IMPROVING THE ERROR REVIEW PROCESS FOR INCIDENT REPORTS AT UT SOUTHWESTERN THROUGH THE USE OF A STANDARDIZED TAXONOMY TOOL

by

CHRISTOPHER CHAN

DISSERTATION

Presented to the Faculty of the Medical School
The University of Texas Southwestern Medical Center
In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF MEDICINE WITH DISTINCTION IN

QUALITY IMPROVEMENT AND PATIENT SAFETY

The University of Texas Southwestern Medical Center
Dallas, Texas

ABSTRACT

IMPROVING THE ERROR REVIEW PROCESS FOR INCIDENT REPORTS AT UT SOUTHWESTERN THROUGH THE USE OF A STANDARDIZED TAXONOMY TOOL

CHRISTOPHER CHAN
The University of Texas Southwestern Medical Center, 2018
Supervising Professor: Isaac Lynch, M.D.

Background: At UT Southwestern and in other hospital systems, medical errors are often the result of latent system errors. These errors may remain unreported due to a variety of reasons. When they are reported, the defect analysis protocols triggered by error review are often inadequate for identification and correction of the underlying system defect. The aim of this project was to improve the process for classifying medical errors at UT Southwestern by testing the use of a standardized taxonomy tool, refining the tool to reduce inter-rater variability, and incorporating its use into the existing workflow for real-time error review. Of particular interest was the subset of medical errors related to OR-ICU handover.

Methods/Intervention: The taxonomy tool used in this project was created by Dr. Isaac Lynch, cardiac anesthesia and intensive care faculty at UT Southwestern. It is a form that allows categorization of several important parameters described by both the World Health Organization's 2005 "Draft Guidelines for Adverse Event Reporting and Learning Systems" and 2014 "Minimal Information Model for Patient Safety Incident Reporting" guidelines. For example, the tool allows for designation of an error as being "administrative," "medication," or "handover-related." Properly trained Clinical Safety personnel can use the taxonomy tool to quickly review incident reports in real-time as they are submitted. For this project, all ICU incident reports submitted through the UT Southwestern Clements University hospital reporting system in 2015, a total of 1317 reports, were reviewed and classified according to the taxonomy tool. The collected data was analyzed using REDCap™ built-in statistical modeling tools for hazard identification and trends. Based on this initial trial run, the taxonomy tool was refined to

improve functionality and reduce inter-rater variability that may occur when different people use the tool. Analysis of the data also showed how taxonomic categorization can highlight system errors that are in need of targeted intervention. Finally, a process map was created to illustrate how use of this taxonomy tool can be incorporated into the existing error review workflow at UT Southwestern.

Results: Analysis of the incident reports from 2015 and classification according to the taxonomy tool revealed some specific areas in need of process improvement. For example, there were a significant number of reports related to inappropriate specimen labeling and also many reports describing patient pressure ulcers. Use of the taxonomy tool enabled identification and classification of these medical errors. Future, in-depth analysis can then be used to inform targeted intervention. Notable statistics from the analysis are that 46% of the reports described errors occurring during perioperative care, 18% were medication errors, and 18% were diagnostic errors. To further classify the factors contributing to these different errors, nearly 52% of the total incident reports were latent system errors, 58% could be attributed to staff error, and 45% had some component of patient cause. Of note, a significant 70% of the latent system errors were equipment-related. Also, only 5.5% of the reports described errors that were probably related to handover, with 78.4% unlikely to be related to handover. Finally, 61% of the medical errors caused temporary harm to the patient, with 1.9% describing errors that contributed to patient death.

Conclusion: This project demonstrated that it is both practical and helpful to use a standardized taxonomy tool for routine, real-time error review. Review of a single incident report and completion of the corresponding taxonomy form takes an average of 7 minutes for trained personnel so there will not be a significant additional work burden for Clinical Safety specialists. The data collected in this project has been helpful for identifying specific system errors, and classification of the types of error has utility in determining the best intervention strategies. The gradual accumulation of data can also be used for trend identification and epidemiological studies. In terms of next steps, the taxonomy tool, which has already been modified based on this initial test run, will undergo further trial testing to improve inter-rater reliability. This will involve training other medical students in the tool's proper use and having them repeat error review on a subset of the 2015 ICU reports, before then continuing the analysis with ICU reports

from subsequent years. Future projects will also focus on integrating use of the taxonomy tool into the existing error review workflow, developing a dashboard module for real-time trend analysis, and enabling human languages algorithm functionality. The vision is for the taxonomy tool to be employed by Clinical Safety personnel in real-time to generate useable classification data for fixing latent system errors.

TABLE OF CONTENTS

PROBLEM DESCRIPTION

According to a Johns Hopkins study that was published in the British Medical Journal in 2016, medical errors are the 3rd leading cause of death in the United States, only ranking behind heart disease and cancer.[1] The traditional and outdated approach to addressing medical errors has been to blame individual providers. However, this blame game is not conducive to continuous quality improvement and has led to a culture in which providers may try to deny and hide perceived errors to avoid punishment. The "Just Culture" model, as explained by David Marx in 2001, is a systems approach that focuses on creating an open, fair culture of shared accountability. This paradigm shift is one that accepts the inevitability of human error and views medical errors as occurring in the context of latent system errors, also called 'unintended events.' Instead of blaming and punishing an individual provider, the systems approach offers anonymity for medical error reporting and focuses instead on determining underlying processes or conditions that lead people to make mistakes. In a just culture, individuals strive to report mistakes, continuously learn, design safe systems, and manage behavioral choices.[2] British psychologist James Reason's "Swiss Cheese model" illustrates the phenomenon in which major accidents often result from multiple, small errors caused by a combination of system flaws, human error, and oversight – the holes in the cheese. The "Swiss Cheese model" illustration is shown below in Figure 1.

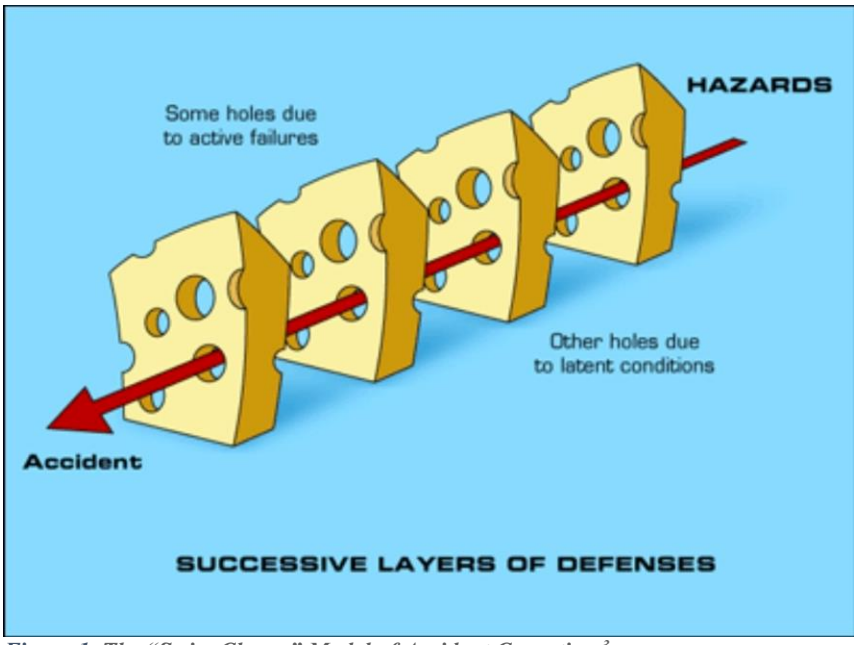

Figure 1. The "Swiss Cheese" Model of Accident Causation.[3]

The goal of a systems approach is to design systems and workflows that make it hard for people to make mistakes. But even with this culture shift that is moving away from individual blame, many hospital systems lack a standardized, robust reporting system for analyzing medical errors and prompting meaningful action. A 2008 study by Farley, *et al.* that evaluated the incident reporting systems of over 1600 U.S. hospitals concluded that most hospitals do not maintain effective incident reporting systems.[4]

When a medical error is reported at UT Southwestern, certain root cause analysis protocols are triggered. Every incident report auto-generates an alert that is sent to a departmental manager, depending on incident location. The alerts only include basic event information such as: individual affected, event location, severity level, and the narrative as written by the person who submitted the report. For majority of the cases, it is the departmental manager's responsibility to review the report narrative and implement a corrective action plan at the unit level. Only those incidents designated by Clinical Safety personnel as sentinel events, or events leading to major patient harm or death, are formally evaluated with RCA or by a performance committee. Monthly patient safety meetings are held to review these major harm events. In this system, departmental managers are tasked with fixing the majority of system errors but are not given useable data to help with devising appropriate interventions. For these mild- to moderate-harm incidents that do not trigger a formal RCA, the response is often inadequate to identify and correct any underlying problems. The result is that fixable system errors persist and continue to cause patient harm.

