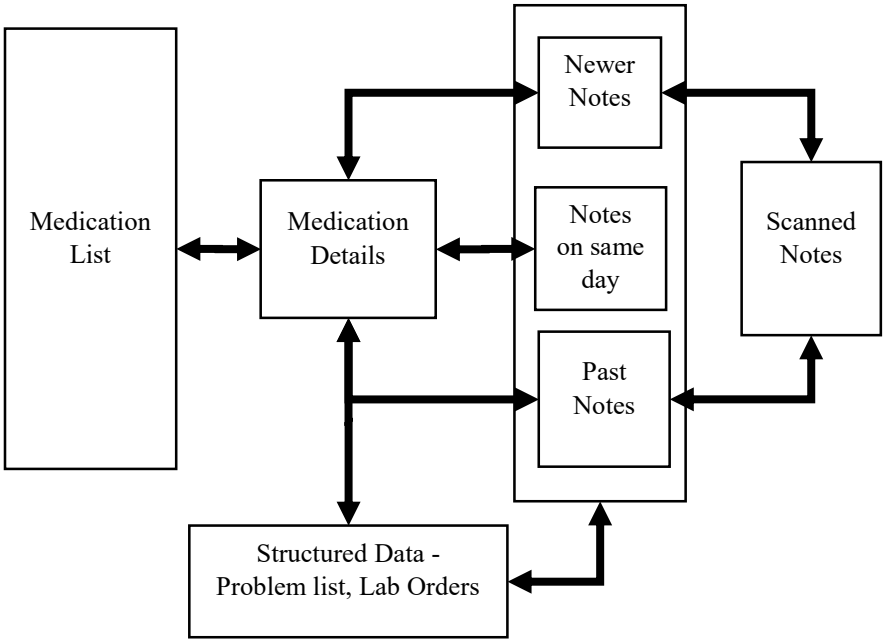



Figure 15 EHR elements accessed in the review process

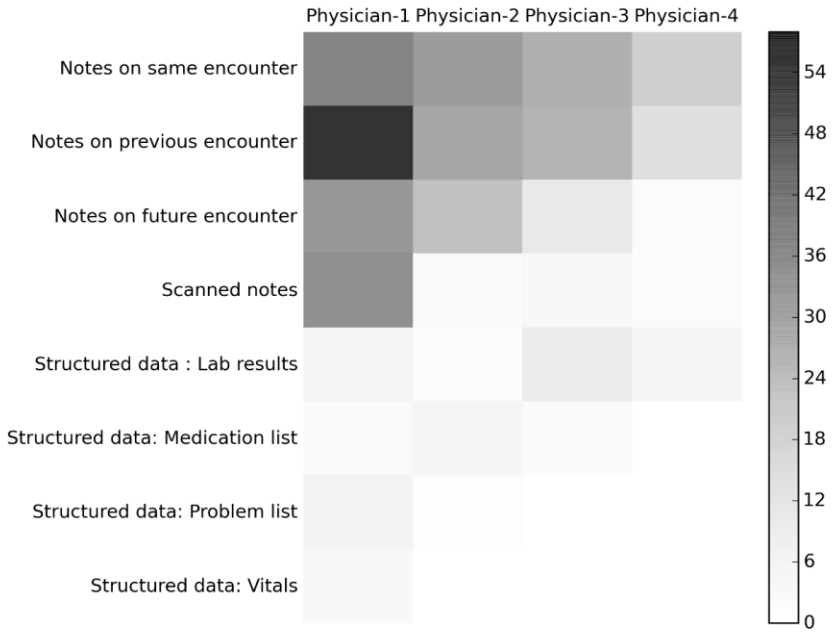



Figure 16 Heat map of count of elements accessed in the EHR.

The review pattern across categories is shown in Figure 17. Keeping in mind the number of cases in each category and that the heat map shows raw counts of frequency in access, there are no differences across the categories. Also, comparing the review of cases across physicians between the first and last five cases reviewed (Figure 18), there are differences.

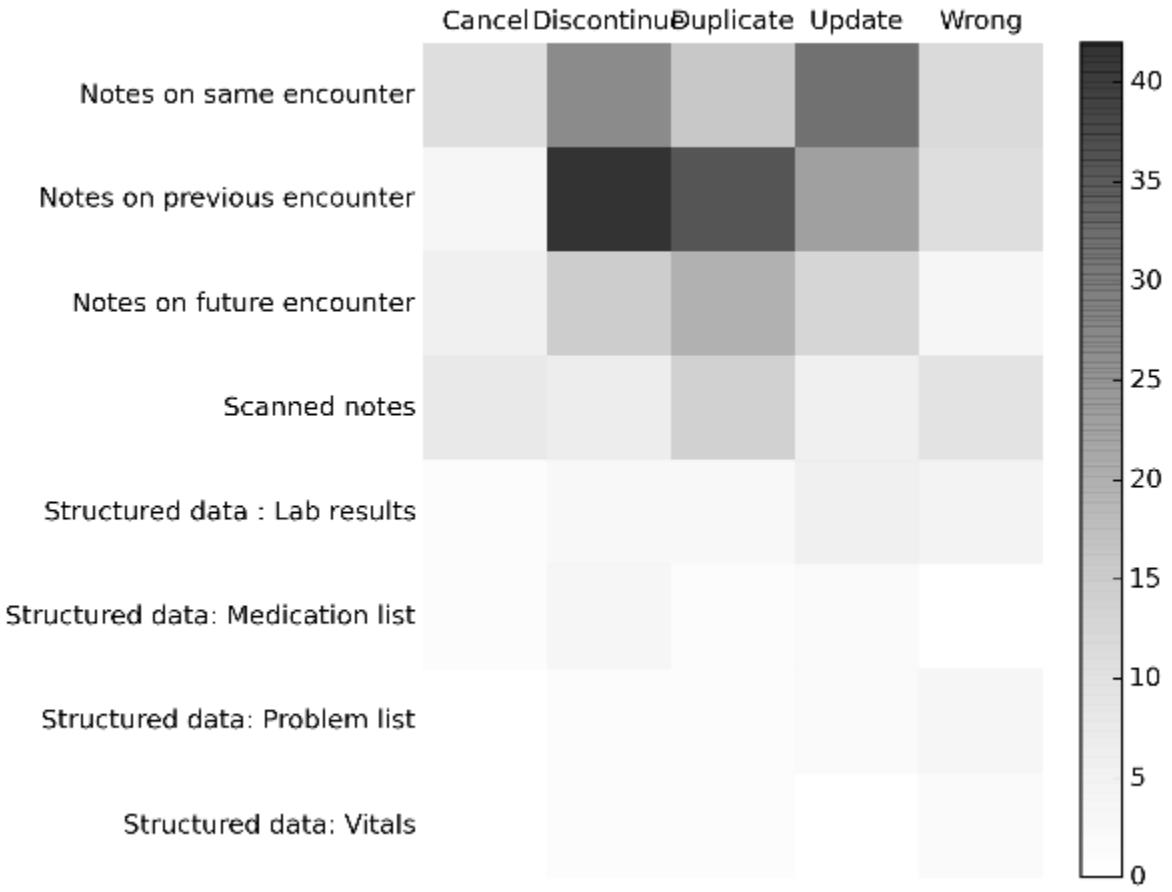
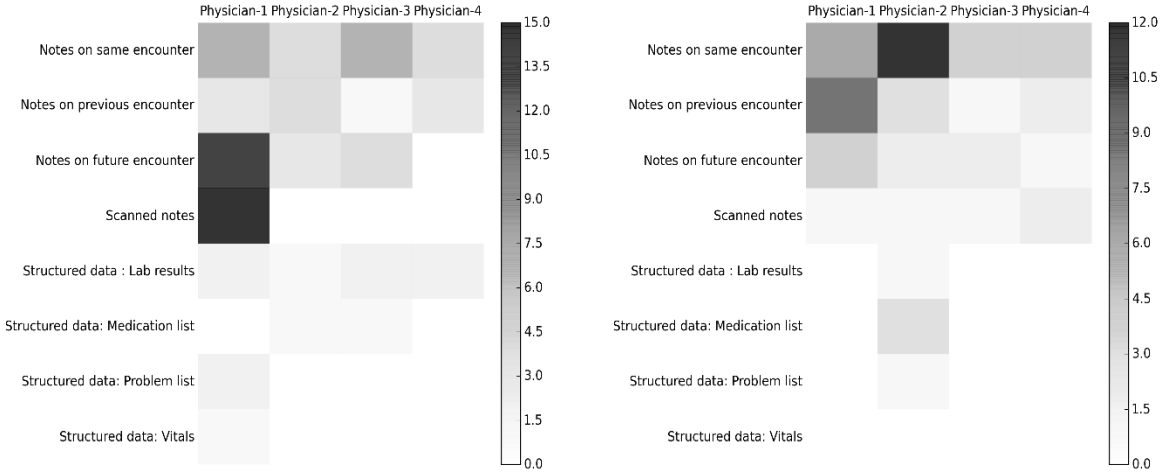



Figure 17 Summary of data elements accessed across categories




**a) Review pattern in first 5 cases** **b) Review pattern in last 5 cases**

Figure 18 Review pattern in the first and last five cases

## Discussion

We showed how expert review of medication revision event can be used to derive categories and descriptions of the event. There is some difference in review patterns across cases, and mostly unstructured data was used in review process. The categories noted were also described in the exploratory data analysis. Due to the small sample size and selection of cases based on one of the factors, a statistical model was not generated. The exploratory data analysis varied somewhat across the categories, but no sufficiently enough. This suggests the need for expert review to differentiate and uncover the factors contributing to medication revisions.

Categorizing events is very useful in situations where no standard definitions of medication revisions exist, and also where there is a need to generate a localized definition that will help in local implementation of process mining. When using external definitions and categories, we will be faced with missing information and ambiguous classification. In aviation or nuclear

industries, the devices, the procedures in use, and the maintenance of the devices are more standardized than the processes of healthcare. Thus in case of accident or failures within those other domains, standard definitions and categories are more readily available and are used for describing the errors (Shappel & Wiegmann, 2000). In healthcare, especially with the use of EHR, there is a difficulty in standardization. So to derive the factors contributing to an EHR event, it is essential to develop locally acceptable definitions and categories.

In this dataset, the medication revisions were classified into five categories: Cancel medication, Discontinue medication, Duplicate medication, Update medication and Wrong medication. When we consider the frequency of these categories within our actual datasets, we find that for the majority (55%) of instances, the medication revision function was used as a system workaround. System workarounds may be used to solve a system issue or to avoid additional effort. Physicians in our study reported using the EIE function to discontinue a medicine, rather than use the discontinuing medication function within this particular EHR system. EIE is preferred over other means as it gives visual feedback by a strikethrough of the text and instant removal of an order from current medications. What may not be known to physicians engaging in this workaround is that following the classification of an order as EIE, it is removed from the medication history. The implications of this removal include loss of medication history, no knowledge if a medication was ever tried by the patient, and not documentation as to why a medication was revised. To find this information, a physician would have to review the free text notes. This may create a gap in the patient history that increases risk and the potential for patient safety issues.

For the other 45 cases, the medical revisions represent potentially good catches of medication errors. These events occur when patient history is entered in the wrong chart, when the patient provided incorrect or incomplete information, when the physician orders medication without reviewing previous medications, or when the medication is from other clinics. Often, the duplicate medications could be attributed to physician reliance on free-text notes. There is very little documentation on why a specific medication is being updated.

Though the categories established here help differentiate good catches and system workarounds, there is in some cases ambiguity. A type of instance such as a duplication of a medication order may be revised in response to an error or a failure in the system. A physician may re-order a medication (a second time) due to an environmental factor, such as an interruption leading to a forgotten previous action (an error). Duplicate orders may also be generated with deliberate intention as a means of solving a system block, such as a problem with insurance, e-prescription transaction, or printing the prescription. Similarly, there may be ambiguity for wrong medication (difficulty resulting from confusion regarding wrong drug/wrong person or an incorrect decision by the provider).

The review process shows the reliability of unstructured data. This informs the difficulty we might face in automating the review process and scaling the process to larger data sets. Also the review process, though different across physicians, showed a common information access pattern, i.e. establish the time sequence of the event and correlating it with the clinical condition. The differences in the review patterns in the first and last five cases suggest there may be learning factors, an area that would be of interest for further exploration. They also point out that

reliability of human review process can vary. A synthesized review process across the experts would help in developing rules and steps for a machine algorithm to mine the process.

