P8

Project Summary / Abstract

Broad Objectives: The long-term objective of this work is to improve technical performance and surgical safety in the operating room.

Specific Aims: This proposal seeks to use several innovative approaches to study intra-operative technical errors, a common but poorly understood threat to patient safety. Our data source contains detailed clinical information on surgical patients treated at community hospitals, an understudied sector of the U.S. healthcare system. We will devise a novel approach to measuring the impact that intra-operative technical errors have on the post-operative clinical course and will apply human factors methodology to identifying common system features that contribute to or mitigate these events. This project has 3 specific aims:

Aim 1. Study the impact of an intra-operative technical error on the clinical severity of a patient's post-operative course.

Aim 2. Determine the institutional rates of APL or POHH with 'clinically-significant post-operative deviation' as an indicator of institutional patient safety.

Aim 3. Use qualitative field observations and human factors analyses in an on-site study of the system factors at hospitals with the highest and lowest rates of clinically-significant intra-operative technical errors.

Research Design and Methods: Our cohort will include over 400,000 patients having an operation between 2008 or 2010 at one of 65 community hospital partners of the [redacted] Available data includes administrative and billing data, as well as detailed clinical information on the hospital course. For patients who do and do not experience an intra-operative technical error, we will first calculate a score that measures the severity of the patient's post-operative course using a previously validated classification system based on detailed clinical data. We will then develop and validate an algorithm based on administrative and billing data that estimates this score and can be widely generalizable. Finally, we will use human factors methodology to investigate the system features that contribute to and mitigate intra-operative technical errors and their post-operative impact.

Potential Impact: Over half of all adverse events experienced by hospitalized patients are surgical in nature. Of these surgical events, 75% occur in the OR and 50% are due to a technical error. For these reasons, targeting technical errors that occur in the OR has the greatest potential to improve surgical safety and save lives.

Project Narrative

This proposal seeks to use several innovative approaches to study intra-operative technical errors, a common but poorly understood threat to patient safety. Our data source contains detailed clinical information on over 400,000 surgical patients treated at community hospitals, an understudied sector of the U.S. healthcare system. We will devise a novel approach to measuring the impact that intra-operative technical errors have on the post-operative clinical course and will apply human factors methodology to identifying common system features that contribute to or mitigate these events.

[redacted] RESOURCES

FACILITIES: Specify the facilities to be used for the conduct of the proposed research. Indicate the project/performance sites and describe capacities, pertinent capabilities, relative proximity, and extent of availability to the project. If research involving Select Agent(s) will occur at any performance site(s), the biocontainment resources available at each site should be described. Under "Other," identify support services such as machine shop, electronics shop, and specify the extent to which they will be available to the project. Use continuation pages if necessary.

Laboratory:

No laboratory facilities will be used in this grant.

Clinical:

No clinical facilities will be used in this grant.

Animal:

No animal facilites will be used in this grant.

Computer:

The PI and her colleagues have access to SAS, MS Access and other analytical applications. Further informational systems (IS) support can be obtained from [redacted] IS support staff and programmers. All PCs are networked and have access to email and Netscape Navigator.

Office:

The [redacted] at the [redacted] occupies approximately 16,000 net square feet in the [redacted] and has ample administrative support and standard office supplies.

Other:

CSPH is located across street from the [redacted] and [redacted]. The [redacted] on campus houses an extensive collection and online resources for books and academic journals.

MAJOR EQUIPMENT: List the most important equipment items already available for this project, noting the location and pertinent capabilities of each.

[redacted] RESOURCES

FACILITIES: Specify the facilities to be used for the conduct of the proposed research. Indicate the project/performance sites and describe capacities, pertinent capabilities, relative proximity, and extent of availability to the project. If research involving Select Agent(s) will occur at any performance site(s), the biocontainment resources available at each site should be described. Under "Other," identify support services such as machine shop, electronics shop, and specify the extent to which they will be available to the project. Use continuation pages if necessary.

Laboratory:

N/A

Clinical:

N/A

Animal:

N/A

Computer:

The production systems are running on IBM x3X50 servers with Red Hat Enterprise Linux operating systems, Oracle 10.2.0.1 and Postgres 8.2.11. We have 15 Terebytes of storage.The [redacted] Data Center has Intrusion Detection/Prevention systems and explicit firewall rules to secure this highly sensitive information.

Office:

[redacted] has 5000 sq ft of office space in [redacted] where staff of experienced data system, engineering and operations and network personnel to maintain the system are located.

Other:

Access to [redacted] offices and data systems are controlled by biometric identification.

MAJOR EQUIPMENT: List the most important equipment items already available for this project, noting the location and pertinent capabilities of each.

Data in the [redacted] data system originates from the Meditech Clinical Information Systems (CIMS) located at each of the participating hospitals. The Meditech system is an operational software system, written in Magic, a proprietary version of MUMPS, and is not a relational data system. [redacted] collects these data and standardizes them, storing the data in an Oracle data system running on LINUX servers. All data is encrypted when sent via VPN to the [redacted] Data Center. Encryption converts data to an unintelligible form called ciphertext; decrypting the ciphertext converts the data back into its original form, called plaintext. Combining explicitly defined nodes, pre-shared keys, 168 to 265-bit encryption and secure password protection standards allows for a more than adequate way of securely transferring sensitive medical data.

Equipment

[redacted]

[redacted] currently has access to [redacted] supported desk top computers, lap tops and other standard office equipment such as printers, scanners, copiers and fax machines. All computer devices are encrypted and protected with fire wall capabilities.

[redacted] is supported in this proposal by the [redacted] which includes full time research coordinators, administrative assistants, research fellows and the center Administrator who also all have access to individual networked computer work stations and printers as well as other shared office equipment such as copiers, faxes, and scanners.

The proposed budget that is included in this proposal requests the purchase of a computer that will physically be located at [redacted] institute for the [redacted] research team to use in order to access [redacted] data on site and off.

Institute for Health Metrics

[redacted] currently has access to standard office equipment provided by the [redacted] including a desk top computer, lap top computers and other standard office equipment such as printers, fax machines and scanners.

The production systems are running on IBM x3X50 servers with Red Hat Enterprise Linux operating systems, Oracle 10.2.0.1 and Postgres 8.2.11. We have 15 Terebytes of storage.The IHM Data Center has Intrusion Detection/Prevention systems and explicit firewall rules to secure this highly sensitive information.

PHS 398 Cover Page Supplement

4. Human Embryonic Stem Cells

* Does the proposed project involve human embryonic stem cells? ☒ No ☐ Yes

If the proposed project involves human embryonic stem cells, list below the registration number of the specific cell line(s) from the following list: http://stemcells.nih.gov/research/registry/. Or, if a specific stem cell line cannot be referenced at this time, please check the box indicating that one from the registry will be used:

Cell Line(s): ☐ Specific stem cell line cannot be referenced at this time. One from the registry will be used.

Tracking Number:GRANT10550226 Funding Opportunity Number:PA-09-070 Received Date:2010-03-05T12:34:59-04:00

OMB Number: 0925-0001

PHS 398 Research Plan

1. Application Type:

From SF 424 (R&R) Cover Page. The response provided on that page, regarding the type of application being submitted, is repeated for your reference, as you attach the appropriate sections of the Research Plan.

*Type of Application:

☐ New ☒ Resubmission ☐ Renewal ☐ Continuation ☐ Revision

2. Research Plan Attachments:

Please attach applicable sections of the research plan, below.

1. Introduction to Application (for RESUBMISSION or REVISION only)	1241-Introduction.pdf	Add Attachment	Delete Attachment	View Attachment
2. Specific Aims	1242-Specific Aims.pdf	Add Attachment	Delete Attachment	View Attachment
3. *Research Strategy	1243-ResearchStrategy.pdf	Add Attachment	Delete Attachment	View Attachment
4. Inclusion Enrollment Report		Add Attachment	Delete Attachment	View Attachment
5. Progress Report Publication List		Add Attachment	Delete Attachment	View Attachment

Human Subjects Sections

6. Protection of Human Subjects	1251-Protection of Human Sub	Add Attachment	Delete Attachment	View Attachment
7. Inclusion of Women and Minorities	1252-Inclusion of Women and	Add Attachment	Delete Attachment	View Attachment
8. Targeted/Planned Enrollment Table	1253-TargetedEnrollmentTabl	Add Attachment	Delete Attachment	View Attachment
9. Inclusion of Children	1254-Inclusion of Children.p	Add Attachment	Delete Attachment	View Attachment

Other Research Plan Sections

10. Vertebrate Animals		Add Attachment	Delete Attachment	View Attachment
11. Select Agent Research		Add Attachment	Delete Attachment	View Attachment
12. Multiple PD/PI Leadership Plan		Add Attachment	Delete Attachment	View Attachment
13. Consortium/Contractual Arrangements	1255-Karcz_StatementofIntent	Add Attachment	Delete Attachment	View Attachment
14. Letters of Support	1256-LOSCombinedFinal.pdf	Add Attachment	Delete Attachment	View Attachment
15. Resource Sharing Plan(s)	1257-Resource Sharing.pdf	Add Attachment	Delete Attachment	View Attachment

16. Appendix Add Attachments Remove Attachments View Attachments

INTRODUCTION:

This proposal is a revision of application 1 RO1 HS018689 entitled "Characterization of surgical complications to improve prevention and quality review". The scope and direction of the project has been significantly altered in response to the critiques provided in the summary statements. The scientific alterations are such that it is not possible to highlight them within the text as nearly the entire grant has been revised. In addition, given the larger scope and increased emphasis on a mixed methods approach, there has been a change in the Principal Investigator. [redacted] from the [redacted] who was a previous consultant, is now the PI in place of [redacted] from the [redacted]. This change reflects [redacted] experience with both the statistical techniques required for large database analyses as well as qualitative research methodology. [redacted] and [redacted] have worked together on the design of both the original and this revised proposal.

The original submission was a 2 year project that focused on characterizing the process and outcomes associated with 3 patient safety indicators that represent complications of surgery using administrative data. The focus of the project remains to understand the process and outcomes associated with 2 intra-operative technical errors identified by these patient safety indicators however the depth, breadth and application of these factors is markedly increased. We have increased the timeline to 5 years to allow us to understand the relationship between the clinical sequelae as well as the system factors that impact the occurrence of these 2 potentially devastating events. In addition, we have broadened our methodology to include complex multivariable modeling and qualitative site visits employing a human factors approach.