AVAILABLE KNOWLEDGE

In 2005, the World Health Organization issued the "WHO Draft Guidelines for Adverse Event Reporting and Learning Systems" and described the framework for establishing an effective incident reporting system. This document is useful because it comes from a credible source and also establishes a standard terminology that forms the basis of discussion. According to these WHO guidelines, having people report medical errors is helpful for identifying hazards and process failures. This is especially true when reports are initiated by health care providers and other individuals working on the front lines. The incident reporting process includes several

steps. First, an easily accessible incident reporting system must be implemented in which people, usually health-care providers, are able to anonymously submit incident reports. Second, the error reports must be reviewed and undergo analysis by trained personnel in a timely manner. Finally, there must be a safety action feedback loop in which the analysis is followed by both a response to fix the underlying problem and also dissemination of the lessons learned[5]. A growing database of aggregated report data will also be helpful in elucidating trends and underlying problems. Studies have shown that incident reporting systems can positively affect patient safety.[6]

There are different methods for measuring patient safety. The method described in the previous paragraph is termed a voluntary incident reporting system, which means detection of system errors is dependent on people reporting. Anybody can submit an incident report, and this will trigger the incident reporting process. Some other measurement strategies include: retrospective chart review, automated trigger tool surveillance, review of administrative/claims data, or patient reports. Each strategy has its unique advantages and disadvantages.[7] The voluntary incident reporting system is used at UT Southwestern because of a combination of cost, ability to collect clinical information, and ability to measure results.

A relatively simple measurement is to focus on increasing the number of incident report submissions that are received. Data collection is a central component of any successful quality improvement initiative. According to the Institute for Healthcare Improvement, measuring the right things and tracking the measurements is critical to determining whether implemented changes are leading to improvement. However, limitations to voluntary incident reporting are well-described in current literature. The nature of voluntary incident reporting makes it subject to selection bias. A 2012 document published by the Department of Health and Human Services states that up to 86% of events were not reported by hospital staff to incident reporting systems.[8] According to a collaborative hospital study published in 2006, staff were asked to self-report on the major barriers to incident reporting. Doctors reported their top self-perceived barriers to incident reporting as: 1) lack of feedback on incident follow-up, 2) form was too long, and 3) the incident seemed trivial. Nurses reported major barriers as: 1) lack of feedback, 2) a belief that there was no point in reporting near misses, and 3) forgetting to make a report when the ward was busy.[9] The previously mentioned study by Farley, *et al.*, in which over 1600 U.S. hospital

incident reporting systems were evaluated, concluded that there was a lack of physician reporting and often inadequate structure for analyzing and acting upon incident reports.

While counting total number of incident reports is relatively easy to measure, it tells an incomplete story. An increase in submitted reports may reflect a positive change in reporting culture, improvement in the actual incident reporting system, or increase in patient safety problems. A reduction in submitted reports may indicate fewer actual events, fewer people submitting reports, natural variability, or something else entirely. However, more important than collecting reports is learning from each reported event and responding appropriately. This is accomplished by conducting appropriate root cause analysis and having a functional safety feedback loop. Unfortunately, some nurses have called these steps a "black hole" because of lack of transparency in the process and inconsistency in its application. In fact, a 2009 systematic review of approximately 2000 studies on incident reporting found that only 13 U.S. reporting systems even described a feedback mechanism[10]. Clearly, there is much room for improvement.

Several different incident reporting systems have been developed to address this problem. In the United Kingdom, the National Patient Safety Agency maintains the National Reporting and Learning System, a nationwide incident reporting system for all hospitals in the country. In the United States, the Agency for Healthcare Research and Quality (AHRQ) encourages use of its 'Common Formats' to standardize the patient safety event information that is collected in an error report.[11] However, a 2015 study to determine reliability of the AHRQ Common Format Harm Scales in rating patient safety events found only moderate inter-rater reliability when different people were using the reporting tool.[12] There is still much work that needs to be done to establish a robust taxonomy tool for reviewing patient safety incident reports.

RATIONALE

High reliability organizations (HRO) are "organizations that operate in complex, high-hazard domains for extended periods without serious accidents or catastrophic failures."[13] The HRO paradigm describes a systems approach in which hospitals should strive to cultivate an environment that prioritizes safety over other performance measures, has leadership commitment

to a positive safety culture, and incorporates robust process improvement. This quality improvement project works towards the HRO goal by attempting to standardize the incident reporting process at UT Southwestern to improve reliability and allow for data comparison between health systems. The Institute of Medicine's seminal report, "To Err is Human," describes the nature of system based errors and recommends standardization as a well-understood safety principle.[14] Process standardization has been successful in achieving enviable quality standards in many industries, including in aviation and in the automobile industry. The health system at UT Southwestern should not be an exception. Standardizing the error review process and establishing a clear safety feedback loop creates a cohesive framework for both identifying patient safety issues and fixing them.

This project also tests the use of a taxonomy tool for classifying medical errors. Using a taxonomy to diagnose medical errors is not a novel concept, and most hospitals have some semblance of it in their incident reporting systems. Our tool's contents are based on both the WHO 2005 "Draft Guidelines for Adverse Event Reporting and Learning Systems" and 2014 "Minimal Information Model for Patient Safety Incident Reporting" (MIM PS) guidelines. Both of these documents provide the framework for creating an effective reporting system, with MIM PS describing the basic information required to be included in any incident report.[15] Our taxonomy tool template, simplified to facilitate ease of use, contains these minimum classification categories and adheres to the World Health Organization recommendations.

SPECIFIC AIMS

The specific aim of this project was to improve the process for classifying and correcting medical errors at UT Southwestern by testing the use of a standardized taxonomy tool, refining the tool to reduce inter-rater variability, and incorporating its use into the existing workflow for real-time error review. The ultimate goal is to organize the analysis of incident reports in a more meaningful way, better identify latent system errors, and subsequently enable targeted intervention to improve patient outcomes. The parameters for analyzed incident reports were those submitted from UT Southwestern Clements University Hospital ICUs during the period January 1, 2015 – December 31, 2015. Of particular interest were the subset of unintended

events caused by handover errors between providers. This quality improvement project was conducted in recognition of the importance of having proper analysis for identifying errors and guiding future targeted intervention. Furthermore, demonstrating a successful pilot use of the taxonomy tool could eventually lead to its integration into the existing UT Southwestern incident reporting system.

CONTEXT

William P. Clements Jr. University Hospital, the site of this QI project, is a 460-bed, academic hospital associated with UT Southwestern Medical Center in Dallas, TX. It is currently undergoing expansion to add more inpatient care services and operating rooms.

Each hospital within the UT Southwestern system, including Clements University Hospital, Parkland Memorial Hospital, Children's Medical Center of Dallas, and the Dallas VA Medical Center, conducts its own separate medical error review. At Clements University Hospital, any health professional can fill out and submit an online incident report by accessing RL Solutions via the UT Southwestern clinical portal. RL Solutions is a contracted healthcare software company that provides a platform for incident reporting. Completing an online incident report takes less than 5 minutes and only requires the submitter to fill out 10 mandatory fields, including general incident information, location, incident severity, and free-text narrative description of the incident. However, physicians do not receive any formal incident reporting training, and nurses only receive a brief overview.

Whenever an incident report is submitted online, an auto-generated alert is sent to a departmental manager, based on incident location. The alert includes basic event information such as: individual affected, event location, severity level, and the narrative as written by the person who submitted the report. It is the departmental manager's responsibility to review the incident report within 72 hours of receipt and implement a corrective action plan at the unit level. At the same time, salaried Clinical Safety personnel also review the submitted report and determine its severity. Mild- to moderate-harm incidents are left to await manager review. Major harm events or sentinel events are escalated for evaluation by a RCA committee. After both

Clinical Safety and departmental manager review, possible resulting action includes: corrective action taken at the unit level, departmental manager to provide closed-loop communication to staff, root cause analysis by performance committee, etc. A process map of the current incident reporting workflow is illustrated below in Figure 2.