From this study, we found that some of the medication revisions may be due to mislabeling of EIE. This arises from system workarounds used by the physicians. This could be because the physicians may be experienced with the system, or they may have not been trained on when and how to use certain features. The majority of training efforts are focused on how to use basic functions and mostly on how to order a medication. The system workarounds point out important caveats in training, such as when to modify the status of medications, and how to update or change status of medication. Also, continuous monitoring of system use could have pointed out that certain functions are not being used correctly. Thus, this study adds to the usability and training value of EHR. Another implication for the mislabeling is that events in the database may not imply the intended activity. This must be taken into account, if a review for patient safety is conducted in the database.

**Conclusion:**

We described and categorized medication revisions in EHR. The cases reviewed also serves as a 'gold standard' of cases for the algorithm. The categories are used in the development of a rule-based classifier. The elements accessed are used to refine the rules. From the review process, we found that physicians mostly rely on free text notes to review the case. Accordingly, this directed our choice of classifier algorithm.

The review process brought out the factors contributing to the medication revisions, and these were used in development of the process model using fault tree analysis described in Chapter 6.

Knowledge about the type and description of revision events will help in detecting them in the system. The decision making process, including the elements accessed in EHR, used in deducing the probable reason for the medication revision, was used to model the algorithm for detection of the medication revisions (Chapter 7).

Limitations of these efforts:

This study primarily described the medication revisions and why they occurred in the system. In addition to gaining understanding about medication revisions, the study also brought out EHR system issues and workarounds, and how physicians review a case using the EHR. It should be kept in mind that this is an exploratory study with primarily a qualitative analysis of medication revisions. The goal was to show how EHR events can be described and classified. So, we evaluated only a selected subset of the total dataset. The developed categories were not tested for validity and reliability with different set of reviewers. The categories were generated from free-text transcription, and the majority of cases were reviewed by only one reviewer. The sample for inter-rater reliability was also limited. This being a retrospective review and a third-person review of events recorded resulted in very limited data for making decisions on probable cause. However, these limitations may not affect the overall goal of demonstrating how process mining must include domain knowledge by expert review.

Extension of these efforts:

Here we showed how categories for EHR event (process) can be described in a given system. Additional work can be directed to generalization of categories across other EHR systems. This will help in identifying differences in EHR use and standardizing the interactions. Such

categories from multiple systems can be later translated to a comprehensive taxonomy in the future. Future direction would be to validate the categories.

## Chapter 6: Analyzing Medication Revisions Using Fault Tree Analysis

Having described the medication revision events, we next generated a process model of the events. The goal was to identify patterns and co-existing factors that could potentially lead to that event. For this analysis, we used fault tree analysis (FTA). FTA is a systems approach used in industries to analyze failure events. We show how FTA can be used for EHR data. We also discuss why this method is suited for EHR data and how this method could be integrated with the EHR. The results of the analysis can be used to improve EHR use.

### Introduction:

Process models are constructed using visual models and in this study, we sought to improve understanding by combining models (Górski, Magott, & Wardziński, 1995; Reza, Pimple, Krishna, & Hildle, 2009). We used the Petri net technique, which is a visual modelling method where each node is mapped in a directed graph (Murata, 1989). We sought to emphasize the interaction between nodes by combining the Petri nets with fault threes. We began with the construction of fault trees, and then synthesize the fault trees into a process model. We focused only on fault trees, as Petri nets are only used as a visualization and not as an analytics medium here.

A report from the Institute of Medicine recommends using systems methods, such as hazard analysis and risk management analysis as strategies for reducing medication errors (Bootman, Wolcott, Aspden, & Cronenwett, 2006). Systems methods for analysis of error events have been developed to ensure highly reliable systems, determine the probability of failures, and determine the causes and sequences leading to failure. This factors reduce errors. They have been used in

many industries including manufacturing, aviation and aerospace, (Li & Harris, 2005; Wiegmann & Shappell, 2001), energy (Singh & Kim, 1988; Yuhua & Datao, 2005) and medicine (Dhillon, 2003; Lyons, Adams, Woloshynowych, & Vincent, 2004).

System analysis can be broadly classified into inductive methods and deductive methods (Vesely, Goldberg, Roberts, & Haasl, 1981). Inductive methods are aimed at confirming the causal effect of a condition. It determines the effect of an individual condition on the overall system. For such analysis, a complete knowledge of all conditions that affect the system must be identified and investigated. Also the effect of each condition must be measurable. Examples include Failure Mode and Effect Analysis (Chiozza & Ponzetti, 2009), and Root Cause Analysis (Friedman et al., 2007).

Deductive methods are aimed at determining what modes/components have contributed to the event. From the available deducted conditions, the system is modeled in a systematic way. No attempt is made to measure or confirm the causal strength of the events. Examples include FTA.

**Fault Tree Analysis:**

FTA is an engineered component-based view which attempts to identify the series and parallel combination of components that lead to an eventual event. This is a systematic approach for understanding a system (W. S. Lee, Grosh, Tillman, & Lie, 1985).

FTA is a deductive process. It is important to note that FTA is not a model of the system, nor does it include all possibilities that could lead to failure.

As FTA is component based analysis, we examined the components of an EHR. EHRs can be considered to be composed of multiple modules. Each module has an input and an output. As

errors can occur on the side of the inputs to module, or in the computation of the inputs. This component based view helps us in defining the combination of components and events that lead to an eventual event. The components can be human action, software action/output, or combinations of both.

FTA was originally developed in mission-critical situations to model system failures in a systematic, traceable fashion. It has been a widely accepted and a proven technique. It is also mandated in the regulations of the Federal Aviation Administration (FAA) (1998) and in the Nuclear Regulatory Commission (NRC) (Vesely et al., 1981) for investigating and analyzing error events. In medicine, it has been applied to study medication order process (Cherian, 1994) and assessment of patient safety risk (Marx & Slonim, 2003).

Some advantages of FTA include:

1. It is a visual model
2. It is based on Boolean algebra, making it easy to automate calculation
3. It can also be used as probability model

From the error analysis, we sought to obtain factors, sequences, and combinations of events contributing to the error event. We only require the presence or absence of factors contributing to the error, and not the strength of association. The objective of the analysis is to identify strategies and tools for error prevention and mitigation as the next step of error management.

In this dissertation, we apply fault tree analysis to EHR events. We discuss how the system is defined, and we identify event nodes. We included factors on the human side (i.e. cognitive

process), the machine side (i.e. modules, functions) and those related to the human-machine interaction. Though we included many factors, it is impossible to precisely identify all possible components and conditions. Furthermore, it is difficult to perform retrospective investigations because very limited data is available with the user activity log and system logs being poor. Thus inductive methods, that require complete knowledge about the system, are not be feasible in EHR data. This lead us to choose FTA.

The fault trees were synthesized into a Petri net process model. From the FTA and process model, we visualized and identified the factors contributing to medication revisions and their relationships.

**Methods:**

Three basic steps are involved in FTA:

1) Describing the system
2) Constructing the Fault Trees
3) Analysis of the Fault Trees.

Describing the system: The first step is defining the undesirable event - the failed state of the system. In our study, the failed state of the system is: marking a medication as "Entered in Error" (EIE). The next step is describing how the various components interact in the system. In our case, the system description is a medication order process - how a medication item is entered into the EHR system. For this, we used physicians' descriptions of the workflow of how medication orders are made in the system. For each case, the physicians described the possible events or

postulated possible conditions that could lead to EIE. From these descriptions, the different stages in medication ordering process were generated. Note that they are list of stages at which medication data is entered and not a sequence of steps or a workflow.

1. Patient fills history sheet - scanned and saved in EHR
2. Provider takes medication history and enters in chart
3. Physician verifies medication history - changes, updates as necessary
    - Physician converts history to prescription, discontinues if drug not required.
    - Physicians orders new medication
4. Physician orders new medications or updates existing prescriptions
    - Uses dropdown pre-loaded medication order sets.
    - Creates new medication order
5. Reconciles medication list - updates status, removes duplicate
6. Physician signs off and sends the prescription to the pharmacy

Constructing the Fault Tree: Fault trees are constructed with a top-down approach. Starting from the terminal event, event nodes are constructed. The view here is that complex systems consist of multiple components, each of which could fail in multiple ways. The path(s) to a specific failure or type of failure can be modeled as a series of interactions modeled in a calculus of logical operations such as AND, OR, and XOR. In our analysis, we defined a “Basic Event” as a basic initiating fault. “Intermediate Event” were events that occurred because of one or more basic

events. The “Basic Event” and “Intermediate Event” were connected by “Gates”. In our analysis, we used two gates: the OR gate and the AND gate.

Fault trees were generated using the symbols shown in Figure 19. The symbols gave a visual overview of the events.

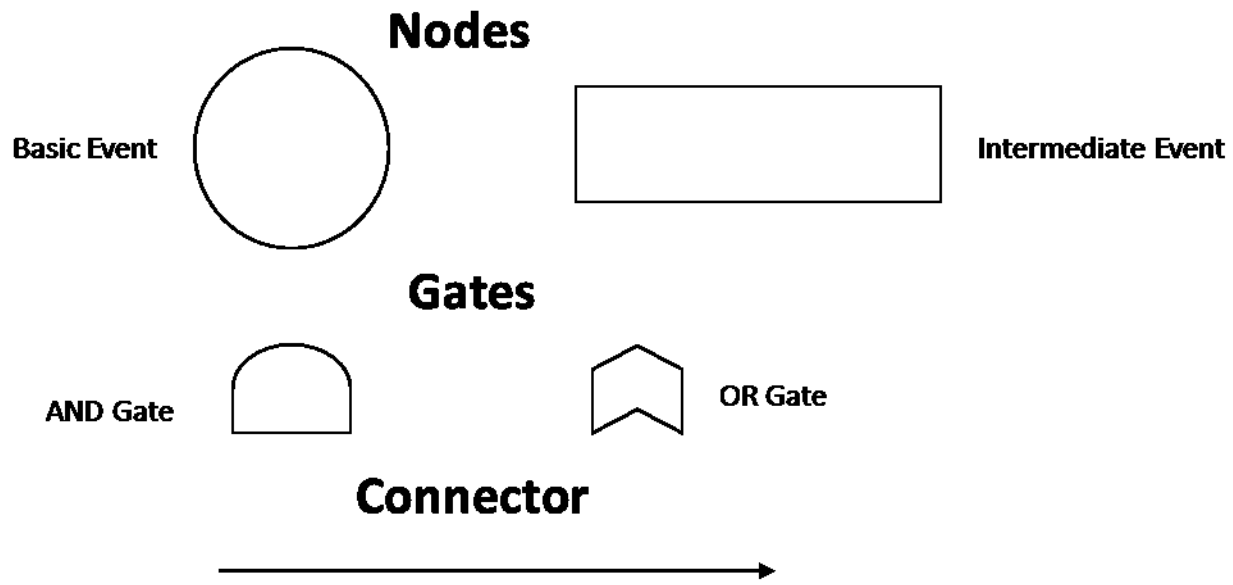



Figure 19 Components of fault tree

Starting from the failure event, EIE in our case, we built a top-down tree of events that could have led to the final events. Basic events were obtained from physician review. Basic events were grouped into Intermediate events. The intermediate events were subjectively defined.

Analysis of Fault Tree: FTA is primarily a qualitative analysis. From FTA, we converted the fault tree to Boolean equations using calculations of the following:

1. Minimal cut sets: the smallest combination of component failures that will cause the top event to occur.

2. Order of cut sets: the number of component failures occurring within a minimal cut set. The smaller the order of cut set, the more important it is in causing the event.

If the probability of events is known, such as failure rate or prevalence, a quantitative calculation can also be performed. But in our case, we do not know the probabilities, and so only qualitative analysis was done.

Development of process model: Fault trees were combined to generate a Petri net like model of the process. Here, all the basic nodes are converted to Petri net events. Petri net states were not coded because of limited data availability.

**Results:**

From the 100 cases reviewed by the physicians, fault trees for error categories were constructed. All of the events in our analysis had an 'OR' relationship. Thus, in the Boolean analysis, the minimal cut set was 1 for all categories.