We feel that the revised proposal has addressed all of the critiques that were provided in the summary statement. However, we would like to specifically address the 2 major concerns raised with the first review of this proposal for consideration as the revised protocol is read. The first concern was that the innovation and impact of the study was limited by the nature of the proposed study design, namely a secondary data analysis. We have expanded this in 2 ways. First, we are proposing to develop a novel approach to estimating the severity of the post-operative course following the occurrence of an intra-operative event. This algorithm will be applicable across the healthcare system and dramatically improve our ability to stratify complications following surgery. Second, a significant component of this grant is based on a human factors analysis of the system features of high and low performing hospitals. As the first review pointed out, this is the logical next step in this project and will lead to a significant advancement in our understanding of the etiology and approach to prevention of intra-operative technical errors.

SPECIFIC AIMS:

Intra-operative technical errors are a poorly understood but critical component of surgical safety. Of surgical adverse events, 75% occur in the OR and 50% are due to a technical error. Despite this, little is known about the systemic features that contribute to or mitigate these events or the clinical sequelae following an intra-operative technical error. This is at least in part due to a lack of a systematic approach to identifying and studying these errors.

Administrative data can be used to identify two markers of intra-operative errors: accidental puncture and laceration (APL) and post-operative hemorrhage and hematoma (POHH). These measures have been previously validated and are endorsed by the Agency for Healthcare Research and Quality (AHRQ) as Patient Safety Indicators (PSI). Once identified, a detailed chart review or root cause analysis is often performed to evaluate the severity of the event and identify potential contributing factors. Such an approach based on individual case review is transactional and poorly equipped to inform system redesign.

The Accordian Severity Grading Scale (ASGS) is the most recent iteration of a classification system for the severity of post-operative complications that was first introduced in 1992. While it is a validated, objective approach to grading the post-operative course, this scale requires detailed clinical detail, which to date has only been available by chart review which is time-consuming, inefficient, and difficult to implement on a large, multi-institutional scale.

The Institute for Health Metrics (IHM) collects real-time electronic data from a cohort of 65 community hospitals, including detailed clinical, pharmacy, and laboratory data from the hospital information system, as well as traditional administrative data. Using this uniquely rich data source, we can identify APL and POHH events and study the details of their post-operative course to calculate a score using the ASGS. Using the ASGS as a gold standard, we can then develop and validate an algorithm based on healthcare cost and other measures of resource utilization to grade the severity of the post-operative course. Once developed, this methodology will provide a generalizable approach that is applicable in diverse healthcare systems to stratify the clinical sequelae of intra-operative technical errors without the need for detailed clinical data.

We will then systematically examine the structure and process of hospitals that are outliers with respect to clinically-significant intra-operative technical errors. We will apply human factors methodology to understand the system factors that contribute to and mitigate intra-operative technical errors and their clinical sequelae with the aim of devising interventions to decrease the occurrence and impact of these harmful and costly events.

Aim 1. Study the impact of an intra-operative technical error on the clinical severity of a patient's post-operative course.

H1: Hospitalizations in which an intra-operative technical error occurs are associated with a more complicated post-operative course.

H2: The severity of the hospital course is correlated with increased healthcare cost.

H3: An algorithm based on administrative data, including resource utilization and cost, can be developed to stratify the clinical severity of the post-operative course.

Aim 2. Determine the institutional rates of APL or POHH with 'clinically-significant post-operative deviation' as an indicator of institutional patient safety.

H1: Patient-level APL or POHH events with 'clinically-significant post-operative deviation' can be aggregated to calculate a hospital rate analogous to the AHRQ PSIs.

H2: Institutions that are high and low outliers can be identified as targets further evaluation.

Aim 3. Use qualitative field observations and human factors analyses in an on-site study of the system factors at hospitals with the highest and lowest rates of clinically-significant technical errors.

H1: Institutional structure and process influences technical performance and can be the target of system improvement initiatives to decrease clinically significant technical errors in the OR.

SIGNIFICANCE

I. Problem or critical barrier to progress in the field to be addressed

A. System and Technical Performance in Surgical Safety

Previous studies have shown that both system factors and individual technical performance in the operating room (OR) influence safety and outcomes. [redacted], using confidential interviews at three teaching hospitals showed that 77% of errors were related to an operation or an invasive intervention [1]. In another study using malpractice claims, system performance was found to contribute to error in 82% of surgical cases that led to patient harm[2]. The leading human factor associated with error was inexperience or lack of technical competence (41%). In both the interview and claims studies, the majority of cases had more than one identifiable system factor that contributed to the error. In a follow-up study, [redacted] found that 53% of injuries to surgical patients were due to technical errors and 21% involved systems failures[3]. These studies suggest that system factors and technical performance in the OR are two critical targets for improving surgical safety and saving lives.

Surgical safety researchers have begun to adapt methodology developed in other high risk domains to the operating room [4-6]. Safety research in other complex work environments, such as nuclear reactor control rooms and aviation, is based on a human factors approach, a theory that latent failures predispose to the occurrence of sentinel events [7-15]. System and process changes can be designed to eliminate these failures. While there has been recognition in surgery that this approach may be useful in improving surgical safety, the complex interplay between system factors and technical performance in the OR has not been investigated. This is at least partially due to a lack of a prospective sampling strategy to target evaluations, which has limited safety research to a transactional, case-based approach.

B. AHRQ Surgical Patient Safety Indicators

The Agency for Healthcare Research and Quality (AHRQ) has published evidence-based patient safety indicators (PSI) that can identify potential adverse events occurring during a hospitalization. Two of the PSI, accidental puncture or laceration (APL) and post-operative hemorrhage or hematoma (POHH) are complications that reflect the technical performance of an operative procedure and are detected during an acute hospitalization. These 2 events, APL and POHH, will be the focus of this proposal.

According to AHRQ, the differences among hospitals in the rates of these complications are low and do not appear to be associated with hospital characteristics [16]. However, there is reason to believe that this lack of institutional variation reflects limitations in the identification of clinically-significant events based on administrative data. The positive predictive value of surgical complication codes have been demonstrated to be in the 60-70% range by [redacted] et al and others[17]. Gallagher et al studied accidental puncture or laceration cases and found that of a sample of 67 cases reviewed, 3% were not truly APLs. They found that an additional 15% of cases were not serious according to expert review[18]. In a study from 1994, Hartz and Kuhn estimated the validity of hemorrhage codes using a gold standard based on transfusion "requirement." They identified only 26% of episodes of bleeding (defined as requiring return to surgery or transfusion of at least six units of blood products) by applying this indicator to Medicare patients who underwent coronary artery bypass surgery[19].

With the recent requirement beginning October 2007 of present on admission (POA) codes attached to secondary diagnoses, the number of these coded complications that result from an event not related to the hospital stay can be estimated. [redacted] et al found that the percentage of these coded surgical complications that occurred in hospital was around 70% when evaluating for present on admission (POA) coding. However, POA coding was not consistent in their samples from California and New York administrative data[20]. Naessens and colleagues found that 87% of POHH were acquired during the hospital stay[21]. According to coding rules, POA coding requires that the physician document the complication. Without physician documentation, the complication cannot be assigned a POA code so excluding cases where a complication is present on admission will likely still over-count coded complications that occur during the inpatient stay.

There is a body of literature to suggest that resource utilization and healthcare cost are increased following the occurrence of a PSI. In undifferentiated populations of inpatient surgical procedures, a statistically significant difference in length of stay (LOS) exists between patients with and without any in-hospital secondary complication diagnosis[21, 22]. [redacted] and colleagues in a study of pediatric patients found a statistically significant increase in length of stay in patients with complications as defined by the AHRQ patient safety indicators[23]. Unlike most other studies, this research examined charges, and found excess charges, including those for laboratory and imaging studies. Encinosa and Hellinger found a significant increase in health care expenditures for Medicare patients up to 90 days following hospitalization in patients who had a coded patient safety complication[24]. However, the mortality rate in Shufelt's study of post-operative hemorrhage or hematoma cases did not differ compared to discharges without POHH[25].

There is also mounting evidence demonstrating poorer clinical outcomes for patients with PSI events. [redacted] et. al. found in reviewing over 2 million discharge abstracts from the New York inpatient database that PSI events were associated with significant increases (2-20 fold) in hospital mortality[26]. This significant association has been further corroborated in more recent studies. Rivard et al. examined relationships in the Veteran's Administration system, showing that in spite of a different financial structure and patient population PSI events increased mortality as much as 20% and LOS more than 4 days[27]. [redacted] evaluated insurance claims to capture both inpatient and outpatient associated outcomes. They found that 11% of all deaths and 2% of readmissions were related to PSI events[28].

In spite of these significant associations, current studies are limited by their reliance on administrative data[26]. Though patient and hospital characteristics have been significantly associated with PSI rates[25, 29, 30], administrative databases lack the significant clinical detail necessary to develop predictive/associative algorithms for PSI events. Previous studies have utilized LOS and overall charges or hospital costs to estimate the financial burden associated with PSIs. With the more detailed clinical information available in the data source that is the basis for this proposal a patient's hospital course can be better associated with a specific event to estimate the direct cost of an error and the resultant impact on the post-operative clinical course.

C. Grading the Severity of Post-operative Complications

Complication rates are an important metric to evaluate the outcomes for specific surgical procedures and more recently hospital performance. Furthermore, these rates have become targets for quality and systems interventions. Although there has been ample

research into preoperative prediction models to appropriately categorize and approximate the risk of complication with surgical intervention[31-35], few generalized models exist to evaluate the impact of complications once they occur[36-41]. Most of the current categorization systems are derived from a model developed by [redacted] 1992 to classify the negative outcomes of operations[42]. This system was unique in its generalizability and objectivity. The severity of an outcome was graded primarily on the type of intervention used to manage the complication and whether it resulted in permanent disability or death[36, 42]. The initial iteration of the classification system was also subdivided into categories of complications, sequelae of the operation and failures to cure. Outcomes were graded as complications carrying minor risks (I), potentially life threatening (II), complications with residual disability (III) and death in a patient with a complication (IV). The grade II definition also included patients with a complication and a length of stay exceeding twice the median and included a subcategory for complications requiring invasive procedures[36, 38].

In 2004, [redacted] expanded the original grading system to 5 grades and 7 levels by adding sub-classifications for ICU care, organ failure and interventions requiring general anesthesia[36, 38]. Additionally, given the variability in length of stay due to factors other than complications it was removed as a defining variable. Sequelae of the operation and failures to cure were also removed from the 2004 classification of complications. Finally, disability changed from a separate complication class to a suffix. This new classification scheme was then validated on 6,336 patients and deemed simple, comprehensive and logical when testing it at 10 different centers across the world[36]. This improvement increased its applicability and it became more broadly used but many of the studies modified the system to fit their own aims, thereby, decreasing its validity and generalizability.