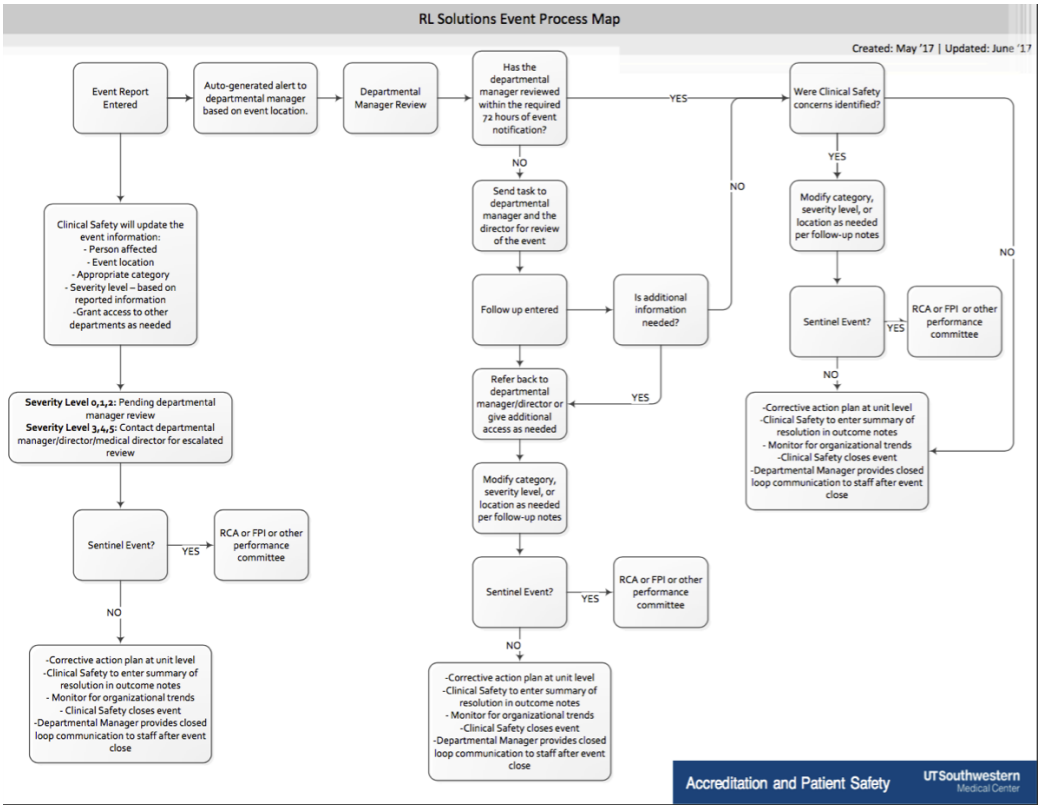

Figure 2. Current Patient Safety Incident Reporting Process Map. Provided by Clinical Safety.

There are several problems that prevent this existing system from having meaningful patient safety impact. First, the auto-generated alerts are not informative enough to really help departmental managers identify latent system errors and devise appropriate corrective action. Second, while there is aggregated basic information data from all previous incident reports, the data is not classified in a way to facilitate quick epidemiological comparisons and big data trend analysis. Currently, this can only be done by tedious manual review. Third, the lack of an established safety action feedback loop for effective follow-up and intervention means that mild- to moderate-harm errors are often left unfixed. These issues result in reduced quality of health care delivery, inadequate data analysis, and perpetuation of latent system errors.

INTERVENTIONS

The centerpiece of this project was a taxonomy tool developed by Dr. Isaac Lynch, cardiac anesthesia and intensive care faculty member at UT Southwestern. The tool was constructed and hosted on REDCap™, a secure, HIPAA-compliant database management software package used by UT Southwestern's Information Resources division and maintained by Academic Information Systems. REDCap™ database content is backed up at a secure off-site location every night. The taxonomy tool itself is a web-based survey that allows Clinical Safety personnel to deconstruct a single incident report into classifications that are conducive to medical error analysis. The classification categories used in the taxonomy tool are described by both the WHO 2005 "Draft Guidelines for Adverse Event Reporting and Learning Systems" and 2014 "Minimal Information Model for Patient Safety Incident Reporting" guidelines. The subclassifications used to describe impact of the medical error are based on the AHRQ 'Medications at Transitions and Clinical Handoffs' MATCH Toolkit recommendations.[16] Additionally, the harm scale used in the taxonomy tool is adapted from the NCC MERP index for categorizing errors, which is shown in Figure 3 and 4.[17]

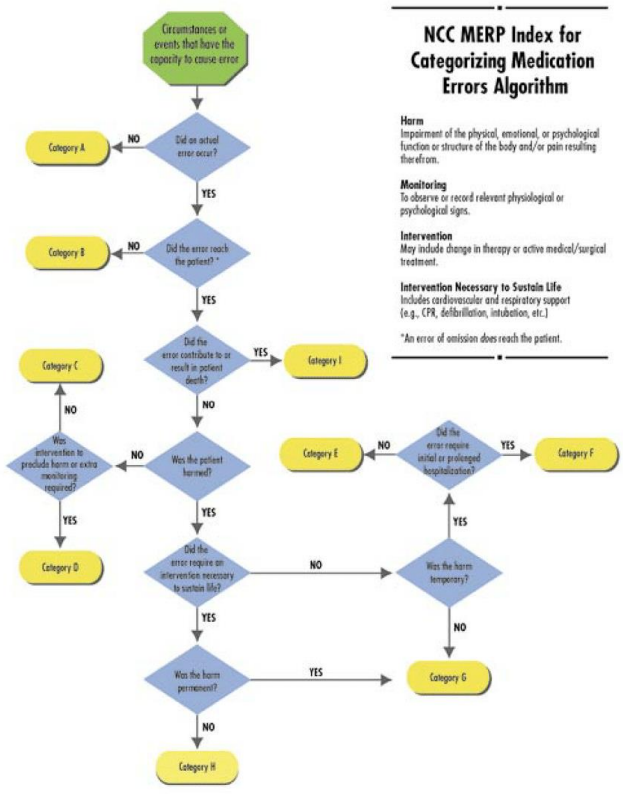

Figure 3. NCC MERP Harm Algorithm. Obtained from the National Coordinating Council for Medication Error Reporting and Prevention website

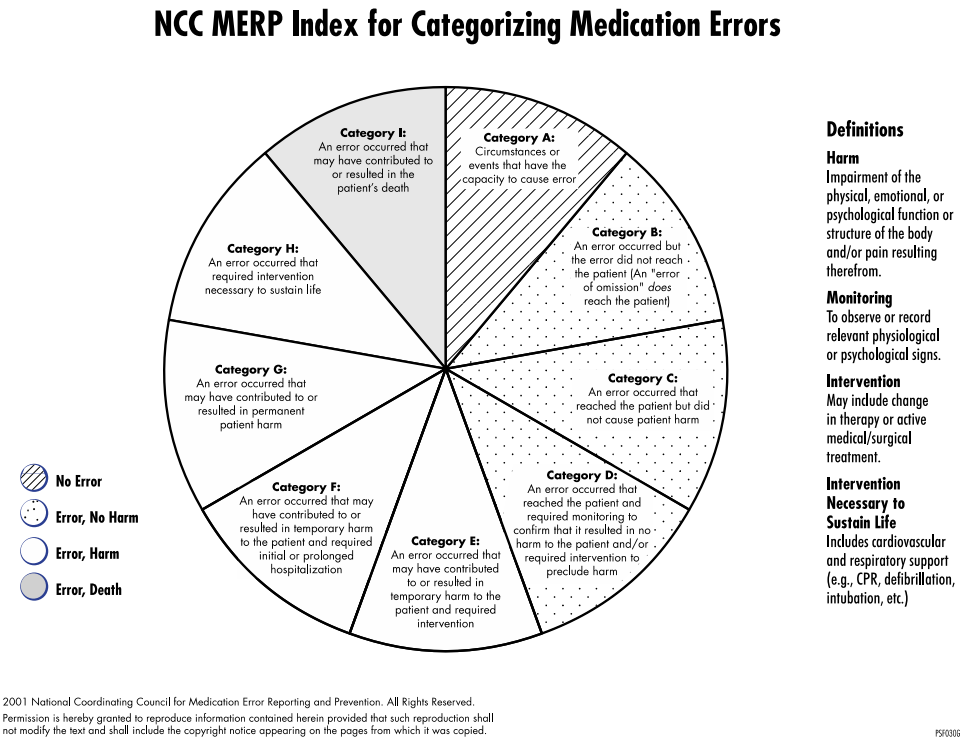

Figure 4. NCC MERP Harm Algorithm. Obtained from the National Coordinating Council for Medication Error Reporting and Prevention website

The purpose of a classification system is to better understand the context of an error. The taxonomy tool was designed for this purpose and also for ease of use. The tool's primary classification categories are: medical error type, medical error domain, contributing factors, impact, and handover. These primary classifications are then divided into additional subclassifications/nodes to allow for more in-depth analysis. Figure 5 shows how each primary classification is subdivided.