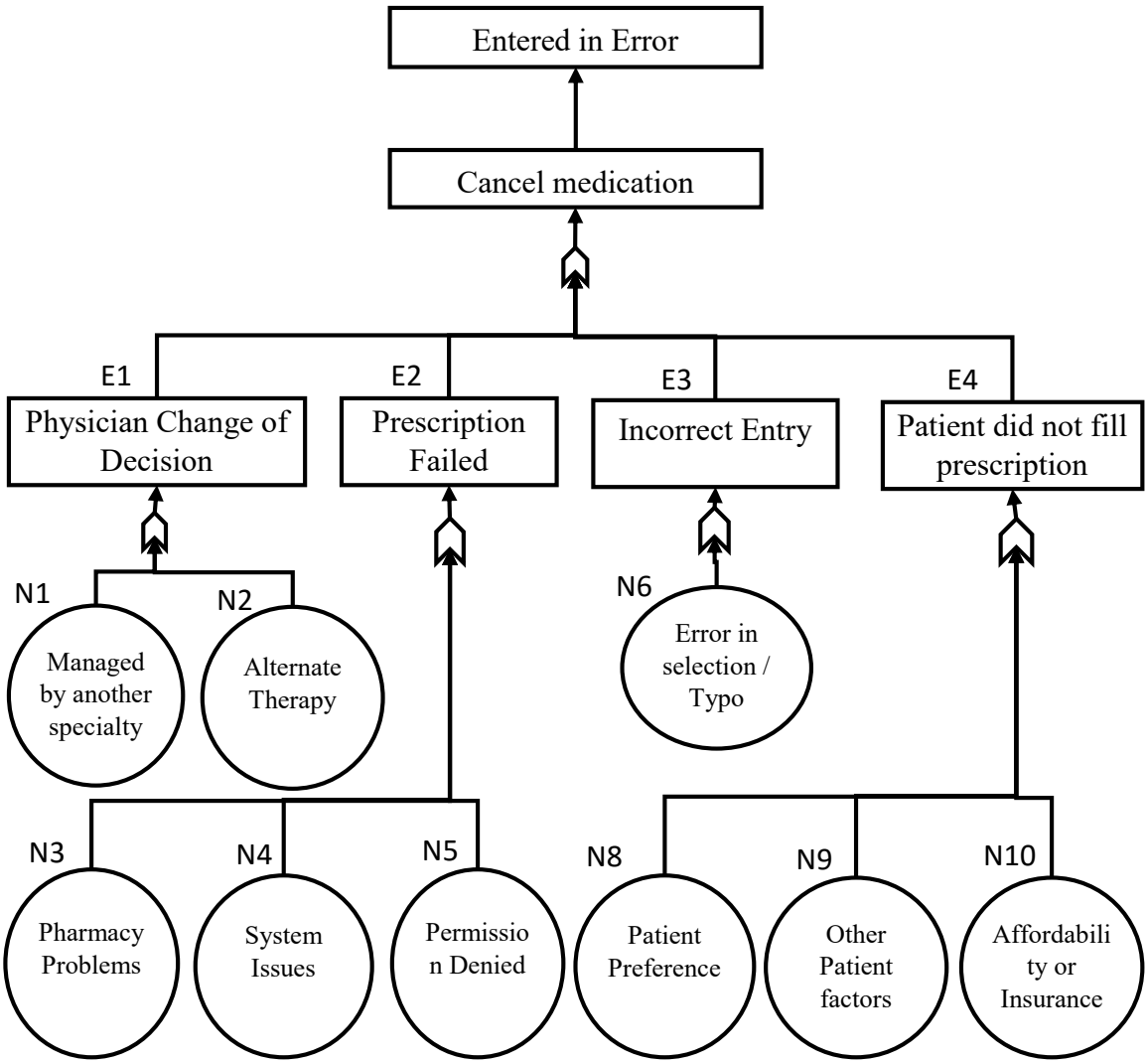
Cancel medication (Figure 20): In the FTA, the factors leading to cancel medication as the reason for the EIE event were patient factors, incorrect entry, problem with prescription and physician decisions.

Discontinue medication (Figure 21): In the FTA, the factors leading to discontinue medication as the reason for the EIE event were patient factors, problem with prescription and physician decisions.

Duplicate medication (Figure 22): In the FTA, the factors leading to discontinue medication as the reason for the EIE event were prescription and incorrect entry.

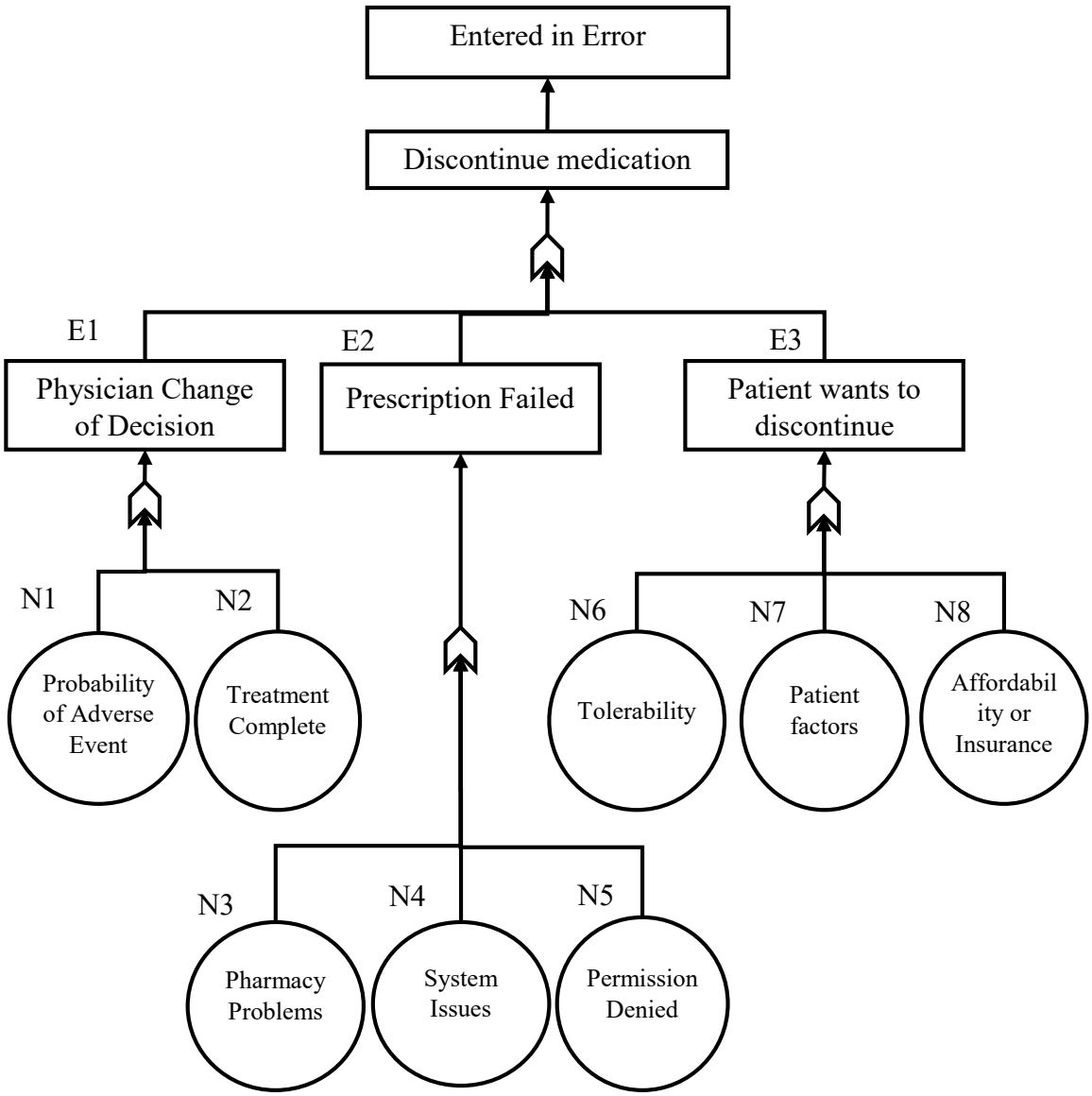
Update medication (Figure 23): In the FTA, the factors leading to discontinue medication as the reason for the EIE event were patient factors, problem with prescription and physician decisions. This was similar to discontinue medications.

Wrong medication (Figure 24): In the FTA, the factors leading to discontinue medication as the reason for the EIE event were incorrect entry and incorrect information.




Top level event
= Entered in Error
= Cancel medication
= E1 OR E2 OR E3 OR E4
= (N1 OR N2) OR (N3 OR N4 OR N5) OR (N6) OR (N7 OR N8 OR N9)
= N1 OR N2 OR N3 OR N4 OR N5 OR N6 OR N7 OR N8 OR N9

Figure 20 Fault tree and Boolean equation for cancel medication




Top level event
= Entered in Error
= Discontinue medication
= E1 OR E2 OR E3
= (N1 OR N2) OR (N3 OR N4 OR N5) OR (N6 OR N7 OR N8)
= N1 OR N2 OR N3 OR N4 OR N5 OR N6 OR N7 OR N8