A recent editorial by the developers of several iterations of this classification system advocates for improved standardization of the grading and reporting of surgical complications. They argue that this would allow studies across disciplines to be compared with a common metric. Additionally, it would allow for more standardized evaluation of surgical performance[36, 43]. The most recent modification of the [redacted] classification system is the Accordion Severity Grading System (ASGS)[38]. The ASGS is based on a similar objective approach to measuring severity based on intervention. One major advantage of the ASGS is a standardized approach to the qualitative description of the complication (mild, moderate, severe, etc.) The following definitions taken from the 2009 publication first describing the expanded ASGS give an overview of the definitions for the grading system as well as the definitions for organ failure. Such clear, objective definitions allow for wide application and comparison across hospitals.

Expanded Accordion Severity Classification of Postoperative Complications[38]:

1. Mild complication
    - Requires only minor invasive procedures that can be done at the bedside such as insertion of intravenous lines, urinary catheters, and nasogastric tubes, and drainage of wound infections.
    - Physiotherapy and the following drugs are allowed-antiemetics, antipyretics, analgesics, diuretics, electrolytes, and physiotherapy.
2. Moderate complication
    - Requires pharmacologic treatment with drugs other than such allowed for minor complications, for instance antibiotics.

- Blood transfusions and total parenteral nutrition are also included.

3. Severe: invasive procedure without general anesthesia
   - Requires management by an endoscopic, interventional procedure or re-operation without general anesthesia.
4. Severe: operation under general anesthesia
   - Requires management by an operation under general anesthesia.
5. Severe: organ system failure (see below)
6. Death

Definitions of Organ Failure for Accordion Classification System[38]:

- Cardiac: Need for any of the following medications in the following doses
  - Norepinephrine >0.1 $\mu g/kg^{-1} \cdot min^{-1}$
  - Epinephrine >0.1 $\mu g/kg^{-1} \cdot min^{-1}$
  - Dopamine >15 $\mu g/kg^{-1} \cdot min^{-1}$
- Central nervous system.
  - Glasgow coma scale equal to or less than 6.
- Hematologic:
  - Platelet count less than $20 \times 10^9/L$
- Liver:
  - Need for FFP to correct INR in patient with serum bilirubin >12 mg/dL
  - INR < 2.5 in patient with serum bilirubin >12 mg/dL
- Renal: Need for dialysis in patient not on dialysis preoperatively
- Respiratory: (Does not include patients already on a mechanical ventilator for respiratory failure)
  - Need for mechanical ventilation for greater than 24 h in a patient who requires reintubation after surgery
  - Need for mechanical ventilation of greater than 72 h in a patient who is not extubated on the day of surgery.

Although the various iterations of this system have become a gold standard in grading post-operative complications, being cited in over 214 articles since its most recent modification in 2004[43, 44], the level of clinical detail necessary for categorization has limited its application to studies that are able to perform detailed chart reviews. Traditional administrative databases lack the clinical detail necessary for appropriately grading complications. One approach to broadening the application of this classification system, which is a critical next step in research related to surgical quality, safety, and outcomes, is to develop an algorithm based solely on administrative or billing data. The development and validation of such an algorithm would require a large data source that contains both rich clinical variables as well as traditional administrative data. This proposal is based on such a database, which contains detailed information on over 400,000 patients undergoing surgery at 65 geographically disparate hospitals.

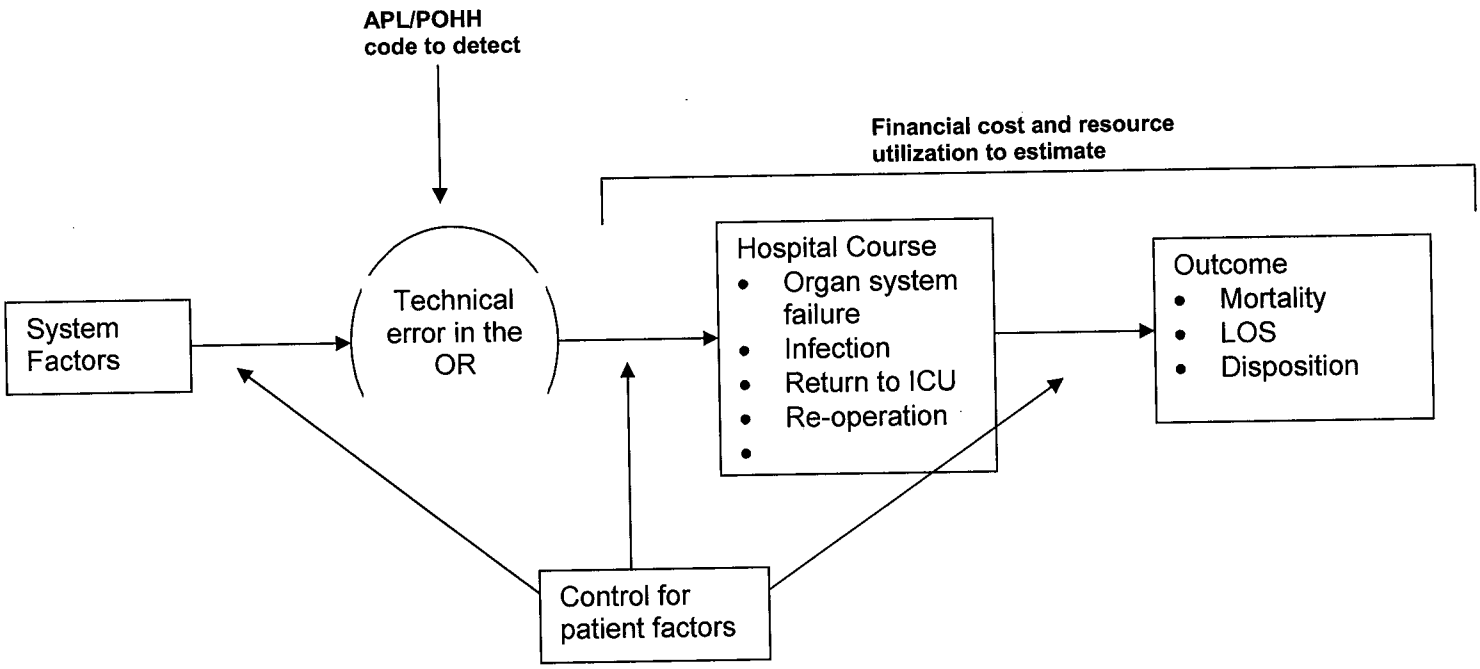
D. Objectives of this Proposal to Address These Gaps

We hypothesize that there is wide variation in the clinical severity of the post-operative course and associated healthcare costs following an intra-operative technical error and that cost can be used as an indicator of the severity of injury. We propose to target inpatient stays during which an APL or POHH event is detected to develop a novel methodology to identify which of these events have serious post-operative consequences at the patient level. This identification methodology will be based primarily on healthcare cost and other information available in administrative data sets

and therefore widely applicable across institutions; however will take advantage of the rich and detailed clinical data available within our data source for validation. Once identified at the patient level, we can study institutional performance to identify high and low outliers based on event rate and severity of the post-operative course. This will provide a prospective sampling strategy to study the role of system factors in technical error and surgical safety. To accomplish these goals we will:

1) Study the impact of an intra-operative technical error on the clinical severity of a patient's post-operative course.
2) Develop and validate an algorithm based on resource utilization and cost to stratify the clinical severity of the post-operative course.
3) Analyze the distribution of clinical severity and healthcare cost associated with the hospital stay of patients experiencing an intra-operative technical error.
4) Use qualitative field observations and human factors analyses for an on-site study of the system factors at hospitals with the highest and lowest rates of clinically-significant APL or POHH.

The following schematic outlines the approach described in this proposal:



II. How will scientific knowledge, technical capability, and/or clinical practice in one or more broad fields be improved

Published studies evaluating these surgical events (APL and POHH) have been performed on administrative data and have either evaluated a heterogeneous group of hospital discharges using large administrative databases or a segment of hospital types, specific procedures or geographic locations. This proposal seeks to advance our knowledge and understanding of surgical complications by utilizing a large geographically diverse sample of community hospitals with electronically available administrative and rich clinical data to characterize intra-operative technical errors and the resulting resource consumption and outcomes in much greater detail than has been previously possible. Using the uniquely rich clinical data available, we will develop and

validate an algorithm for identifying clinically-significant APL or POHH utilizing billing and other administrative data that is widely generalizable across the U.S. healthcare system. Our study cohort includes over 400,000 surgical patients treated between 2008 and 2010 at 65 U.S. community hospitals and therefore reflects practice as it is occurring in hospitals everyday.

Because this proposal is based in administrative data, coding accuracy is an important consideration for this proposal as well as for other research in this area. While coding is an imperfect method for identification of clinical events, it is widely used and will continue to be used for physician evaluation and public reporting. Since hospitals use ICD9 codes for case finding for quality review, our samples will use the ICD9 codes applied by hospital coders outside research settings. This will insure that our research is congruent with actual hospital practice and can be generalized broadly to U.S. hospitals.

The other advantage to such a large, geographically diverse cohort of hospitals is the ability to identify patterns across hospitals. This approach will allow us to identify the role of system factors in the occurrence of these surgical complications thereby informing the development of strategies for preventing them. For example, if a shorter post-operative critical care unit length of stay is associated with return to the critical care unit and/or longer length of stay, a change in clinical practice to lengthen that initial critical care unit length of stay may be beneficial. Such patterns can only be identified with a large, diverse set of hospitals and detailed rich clinical data.

III. How will the concepts, methods, technologies, treatments, services, or preventive interventions be changed

Currently little is known about the role of system performance in the occurrence of technical errors in the operating room. Studying performance in the OR using a mixed method of quantitative structured data analysis followed by qualitative site visits will enable us to identify which factors play the most significant role in the causation of intra-operative technical errors. This will allow for a better understanding of the system and provider factors that contribute to adverse events and impact surgical outcomes both positively and negatively. Those factors that positively impact surgical outcomes can be emphasized and disseminated as best practices while evidence-based interventions can be developed to reduce the consequences of those that impact surgical outcomes negatively. The employed methodology will allow for generalizable results and will ultimately improve surgical patient safety and save lives.

Examining a large sample of cases (>400,000) beyond administrative data, as we propose to do in this study, would not be possible without the depth of electronic clinical data we have available.