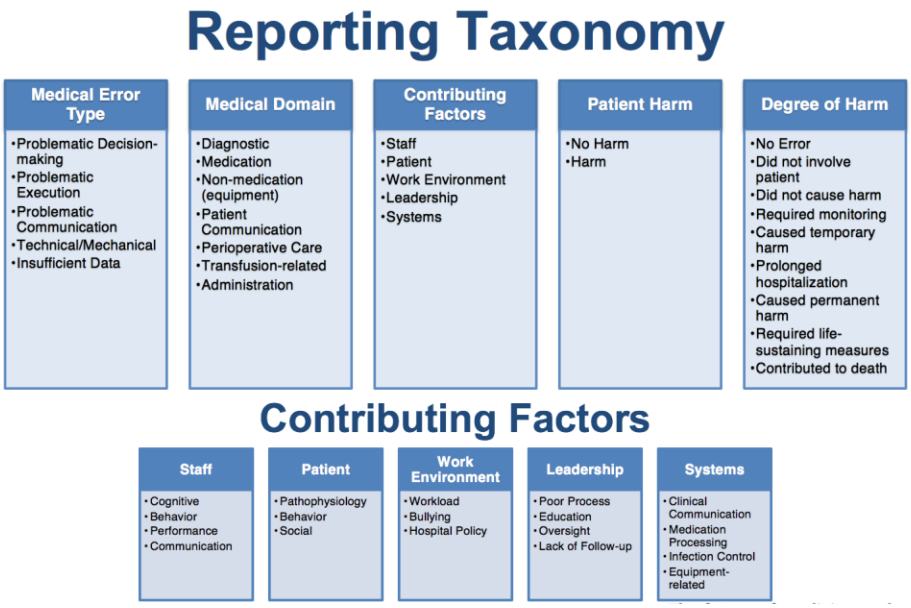

Figure 5. Taxonomy Tool Primary Classification Category Breakdown.

The taxonomy tool itself has a plain design and is comprised of several drop-down lists and checkbox items. For some items on the survey, multiple selections are possible. For instance, a single incident report can be classified as having both staff and patient contributing factors. A sample screenshot of the tool is shown in Figure 6.

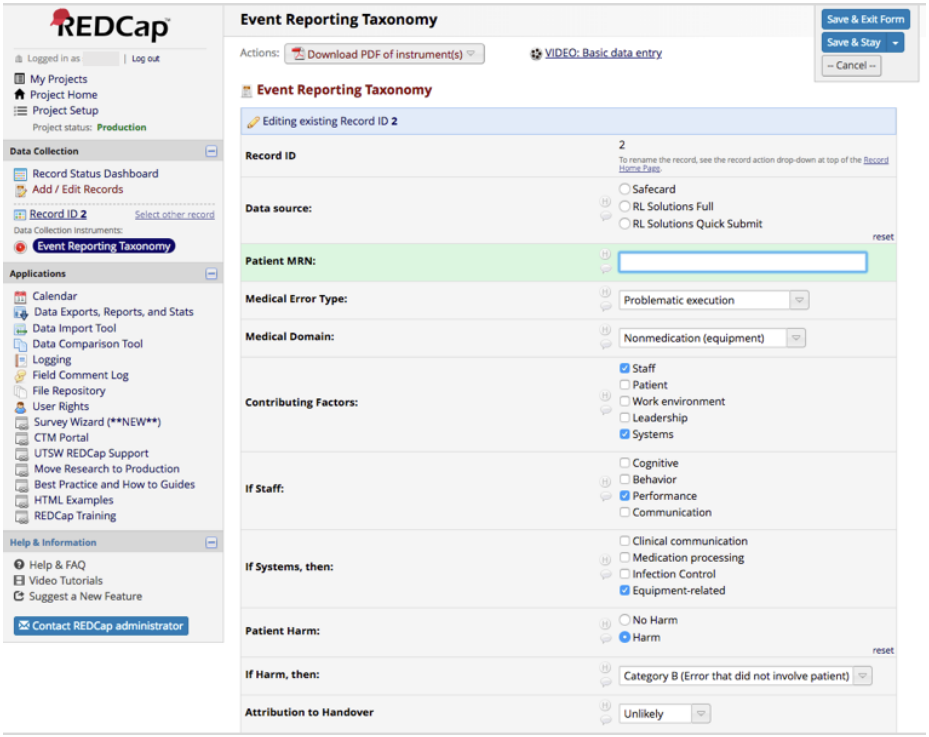

Figure 6. Screenshot of taxonomy tool as hosted on REDCap™.

Before starting the project, an IRB waiver was approved, such that a proposal did not need to be submitted. The Clements University Hospital clinical safety team generated an excel sheet with information from all ICU patient safety incident reports submitted through the UT Southwestern Clements University Hospital incident reporting system in 2015. There were a total of n = 1317 incident reports that satisfied the pre-designated parameters of the study. The specific ICUs included SICU, MICU, and CVICU. This QI project was divided into three phases. Phase I involved performing a retrospective root cause analysis of the incident reports as a trial run to test the taxonomy tool's usefulness and practicality. Each of the incident report narratives were reviewed and classified according to the taxonomy tool. After becoming proficient in filling out the surveys, it would take an average of 7 minutes to review and classify each report.

Phase II started after completion of the pilot error review. This phase included analysis of the data and revision of the taxonomy tool. The data was analyzed with REDCap™ built-in statistical modeling tools, and the tool was simplified based on insight gained from the trial run.

The focus of Phase III was preparing for possible future incorporation of the taxonomy tool into the existing error review workflow. Figure 7 shows a modified process map illustrating the proposed change. The additional step requires Clinical Safety personnel to use the taxonomy tool in real-time to classify incident reports as they are submitted. This classification data will then be included in the alert sent to departmental managers. Additionally, three outputs were created as part of Phase III. The first document was an adverse event reporting log to record thought process and reasoning when classifying each report. The second was an instruction manual to teach others how to use the taxonomy tool. Third was a best-use guide describing the most frequently reported medical error scenarios. The goals of producing all these outputs were to better create a standardized, user-friendly review process tool and also to help reduce inter-rater variability when different personnel conduct analyses in the future.

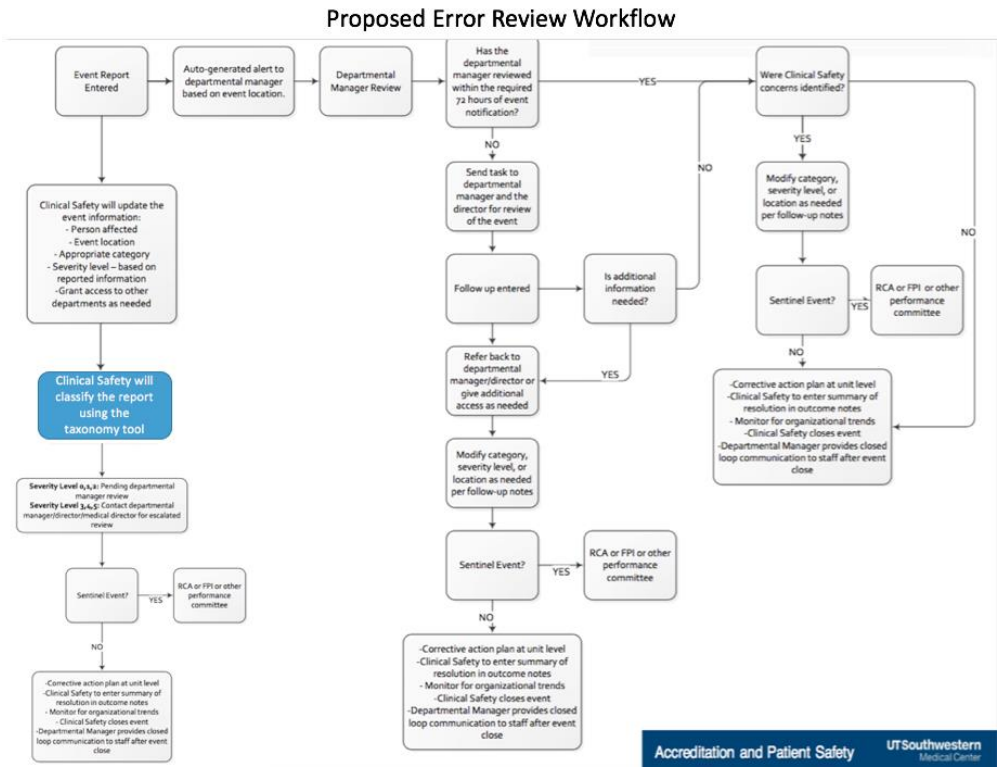

Figure 7. Proposed Patient Safety Incident Reporting Process Map.