Figure 21 Fault tree and Boolean equation for discontinue medication

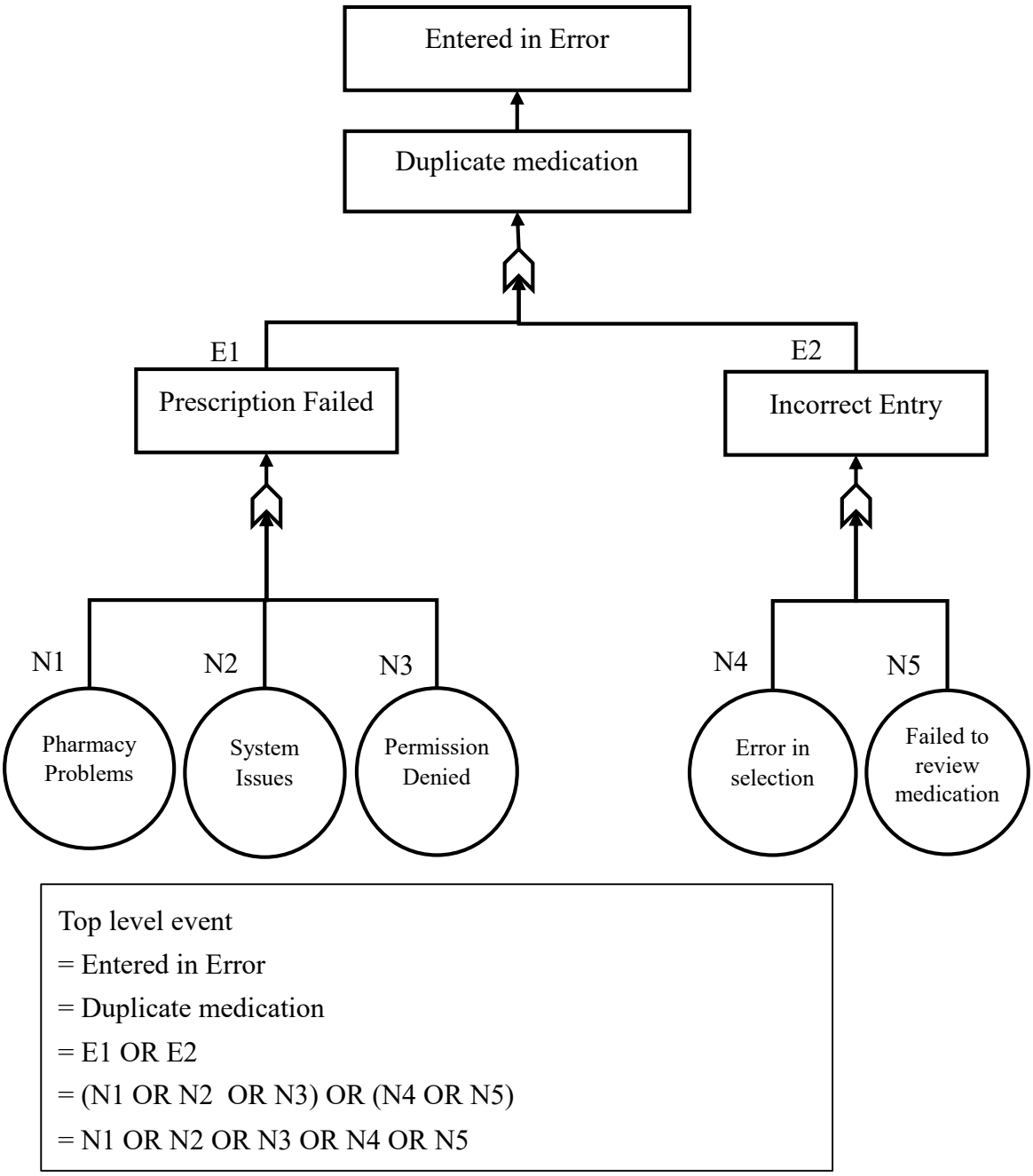
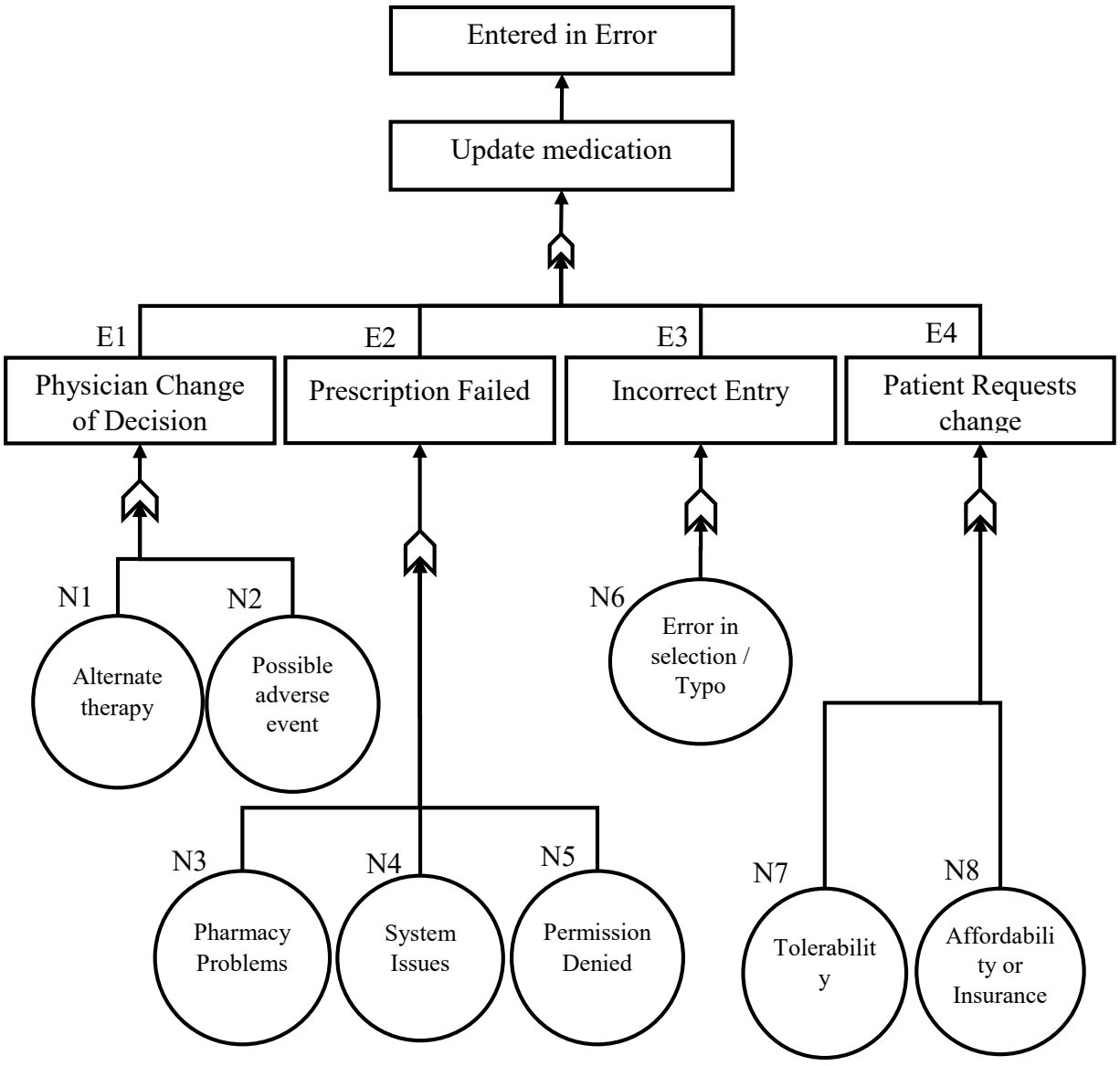
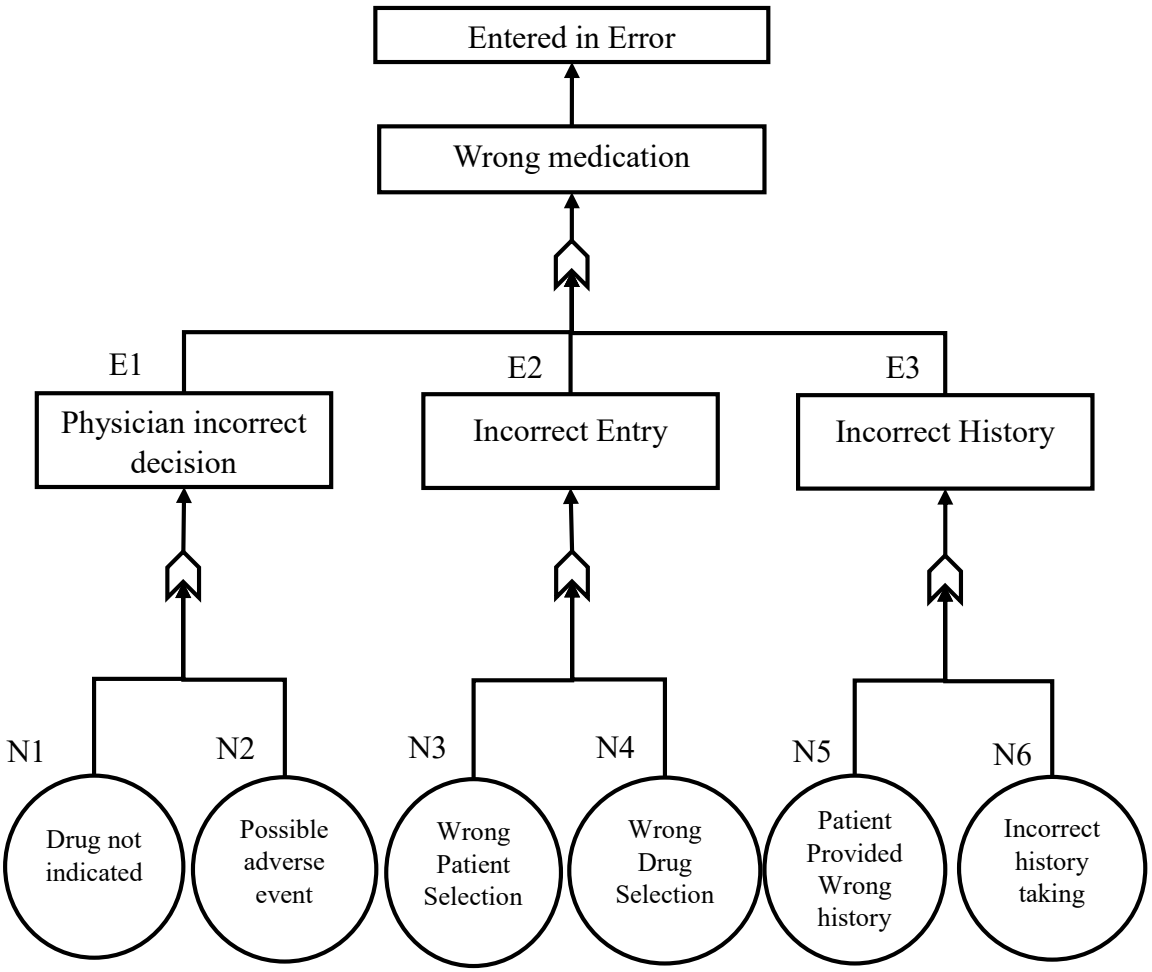



Figure 22 Fault tree and Boolean equation for duplicate medication




Top level event
= Entered in Error
= Update medication
= E1 OR E2 OR E3 OR E4
= (N1 OR N2) OR (N3 OR N4 OR N5) OR (N6) OR (N7 OR N8)
= N1 OR N2 OR N3 OR N4 OR N5 OR N6 OR N7 OR N8

Figure 23 Fault tree and Boolean equation for update medication




Top level event
= Entered in Error
= Wrong medication
= E1 OR E2 OR E3
= (N1 OR N2) OR (N3 OR N4 ) OR (N5 OR N6)
= N1 OR N2 OR N3 OR N4 OR N5 OR N6

Figure 24 Fault tree and Boolean equation for wrong medication

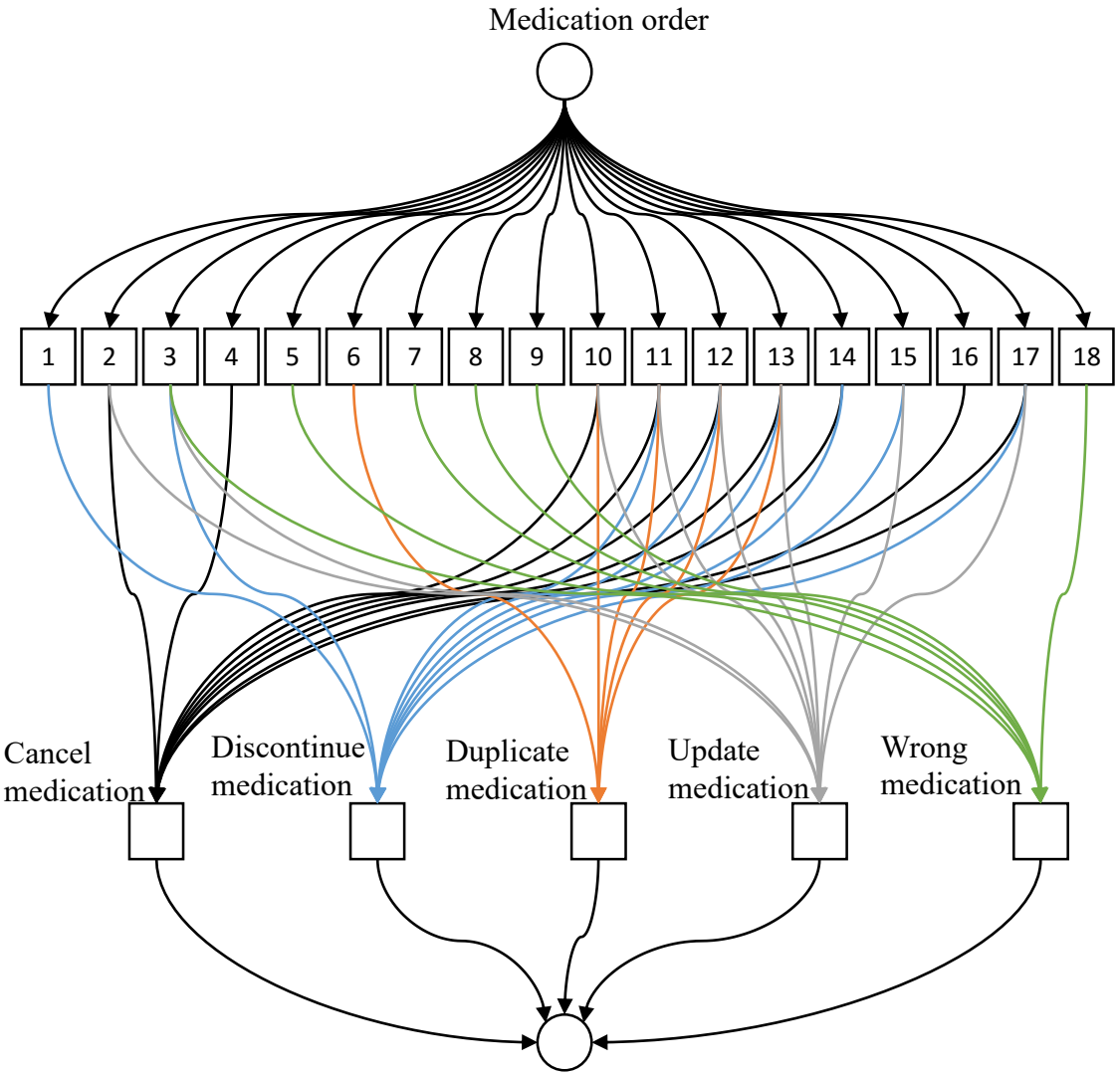
Minimal cut sets for all the categories was 1. Thus, the order of cut sets in each case was the total number of basic events.