INNOVATION

I. How the application challenges and seeks to shift current research or clinical practice paradigms

The ability to objectively discern the clinical detail surrounding coded surgical complications will provide the basis for a more rigorous evaluation of clinical processes in hospitals and of physician performance. The case review process in hospitals is often fraught with over- and under-identification of cases for review, and frequently assessed with implicit criteria applied on a case by case basis. Individual case review takes

precedence over a systematic review of complications and strategies for prevention. Improvement in patient safety requires moving from this transactional approach to a more systematic approach. The ongoing incidence of these particular events (APL and POHH) tends to be low in a single institution, making it difficult to draw valid conclusions regarding addressable causes of these complications. Our large sample will allow us to identify patterns and causes across and within institutions.

This application seeks to shift current research paradigms from a transactional, case-based to a more systematic approach in the following ways:

1) Increase the accuracy of identification of intra-operative technical errors with significant impact on the post-operative clinical course within administrative data. This will decrease the "noise" currently introduced by the identification of events that do not have clinical sequelae for the patient.
2) Use the occurrence of events defined at the patient level to describe performance at the institutional level. By targeting high and low performing hospitals for evaluation rather than targeting specific events once they have occurred, we can identify patterns across hospitals rather than aspects of a particular case making the findings significantly more generalizable.

II. Novel theoretical concepts approaches or methodologies, instrumentation or interventions to be developed or used, and any advantage over existing methodologies instrumentation, or interventions

[redacted] offers the unique opportunity to further study several of the markers of intra-operative technical error that can be identified using the PSI definitions developed by AHRQ. We will be able to study the complex interplay of systemic factors and technical performance that contribute to and mitigate surgical adverse events, as well as the clinical cascade that leads to poorer clinical outcome. **The following combination of characteristics of IHM data does not exist in any other currently available data source:**

- Includes both rich clinical and administrative data
- Is from a geographically diverse sample of 65 U.S. community hospitals, reflecting clinical practice as it exists today in most of the U.S. health system.
- Is based on current data, from 2008 through 2010
- Is all available in electronic format, including operative duration and return to intensive care among other rich details of the clinical course.

The following table illustrates the types of data available within the IHM system as well as a description of how the data is coded and examples of potential uses.

Data type	Description	Examples of data use
Procedures	All ICD9 coded procedures	Coded post-operative complications requiring procedure
Diagnoses	All ICD9 coded Diagnoses	Adjustment for Comorbidity
Laboratory	Test results and time collected	Serial hemoglobin measurements to validate hemorrhage, perioperative glycemic control
Blood bank	Blood product name, release time	Ordering physician, time and amount given in relation to surgery times

Medications	Dosage, route and time given	Time, name, dose of medication in relation to surgery times
Radiology	Radiology test name, time performed and radiology report	Define complications
Nursing therapy and assessments	Time stamped data field documentation	Documentation and timing of therapies such as leg compression, ventilator duration, central line insertion and duration
Location	Locations of patient through stay	Return to ICU, critical care unit length of stay
Text notes	All typed, dictated, transcribed electronic notes	Admission note, operative notes, consultant notes, nurses' notes, discharge summary
Surgery	Surgery times	Incision time and end time

With such detailed clinical information, we will be able to perform a detailed analysis of the clinical course, resource utilization and outcomes that result when a APL or POHH occurs in a typical U.S. hospital today. By validating the relationship between healthcare cost and these more proximal measures of clinical severity, we can develop a novel approach to identifying clinically-significant APL or POHH that can be employed in any hospital in America. Furthermore, by performing site visits and detailed analyses of high and low performing hospitals, we can develop and disseminate best practices with the goal of ameliorating intra-operative technical errors, namely clinically-significant, accidental puncture or laceration (APL) and post-operative hemorrhage or hematoma (POHH).

APPROACH

I. Data Source

The [redacted] is a not-for-profit organization that has developed an electronic data collection, communication and analysis system to support quality and operational improvement in hospitals. This data system is unlike any other healthcare data system in that it combines the entire contents of the clinical information system used daily in hospitals, reaching well beyond claims data and including thousands of data elements per patient, such as lab results, drugs given, and radiology studies performed. Automated extraction of required data, such as Joint Commission on Accreditation of Healthcare Organizations (TJC) core measures [45], Center for Medicare and Medicaid Services (CMS) criteria[46] and Center for Disease Control (CDC) [47] infectious disease information, can be more readily automated by a central resource such as [redacted] rather than by the individual hospitals. This centralized data system allows [redacted] to provide hospitals with a level of statistical analysis and reporting not feasible at an individual hospital level, due to constrained resources at community hospitals, particularly small and rural hospitals.

[redacted] is a HIPAA "business associate" of its hospital partners. As part of their agreement (please see Appendix), they have agreed to make the de-identified data available to health services researchers. [redacted] offers a unique opportunity to study healthcare in community hospitals, an understudied sector of the U.S. healthcare system where the

majority of surgery is performed. [redacted] works with 65 community hospitals geographically dispersed across the U.S.

	[redacted] data system characteristics
Total inpatient beds	13,840
Annual inpatient discharges	525,349
Hospital locations	25 States: AK,CA,CT,GA,ID,IL,IN,KS,KY, LA,MA,MD,ME, MN,MO,NC,NY,OH,OK,PA,SD,TX,UT,VA,WA
Hospital bed size	Range from 25 to 1100 beds

Data in the [redacted] system originates from the Meditech Clinical Information Systems located at each of the participating hospitals. The Meditech system is an operational software system, written in Magic, a proprietary version of MUMPS, and is not a relational data system. [redacted] collects these data and standardizes them, storing the data in an Oracle data system running on LINUX servers.

II. Preliminary data

We examined the population of inpatient surgical cases from 2006 through 2008 in order to assess feasibility of this proposal. The AHRQ list of surgical ICD9 procedure codes as described in the Patient Safety Indicator specifications [48] was used to define the total sample. This sample included 196,864 inpatient surgical procedures.
Accidental puncture or laceration (APL) and post-operative hemorrhage or hematoma (POHH) were identified using the AHRQ diagnosis code definitions as described in the Patient Safety Indicators specification [48].

Complication code group	ICD9 diagnosis codes
Accidental puncture or laceration	998.20, E870.0,E870.1,E870.2,E870.3, E870.4,E870.5,E870.6,E870.7,E870.8,E870.9
Hemorrhage/hematoma	998.11,998.12

We found the following incidence of coded complications in our sample of 196,864 inpatient discharges with surgical procedures:

	Number of cases	Percent of inpatient surgical procedures
Accidental puncture or laceration	5873	2.98%
Hemmorrhage/hematoma	4798	2.44%

While these percentages are higher than those found in the literature, our denominator includes only inpatient discharges with surgical codes as listed in the AHRQ surgical procedure group, the recommended denominator for PSIs. [redacted] et al demonstrated an APL rate of 1.93% and a POHH rate of 1.73% using all discharges as a denominator [21]. Shufelt et al found an incidence rate of 1.43% of hospital discharges with POHH in over 2 million discharges in New York State [25] [redacted] et al studied accidental puncture or laceration cases in New York administrative data and found a rate of 2.4% using a denominator of all discharges [18].

Our preliminary frequency analysis of all surgical procedures between Jan 1, 2006 and Dec 31, 2008 demonstrates that the following procedures/groups had the highest frequency of complications.

Highest APL rates	
	Cases with APL
Open cholecystectomy	6.4%
Colon surgery (TJC/CMS)	4.2%
Spine surgery	4.1%
Hysterectomy (TJC/CMS)	3.3%
Highest POHH rates	
	Cases with POHH
CABG (TJC/CMS)	6.0%
Vascular surgery (TJC/CMS)	5.1%
Colon surgery (TJC/CMS)	3.5%
Hysterectomy (TJC/CMS)	2.0%

We examined length of stay in patients with and without these events in two surgical code groups as part of this preliminary analysis. Our hypothesis was that we could demonstrate differences in the incidence of complications between two sample surgical groups and could also examine length of stay differences within these groups in sub-groups with and without these complications. The groups chosen were laparoscopic cholecystectomy (inpatients only) and colon surgery (inpatients only). We defined colon surgery using the same procedure codes for the antibiotic prophylaxis population defined by TJC and CMS. Laparoscopic cholecystectomy patients are often treated as same day surgery cases rather than as inpatients, so our sample may be skewed toward higher risk patients.

Procedure code group	ICD9 procedure codes
Laparoscopic cholecystectomy	051.23, 051.24
Colon surgery	045.00, 045.03,045.49,045.50,045.71,045.72, 045.73,045.74,045.75,045.76,045.79,045.80, 045.90,045.92,045.93,045.94,045.95,046.03, 046.04,046.10,046.13,046.75,046.76,046.91, 046.92,046.94,048.50,048.61,048.62,048.63, 048.64,048.65,048.69

We have applied in this preliminary analysis, and plan to apply in the proposed research, procedure codes using the TJC/CMS surgical procedure groups. These groups would be large enough for statistical analysis, would be clinically meaningful for hospital quality review and would encourage generalized application of research findings from this study. We also plan to augment the TJC/CMS surgical procedure groups by analyzing common clinical groups with these complications and in the general surgical population.

The following preliminary analysis was performed on inpatients discharged between Jan 1, 2006 and Dec 31, 2008. There were no laparoscopic cholecystectomy or colon surgery cases with both APL and POHH diagnosis codes.

Laparoscopic cholecystectomy Inpatients	No APL	APL	No POHH	POHH

Number of cases	9363	65	9338	90
Average LOS days	4.43	6.47	4.41	9.87
% of cases with complication		0.69%		0.95%
LOS difference		p<.005		p<.000001
Colon surgery	No		No	
Inpatients	APL	APL	POHH	POHH
Number of cases	8217	360	8278	299
Average LOS days	11.3	12.4	11.2	16.2
% of cases with complication		4.20%		3.49%
LOS difference		p=.128		p<.000001

III. Methodology

Aim 1. Study the impact of an intra-operative technical error on the clinical severity of a patient's post-operative course.

A. Data collection

1. Cohort definition and identification of intra-operative technical errors

The AHRQ list of ICD-9 procedure codes as described in the Patient Safety Indicator specifications [48] will be used to define the cohort. We will examine all inpatient discharges with a procedure code in this AHRQ defined code group occurring between January 1, 2008 and December 31, 2010. By limiting our analysis to this time period, we will only examine cases after the present on admission (POA) designation was adapted.