Two individuals made up the core team involved in this QI project: Christopher Chan, 4th year medical student at UT Southwestern, and Dr. Isaac Lynch, Assistant Professor of Anesthesiology at UT Southwestern. Dr. Philip Greilich, S.T. "Buddy" Harris Distinguished Chair in Cardiac Anesthesiology at UT Southwestern, was a member of the thesis committee and project champion. Members of the QI team with peripheral involvement in this project were: Dr. Gary Reed, S.T. Harris Family Distinguished Chair in Internal Medicine, Ms. Eleanor Phelps, Director for Nursing Quality Improvement, and Ms. Patty Brown, Quality Improvement Program Coordinator. Other contributing members included: Ms. Mandy McBroom, clinical research professional for the UT Southwestern Department of Anesthesiology and Pain Management, Patrick Roberts, 3rd year medical student at UT Southwestern, Nachae Wren, Ruth Kubajak, and Ripple Chokshi. In the future, other medical students will be involved in continuing the project and testing the taxonomy tool's inter-rater variability.

STUDY OF THE INTERVENTIONS

An important aspect of any quality improvement project is being able to credibly assess its impact. For this specific project, one tangible measure is the value created from accumulating classification data for each incident report. In the existing error review workflow, mild- to moderate-harm incidents at Clements University Hospital do not undergo in-depth analysis. The narratives submitted in each incident report are not dissected in a way to assist in detecting system errors. As a result, fixable errors are allowed to persist. Using the taxonomy tool to review ICU incident reports from 2015 has produced aggregated data that can be used by Clinical Safety personnel to detect epidemiological trends, stratify incidents by different criteria, and perform other public health comparisons. Building this kind of database allows for precise analysis and the ability to more easily determine underlying causes.

Another way to assess the impact of using this taxonomy tool is by measuring its inter-rater variability. The tool will have decreased utility if different people using the same tool obtain different results. The need to maintain fidelity of the analytical process highlights the importance of having outputs like an instruction manual and best-use guide to teach future personnel how to properly use the tool. The trial run of reviewing ICU incident reports provides a baseline analysis. The next step is to have another UT Southwestern medical student review a subset of the same incident reports and measure the differences in results. This will provide insight helpful for further refining the tool and teaching materials.

Finally, a way to gauge the efficacy of the taxonomy tool in impacting patient safety is to observe a decrease in the number of submitted incident reports pertaining to a specific system error after intervention is implemented. It is expected that the classification data generated from analyzing incident reports will allow for better understanding of the types of latent system errors contributing to patient safety problems. Targeted interventions can then be designed to fix the identified system errors. As each patient safety issue is resolved, a subsequent decrease in submitted incident reports pertaining to that issue would be expected. Unfortunately, obtaining this kind of time-intensive, post-intervention analysis was not possible in the timeframe of this QI project.

MEASURES

The taxonomy tool is useful for reviewing incident reports because of what it measures. To quote directly from WHO's "Draft Guidelines for Adverse Event Reporting and Learning Systems," some of the most important factors for classifying a medical error include: "error type (wrong dose, wrong diagnosis, etc.), patient outcome (level of harm), setting, personnel involved, product or equipment failures… underlying causes (lack of knowledge, information, skills, etc.), contributing factors (organizational factors, environmental factors, etc.)… and mechanism of error (knowledge-based, rule-based, skill-based)." The tool's design accounts for all these factors.

Phase III of this project was about planning how to incorporate use of the taxonomy tool into the existing error review workflow. During the trial run, it took an average of 7 minutes to review an incident report. This is a reasonable additional work burden for properly trained Clinical Safety personnel. However, in order to obtain feedback from these individuals who will be most directly affected by this QI intervention, all Clinical Safety personnel will have a post-intervention interview after they have had ample time to work with the tool. The focus of the interview will be to determine the burden of this additional responsibility, whether the proper classification categories are being used in the tool, if it is even useful for identifying latent system errors, and if real-time error review is feasible.

Additionally, it will be worthwhile to conduct a survey of those frontline health providers who are most responsible for submitting patient safety incident reports. This population includes nurses and departmental managers. The survey will assess their perceptions of improvements in outcomes of care and also gauge how well the safety feedback loop is functioning. If system errors revealed by submitted reports are being fixed in a timely manner and those changes are communicated appropriately, then post-intervention survey of frontline health providers will reflect this.

ANALYSIS

Analysis was performed using REDCap™ built-in statistical modeling tools for hazard identification and trends. However, data from these studies could have also been exported to either Microsoft Excel or a statistical software package such as SAS for further analysis. REDCap™ analysis is convenient because it automatically aggregates data points and renders it in the form of bar graphs and pie charts. Data can also be stratified according to criteria of interest. Percentages and count data are displayed in the default statistical modeling, and the analysis can be updated in real-time as data points are added. Taxonomy classifications given for individual incident reports can also be retrieved for immediate review because both the analytical and error review components are on a shared REDCap™ hosting database. The only limitation is access to computer and Internet. Because the data is based on a retrospective root cause analysis model of data collection, it offers a fixed snapshot in time that does not capture effects of time as a variable. REDCap™ statistical modeling makes it very straightforward to collect the data needed to construct a Pareto chart, which can illustrate the most common types of errors and magnify areas most in need of improvement. After classification according to the taxonomy tool, individual incidents can also be plotted on risk assessment matrices, providing a visual comparison of severity vs. likelihood of event occurrence.

When REDCap™ analysis identifies a potential latent system error, the next step is developing an intervention to fix the problem. In order to develop an effective intervention, additional information must be collected to further evaluate the system error. Theoretically, this will likely include creating a stakeholder analysis matrix and determining the parties involved. Then, these interest groups will be interviewed to obtain their support and feedback. Clinical Safety personnel can use a fishbone diagram to aid in root cause analysis. Finally, an appropriate intervention can take many forms. Some examples include: introducing checklists, implementing forcing functions, removing variation, redesigning protocols and work environments, education, or revising policy.

ETHICAL CONSIDERATIONS

No patient identifiers were ever stored or transported on portable computing devices (Laptops, USB drives, CD, DVD, etc.) that could have been lost or misplaced. The statisticians, clinical researchers, and manuscript writers did not have access to the protected health information. Only the individuals who were designated for data collection had temporary access, and they used campus computers with password protection to collect variables and upload into the secure REDCap™ database.

RESULTS

The timeline shown in Figure 8 outlines the 3 phases of this QI project.