The compiled list of basic events across categories is shown in Table 10. The most common factors for EIE events were system use issues - problems in prescription error in data entry, and error in system use.

Table 10 *Factors responsible for medication revisions*

| **Clinical Factor** | **Provider Factor** | **System Factor** | **Patient Factor** |
|---|---|---|---|
| Treatment complete | Drug not indicated | Wrong patient selection | Patient Factors |
| Alternate therapy | Failed to review medication list | Wrong drug selection | Tolerability |
| Possible adverse event | Incorrect history taking | Error in selection or typo | Patient Preference |
| Managed by another specialty | | System Issues | Affordability or Insurance |
| | | Permission Denied | Patient provided wrong history |
| | | Pharmacy Problem | |

Based on the fault tree, a Petri net like process model of medication revisions was developed (Figure 25).




**Clinical**

1. Treatment complete
2. Alternate therapy
3. Possible adverse event
4. Managed by another specialty

**Provider**

5. Drug not indicated
6. Failed to review medication list
7. Incorrect history taking

**System**

8. Wrong patient selection
9. Wrong drug selection
10. Error in selection or Typo
11. System Issues
12. Permission Denied
13. Pharmacy Problem

**Patient**

14. Patient Factors
15. Tolerability
16. Patient Preference
17. Affordability or Insurance
18. Patient provided wrong history

Figure 25 Process Model of Medication Revisions in EHR

**Discussion:**

We described a method for constructing fault trees for EHR events. Fault trees were constructed for the 5 categories identified in the dataset. Synthesizing the 5 fault trees, we developed a Petri net like model of the medication revision process. From this, we can see that factors common across the categories were system and workflow issues. The process model shows only the reasons but not the resulting states of the events. This is because of limited data available at retrospective review, and the deductive nature of the initial manual review. As a result, very few factors were identified. As the fault trees were constructed on a human review of limited data set, the basic and intermediate events are subjective and require validation.

From the process model, we find that system issues are common across the categories. These issues may include problems such as issues with connectivity, rejection by the pharmacy system with no reason mentioned, and the inability to print prescriptions. In these cases, the user tends to use the EIE function to rectify the orders.

The next most common cause is adverse event related. In our study, the physicians reviewing the cases were only able to postulate that adverse events could have been a reason for making the revisions. The type and nature of adverse event, or other details about possible adverse events were not provided by the physicians. This could be because of the difficulty in identifying a documented adverse event. Also as the revised medications may not have been prescribed, the actual occurrence of adverse event may be absent.

It has been reported that more than 40% of adverse event occur due to preventable medication errors (Bates et al., 1995). Due to the complexity of detection of adverse event in EHR, this

domain is still evolving (Haerian et al., 2012; Trifirò et al., 2009). A future direction would be exploring automated detection of adverse events and including them in the fault trees.

The other system issues described, such as typo and selection issues, have also been found in other studies (Khajouei & Jaspers, 2008; Walsh et al., 2006). Those studies have obtained their data from observation and interviews. These factors require further exploration, for they inform us on where the problem lies, such as in system usability or in user training. For a comprehensive analysis of these system factors, logging of system use is essential. The log data must go beyond the timestamps, and include the information was provided to the user in the front-end, such as the drop down menus, search terms, etc.

The patient factors which included patient non-compliance, patient demanding change in therapy, and patient not satisfied with therapy have also been shown to be associated with medication errors (Hulka, Cassel, Kupper, & Burdette, 1976). These factors can also be improved if medication compliance data is included. Such data collection tools are upcoming.

The minimal cut sets used here were found to be similar to those of another study on the construction fault trees for medication orders (Cherian, 1994). The similarity between the studies is that they were generated based on human review, thus limiting the objectivity and the exploration of factors. To improve this, the factors can be further broken down if they are combined with a more structured task analysis and workflow analysis (Doytchev & Szwillus, 2009). For this, a standard task analysis will need to be generated from the system perspective and the workflow perspective. Also, additional data logging will improve the depth of the fault tree generated.

In engineering domains, FTA has been automated (Dugan, Bavuso, & Boyd, 1992; W. S. Lee et al., 1985). In medicine, there is no standard definition of the input and output of tasks. However, in the case of EHR data, it is possible to define the inputs and outputs for each module and track the status. This can then be used in automated generation of FTA. When FTAs are synthesized for larger data and over time, a database consisting of basic events and probabilities could be generated. This would be useful for obtaining quantitative results, similar to human error databases in aviation and nuclear power industries (International Atomic Energy Agency, 1998; Kirwan, Gibson, & Hickling, 2008).

**Conclusion:**

We constructed fault trees for the medication revisions reviewed. The FTA provided the contributing factors leading to the error event. The fault trees were later synthetized into a single Petri net like model of the medication revision process. From the results of the FTA, strategies to improve the EHR and workflow can be determined. The FTA done here is only a prototype to illustrate how they can be done for EHR events. Future work is required to standardize and validate the fault trees.

## Chapter 7: Application of Processing Mining: Detecting Medication Revisions

Process mining can be applied to check for conformance, i.e. detect deviations. Learning from the expert review of medications revisions, we developed an algorithm to detect the events. The algorithm was trained on the 100 cases that were manually reviewed. The algorithm will predict if a given order would be labelled as 'Entered in Error' (EIE). To evaluate the algorithm, a dataset without any EIE medication order labels was used. The sensitivity and specificity of algorithm measured.

### Introduction

As the volume of electronic patient data is increasing, the reliance on automated methods becomes inevitable. In the industries of banking and IT security, automated processes are available for detection of anomalies and outliers (Hauskrecht et al., 2010; Hodge & Austin, 2004), to perform such tasks as fraud detection and security intrusion detection (W. Lee & Stolfo, 1998; Mukkamala, Janoski, & Sung, 2002; Phua, Lee, Smith, & Gayler, 2010). These methods show how analyzing system use data can be used for detecting deviations - either from final outcome (in case of fraud detection) or from usage pattern (in case of security intrusion).

Here, our objective is to detect medication revisions. This can be considered as a classification problem - medication revisions classified as yes/no. For this, we can either use a rule-based classifier or use machine learning classifiers, such as decision trees, probabilistic classifiers (Bayes, logistic regression), and support vector machines. In the context of big data, classifier algorithms have been proved useful in marketing (Kim, Kim, & Lee, 2003; Schafer, Konstan, & Riedl, 2001), and finance (Huang, Nakamori, & Wang, 2005; Olmeda & Fernández, 1997).

Some algorithms have also been used in EHR data for predicting patient prognosis, analyzing clinical practice patterns, providing business support and providing clinical decision support (Balas et al., 1994; Brown et al., 2005; Kohli et al., 2001; O'Reilly, Talsma, VanRiper, Kheterpal, & Burney, 2006; Pochet & Suykens, 2006). Longitudinal EHR data with presence of structured and unstructured data require modifications to these existing algorithms when applied to EHR (Wu, Roy, & Stewart, 2010).

In one study, machine learning was used to detect inappropriate medications (Hauskrecht et al., 2013). They analyzed medications from ICU patients and developed an algorithm that would predict if the appropriate medication was missing from the patient orders. This was based on the statistical (probabilistic) model built for the specific ICU and on prescription practices in that unit. Another major domain for using machine learning is in detecting adverse events from EHR data (Bates et al., 2003; Melton & Hripcsak, 2005). In these studies, there is a reliance on mining for information on outcome of tasks, such as adverse event, discovery of adverse events. These studies show that medication orders and EHR data can be modeled for various purposes.

In our study, we began with a rule-based classifier. Rule-based classifier can be developed from expert review and can serve as a descriptive model. It is also best suited for working with a limited data set for training and for unfamiliar data sets (Qin, Xia, Prabhakar, & Tu, 2009). It is easy to generate and performance is comparable to machine learning methods like decision trees (Entezari-Maleki, Rezaei, & Minaei-Bidgoli, 2009). We did not use machine learning methods because the training data set was limited, some of the data required was in free-text notes, and the relationship of event with other elements of EHR was unknown.

**Methods**

We developed an algorithm to classify if a medication order would be labeled as EIE.

Defining the problem: The problem is a 2-classfier problem. Medication orders with an EIE label will be assigned a positive value and used for learning rules. The negative cases will be the default.

Development of the classifier: We developed a direct rule-based classifier. First, rules were learned from the 100 medication orders manually reviewed. The rules were manually defined. To prune the rules, we used all the medication orders from these 100 patients. Together this formed the training set for the classifier algorithm (Table 1).

Testing and evaluation: The testing set consisted of medication orders from the remaining 904 patients in the sample data. All instances of EIE were removed from the testing data. The algorithm was applied to the testing set and was evaluated by measuring the sensitivity, specificity and accuracy of classification.

**Results**

A rule-based classifier algorithm was developed. The overall flow of the algorithm is show in Figure 26 and the pseudocode is illustrated in Figure 27.

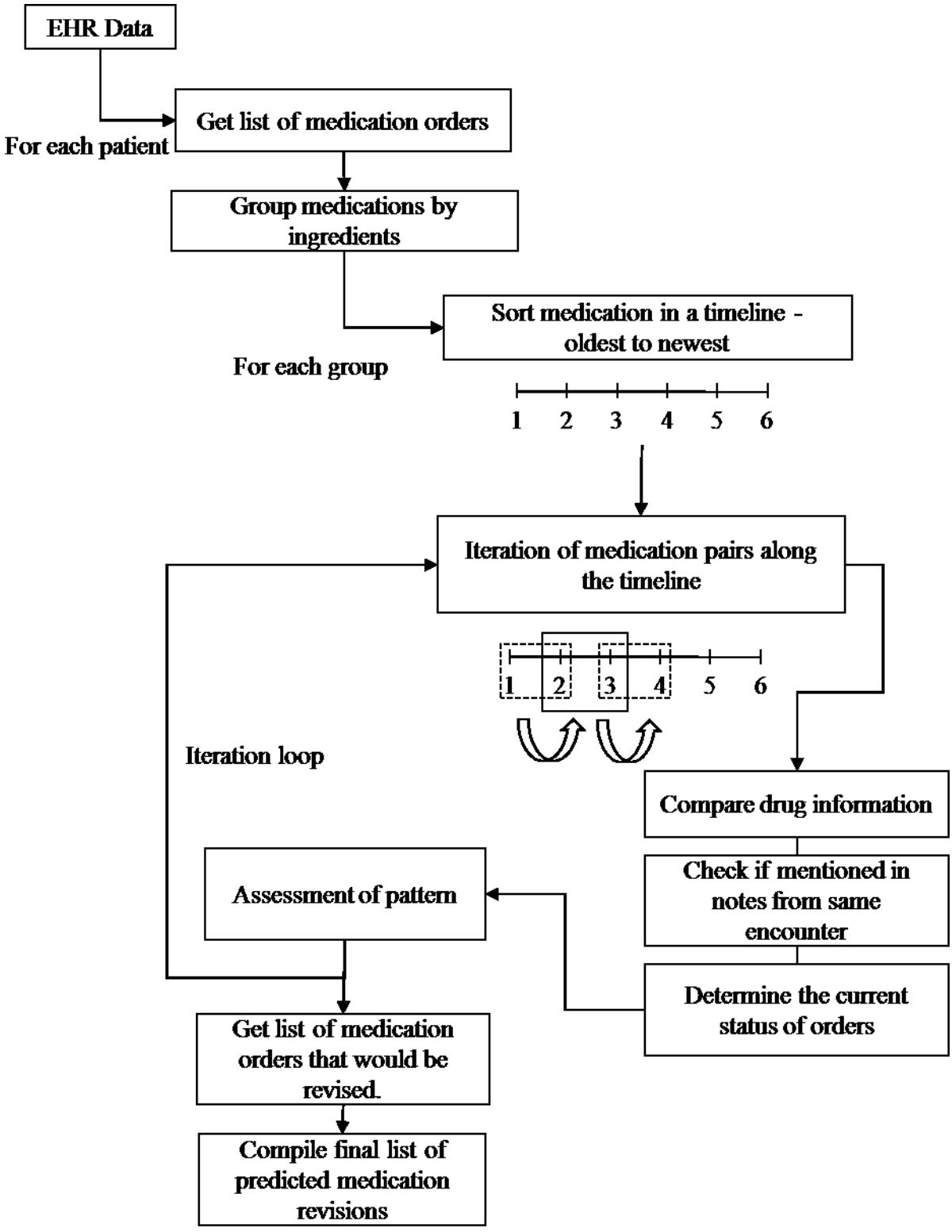



Figure 26 Overview of algorithm for detection of medication revisions

```
if number of orders in group = 1
{
        if difference in date > 0
                {
                        if final status is a incomplete status:
                                then error = discontinue medication / cancel medication
                }
        else
                {
                        check if drug mentioned in notes on same encounter:
                                if no, then error = wrong medication
                }
}
else
{
        for each pair of orders
                {
                        check if medication information is different
                        check if type of order is different
                        check if status of order is different
                        check which order is current

                        if either medication status is incomplete
                                {
                                        if difference in date > 0
                                                {
                                                        if change in medication information:
                                                                error = update medication
                                                        if order information is same:
                                                                error = duplicate medication
                                                }
                                        else
                                                {
                                                        if order information is same:
                                                                error = duplicate medication
                                                        else:
                                                                error = update medication
                                                }
                                }
                }
}
```