Accidental puncture or laceration (APL) and post-operative hemorrhage or hematoma (POHH) will be identified using the AHRQ diagnosis code definitions as described in the Patient Safety Indicators specification and listed in the preliminary analysis section.

So that the results of this proposal may be applied efficiently and effectively by hospitals, surgical groups such as "Colon surgery" will be defined by the ICD9 procedure code groups defined by TJC/CMS in their specification manual [45]. This table summarizes the core procedure groups to be included in the analysis. A full description of the procedure codes used to identify these procedures can be found in the Appendix.

Core measure procedure code groups
Coronary bypass surgery
Cardiac surgery
Colon surgery
Hysterectomy
Total hip replacement
Total knee replacement
Vascular surgery

We will expand the subsets of cases to be studied beyond these groups based on frequency analysis of the groups of surgical procedures with the highest incidence of these surgical complications. In our preliminary analysis, we have identified some non-TJC/CMS procedure groups with high APL or POHH rates and will perform this frequency analysis on the entire study sample as part of this proposal. These defined code groups have changed over the time span of this study. In our analysis, we will apply the retrospective code groups that were active at the time that the procedures

were performed. The changes during our study period are summarized in the following table. These are included in the detailed procedure list in the appendix.

Core measure procedure code group changes 2008-2010

Code group	Date of change	Change made
Cardiac surgery	2008-Q2	Added these codes
		035.00
		035.01
		035.02
		035.03
		035.04
Cardiac surgery	2008-Q4	Deleted this code
		035.95
Colon surgery	2009-Q4	Added these codes
		017.31
		017.32
		017.33
		017.34
		017.35
		017.36
		017.39
Hysterectomy	2008-Q4	Deleted this code
		068.39
Hysterectomy	2009-Q4	Added this code
		068.79

2. Study variables

Given the rich clinical data maintained within the IHM database, these variables are all electronically available.

Clinical Outcomes:

- Date of death
- Length of stay (LOS)
- Discharge disposition (expired, skilled nursing facility, home, etc).

Hospital course / resource consumption:

- Surgery duration
- Blood products given
- Return to or admission to intensive care unit after immediate post-operative period
- Return to operating room
- Re-intubation
- Cardiac or respiratory arrest
- Microbiology such as positive post-operative cultures
- Laboratory results
- Medication and time of administration (e.g. antibiotics given beyond the initial post-operative period or vasopressor agents)
- Laparoscopic to open conversion
- Initial post-operative critical care unit LOS
- Organ system failure or complication including cardiac, renal or respiratory failure

- Associated POA code
- All hospital charge data

Patient and procedure characteristics:

- Race
- Sex
- Age
- Ethnicity
- Type of insurance
- Charlson co-morbidity score
- Body mass index
- Smoking status
- Admission status (emergent, urgent, routine)
- Location of origin (long term care, home, etc)
- Day surgery versus inpatient (where appropriate)

Provider characteristics:

- Hospital procedure volume
- Physician procedure volume
- Hospital size
- Hospital location (urban, suburban, rural)
- Hospital teaching status
- Encrypted NPI of surgeon

We will also link the [redacted] data to the American Hospital Association (AHA) file and the National Provider Index (NPI) to get more structural details about the hospital and the specialties of the providers. The initial phase of this project will include identifying additional measures of resource consumption and additional population characteristics beyond those listed above.

3. Data quality/frequency analysis

[redacted] routinely performs logical checks on the data (for example, whether male patients are designated as receiving hysterectomies or a patient is discharged before arrival). While these errors are possible, we have found them to be almost non-existent due to the multiple individuals rechecking data during the registration, clinical and billing procedures in hospitals. Despite this rarity, we continue to perform logical checks as part of our standard operating procedure.

All data in the system is automatically time stamped. We know when tests were ordered, when blood was drawn and when results were completed. The data includes patient locations and arrival and departure from each clinical area during the hospital stay, so that we can track returns or new admissions to critical care units and length of stay in critical care units as well as return to the operating room.

As part of [redacted] quality control procedures in the data extraction process, we filter for anomalies, catalog and exclude these values from the working data. Our data screens include:

- Negative number results
- Lower than specified digit range
- Higher than specified digit range
- Alpha character in numeric field

- Numeric character in alpha field

4. Sample size calculation

Based on volumes from 2008-2009, we project the following sample size for our study period of 2008 - 2010:

Projected sample for study	Denominator	APL	%	POHH	%
Total surgical cases	**407,097**	**3,359**	**0.8%**	**6,713**	**1.6%**
Inpatient only					
Laparoscopic chole.	15,774	122	0.8%	146	0.9%
Open cholecystectomy	3,878	140	3.6%	126	3.2%
Colon surgery	16,979	843	5.0%	449	2.6%
Hysterectomy	18,552	354	1.9%	344	1.9%
Hip replacement	16,673	15	0.1%	203	1.2%
Knee replacement	26,856	14	0.1%	251	0.9%
Vascular surgery	4,068	44	1.1%	149	3.7%
CABG surgery	14,301	98	0.7%	954	6.7%
Cardiac surgery	3,252	27	0.8%	273	8.4%

B. Data analysis

AIM 1 - H1: **Hospitalizations in which an intra-operative technical error occurs are associated with a more complicated post-operative course.**

The data for this aim consists of the entire surgical cohort for the study period. We will search the record of each patient in the cohort for evidence of a post-operative complication and then assign a severity score using the expanded Accordion Severity Grading Scale (ASGS) as described above.[38, 42]

[redacted] contains all variables necessary to categorize complication grades 2 and higher. A subset of the hospitals contains sufficient detail to also categorize minor complications (Grade 1); however this is less reliable. Given that the goal of this study is to associate the occurrence of an intra-operative technical error with significant post-operative consequences, this is unlikely to be an issue. Given the nature of the complications that are categorized as Grade 1, it is unlikely that the incidence of these events will vary between cases that do and do not involve an APL or POHH. We will use the subset of hospitals that do have this data available to investigate Grade 1 complications; however our main analysis will focus on Grade 2 and higher.

Each patient will then be assigned a composite score of an integer from 1 to 6 that corresponds with the most severe post-operative complication (as defined in the above ASGS table) that occurred, or a score of "0" if no post-operative complication occurred. This score will be treated as a new categorical variable (Y = 0, 1, ..., 6) termed "clinically-significant post-operative deviation", which will serve as the main outcome of our analysis. The main covariate (PSI) is whether or not the patient experienced an APL or POHH as defined by the PSI definitions as described in the previous section. *A priori* covariates that we expect to be related to our outcome, Y, include surgical procedure, age, gender, and co-morbidity score.

Rather than dichotomize Y in order to perform traditional logistic regression, we will model the full spectrum of the composite score Y using a cumulative logistic regression

model for ordinal data. In the cumulative logistic regression model, if we dichotomize Y at any integer (cut point), the dichotomized outcome follows a logistic regression, with a different intercept at each cut point. In particular, instead of having one logistic regression model at the single cut point, we will have 6 logistic regression models at the cut points k=1,2,…,6 (we do not need to cut at 0 since the probability that Y≥0 equals 1). At each of the six cut points, the ordinal cumulative logistic regression model will have the following general structure

$$\text{logit}[\text{Prob}(Y \geq k)] = \beta_0 + \beta_1 PSI + \beta_2 \mathbf{X} \qquad (1)$$

The logit of the probability that Y ≥ k is modeled as a linear function of covariates: PSI (equals 1 if patient has a code for APL or POHH, 0 if not), and **X**, a vector containing all the other control variables including patient characteristics, surgical procedure, etc. Our main interest is estimating β_1 (the effect of PSI) and obtaining a 95% confidence interval for it, with the β_2 (the effects of the confounders) of less interest. Generalized estimating equations (GEE) via SAS Proc Genmod will be used to estimate model (1), accounting for clustering within hospital.

As discuss in Simon (2005), we will randomly split the data into a 2/3 training (derivation) sample and a 1/3 validation sample [49]. One problematic issue with this split-sample method is that often there are few subjects in the validation set, and thus the validation can be meaningless. However, our total sample size of our study is approximately 400,000, so that a 267,000 training sample and a 133,000 validation sample should be sufficiently large, even if the rate of APL and POHH is less than 3%, as expected in our dataset. The model derived from the training sample will be applied to the validation sample, and the c-statistic for ordinal categorical variables will be used to quantify the fit of the model in the validation sample [50].The c-statistic will also be used to evaluate the discriminative ability of the model to distinguish between patients who will have a "clinically-significant post-operative deviation", and those who will not. Calibration of the model in the training and validation samples will also be assessed using plots of observed vs. predicted values and Hosmer-Lemeshow goodness-of-fit statistics. Further, we will calculate a 'pseudo R^2' [51] to examine goodness of fit of the logistic regression model. Besides the split-sample approach, we will also use the full dataset in a 'leave-one out', cross-validation, in which each subject is dropped out, the model is re-estimated, and a cross-validation c-statistic will be used to quantify the fit of the model.

We will perform a similar analysis using surgical procedure groups such as TJC/CMS defined hysterectomy, coronary artery bypass graft, other cardiac surgery, colon surgery and major vascular procedures that we have previously found to be associated with increased rates of APL and POHH (see preliminary data). We will also examine other common procedures, such as laparoscopic cholecystectomy and other procedures seen in frequency analysis of these complications.

As a secondary analysis, we will also plan to dichotomize Y into no complication or only mild complications (0 – 1) and at least one moderate complication or higher (2 – 6).

AIM 1 - H2: **The severity of the hospital course is correlated with increased healthcare cost.**

We will use hospital charges as a proxy for healthcare cost. We recognize that this is a limitation because charges vary across both region and institution; however, true cost data

is not accessible for most hospitals. In order to control for this variation, we will normalize charges at the institutional level. The distribution of hospital charges will first be determined for all patients that undergo a specific procedure and do not experience a APL or POHH. We expect this to be a skewed distribution and will therefore either take the log transformation or use the median charges as our "base case". We will then convert charges to % of average charges at the hospital level for this analysis.

We will add healthcare charges to the logistic regression model in (1); we expect the discriminative ability of the model (in terms of the c-statistic) to increase dramatically with the addition of charge data. We will also add an interaction term between charges and PSI to the model, as we expect that the odds ratio of having a 'clinically-significant post-operative deviation' (Y) for patients with and without an APL or POHH will depend on the charges i.e., if the charges are low, then the relative risk for Y for patients with and without a APL or POHH will be lower then if the charges are higher. Generalized estimating equations (GEE) will again be used to estimate the cumulative logistic regression model, accounting for clustering within hospital.

$$\text{logit}[\text{Prob}(Y \geq k)] = \beta_0 + \beta_1 \text{PSI} + \beta_2 \text{cost} + \beta_{12} \text{PSI*cost} + \beta_3 \mathbf{X} \qquad (2)$$

for k = 1,...,6. The statistical approach to determine the fit of the model with the addition of costs as a covariate will be similar to ***AIM 1 - H1***. We will use the same random split of the data into a 2/3 training (derivation) sample and a 1/3 validation sample as in ***AIM 1 - H1***. Similarly, we will also perform leave-one-out cross-validation. The c-statistic will again be used to determine the fit of the model in the training and validation samples.