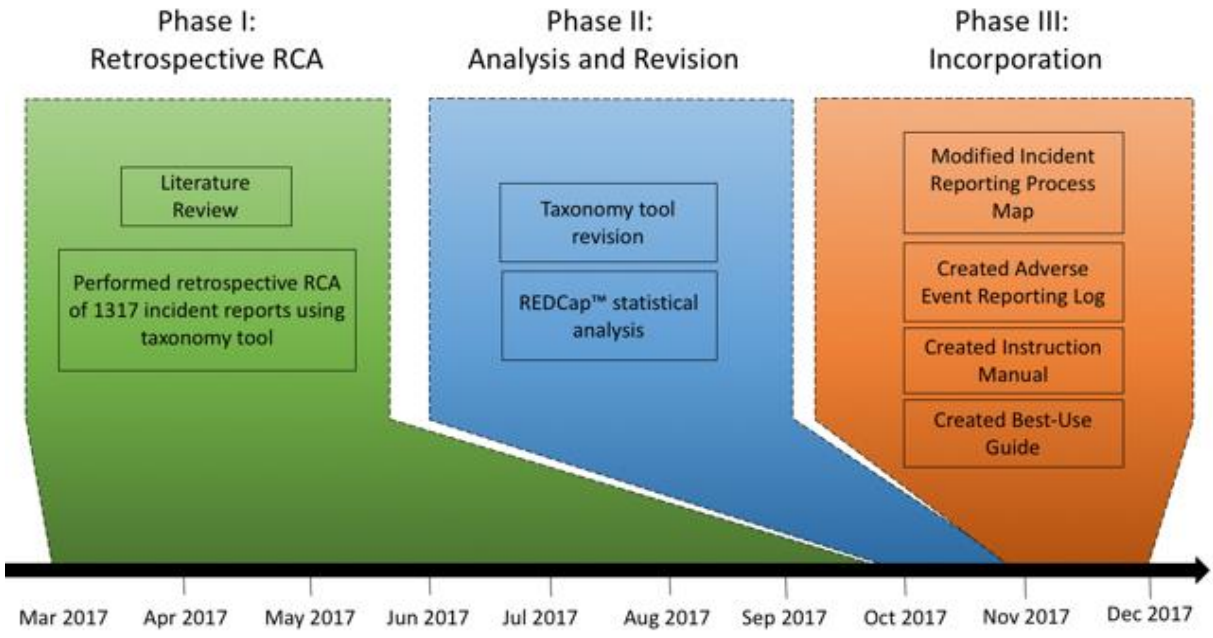

Figure 8. Timeline of QI project.

Phase I was the trial run in which the taxonomy tool was used to perform retrospective root cause analysis of 1317 ICU incident reports. The review was started in May 2017 and completed in October 2017.

Phase II focused on data analysis and also tool revision. While data analysis was completed using REDCap™ statistical analysis, the tool revision was based on issues that arose during the

trial run. Some of the issues included redundancy in classification selections, difficulty in classifying certain incident reports, and ambiguity in determining harm severity. Overall, the taxonomy tool modifications were made to improve functionality and reduce inter-rater variability. Phase II was completed in November 2017.

In phase III, a plan was proposed for incorporating use of the taxonomy tool into standard error review practice. Implementation also meant ensuring proper use of the taxonomy tool in future analyses. Documents like an adverse event reporting log, instruction manual, and best-use guide, were created to teach others how to use the tool. Phase III was completed in December 2017.

Phase II analysis used the classification data collected from incident reports in Phase I to characterize existing system errors. Under the 'medical error type' primary classification, a 55.3% majority of the incident reports provided 'insufficient information' for proper characterization. 'Problematic execution' accounted for 25.6% of the error types, followed by 'technical/mechanical' error at 8.6%, 'problematic decision' making at 5.8%, and 'problematic communication' at 4.7%. As for the medical domains in which medical errors occurred most often, 46.5% of the reports described situations occurring during 'perioperative care.' For incident reports originating from the SICU and CVICU, the 'perioperative care' classification was used for all incidents occurring immediately before, during, and after a surgical operation. In the case of MICU incident reports, the same term was used to describe direct patient care incidents occurring while the patient was on the MICU floor. The second and third most frequent subclassifications were 'medication' domain at 17.9% and 'diagnostic' domain at 17.8%. 8.7% of the reports described 'equipment' errors, with a much smaller percentage related to 'patient communication,' 'transfusion-related,' and 'administrative.' Figure 9 shows a Pareto chart representation of the medical domains in which medical errors were found to have occurred.

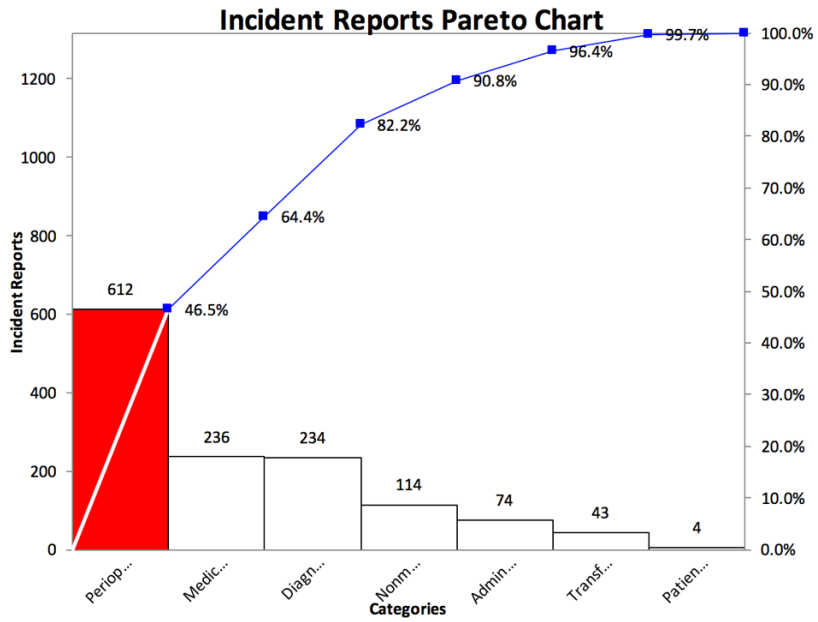

Figure 9. Pareto Chart with Frequency of Medical Domain Reports.

Each incident report was also classified according to perceived contributing factors. The top 3 contributing factors were 58.3% from staff, 52.1% for system errors, and 45.3% attributed to the patient. Only 18.6% reports described work environment factors, and 7.1% described leadership issues.

Among those reports with some component of staff contribution, 74.7% were deemed to be related to performance, 43.2% to cognitive, 19.9% to communication, and 5.6% to behavior. Patient-related medical errors were predominantly related to patient pathophysiology at 91.2%. Of the events with work environment as a contributing factor, 83.8% were from hospital policy and 20.4% from workload. When leadership was a contributing factor, 41.9% were because of a poor process, 26.9% due to lack of oversight, 21.5% were education issues, and 11.8% due to lack of follow-up. Finally, system errors were predominantly equipment-related issues at 69.8%, with clinical communication, medication processing, and infection control occurring less frequently at 23.5%, 12.2%, and 6.3%, respectively.

While most medical errors described in the ICU incident reports did cause patient harm, there was variability in severity of harm. Having a risk stratification system is useful for determining which specific areas can benefit most from targeted intervention and usage of

limited resources. From the analysis, 61% of the reports described errors that caused temporary harm to the patient, while 21.5% did not cause any patient harm. 4.7% errors involved the patient and required life-sustaining measures, and 1.9% of the reports described errors that contributed to patient death. A benefit of building this taxonomy database is how each specific incident can be plotted on a risk assessment matrix as a visual representation of its severity vs. likelihood. This can help direct attention and limited resources towards the most high-priority system errors. Figure 10 shows an example risk assessment matrix by The National Patient Safety Agency, and Figure 11 is its color key.

Risk Grading Matrix		Likelihood score				
		5	4	3	2	1
Consequence score		Almost certain	Likely	Possible	Unlikely	Rare
	5 Catastrophic	25	20	15	10	5
	4 Major	20	16	12	8	4
	3 Moderate	15	12	9	6	3
	2 Minor	10	8	6	4	2
	1 Negligible	5	4	3	2	1

Figure 10. Risk Assessment Matrix. Adapted from the NHS National Patient Safety Agency.[18]

Green	1 - 3	Low risk
Yellow	4 - 6	Moderate risk
Amber	8 - 12	High risk
Red	15 - 25	Extreme risk

Figure 11. Assessing and Grading the Risk Assessment Matrix.

Because of the propensity for medical errors to occur during the window of patient care transfer between health professionals, there was particular interest in errors related to handover. The analysis showed that 78.4% of the errors were unlikely related to problematic handover, with 16.2% possibly related to handover, and only 5.5% probably related to handover. Currently, there are ongoing UT Southwestern quality improvement projects seeking to improve patient handover in the OR-ICU and OR-OR settings. Using the taxonomy tool to analyze patient safety incident reports adds value because of its ability to generate data that can help guide and inform other QI initiatives.

99.7% (1314/1317) of the incident reports that fit the parameters of this project were reviewed and classified using the taxonomy tool. The remaining 3 reports were submitted with insufficient information such that they could not be properly classified.