Figure 27 Pseudocode of rule-based algorithm

The performance of the algorithm in the test set is show in Table 11. The algorithm has a 65.72% sensitivity, 85.49% specificity and 85.27% accuracy when tested in a dataset with prevalence of 0.011. The Positive Predictive Value (PPV) is 4.8%

Table 11 *Performance of classifier algorithm*

| | Case Positive | Case Negative | |
|---|---|---|---|
| Algorithm Positive | 1963 | 38402 | 40365 |
| Algorithm Negative | 1024 | 226329 | 227353 |
| | 2987 | 264731 | 267718 |

**Discussion**

A rule-based classifier algorithm was developed to detect medication revisions in the dataset. The algorithm has moderate sensitivity, but high specificity. The sensitivity is affected by the presence of workarounds and inappropriate use of EIE in the datasets. The algorithm shows that the medication revisions can be detected based on the process mining. The algorithm is a prediction algorithm and thus can be used to uncover previously unknown events, and also to notify or warn of an impending event. The algorithm is a standard rule-based algorithm. The

rules were manually learnt from a limited set of data. For generalization, more samples might be required. We only tested the prediction ability and not the classification ability of the algorithm.

The classifier's moderate sensitivity can be attributed to presence of workarounds in the data. From the manual review, we found that 55% of the revisions are workarounds/inappropriate use. Some of the workarounds, such as detecting medication updates were incorporated in the rules. However, workarounds on discontinue medications were difficult to incorporate. To balance the specificity, the rules to detect these workarounds were not included. The higher specificity suggests that the algorithm is more of a confirmatory test rather than a screening test. This is particularly important when implementing in real-time systems - balancing the cost of missing an event and the cost of false positive events.

In this study, we faced with problem of noise in the clinical data set. In this system, the patient problem list in the EHR was not always current and reliable. It was also noted that when a wrong medication is added to record and later removed using EIE, the associated problem still remains in the patient record. This made it difficult to assess the appropriateness of a medication based on the patient problem list. Thus, we used the presence of medication order in physician signed note as an indicator for an appropriate drug. This idea was conceived from the physician review process. This rule was satisfied in the majority of the wrong drug cases reviewed. However, in real implementation, the system should be detecting the wrong drugs independent of notes. For this, methods are required to improve the clinical knowledge from the data to find if given drug is indicated for patient.

In the algorithm, the question posed was "will this medication order be labeled as EIE?". But in the human review process, the question posed was "why is the given medication order labeled as

EIE?". The difference in the two processes is that the latter is an investigative process, and former is a recognition of patterns. The algorithm's performance would improve if it were to reason why a medication was labeled as EIE. In this case, the effect of workarounds will be minimal. This will be useful when expanding the algorithm for categorizing the errors. Another difference in the review process is that the algorithm relied on structured data, whereas the clinicians relied on free text notes. Based on the algorithm's performance, we can infer that the structured data has adequate information to detect the errors, but may not be available in a usable format in the front-end for the physicians.

The algorithm is a classifier and can only predict if a revision event has occurred or not. For future direction, given a set of definitions to categorize (from expert review process), the algorithm should be able to categorize the detected events. For this, additional gold standard data on the categories is required to develop and validate the categorization. Additional manual review process for evaluating the test results is also required. Categorizing the error events will also help in analyzing them.

The detection algorithm has a low PPV of 4.8%. This must be interpreted with the 1.1% prevalence of revision events in the dataset. The low PPV is probably due to unknown revision events in the false positives. Also, rules for update medication and discontinue medication may be identifying more revisions, as these actions are more prevalent in clinical processes. These will affect the implementation in real-time systems, for the algorithm may invoke more alerts and produce alert fatigue. To improve and validate the PPV, review of false positive and gold standard data is required.

The algorithm can be automated for rule extraction. For this, based on the manually defined rules, a set of factors can be defined and machine learning methods can be used to uncover rules and patterns. This can be further evolved to fully automate an unsupervised learning algorithm. For best performance of such automated algorithms, the data quality must be improved. Full use of structured data in documenting medication orders, additional logging of user access and logging of user activity in the EHR will be beneficial. In addition to pattern based detection, knowledge based detection will also improve the accuracy.

**Conclusion:**

As an application of process mining, we developed and evaluated a rule-based classifier algorithm for detecting medication revision events. The algorithm can predict medication revision events by learning from patterns. The algorithm can be used to monitor user activity and provide user warning on potential revisions and detect previously undetected revision events. Future work is needed to extend the algorithm for automated rule extraction and unsupervised learning.

## Chapter 8: Process Mining of Medication Revisions: Discussion

In this thesis, we described process mining of medication revision events in the EHR. Process mining is the generation of structured process definitions from real time executions. Process mining has been applied at the task level or outcome level to monitor performance and assess workflows. However, with EHRs, the tasks contain steps within, each of which can be considered as process. For example, a task of ordering medication in EHR, contains process within, such as selecting a medication, specifying medication details, signing off medication, updating status of medications, etc. Here, we study medication revision as a process. This is an initial effort of applying process mining on EHR data. By applying process mining to the medication revisions, we seek to understand how and what causes the revisions. We generated a process model. In the model, we identified clinical factors, user factors, system factors and patient factors.

We first began by exploring the medication revision data using data mining. We learned that medication revision events are rare events (prevalence 1.7%), but affect 5.9% of patients. From the exploratory data analysis, we learned that the majority of the revisions are not made on the same day (59%) but instead are made on medications that have been on the record for more than a month (62%). Additionally, most revisions were made by physicians other than the original prescriber (55%).

Next, we conducted expert review on selected cases (n=100). From the review, we defined 5 categories: cancel medication, discontinue medication, duplicate medication, update medication and wrong medication, as reasons for medication revisions. Also from the review, we learned

that 55% of the cases are workarounds. This is a system factor impacting performance and impeding the intended use of EIE as a recognition of an error. Based on the review, fault tree analysis was done to model the process of medication revisions. The most common factors across all the categories were system related factors, including failure of the prescription and incorrect entry. Other common factors included risk of adverse event, patient compliance and patient preference. Based on the review process, a rule-based algorithm was developed to detect medication revision events. The algorithm had moderate performance with a sensitivity of 66% and a positive predictive value of 4.8% (given a prevalence 1.1%). Thus, in this initial effort, we demonstrated how process mining can be applied to medication revisions in EHR.

**Implications**

For Clinicians: In our study, we find that system workarounds impact the quality of data and performance of automated algorithms. This indicates both a system failure (usability) as well as a need to train providers on use and functions within the EHR. Current training appears to focus on how to complete a task, says steps involved in ordering a medication. But the training does not cover when and why to use certain functions within the tasks, like use of different status updates during ordering process. We can use process mining to identify training needs and evaluate training efforts. We found that the majority of clinical information that was reviewed is from unstructured data. Training efforts must also be directed towards the appropriate and balanced use of structured and unstructured data elements in the EHR. Our detection algorithm that detects potential medication revisions can also be considered as a supplement to clinical decision support tools.

For EHR vendors: EHR developers are required to adopt user centered design principles. In that process, they define EHR tasks and evaluate performance. But all of these are done in a laboratory setup, with limited users and in pre-defined scenarios. In actual implementation of an EHR, the socio-technical factors in implementation and use may result in unintended consequences. To monitor for such issues post-implementation, process mining can be used. Process mining can be defined from the developer perspective to highlight the system factors responsible for deviations or failures. Additionally, we found in this study that a process model can be improved by better activity logs. EHR vendors must collaborate with the federal agencies, professional organization and research institutions to develop and implement a comprehensive user activity log for EHRs.

For Informaticians: For process mining, conventional methods depend on analysis of raw data with event timestamps. However, in clinical data, such audit trails are of limited value. This was evident in our exploratory data analysis. To add information to the raw data, we conducted expert review. The expert review brought out the relationship between the EHR activity and the clinical data, giving meaning to the events. This enabled us to categorize and define a rule for the events. This was useful in developing the algorithm for detection of such events.

The expert review process revealed that 55% of medication revisions are workarounds. This shows that data quality assessment is a must, especially for EHR related data. Though the data is created through same medium (EHR), user interactions have high variability. To detect such outliers, local implementation and workflow must be explored. Some of the methods for uncovering such findings include assessment of training materials used for providers, observation and interview of users, and expert review of retrospective data.

User interactions and steps in the EHR constitute a task. When a task results in an unfavorable or sub-optimal outcome, we seek sources or potential risks. This has directed many outcomes-based studies to uncover and monitor EHR use. However, not all actions or interactions have an evident form of outcome, and not all outcomes be traced back to specific steps. This makes it difficult to evaluate the potential of an EHR action to impact patient safety. Thus, process mining of EHR activity helps in understanding events. This can potentially be applied to understand clinical workflow and monitor the workflow in EHR. The methods can be extended to known instances of EHR actions that have potential to impact patient safety. When applied to clinical decision nodes, the methods can be extended for normative decision analysis.

Process mining of EHR activity can potentially be automated. The limitation towards such an effort is the quality of the data and the technology to process clinical information. Applying domain knowledge to audit trails is critical for process mining in EHR. Thus, improving methods and tools for understanding clinical information, especially from clinical notes, would help in improving the methods. For audit trails, the quality of system use data should be improved and work should be directed towards logging user intent logging related clinical activity. Such comprehensive data can help in efficient process mining. This in turn will improve detection of deviations and good catches. This would allow for assessment of the appropriate use of EHR functions and for development of motivational suggestions to users.

**Future Directions**

This dissertation demonstrates initial efforts in process mining in EHR. The next steps in the study would be to extend the methods for other EHR events, such as cancel medications and discontinue medications. Based on the results for other events, a standardized procedure for

process mining can be defined. Future directions specific to the methods include automating the categorization and fault tree analysis. The data mining methods can be extended to detect clusters and patterns, and thus aide in categorization. The fault trees can be extended to include quantitative data, such as probability trees, and also to automate generation of nodes and interactions. Additional work is required to extend the detection algorithm to categorize and improve performance. To improve performance, more gold standard data would be required. The detection algorithm can also be updated with automated rule-generation and other machine learning methods.

The future vision is to extend the methods to develop a system that would record EHR activity, in a manner similar to a flight recorder, and would monitor the activity for deviations or risks, similar to antivirus software. This would require changes to be made to the EHR system and the logging of activities. For comprehensive logging, we recommend that EHR systems include the software modules, widgets and data represented in the log. At the time of development and deployment, a full structured task, user, representation and functional analysis (TURF) can be included with every module and widget in the EHR system (Zhang & Walji, 2011). As EHR systems now require user centered design process for certification, such analysis will not be additional work for the developers. The next step will be to relate the TURF analysis to the workflow and interactions of various elements in the system. This should include relationships and input-output definitions. Once such meta-data is available with the system, the system must log every activity in the system. This should include time logs, and also the meta-data from widgets such as begin and end stages of elements, data values, etc. Additionally, based on the expert review and modeling of fault trees, we can identify clinical data that adds value to the

audit trail. Efforts must be directed to defining a clinically oriented audit trail that relates to the system oriented audit trail. This would involve more work in generating clinical knowledge from the patient data, making meaningful relationships to actions when an EHR event occurs.