We will use an ROC (Receiver Operating Characteristics) curve to determine the optimal charge cut-off for cases with a 'clinically-significant post-operative deviation'. The area under the ROC curve is a measure of diagnostic accuracy such that values between 0.5 and 0.7 indicate low discriminative ability, values between .7 and .9 indicate moderate discriminative ability and values greater than .9 indicate high discriminative ability.

Using the 2/3 training sample, we will fit similar logistic regression models as above except we will dichotomize cost at each distinct observed value from 0 to the maximum value of cost. For each value of cost, we will calculate the c-statistic (estimate of the area under the ROC curve). The value of cost that maximizes the c-statistic will be deemed the 'optimal cutoff' that maximizes the discriminative ability of costs. The model derived from the training sample with this cutoff will be applied to the validation sample, and the c-statistic will be used to evaluate the discriminative ability of the cost cutoff (along with APL or POHH) to distinguish between patients who will have a 'clinically-significant post-operative deviation', and those who will not.

As a secondary analysis, given the variation in cost based on type of procedure, we will do a stratified analysis according to the 6 procedures identified to have the highest rate of APL or POHH on our preliminary analysis: colectomy, hysterectomy, CABG, vascular surgery, cholecystectomy, and spine surgery.

Aim 1 - H3: **An algorithm based on administrative data, including resource utilization and cost, can be developed to stratify the clinical severity of the post-operative course.**

We will add other data elements that are generally available in administrative data to the logistic regression model in (1) to maximize the discriminative ability of the model (in terms of the c-statistic). Examples include length of stay (LOS) or discharge to a higher level of care than admission (DIS). This will be an iterative process to identify the most discriminative model. Generalized estimating equations (GEE) will again be used to estimate the cumulative logistic regression model, accounting for clustering within hospital.

$$\text{logit}[\text{Prob}(Y \geq k)] = \beta_0 + \beta_1 PSI + \beta_2 cost + \beta_{12} PSI*cost + \beta_3 LOS + \beta_4 DIS + \beta_5 \mathbf{X} \quad (3)$$

for k = 1,...,6. The statistical approach to determine the fit of the model in this aim will be similar to ***AIM 1 - H1***. We will use the same random split of the data into a 2/3 training (derivation) sample and a 1/3 validation sample as in AIM 1 - H1. Similarly, we will also perform leave-one-out cross-validation. The c-statistic will again be used to determine the fit of the model in the training and validation samples.

Using the GEE estimates for the model in (3), we can estimate the probability that a patient has each value of the Y, i.e., we can estimate the probabilities Prob(Y=0), Prob(Y=1),..., Prob(Y=6). Our 'algorithm to classify patients into a clinical severity score' will be based on these probabilities. There are two possible ways to assign severity scores based on these probabilities.
1) We could pick the mode, i.e., the score which corresponds to the largest probability (e.g., if Prob(Y=1) is the larger that the other 6 probabilities, then we assign score '1').
2) We could also assign the score to equal the mean, which equals 0*Prob(Y=0)+1*Prob(Y=1)+2* Prob(Y=2)+...+6*Prob(Y=6).

The c-statistic will be used to determine the association between the ASGS score and our assigned 'score', as well as to determine whether the mean or the mode gives better discrimination. We will also apply the algorithm at each cut point (i.e., for the six dichotomized versions of Y); however, for dichotomized outcomes, the c-statistics based on the mean and the median are identical

C. Power Calculation

In order to preserve an overall Type I error rate of 5%, the 3 important tests (H1 regarding APL/POHH and H2 regarding cost and H2 regarding the interaction between APL/POHH and cost) will each be tested at the 1.67% level of significance level. All tests below are two-sided. As discussed above, we expect there to be approximately 267,000 total patients in our training sample. We expect the percentage of patients with a APL or POHH to be 2.5%. Thus, out of the 267,000 total patients in our training sample, we expect that 6,675 patients will have a APL or POHH, and 260,325 patients will not have a such an event.

Although the statistical analysis will be based on the ordinal cumulative logistic regression model for the expanded Accordion Severity Grading Scale (Y), for purposes of the power calculations, we will use the logistic regression model for the dichotomous outcome for Y≥2versus Y<2 Dichotomizing Y in this form will give a conservative power calculation. In our training sample with 267,000 patients, assuming the rate of Y≥2 is 15% in patients without a APL or POHH, this study will have at least 80% power (with a significance level of 1.67%) to detect an increase in the rate of Y≥2 from 15% in patients without a APL or POHH to at least 17% in patients with a APL or POHH using a GEE z-

test statistic which accounts for clustering within hospital. In this power calculation using the GEE z-test statistic, the intra-class correlation coefficient (ICC) for patients in the same hospital is assumed to be at most 0.0001. An intra-class correlation of 0.0001 led us to conservatively discount the sample size (267,000 in the training sample by 50% before performing the power calculations. Of course, we expect the rate of Y≥2 to be much higher than 17% in patients with a APL or POHH (actually, over 50%), so we expect to have much greater than 80% power to detect the true difference.

For continuous cost data (using the appropriate transformation such as the log), the GEE logistic regression test statistic for no association between cost and 'clinically-significant post-operative deviation' is equivalent to testing if the mean (of transformed cost) is the same in patients with and without a 'clinically-significant post-operative deviation' (adjusted for the other confounders). As discussed in Cohen, an effect size of one-half a SD is clinically important to detect [52].

Overall, in the sample, we expect the overall 'clinically-significant post-operative deviation' rate to be at least 15%. With approximately 267,000 total patients in our training sample, 40,500 whom we expect to have 'clinically-significant post-operative deviation', when using logistic regression, we will have 99% power (with alpha=.0167) to detect a ½ standard deviation difference in the transformed costs between patients with and without 'clinically-significant post-operative deviations'.

The final test will be for an interaction between cost and APL or POHH. An interaction can be interpreted as meaning the difference in costs between patients with and without a 'clinically-significant post-operative deviation' will depend on whether a patient has (or does not have) a APL or POHH. Again, to be meaningful, we should have power to detect a ½ standard deviation in the 'difference in differences'. A conservative estimate of power can be obtained by using an 'effective sample size' as the minimum sample size of the four groups formed by the cross-classification of APL or POHH and 'clinically-significant post-operative deviation'. We expect the smallest groups will be patients with APL or POHH and no 'clinically-significant post-operative deviation' or patients without a APL or POHH but with a 'clinically-significant post-operative deviation'. We expect there to be at least 1300 patient in both of these groups, and discount this by 50% for clustering to get an effective sample size for interactions of 650. With at least 650 patients in each of the four groups formed by the cross-classification of APL or POHH and 'clinically-significant post-operative deviation', we will still have 99% power (with alpha=.0167) to detect a ½ standard deviation difference in the transformed costs between patients with and without 'clinically-significant post-operative deviations'.

To determine the discriminative ability of the models, we will use a GEE Wald test [53] for the c-statistic with significance level of 0.05. Using the c-statistic by applying the logistic regression model from the derivation dataset to the validation dataset (133,000 patients), this study will have 99% power to detect a model that is highly accurate in discriminating the severity of the post-operative complication (c-statistics .91 or greater) versus an AUC that is moderately to highly accurate (c-statistic that is .85 or less).

Further, we also want to obtain a 95% confidence interval for the c-statistic for the validation dataset. Given the 133,000 patients in the validation dataset, the resulting 95% confidence interval [53] will be at most within ±0.01 of the true c-statistic. For example, if the

observed c-statistic is .65, then the 95% confidence interval is [0.640,0.660]. If the observed c-statistic is .95, then the 95% confidence interval is [0.944,0.955].

D. Data interpretation

We expect to find that patients with an APL or POHH are more likely to experience a 'clinically-significant post-operative deviation' than patients who do not experience an APL or POHH. Furthermore, we expect that these patients will also have higher rates of the measures of resource consumption and cost for the hospital course than their counterparts who do not experience an APL or POHH.

Within the APL/POHH group, we expect that patients that have a 'clinically-significant post-operative deviation' will have higher hospital charges while the charges for patients that do not have a 'clinically-significant post-operative deviation' will approximate the control group that did not experience an APL or POHH.

If this is the case, we will have demonstrated a direct link between healthcare cost and the existence of a 'clinically-significant post-operative deviation', validating its use as the basis for our algorithm as described above.

Aim 2. Determine the institutional rates of APL or POHH with 'clinically-significant post-operative deviation' as an indicator of institutional patient safety.

A. Data collection

The analysis in Aim 2 will use the same data that was collected in Aim 1. Please refer to Aim 1 for a detailed description of the method of case identification, study variables, and approach to quality assurance.

B. Data analysis

***Aim 2 - H1:* Patient-level APL or POHH events with 'clinically-significant post-operative deviation' can be aggregated to calculate a hospital rate analogous to the AHRQ PSIs.**

The AHRQ Patient Safety Indicators are institutional rates of potentially-preventable adverse events. We want to determine whether consideration of their impact on the post-operative clinical course improves the ability to detect clinically-significant events. To do so, we will calculate the institutional rate of APL or POHH events without consideration of the severity of the post-operative clinical course. These rates will be calculated using the software available for download from AHRQ [48]. This software includes risk-adjustment for age, sex, DRG, and comorbidity. Next, we will calculate the rate of having an APL or POHH that leads to a 'clinically-significant post-operative deviation'. In this rate, the denominator remains the same as defined by the AHRQ PSI above; to be a 1 (versus a 0) in the numerator, a patient must have both an APL or POHH and a 'clinically-significant post-operative deviation'. We calculate 'clinically-significant post-operative deviation' in 2 ways. The first approach will use a dichotomized classification derived from the gold standard ASGS of a 'clinically-significant post-operative deviation', defined as $Y \geq 2$ or not. Next, we will use our algorithm based solely on administrative data in an analogous manner. Thus, we will have three rates for each of the 65 hospitals: 1) the AHRQ derived rate of PSI (APL or POHH), 2) the rate of APL or POHH leading to ASGS 'clinically-significant post-operative deviation'; 3) the rate of APL or POHH leading to 'clinically-

significant post-operative deviation' as defined by our algorithm. We will use rank correlation coefficients (e.g. the c-statistic) to determine how well the rates from our algorithm agree with the rates from the gold standard. We will also correlate the rates determined by our algorithm with the rates detected using the AHRQ PSI software.