INTERPRETATION

This project was designed with the dual purposes of: 1) developing and testing an in-process classification tool to aid in identifying trends in medical errors, and 2) organizing the analysis in a way to facilitate greater understanding of underlying factors and formation of solutions. The main interventions were the introduction, testing, and modification of a taxonomy tool for reviewing patient safety incident reports. From these interventions, value was created in the form of baseline data on system errors, and a realistic framework was proposed for integrating the tool's use into UT Southwestern's existing incident reporting system. However, seeing improvements in patient safety and fixing latent system errors takes time. The various types of system errors identified in this project will benefit from future QI interventions tackling them individually. Only after post-intervention data is collected will the true impact of this project become apparent.

Although there are some similarities in the baseline data obtained in this study and the findings from published literature, there are also areas of discrepancy. According to the Harvard Medical Practice Study, diagnostic errors account for 17% of preventable errors in hospitalized patients,[19] and analysis from this project found diagnostic errors to constitute 17.8% of the total. However, at least one 2006 study of electronic hospital incident reporting systems found medication errors to account for 33%, nonmedication errors for 34%, and administrative errors for 13% of total errors.[20] In contrast, the REDCap™ analysis found only 17.9% medication errors, 8.7% nonmedication errors, and 5.6% administrative errors. This is a sizable disparity, but it is one that can be partly explained by the different classification categories used in the respective studies. The REDCap™ tool includes more classification category options, including a catch-all category for errors occurring during perioperative care. 46.5% of the medical errors reviewed in this QI project fell into this category, essentially diluting the results for the remaining categories.

Pertaining to patient harm, there is continued debate on the actual number of patients who experience harm as a result of medical errors. A 2013 paper published in the Journal of Patient Safety looked at four studies that analyzed medical death rate data from 2002 to 2008. It estimated that death occurred in up to 1.4% of medical error cases, with serious harm occurring 10- to 20-fold more than that.[21] REDCap™ analysis found 61% medical errors causing temporary harm, 4.7% requiring life-sustaining measures, and 1.9% resulting in death. The death rates are comparable, but variation should be expected because harm classifications are very dependent on the judgment of the person reviewing the report. Furthermore, while the 9-category NCC MERP harm scale was chosen for use in this project, use of different algorithms can also cause discrepancies. Another REDCap™ statistic is that 21.6% of medical errors were determined to be possibly or probably related to handover. This is consistent with findings from a 2014 study examining the change in rates of medical errors following implementation of I-PASS handoff bundles, which reported a pre-intervention handoff error rate of 24.5%. The I-PASS handoff bundle standardized communication and handoff training, employed computerized tools and supervisors, and included a promotional campaign to influence institutional process and culture. In the 9 academic hospitals involved in the study, overall rate of medical errors decreased from 24.5% to 18.8%, perhaps indicating the potential for similar patient handover improvements at UT Southwestern.[22]

Classification data obtained from using the taxonomy tool is prone to having significant variability, especially when different people are performing the error review. This inter-rater variability can be caused by disparity in medical understanding, interpretation of events, understanding of the taxonomy tool itself, or bias from recent incident report classifications. Outputs from this QI project, including the adverse event reporting log, instruction manual, and best-use guide, were created with the explicit purpose of helping to reduce this variability. In the future, further reducing variability should be the main priority because it is a prominent obstacle that can severely reduce the utility of this taxonomy tool.

LIMITATIONS

This QI project has a number of important limitations. Inter-rater variability presents as one of the more prominent limitations as it will likely persist in some degree, regardless of efforts to reduce it. Retrospective root cause analysis used in this project is possible because there is a human person manually reviewing incident reports and deciding how to classify them. Despite drawbacks of this approach, the author of this paper still believes this to be the best option. Incident reports are purposely free text narratives so that submitters have liberty to explain and give details about the patient safety event. Automated systems would be unable to read between the lines, detect subtleties and nuances in written language, or make interpretations using clinical judgment. Two strategies that can be used in this QI project to decrease inter-rater variability are to maintain rigorous training requirements for Clinical Safety personnel and to simplify the tool as much as possible.

Relying on a voluntary error reporting system is inherently limiting. It can be useful for highlighting patient safety events that health providers perceive as important, but it also captures a nonrepresentative fraction of adverse events. One notable absence is the voice of the patient. Currently, there is no structure at UT Southwestern for patients to personally submit their own incident reports. Unfortunately, other safety measurement strategies also have their own inherent advantages and disadvantages. Additionally, patient safety event reports only give a snapshot in time. The incident reporting system advocated in this QI project is not able to follow-up with patients and precisely determine the harm they suffer as a result of the medical error. Practically, this means that Clinical Safety personnel will often be forced to make speculations on patient outcome during the error review.

Another limitation of this project is its use of retrospective root cause analysis for identifying system errors. The retrospective nature means that Clinical Safety personnel will always be fixing errors instead of preventing them. More ideal for reducing patient safety events would be a prospective method such as failure modes effect analysis (FMEA). Retrospective root cause analysis was used in this QI project because it remains an effective method for identifying underlying factors contributing to medical errors.

The taxonomy tool itself is a potential limitation. The tool used in this QI project was created by Dr. Isaac Lynch as a version of previous models adapted for use in real-time monitoring and in the ICU setting. However, there is continued disagreement in the medical community on ideal design, classification categories, and terminology. Nobody has yet to create a standardized taxonomy tool robust enough to encompass the diversity of health care settings. For instance, classification categories may be different depending on whether the tool is being used to study medical errors in the large, academic inpatient setting vs. an outpatient family practice clinic. Development of a common terminology and taxonomy for classifying event reports remains a persistent obstacle to achieving standardization.

Finally, generalizability of the results from this project is limited by context. This QI project was conducted using ICU incident reports from a large, academic hospital in a major Texas metropolitan area. ICU incident reports will likely be different than those submitted from general inpatient floors. There may also be differences depending on public vs. private hospitals, rural vs. urban, and depending on region of the world. For example, Clements University Hospital is a tertiary referral center that receives many patients transfers from other health centers. During the trial error review conducted in this project, a sizable proportion of event reports were submitted by nurses who noticed patient pressure ulcers during admission from outside facilities. The frequency of this type of patient safety incident will likely differ based on hospital context. Other limiting context factors include hospital culture of error reporting and ease of reporting. At UT Southwestern, there is a system-wide effort to shift the culture away from individual blame, and there is also an ongoing quality improvement project focused on making it easier to report patient safety incidents.

CONCLUSIONS

The distinctive feature of this QI project was its demonstration of the potential of using a formal taxonomy tool for reviewing patient safety incident reports, something that is not currently part of the error review process at UT Southwestern. The retrospective root cause analysis performed in this project was one of the largest systematic reviews of incident reports to date at UT Southwestern. It succeeded in providing quantifiable data on the different types of

medical errors, underlying contributing factors, and impact on patient outcomes. It also provided useful baseline data by which the effects of future intervention can be observed.

Creating an ideal taxonomy tool using common terminology has remained an elusive goal for government agencies, health care groups, and private sector organizations. Not only does this present a patient safety problem, but also it has serious health policy implications. As the world's aging population continues to grow, use of health care services will increase in tandem. This further emphasizes the need for organization of patient safety information systems and efforts to reduce preventable medical errors. Health care costs continue to rise as well, and policymakers will be expected to make judicious allocation of limited health care dollars. Use of the taxonomy tool has the potential to assist with financial stewardship because collected data can be plotted on risk assessment matrices, allowing for triaging of resources towards those system errors that are identified as causing greatest patient harm or that are occurring most frequently.

This project only tested the tool at Clements University Hospital, but a strength of the project is its potential to become the standardized error review process for the entire UT Southwestern health system. This is possible because of the taxonomy tool's simple design and transferability to multiple health care settings. In this project, a modified error review workflow has already been proposed. Incorporating use of the taxonomy tool in all UT Southwestern-associated hospitals would ensure quality control in the error review process because each event report would at least be subjected to a minimum level of review/analysis. Standardization would also enable comparison of data between hospitals and be useful for future public health studies. There is potential for far-reaching positive impact on patient safety.

Part of the original vision for creating this new taxonomy tool was for Clinical Safety personnel to utilize it for real-time error review of patient safety incident reports. There are a number of steps that need to be taken to realize this vision. For instance, further refinement of the tool is needed to minimize inter-rater variability. This means having other medical students perform more trial analyses using the taxonomy tool and monitoring the variability in results. It will also require buy-in from hospital administrators and the personnel who will be using the tool. In the near future, implementation of a dashboard module may allow Clinical Safety

personnel to perform analysis with the taxonomy tool and be able to immediately detect medical error trends in real-time. Once a sizeable database of past error events is accumulated from multiple rounds of taxonomy analysis, another future goal would be to integrate use of human language algorithms into the error review process. These algorithms, which have already been developed, would teach computers to process and understand human languages. Eventually, it would be possible for Clinical Safety to use verbal keywords or phrases to easily search through the entire UT Southwestern error review database for specific incidents of interest. Clearly, use of the taxonomy tool has the potential to significantly improve the error review process for incident reports, and it will be the responsibility of future studies to further develop the tool and its integration into existing processes.