With such comprehensive audit logs and meta-data available, the EHR events can be monitored using rule-based methods and pattern analysis for deviations. From them, a generalization of trends and patterns can be performed. The deviations can be later human-reviewed to assign a relatable category and review of labeled errors. With categories available, detailed pattern analysis can be done to generate the fault trees. With all widget states and action logged, the fault trees would be in-depth and detailed. Learning from these analyses, the system would be able to generate rules for the detection algorithm and alert on potential deviations. Thus, the key for automation is improving the audit log and strengthening the meta-data.

**Limitations:** The study is location specific and EHR system specific, and so the results cannot be generalized. The study used a small sample of cases to describe categories and develop algorithm and fault trees. Because of the small sample size and the limitation to Internal Medicine and adult patients under 60, there may be other categories or reasons for medication revisions not captured in the study. This being a study of possible deviations in medication ordering process, the physician reviewers were hesitant or unable to provide definitive answers in the review process. The categories and definitions were subjective and require validation. Being a retrospective review, there was limited data available. The data logged in the EHR system was also incomplete.

Despite these limitations, the study shows a valid method for process mining in EHRs. These limitations should be addressed to generalize the results and extend the methods to other EHR systems.

# Chapter 9: Conclusion

This dissertation shows how process mining can be applied to medication revisions events in EHR. We described the process of medication revisions in EHR – the types and the factors responsible. We found human factors and system factors that were responsible, which could be used to direct improvements in EHR and EHR use. We showed how process mining can be applied in the EHR, specifically to EHR use. This is a distinction from existing studies on process mining that were applied to tasks and outcomes. This is a novel application to understanding individual process in the EHR tasks.

We demonstrated how process mining can be applied to medication revisions, a deviation in medication ordering process. We showed how data mining alone would not provide insight into the process, and how domain knowledge can be applied, using expert review, to reveal factors related to the process. Based on the review, we modeled the factors and their interactions using fault tree analysis. We also demonstrated the utility of process mining to predict potential revisions.

**Contributions:**

Application contributions:

This study shows how process mining can be applied to EHR activity to reveal knowledge about the process. We applied process mining to medication revisions to understand human and system factors leading to medication revisions. We identified physician factors, system factors and patient factors resulting in medication revisions. These insights can inform EHR vendors and clinical informatics scientists on areas to improve EHR use. We also demonstrated development

of a rule-based algorithm to predict occurrence of medication revisions. The algorithm can be used in real-time monitoring of user activity to detect potential deviations.

The process mining demonstrated in this dissertation can be applied to other EHR activities. The methods can be used to develop a normative model of the process and detect and study deviations in process. Understanding the process in EHR will help in uncovering covert factors that can impact patient safety.

Methods Contributions:

The methods showed how medication revisions events in EHR can be described, their characteristics and the factors responsible for such events. From this we identified how workarounds exist for medication revisions and how they impact the detection of true deviations or failures. The descriptions are purely based on the EHR activity rather than the outcome. As adverse outcomes of an EHR activity is difficult to associate, such methods will be useful to study and characterize EHR activity irrespective of outcome. This will help in improving the EHR system.

We also demonstrated novel use of fault trees for EHR activity. Fault trees provide a visual model and also a probabilistic model of factors responsible for an event. Fault trees have been applied to clinical tasks in health care, but not to individual process within the tasks, especially not to those in an EHR. We identified that limitations in data that must be addressed to improve the utility of fault trees and to allow for the automation of the generation of fault trees.

Learning from the description of medication revisions by exploratory data analysis and expert review, we developed a rule-based algorithm for detection of such events in the EHR data. The

algorithm described in the dissertation can be used to detect error events in EHR. We discussed the problems of limited data and noise in the data affecting the performance of the algorithm.

The methods of process mining can be extended to any event in EHR, such as cancelling orders and discontinuing orders. From the process mining, the methods can be used for real-time monitoring to detect potential deviations in the process. The study also discusses how the methods can be automated for process mining. The vision is to develop a monitoring tools for the EHR similar to flight recorders and antivirus software.

Informatics Contribution:

This dissertation shows how process mining can be applied in EHR. We used clinical activity in an EHR as an audit trail for the mining. In this process, we identified the limitations in the EHR's current level of logging of user activity. The log is restricted to person and time stamp only. We learnt from the exploratory data analysis and expert review that additional information had to be reviewed from clinical data to complete the activity log. This was evident from the fault trees developed that showed nodes that require additional information. Based on such nodes, additional activity logging must be defined for the EHR. Such comprehensive logging will help in automating the process mining.

EHR activity at the task level is documented as when each step was completed and who did it. Such audit data can help in assessing performance of tasks, but not in assessing the processes within the task. As demonstrated in this dissertation, we showed that methods of data mining may not be useful as a standalone. The processes in the EHR are tightly linked to the clinical

activity, and the raw audit data fails to bring out the relationships. We showed how domain knowledge added to the process mining in categorizing and uncovering factors.

## References


Ammenwerth, E., Schnell-Inderst, P., Machan, C., & Siebert, U. (2008). The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. *Journal of the American Medical Informatics Association, 15*(5), 585-600.

Ash, J. S., Berg, M., & Coiera, E. (2004). Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *Journal of the American Medical Informatics Association, 11*(2), 104-112.

Baker, D. W., Persell, S. D., Thompson, J. A., Soman, N. S., Burgner, K. M., Liss, D., & Kmetik, K. S. (2007). Automated review of electronic health records to assess quality of care for outpatients with heart failure. *Annals of internal medicine, 146*(4), 270-277.

Balas, E. A., Li, Z., Mitchell, J., Spencer, D., Brent, E., & Ewigman, B. (1994). *A decision-support system for the analysis of clinical practice patterns.* Paper presented at the Proceedings of the Annual Symposium on Computer Application in Medical Care.

Barach, P., & Small, S. D. (2000). Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. *BMJ, 320*(7237), 759-763.

Barker, K. N., Flynn, E. A., Pepper, G. A., Bates, D. W., & Mikeal, R. L. (2002). Medication errors observed in 36 health care facilities. *Archives of internal medicine, 162*(16), 1897-1903.

Bates, D. W., Cohen, M., Leape, L. L., Overhage, J. M., Shabot, M. M., & Sheridan, T. (2001). Reducing the frequency of errors in medicine using information technology. *Journal of the American Medical Informatics Association, 8*(4), 299-308.

Bates, D. W., Cullen, D. J., Laird, N., Petersen, L. A., Small, S. D., Servi, D., . . . Hallisey, R. (1995). Incidence of adverse drug events and potential adverse drug events: implications for prevention. *Jama, 274*(1), 29-34.

Bates, D. W., Evans, R. S., Murff, H., Stetson, P. D., Pizziferri, L., & Hripcsak, G. (2003). Policy and the future of adverse event detection using information technology. *Journal of the American Medical Informatics Association, 10*(2), 226-228.

Behrens, J. T., & Yu, C. H. (2003). Exploratory data analysis. *Handbook of psychology*.

Blendon, R. J., DesRoches, C. M., Brodie, M., Benson, J. M., Rosen, A. B., Schneider, E., . . . Steffenson, A. E. (2002). Views of practicing physicians and the public on medical errors. *New England Journal of Medicine, 347*(24), 1933-1940.

Bobb, A., Gleason, K., Husch, M., Feinglass, J., Yarnold, P. R., & Noskin, G. A. (2004). The epidemiology of prescribing errors: the potential impact of computerized prescriber order entry. *Archives of Internal Medicine, 164*(7), 785. Retrieved from http://archinte.jamanetwork.com/data/Journals/INTEMED/5481/IOI21119.pdf

Bootman, J. L., Wolcott, J., Aspden, P., & Cronenwett, L. R. (2006). *Preventing Medication Errors:: Quality Chasm Series*: National Academies Press.

Brown, A. W., Malec, J. F., McClelland, R. L., Diehl, N. N., Englander, J., & Cifu, D. X. (2005). Clinical elements that predict outcome after traumatic brain injury: a prospective multicenter recursive partitioning (decision-tree) analysis. *Journal of neurotrauma, 22*(10), 1040-1051.

Campbell, E. M., Sittig, D. F., Ash, J. S., Guappone, K. P., & Dykstra, R. H. (2006). Types of unintended consequences related to computerized provider order entry. *Journal of the American Medical Informatics Association, 13*(5), 547-556.

Carroll, J., Rudolph, J., & Hatakenaka, S. (2002). Lessons learned from non-medical industries: root cause analysis as culture change at a chemical plant. *Quality and Safety in Health Care, 11*(3), 266-269.

Cherian, S. M. (1994). *Fault tree analysis of commonly occurring medication errors and methods to reduce them.* Texas A&M University.

Chiozza, M. L., & Ponzetti, C. (2009). FMEA: a model for reducing medical errors. *Clinica Chimica Acta, 404*(1), 75-78.

Cho, Y. H., & Kim, J. K. (2004). Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce. *Expert systems with Applications, 26*(2), 233-246.

Dean, B., Schachter, M., Vincent, C., & Barber, N. (2002). Causes of prescribing errors in hospital inpatients: a prospective study. *The Lancet, 359*(9315), 1373-1378.

Dhillon, B. (2003). Methods for performing human reliability and error analysis in health care. *International Journal of Health Care Quality Assurance, 16*(6), 306-317.

Doytchev, D. E., & Szwillus, G. (2009). Combining task analysis and fault tree analysis for accident and incident analysis: a case study from Bulgaria. *Accident Analysis & Prevention, 41*(6), 1172-1179.

Dugan, J. B., Bavuso, S. J., & Boyd, M. (1992). Dynamic fault-tree models for fault-tolerant computer systems. *Reliability, IEEE Transactions on, 41*(3), 363-377.

Edwards, P. J., Moloney, K. P., Jacko, J. A., & Sainfort, F. (2008). Evaluating usability of a commercial electronic health record: A case study. *International Journal of Human-Computer Studies, 66*(10), 718-728.

Entezari-Maleki, R., Rezaei, A., & Minaei-Bidgoli, B. (2009). Comparison of classification methods based on the type of attributes and sample size. *Journal of Convergence Information Technology, 4*(3), 94-102.

Fiumara, K., Moniz, T., Churchill, W., Bane, A., Luppi, C., Bates, D., & Gandhi, T. (2007). Case study on the use of health care technology to improve medication safety. *Medication Use: a system approach to reducing errors*, 103-114.

Friedman, A. L., Geoghegan, S. R., Sowers, N. M., Kulkarni, S., Formica, R. N., & Jr. (2007). Medication errors in the outpatient setting: Classification and root cause analysis. *Archives of Surgery, 142*(3), 278-283. doi:10.1001/archsurg.142.3.278

García, A. O., Alfonso, D. P., & Armenteros, O. U. L. (2015). *Analysis of Hospital Processes with Process Mining Techniques.* Paper presented at the MEDINFO 2015: EHealth-enabled Health: Proceedings of the 15th World Congress on Health and Biomedical Informatics.

Górski, J., Magott, J., & Wardziński, A. (1995). Modeling fault trees using petri nets *Safe Comp 95* (pp. 90-100): Springer.

Guo, M., Wagner, M., & West, C. (2004). *Outpatient clinic scheduling: a simulation approach.* Paper presented at the Proceedings of the 36th conference on Winter simulation.

Haerian, K., Varn, D., Vaidya, S., Ena, L., Chase, H., & Friedman, C. (2012). Detection of Pharmacovigilance-Related Adverse Events Using Electronic Health Records and Automated Methods. *Clinical Pharmacology & Therapeutics, 92*(2), 228-234.

Hasan, L., Morris, A., & Probets, S. (2009). Using Google Analytics to evaluate the usability of e-commerce sites *Human centered design* (pp. 697-706): Springer.

Hauskrecht, M., Batal, I., Valko, M., Visweswaran, S., Cooper, G. F., & Clermont, G. (2013). Outlier detection for patient monitoring and alerting. *Journal of biomedical informatics, 46*(1), 47-55.