Aim 2 - H2: **Institutions that are high and low outliers can be identified as targets for site visits and further evaluation.**
Based on the distribution of hospitals rate of APL or POHH that leads to 'clinically-significant post-operative deviation', we will identify hospitals that are high and low outliers. Our initial goal will be to identify the top and bottom 10% of hospitals as our institutions of interest (with 65 institutions, this corresponds to the top 6 and the bottom 6 institutions). Thus, an outlier for this aim is any hospital in the top 10% or bottom 10%. The distribution of hospitals will be depicted with a box-and-whisker plot.
The 75th percentile (Q3) and the 25th percentile (Q1) are top and bottom of the box, respectively. Thus, the interquartile range (IQR=Q3-Q1) is the length of the box in the box-and-whisker plot. Without assuming normality, using the adjusted box-plot approach of Hubert and Vandervieren [54], we will consider an outlier in the box-and-whisker plot to be any value that lies more than 0.5 times the length of the box (from either end of the box). That is, if a hospital is below Q1 – 0.5*IQR or above Q3 + 0.5*IQR, it is viewed as being an 'outlier'. This approach does not require normal or symmetric data. Note, however, if the data are approximately normal, then Q1 – 0.5*IQR and Q3 + 0.5*IQR correspond to approximately the top and bottom 10% of hospitals, respectively. We will also identify 2 hospitals with "average" performance for pilot visits.

C. Data interpretation
Given the previous literature documenting the lack of institutional variation in rates of APL or POHH, we anticipate that there will not be significant institutional variation in the rate of APL or POHH depicted using AHRQ PSI software. However, we predict that there will be significant variation in the rate of APL or POHH associated with a 'clinically-significant post-operative deviation' as detected by the ASGS gold standard. We anticipate that our algorithm will more closely approximate the rate of APL or POHH with a 'clinically-significant post-operative deviation' as detected by the ASGS than the rate detected by the AHRQ PSI software.

This means that the addition of our algorithm based on cost and resource utilization to the current approach to identify cases of APL and POHH within administrative data can accurately detect those APL and POHH events that lead to 'clinically-significant post-operative deviation'. The accurate detection of these events is critical to our ability to advance surgical safety.

D. Power Calculation
Our goal is to determine agreement (via the c-statistic rank correlation coefficient) of the rate of APL or POHH leading to 'clinically-significant post-operative deviation' as defined by our algorithm with the rate of APL or POHH leading to ASGS 'clinically-significant post-operative deviation'. With 65 hospitals in this aim, we will have at least 80% power to detect a Rank Correlation Coefficient (c-statistic) between of at least 0.34 using a Wald Chi-Square test statistic with a significance level of 5%.

Aim 3. Use qualitative field observations and human factors analyses for an on-site study of the system factors at hospitals with the highest and lowest rates of clinically-significant APL or POHH.

Aim 3 - H1: **Institutional structure and process influences technical performance and can be the target of system improvement initiatives to decrease clinically-significant technical errors in the OR.**

A. Data collection

1. Subject Recruitment

Once the study institutions are identified as described in Aim 2 – H2, the institutional administration will be contacted and approached for voluntary participation. This will be facilitated through the established relationship that [redacted] at [redacted] has with each of the participating hospitals. Once all study hospitals have been identified and agreed to participate, we will obtain appropriate approval (including human subjects) at each institution. The hospitals will then be randomly sorted using computer-generated block randomization scheme to determine the order of the site visits. The randomization will be done in blocks of 2 to prevent observations of multiple high or low performers in a row. One 2-day visit will be made to each site.

2. Study Team

A multi-disciplinary observational team will perform a site visit to each hospital in the randomly generated order. The study team will be blinded to the status of the institution as a high or low outlier.

A critical component of field observations is the training and expertise of the observational team.[55] Given the complexity of the surgical domain, it is critical to have both surgical expertise as well as a human factors expertise. The PI, [redacted], will serve as the surgical domain expert and the human factors consultant, [redacted], will serve as the content expert in developing the methodology and perform the pilot visit at a minimum.

3. Site Visit Structure

The site visit will have 3 phases.

a. *Safety Attitudes Questionnaire (SAQ)* – The SAQ is a previously validated instrument that measures safety culture at an institution.[56, 57] We will administer the SAQ to all hospital personnel in advance of our site visit and derive a safety score.

b. *Semi-structured observations* - The next phase will consist of semi-structured observations of each phase of peri-operative care. On one day of the visit, the team will identify patients undergoing one of our 6 targeted procedures: cholecystectomy, colectomy, CABG, vascular procedure, hysterectomy or spine surgery. The study team will accompany the patient through the peri-operative course, making observations about the structure and process of the system. Examples of factors to be targeted are included in Table 1. These are adapted from our own experience as well as that described by Daley, et al.[58] This coding scheme will be refined during the pilot visit. We will devise a quantitative rating scale for each system dimension to facilitate discrimination in an objective manner. In addition, general qualitative observations about the system of care will be made. These observations will be supplemented with impromptu interviews with staff members as appropriate throughout the course of the day. The goal is to understand the way in which the pre-operative assessment area,

operating room, and recovery room function as a system. The team will debrief immediately after each day of the site visit. In the weeks that follow, they will collate, document and code the observations made on-site.

Structure	*Process*

Team Factors
• Training / education
• Clear lines of responsibility for patient-centered tasks
• Interface between other clinical and support services in the hospital
• Quality monitoring
• Safety culture
• Surgical leadership
• Nursing leadership
• Surgeon staffing
• Nurse staffing
 	
System Factors
• Physical layout
• Environmental factors
  o Ergonomics
  o Lighting
  o Room temperature
• Technology and equipment
  o Availability of appropriate technology
  o Team familiarity
  o Functioning
  o Effective allocation
	Team Factors
	• Communication
	  o Team familiarity
	  o Tone
	• Collaborative decision-making
	• Coordination
	  o Shared mental model
	  o Cooperation
	  o Anticipation
	  o Workload
	• Technical competence of staff
	• Organizational factors
	  o Scheduling
	  o Handoffs
	  o Breaks
	• Use of error-catching protocols
	• Readbacks / confirmation across transfer of information
	
	System Factors
	• Disruptions / interruptions
	  o Phone/pager
	  o Visitor
	  o Chit-chat
	• Organizational factors
	  o OR scheduling
	  o Turn over times
	  o Coordination between OR and pre-op and recovery room
	  o Time pressure / overbooking

C. *Meetings with Staff* - The second day of the site visit will consist of pre-scheduled one hour meetings with representatives of each of the three disciplines: nursing, surgery, and anesthesia. A semi-structured interview will be designed to target system factors and organizational safety culture factors. These meetings will be audiotape and transcribed. We anticipate the results of these meetings will be complementary to our field observations.

B. Data analysis

Once all site visits are complete, the documentation, coding, and rating of the observations are finalized, and all audio tapes transcribed, the status of the institution as a high or low outlier will be revealed to the observational team and [redacted] who will work with the observational team on the data analysis. Both a quantitative and qualitative analysis will be performed.

1. Quantitative Analysis – We will start by correlating safety culture as measured by the SAQ with high or low outlier status. We will then correlate our observations from our site visits with outlier status and SAQ score. Our quantitative ratings will allow us to identify which system features have the strongest impact on safety culture and the rate of clinically-significant APL or POHH.

2. Qualitative Analysis - Grounded theory analysis will then be used to review the records from the site visits. Using the finalized dataset, we will identify themes across the hospitals. The goal is to provide a qualitative description of the mechanism by which the system features identified in the quantitative analysis contribute to or mitigate the occurrence of clinically-significant APL or POHH and ultimately surgical safety. Specific examples and contrasting cases will help to support our results.

C. Data interpretation

We hypothesize that we will be able to identify patterns across institutions that make them more or less resilient to the occurrence of a technical error in the operating room. First, we believe that SAQ scores will correlate with rates of clinically-significant APL or POHH for the entire cohort. We also believe that we will be able to identify similar structural and process characteristics that contribute to system performance. These components of the structure and process of the system will be the basis for developing safety improvement interventions.

D. Member Checking and Survey

A critical component of qualitative research is the evaluation of the credibility of the results. Member checking is a process by which the research findings are relayed back to the source population to validate the findings and interpretation. Using focus groups culled randomly from the staff at the hospitals in our sample population, we will develop and validate a survey that would include both safety culture questions as well as items addressing the "system" of care. The goal is to utilize the entire cohort of 65 hospitals to establish a statistically significant correlation between safety culture and system factors with the presence of clinically significant PSIs as guided by our obervations.

IV. Timeline and Benchmarks for Research Proposal

A. Timeline

Activity	Responsible Team Member	Timeframe
- Data quality/frequency analysis - Data element definition	[redacted] study team	Year 1
- Definition of data elements finalized - Study populations and dataset finalized for Aim 1 and 2	[redacted] study team	Year 1
- Aim 1 Analysis	[redacted] study team	Year 1 - 2
- Aim 2 Analysis	[redacted] study team	Year 2

- Aim 3 Hospital identification and recruitment - Human subjects and other permissions obtained - Randomization and site visit preparation	[redacted] study team Human Factors consultant	Year 3
- Final reporting and clinical articles created for publication from Aim 1 and 2	[redacted] study team	Year 3
- Aim 3 Site visits and analysis	[redacted] study team Human Factors consultant	Year 4
- Aim 3 Member checking focus groups - Final reporting and publications from Aim 3	[redacted] study team Human Factors consultant	Year 5

B. Benchmarks

Success will be measured according to the following deliverables:

- The cohort will be identified and final dataset ready for analysis by end of Year 1
- The analysis described in Aims 1 and 2 will be complete by end of Year 2
- Preparatory work including institution identification, recruitment and approvals will be completed by end of Year 3
- Manuscripts from Aim 1 and 2 will be submitted by end of Year 2
- Site visits will be completed by end of Year 4
- Manuscripts from Aim 3 will be submitted by end of Year 5

V. Potential Problems and Alternative Strategies

A. Potential Technical Issues with Data Acquisition

We have demonstrated the feasibility of accessing hospital data on a daily basis via a secure on-line link and loading it into a large relational data system. While systems problems may arise, we are skilled in maintaining and troubleshooting such systems.