REFERENCES

[1] **Allen, Marshall, and Olga Pierce**. "Medical Errors Are No. 3 Cause Of U.S Deaths, Researchers Say." *NPR*, NPR, 3 May 2016, www.npr.org/sections/health-shots/2016/05/03/476636183/death-certificates-undercount-toll-of-medical-errors.

[2] **Marx, David**. "Patient Safety and the "Just Culture": A Primer for Health Care Executives." *PSNet: Patient Safety Network*, Agency for Healthcare Research and Quality, 17 Apr. 2001, psnet.ahrq.gov/resources/resource/1582/patient-safety-and-the-just-culture-a-primer-for-health-care-executives--.

[3] **Reason, J, J Carthey, and M R de Leval**. "Diagnosing 'Vulnerable System Syndrome': An Essential Prerequisite to Effective Risk Management." *Quality in Health Care : QHC* 10.Suppl 2 (2001): ii21–ii25. *PMC*. Web. 20 Feb. 2018.

[4] **Farley DO, Haviland A, Champagne S, Battles JB, *et al***. Adverse-event-reporting practices by US hospitals: results of a national survey. *Qual Saf Health Care*. 2008; 17(6): 416-423. https://www.ncbi.nlm.nih.gov/pubmed/19064656.

[5] **"World alliance for patient safety: WHO draft guidelines for adverse event reporting and learning systems: from information to action."** *Institutional Repository for Information Sharing*, Geneva: World Health Organization, 2005, www.who.int/iris/handle/10665/69797.

[6] **Hutchinson A, Young TA, Cooper KL, *et al***. Trends in healthcare incident reporting and relationship to safety and quality data in acute hospitals: results from the National Reporting and Learning System. *Qual Saf Health Care*. 2009; 18(1):5-10. https://www.ncbi.nlm.nih.gov/pubmed/19204125.

[7] **"Measurement of Patient Safety."** *PSNet: Patient Safety Network*, Agency for Healthcare Research and Quality, June 2017, psnet.ahrq.gov/primers/primer/35/measurement-of-patient-safety.

[8] **Levinson, DR**. "Hospital Incident Reporting Systems Do Not Capture Most Patient Harm." *PSNet: Patient Safety Network*, US Department of Health and Human Services, Office of the Inspector General, Jan. 2012. Report No. OEI-06-09-00091. psnet.ahrq.gov/resources/resource/23842/hospital-incident-reporting-systems-do-not-capture-most-patient-harm.

[9] **Evans, S M et al**. "Attitudes and Barriers to Incident Reporting: A Collaborative Hospital Study." *Quality & Safety in Health Care.* 2006; 15.1: 39-43. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2563993/.

[10] **Benn J, Koutanji M, Wallace L, Spurgeon P, Rejman M, Healey A, Vincent C**. "Feedback from incident reporting: information and action to improve patient safety." *Qual Saf Health Care.* 2009; 18(1): 11-21. https://www.ncbi.nlm.nih.gov/pubmed/19204126.

[11] **Nuckols, Teryl K**. "Incident Reporting: More Attention to the Safety Action Feedback Loop, Please." *PSNet: Patient Safety Network*, Agency for Healthcare Research and Quality, Sept. 2011, psnet.ahrq.gov/perspectives/perspective/108.

[12] **Williams, Tamara, et al**. "The Reliability of AHRQ Common Format Harm Scales in Rating Patient Safety Events." *Journal of Patient Safety*, Journal of Patient Safety, 1 Mar. 2015, insights.ovid.com/pubmed?pmid=24080718.

[13] **"High Reliability."** *PSNet: Patient Safety Network*, Agency for Healthcare Research and Quality, November 2017, https://psnet.ahrq.gov/primers/primer/31/high-reliability.

[14] **"Read "To Err Is Human: Building a Safer Health System" at NAP.Edu."** *National Academies Press: OpenBook*, The National Academies of Sciences, Engineering, Medicine, 2000, www.nap.edu/read/9728/chapter/10#156.

[15] "**Minimal Information Model for Patient Safety."** *World Health Organization*, World Health Organization, www.who.int/patientsafety/implementation/information_model/en/.

[16] **"Table 6: Categories of Medication Error Classification."** *AHRQ--Agency for Healthcare Research and Quality: Advancing Excellence in Health Care*, U.S. HHS: Agency for Healthcare Research and Quality, 1 Aug. 2012, www.ahrq.gov/professionals/quality-patient-safety/patient-safety-resources/resources/match/matchtab6.html.

[17] **Forrey, Ryan A, et al**. "Interrater agreement with a standard scheme classifying medication errors." *American Journal of Health-System Pharmacy*. Jan. 2007; 64(2): 175-81. www.ncbi.nlm.nih.gov/pubmed/17215468.

[18] **"A risk matrix for risk managers."** *NHS*, The National Patient Safety Agency, Jan. 2008, www.nrls.npsa.nhs.uk/EasySiteWeb/getresource.axd?AssetID=60149.

[19] **"Diagnostic Errors."** *PSNet: Patient Safety Network,* Agency for Healthcare Research and Quality, June 2017, https://psnet.ahrq.gov/primers/primer/12.

[20] **Milch, C E, et al**. "Voluntary electronic reporting of medical errors and adverse events. An analysis of 92,547 reports from 26 acute care hospitals." *Journal of general internal medicine.*, U.S. National Library of Medicine, Feb. 2006; 21(2): 165-70. www.ncbi.nlm.nih.gov/pubmed/16390502.

[21] **James, John T**. "A New, Evidence-Based Estimate of Patient Harms Associated with Hospital Care: Journal of Patient Safety." *Wolters Kluwer*, Journal of Patient Safety, Sept. 2013; 9(3): 122-128. doi: 10.1097/PTS.0b013e3182948a69. journals.lww.com/journalpatientsafety/Fulltext/2013/09000/A_New,_Evidence_based_Estimate_of_Patient_Harms.2.aspx.

[22] **Fliesler, Nancy**. "Safer patient Handoffs." *Harvard Medical School*, 18 Nov. 2014, hms.harvard.edu/news/safer-patient-handoffs.

VITAE

Christopher Chan is a fourth year medical student at UT Southwestern who will graduate from the institution in 2018 with an M.D. with Distinction in Quality Improvement and Patient Safety. Afterwards, he intends to pursue a residency in Anesthesiology. A native of Houston, Texas, born in 1989, he obtained his B.A. degree in Public Policy Studies from Duke University in 2012.

As an undergraduate, Christopher discovered a passion for quality improvement in 2009 while working as a social entrepreneur with the non-profit TROSA. His market research and creation of a complete business plan was instrumental in the eventual establishment of the TROSA grocery store to combat food insecurity in a designated food desert. He was also involved in student-led initiatives to address perceived shortcomings in the student experience. Christopher launched the Civic Engagement initiative to help Duke students reflect on and make meaning out of high-impact experiences, and his work on the Duke Way initiative stimulated student reflection on the discrepancy between personal values vs. Duke student culture values. In addition, he consulted for the non-profit Clean Energy Durham and helped develop an effective neighbor-to-neighbor energy savings rental protocol.

He also engaged in the policymaking process and worked to inform health policy at the state, municipal, and school levels. With the sponsorship of a state Congressman, Christopher co-produced a memo with recommendations for combating adult obesity that was distributed in the North Carolina General Assembly. In 2011, he interned at the Dept. of Community Development and produced an operational plan that led to the implementation of a coordinated intake system to streamline all City of Durham homeless services.

Christopher continued his passion for quality improvement in medical school. In 2015, he worked on a QI project under the mentorship of Dr. Pranavi Sreeramoju and created a defect analysis tool for reviewing surgical site infections at Parkland Memorial Hospital as part of Project RITE: Reducing Infections Together in Everyone. He presented his project poster at several conferences, including the IHI National Forum on QI in Orlando and the UT Shared Visions conference in San Antonio. From 2017-2018, as his capstone project for the M.D. Distinction in QI/PS, Christopher worked with Dr. Isaac Lynch to improve the error review process at UT Southwestern through the use of a taxonomy tool.