Hauskrecht, M., Valko, M., Batal, I., Clermont, G., Visweswaran, S., & Cooper, G. F. (2010). *Conditional outlier detection for clinical alerting.* Paper presented at the AMIA annual symposium proceedings.

Hodge, V. J., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review, 22*(2), 85-126.

Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research, 32*(10), 2513-2522.

Hulka, B. S., Cassel, J. C., Kupper, L. L., & Burdette, J. A. (1976). Communication, compliance, and concordance between physicians and patients with prescribed medications. *American Journal of Public Health, 66*(9), 847-853.

International Atomic Energy Agency. (1998). *Collection and classification of human reliability data for use in probabilistic safety assessments*. Retrieved from http://www-pub.iaea.org/mtcd/publications/pdf/te_1048_prn.pdf

Jun, J., Jacobson, S., & Swisher, J. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the operational research society*, 109-123.

Khajouei, R., & Jaspers, M. (2008). CPOE system design aspects and their qualitative effect on usability. *Studies in Health Technology and Informatics, 136*, 309.

Kim, E., Kim, W., & Lee, Y. (2003). Combination of multiple classifiers for the customer's purchase behavior prediction. *Decision Support Systems, 34*(2), 167-175.

Kirwan, B., Gibson, W. H., & Hickling, B. (2008). Human error data collection as a precursor to the development of a human reliability assessment capability in air traffic management. *Reliability Engineering & System Safety, 93*(2), 217-233.

Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). Emerging trends in business analytics. *Communications of the ACM, 45*(8), 45-48.

Kohli, R., Piontek, F., Ellington, T., VanOsdol, T., Shepard, M., & Brazel, G. (2001). Managing customer relationships through E-business decision support applications: a case of hospital–physician collaboration. *Decision Support Systems, 32*(2), 171-187.

Kolker, A. (2008). Process modeling of emergency department patient flow: Effect of patient length of stay on ED diversion. *Journal of Medical Systems, 32*(5), 389-401.

Koppel, R., Metlay, J. P., Cohen, A., Abaluck, B., Localio, A. R., Kimmel, S. E., & Strom, B. L. (2005). Role of computerized physician order entry systems in facilitating medication errors. *Jama, 293*(10), 1197-1203.

Leape, L. L. (1994). Error in medicine. *Jama, 272*(23), 1851-1857.

Lee, W., & Stolfo, S. J. (1998). *Data mining approaches for intrusion detection.* Paper presented at the Usenix Security.

Lee, W. S., Grosh, D. L., Tillman, F. A., & Lie, C. H. (1985). Fault Tree Analysis, Methods, and Applications - A Review. *Reliability, IEEE Transactions on, R-34*(3), 194-203. doi:10.1109/TR.1985.5222114

Lesar, T. S., Briceland, L. L., Delcoure, K., Parmalee, J. C., Masta-Gornic, V., & Pohl, H. (1990). Medication prescribing errors in a teaching hospital. *Jama, 263*(17), 2329-2334.

Li, W.-C., & Harris, D. (2005). HFACS Analysis of ROC Air Force Aviation Accidents: reliability analysis and cross-cultural comparison. *International Journal of Applied Aviation Studies, 5*(1), 65-81.

Lyons, M., Adams, S., Woloshynowych, M., & Vincent, C. (2004). Human reliability analysis in healthcare: a review of techniques. *International Journal of Risk and Safety in Medicine, 16*(4), 223-237.

Maruster, L., Weijters, A. T., Van der Aalst, W. W., & van den Bosch, A. (2002). *Process mining: Discovering direct successors in process logs.* Paper presented at the Discovery Science.

Marx, D. A., & Slonim, A. D. (2003). Assessing patient safety risk before the injury occurs: an introduction to sociotechnical probabilistic risk modelling in health care. *Quality and Safety in Health Care, 12*(suppl 2), ii33-ii38. doi:10.1136/qhc.12.suppl_2.ii33

Melton, G. B., & Hripcsak, G. (2005). Automated detection of adverse events using natural language processing of discharge summaries. *Journal of the American Medical Informatics Association, 12*(4), 448-457.

Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science, 23*(4), 579-595.

Mukkamala, S., Janoski, G., & Sung, A. (2002). *Intrusion detection using neural networks and support vector machines.* Paper presented at the Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on.

Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE, 77*(4), 541-580.

Nanji, K. C., Rothschild, J. M., Salzberg, C., Keohane, C. A., Zigmont, K., Devita, J., . . . Poon, E. G. (2011). Errors associated with outpatient computerized prescribing systems. *Journal of the American Medical Informatics Association, 18*(6), 767-773.

O'Reilly, M., Talsma, A., VanRiper, S., Kheterpal, S., & Burney, R. (2006). An anesthesia information system designed to provide physician-specific feedback improves timely administration of prophylactic antibiotics. *Anesthesia & Analgesia, 103*(4), 908-912.

Olmeda, I., & Fernández, E. (1997). Hybrid classifiers for financial multicriteria decision making: The case of bankruptcy prediction. *Computational Economics, 10*(4), 317-335.

Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*.

Pochet, N., & Suykens, J. (2006). Support vector machines versus logistic regression: improving prospective performance in clinical decision-making. *Ultrasound in Obstetrics & Gynecology, 27*(6), 607-608.

Poelmans, J., Dedene, G., Verheyden, G., Van Der Mussele, H., Viaene, S., & Peters, E. (2010). Combining business process and data discovery techniques for analyzing and improving integrated care pathways *Advances in Data Mining. Applications and Theoretical Aspects* (pp. 505-517): Springer.

Qin, B., Xia, Y., Prabhakar, S., & Tu, Y. (2009). *A rule-based classification algorithm for uncertain data.* Paper presented at the Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on.

Reason, J. (2000). Human error: models and management. *BMJ : British Medical Journal, 320*(7237), 768-770. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1117770/

Reckmann, M. H., Westbrook, J. I., Koh, Y., Lo, C., & Day, R. O. (2009). Does computerized provider order entry reduce prescribing errors for hospital inpatients? A systematic review. *Journal of the American Medical Informatics Association, 16*(5), 613-623.

Rex, J. H., Turnbull, J. E., Allen, S. J., Voorde, K. V., & Luther, K. (2000). Systematic root cause analysis of adverse drug events in a tertiary referral hospital. *Joint Commission Journal on Quality and Patient Safety, 26*(10), 563-575.

Reza, H., Pimple, M., Krishna, V., & Hildle, J. (2009). *A safety analysis method using Fault Tree analysis and Petri Nets.* Paper presented at the Information Technology: New Generations, 2009. ITNG'09. Sixth International Conference on.

Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications *Applications of Data Mining to Electronic Commerce* (pp. 115-153): Springer.

Schiff, G. D., Galanter, W. L., Duhig, J., Lodolce, A. E., Koronkowski, M. J., & Lambert, B. L. (2011). Principles of conservative prescribing. *Archives of internal medicine, 171*(16), 1433-1440.

Shamliyan, T. A., Duval, S., Du, J., & Kane, R. L. (2008). Just What the Doctor Ordered. Review of the Evidence of the Impact of Computerized Physician Order Entry System on Medication Errors. *Health Services Research, 43*(1p1), 32-53. doi:10.1111/j.1475-6773.2007.00751.x

Shappel, S. A., & Wiegmann, D. A. (2000). *The human factors analysis and classification system--HFACS*. Retrieved from

Shulman, R., Singer, M., Goldstone, J., & Bellingan, G. (2005). Medication errors: a prospective cohort study of hand-written and computerised physician order entry in the intensive care unit. *Critical Care, 9*(5), R516 - R521. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1297620/pdf/cc3793.pdf

Singh, C., & Kim, Y. (1988). An efficient technique for reliability analysis of power systems including time dependent sources. *Power Systems, IEEE Transactions on, 3*(3), 1090-1096.

Sittig, D. F., Ash, J. S., Zhang, J., Osheroff, J. A., & Shabot, M. M. (2006). Lessons from "Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system". *Pediatrics, 118*(2), 797-801.

Tarkan, S., Plaisant, C., Shneiderman, B., & Hettinger, A. Z. (2011). *Reducing missed laboratory results: defining temporal responsibility, generating user interfaces for test process tracking, and retrospective analyses to identify problems.* Paper presented at the AMIA Annual Symposium Proceedings.

Tesh, D., & Beeley, L. (1975). Errors of drug prescribing. *British journal of clinical pharmacology, 2*(5), 403.

The Office of the National Coordinator for Health Information Technology (ONC). (2014a). *Electronic Prescribing Adoption and Use by State*. Retrieved from: http://dashboard.healthit.gov/datadashboard/data.php

The Office of the National Coordinator for Health Information Technology (ONC). (2014b). *Report to Congress : Update On The Adoption Of Health Information Technology And Related Efforts To Facilitate The Electronic Use And Exchange Of Health Information*. Retrieved from https://www.healthit.gov/sites/default/files/rtc_adoption_and_exchange9302014.pdf

Trifirò, G., Pariente, A., Coloma, P. M., Kors, J. A., Polimeni, G., Miremont-Salamé, G., . . . Moore, N. (2009). Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiology and drug safety, 18*(12), 1176-1184.

Tukey, J. (1977). Exploratory data analysis.

U.S Department of Transportation Federal Aviation Administration. (1998). *Safety Risk Management Policy*. Retrieved from

Van Der Aalst, W. (2011). *Process mining: discovery, conformance and enhancement of business processes*: Springer Science & Business Media.

van der Aalst, W. (2013). *" Mine your own business": using process mining to turn big data into real value.* Paper presented at the Keynote 21st European Conference on Information Systems, ECIS, Utrecht, The Netherlands.

van der Aalst, W. M., Reijers, H. A., Weijters, A. J., van Dongen, B. F., De Medeiros, A. A., Song, M., & Verbeek, H. (2007). Business process mining: An industrial application. *Information Systems, 32*(5), 713-732.

Vesely, W. E., Goldberg, F. F., Roberts, N. H., & Haasl, D. F. (1981). *Fault tree handbook*: Systems and Reliability Research, Office of Nuclear Regulatory Research, U.S. Nuclear Regulatory Commission, Washington, D.C.

Walsh, K. E., Adams, W. G., Bauchner, H., Vinci, R. J., Chessare, J. B., Cooper, M. R., . . . Landrigan, C. P. (2006). Medication Errors Related to Computerized Order Entry for Children. *Pediatrics, 118*(5), 1872-1879. doi:10.1542/peds.2006-0810

Westbrook, J. I., Reckmann, M., Li, L., Runciman, W. B., Burke, R., Lo, C., . . . Day, R. O. (2012). Effects of Two Commercial Electronic Prescribing Systems on Prescribing Error Rates in Hospital In-Patients: A Before and After Study. *PLoS Med, 9*(1), e1001164. doi:10.1371/journal.pmed.1001164

Wiegmann, D. A., & Shappell, S. A. (2001). Human error analysis of commercial aviation accidents: Application of the Human Factors Analysis and Classification System (HFACS). *Aviation, space, and environmental medicine, 72*(11), 1006-1016.

Wu, J., Roy, J., & Stewart, W. F. (2010). Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. *Medical care, 48*(6), S106-S113.

Yuhua, D., & Datao, Y. (2005). Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis. *Journal of loss prevention in the process industries, 18*(2), 83-88.

Zhan, C., Hicks, R. W., Blanchette, C. M., Keyes, M. A., & Cousins, D. D. (2006). Potential benefits and problems with computerized prescriber order entry: analysis of a voluntary

medication error-reporting database. *American Journal of Health-System Pharmacy, 63*(4), 353-358. doi:10.2146/ajhp050379

Zhang, J., & Walji, M. F. (2011). TURF: Toward a unified framework of EHR usability. *Journal of biomedical informatics, 44*(6), 1056-1067. doi:http://dx.doi.org/10.1016/j.jbi.2011.08.005

Zheng, K., Padman, R., Johnson, M. P., & Diamond, H. S. (2009). *An Interface-driven Analysis of User Interactions with an Electronic Health Records System* (Vol. 16).

Zhu, M., Rogith, D., Franklin, A., Walji, M., & Zhang, J. (2014). Turf Usability Tool Suite *Better EHR: Usability, workflow and cognitive support in electronic health records* (Vol. 1, pp. 111-120): The National Center for Cognitive Informatics & Decision Making in Healthcare.