We are experienced in healthcare operations and data analysis, and therefore should be able to identify delays and/or gaps in clinical data acquisition due to input omissions, standalone systems or operational issues.

B. Missing Data

Despite the fact that the data elements we will be using are very common and we are familiar with the data within our system, we anticipate that there will be missing data. We will conduct comparative analyses to determine if cases with and without missing data are similar with respect to variables which will be observed on all patients, such as age, sex, and hospital. If subjects with and without missing data are similar on characteristics observed on all subjects, then subjects without missing data are likely a representative sample, and our analyses can be based on these subjects. However, we will perform sensitivity analyses in order to determine if our results change under different missing data mechanisms. We will use established methods for conducting regression analyses for clustered data with random or non-random missing data. If the outcome data are missing at random, we will get unbiased regression coefficient estimates using modified GEE (MGEE) for clustered data [59]. Next, we will posit various

non-ignorable selection mechanisms, which typically lead to the most bias, and see if our results change much under these scenarios. If the results do change under non-ignorable missingness, then we will use an approach proposed by Ibrahim et al. to analyze these data [60].

C. Cost does not predict 'clinically-significant post-operative deviation'
It is possible that our hypothesis is incorrect and cost cannot serve as a proxy for experiencing a 'clinically-significant post-operative deviation'. If this is the case, we will look to develop an alternative approach to stratifying clinical significance based on other outcomes and resource utilization measures that are available within administrative data. Examples of such potential markers include return to the operating room, LOS, or discharge to a higher level of care. Such an algorithm will still achieve our goal of providing a widely generalizable approach to identifying APL and POHH events occurring in the operating room but leading to 'clinically-significant post-operative deviation'.

D. No variation across hospitals
Another potential challenge is that there may not be significant variation in the rate of APL or POHH leading to 'clinically-significant post-operative deviation'. If this is the case, we will still be able to complete Aim 3 of the study. Our first approach will be to look at the secondary analysis for specific procedures, such as colon surgery or cardiac surgery, with higher rates of APL or POHH to determine if there is significant variation. If so, we will perform procedure-specific site visits and evaluations. Even if this is not the case, there is much still to be learned about the system features that influence surgical safety. To date, much of the work in this area has been single-institution studies and almost exclusively only involves academic medical centers. [redacted]still offers us a convenience sample of U.S. community hospitals. We will select 12 hospitals at random from the sample of 65 and perform site visits as described in Aim 3.

E. No identifiable system features to target for intervention
It is also possible, but highly unlikely, that there will be no discernable patterns regarding the system factors that contribute to the occurrence of APL or POHH events upon our analysis described in Aim 3. This would be important and valuable information that could be used to inform future safety improvement initiatives. If this were to be the case, rather than best practices for dissemination, our end result would potentially represent a process for reporting high and low outlier status to hospitals and a methodology for self-assessment regarding current structure and process characteristics.

Protection of Human Subjects

Human subjects approval for Aim 1 and 2 of this study is pending and we anticipate exemption. All identifiable patient or provider data will be maintained by [redacted] personnel. Investigators at the primary institution, [redacted] will only be provided de-identified data.

As detailed in Aim 3 of the Research Plan, Human Subjects approval will be obtained from each institution prior to the site visit. Approval is not feasible until completion of Aim 2 as it is unclear which of the hospitals in the cohort will be used for this analysis.

The primary risks involved in this analysis relate to issues of data management as it pertains to data security, confidentiality and privacy. We therefore describe our technical specifications and measures to taken to protect the rights of the subjects in detail here.

A. Technical Specifications

Data in the [redacted] data system originates from the Meditech Clinical Information Systems (CIMS) located at each of the participating hospitals. The Meditech system is an operational software system, written in Magic, a proprietary version of MUMPS, and is not a relational data system. [redacted] collects these data and standardizes them, storing the data in an Oracle data system running on LINUX servers.

We are currently running our production systems on IBM x3X50 servers with Red Hat Enterprise Linux operating systems, Oracle 10.2.0.1 and Postgres 8.2.11. We have 15 Terabytes (TB) of storage. We utilize industry standard message formats and use standard coding conventions as appropriate, for example, LOINC and SNOMED codes for laboratory test names and results.

The [redacted] system utilizes an XML compliant service that can be reached via an HTTPS connection after appropriate authentication. This system supports future extensions for bi-directional messaging. The IHM system sits behind a firewall where a traditional HTTPS port is open (as is normal for secure web access). Messages are in the industry standard XML format and include standardized HL7 Version 2.4 message content to support secure bi-directional communications with our system.

Whether Meditech systems are MAGIC or Client-Server (CS) based determine the method of how information is transmitted back to [redacted] CS systems require [redacted] to send a 1U rack-mountable machine running WindowsXP to the hospital, where hospital IT staff are instructed to secure the machine in accordance with hospital IT policy. All extraction is done locally on the [redacted] CS Machine and transmitted over the VPN via FTP to the [redacted] Data Center. MAGIC systems do not require [redacted] hardware to be sent on-site; however, data is still transferred over the VPN via FTP to the [redacted] Data Center.

The [redacted] Data Center has Intrusion Detection/Prevention systems and explicit firewall rules to secure this highly sensitive information. As security is a paramount concern, the system, network and security events are monitored on a 24x7x365 basis.

Only two (2) [redacted] located devices are allowed to participate in the VPN tunnel between the [redacted] Data Center and the Hospital network. Any number of explicitly defined hospital

devices are allowed to participate in the VPN tunnel, after their necessity is validated. All data is encrypted when sent via VPN to the [redacted] Data Center.

[redacted] requires pre-shared keys and a defined endpoint prior to VPN establishment. [redacted] is able to entertain any number of encryption methods for VPN establishment, primarily recommending AES encryption. [redacted] also seamlessly handles various NAT and RFC addressing issues that usually pose problems during large scaled deployments. Hospital policies and capabilities may be unable to support AES encryption; however the only other method allowed by [redacted] is 3DES (TDES).

AES encryption meets the requirements of the National Institute of Standards and Technology (NIST) and the Federal Information Processing Standard (FIPS), FIPS-197[41]. This standard specifies a FIPS-approved cryptographic algorithm that is used to protect electronic data. AES, approved in 2001, is a symmetric block cipher that can encrypt (encipher) and decrypt (decipher) information.

3DES encryption is less used today, as AES has superseded it with better security and performance benefits. Nonetheless, 3DES meets the requirements of the National Institute of Standards and Technology (NIST) and Federal Information Processing Standard (FIPS), FIPS-46-3 [41].

Encryption converts data to an unintelligible form called ciphertext; decrypting the ciphertext converts the data back into its original form, called plaintext. The AES algorithm is capable of using cryptographic keys of 128, 192 and 256 bits to encrypt and decrypt data in blocks of 128 bits. Stored data from messages are protected using strong authentication and role based access control. Further, in the document CNSS Policy No. 15, FS-1, June 2003[42], NSA has approved use of AES with 192 and 256 bit key lengths for materials classified TOP SECRET. The 3DES algorithm has three cipher operations, where three keys are used, allowing for 168-bit encryption (56bits x 3keys). 3DES is the predecessor of DES, where only one key was used & was less secure. Combining explicitly defined nodes, pre-shared keys, 168 to 265-bit encryption and secure password protection standards allows for a more than adequate way of securely transferring sensitive medical data.

We are experienced in large information technology production systems and processing and have created systems that can be monitored and expanded as required. We use standard system tools (Solarwinds; IBM Director; syslog-ng, etc) to continuously monitor operations and security (e.g. Tipping Point, Checkpoint, VeriSign/ Thawte). We have implemented standard operations controls to document modification of products systems in order to minimize unexplainable system events that may be caused by configuration changes. These are done in accordance with best practices as outlined in the Systrust and ISO 17799 specifications.

[redacted] has a staff of experienced data system, engineering and operations and network personnel to maintain the system described in this proposal.

B. HIPAA:

[redacted] serves as a "business associate" of its hospital partners, in accordance with §164.501 of HIPAA. As a business associate, [redacted] binds itself by contract to meet all of the federal regulatory requirements for "covered entities" under HIPAA, 45 CFR 160 and 164. A copy of [redacted]'s standard business associate agreement is included in the appendix

to this proposal.

As a business associate of member hospitals, [redacted] aggregates patient-level data for purposes of supporting member hospitals in the specific functions defined by HIPAA as "health care operations" including "quality assessment and improvement activities, outcomes evaluation, development of clinical guidelines...population-based activities relating to improving health or reducing health care costs and protocol development." Research activities are conducted with IRB review in accordance with § 164.516 of the HIPAA regulations and the federal Common Rule 45 C.F.R. § 46.107 or on a "limited data set" as defined by 164.514. Any other uses of data will only be performed on de-identified data as defined by § 164.514. [redacted] maintains data in full compliance with the technical security requirements set out in 45 CFR 160, 162 and 164.

C. Confidentiality:

[redacted] has made a commitment as part of its mission to build data security and confidentiality in keeping with the highest ethical and legal standards. Our goal is to exceed all HIPAA requirements by setting the highest standards for data privacy and security. The [redacted] Ethics Advisory Board oversees the activities of [redacted] with regard to all research conducted with patient data and has the ability to veto any research protocols that may violate ethical or legal standards.

D. Privacy:

Policies and procedures regarding use within the [redacted] system surpass all applicable state and federal privacy laws, including the HIPAA privacy regulations.

Our model is based upon the following principles:

- All [redacted] personnel undergo a background check prior to employment and are required to sign confidentiality statements upon hiring. Violation of those statements is grounds for termination;
- All [redacted] staff are trained in the appropriate use of data under the [redacted] policies and under applicable federal and state law;
- Only those [redacted] personnel who have a legitimate purpose have access to any data held by [redacted];
- [redacted] staff only have access to the minimal amount of data necessary to fulfill a particular purpose;
- All access to the [redacted] repository by [redacted] personnel is made by biometric identifier;
- Every purpose for which access is granted is documented;
- Patterns of use within [redacted] are randomly audited on an on-going basis to detect occurrences outside of normal tolerances.

E. Security:

[redacted] data security system meets all the requirements of Office of Management and Budget (OMB) in OMB Circular No. A-130, Appendix III - Security of Federal Automated Information Systems [43]. The system is secure in compliance with normal protocols for handling of sensitive data exceeding present HIPAA requirements for information security. Information extracted from this system complies with HIPAA privacy requirements as specified in 45 CFR 164.

Data in the [redacted] data system is structured as follows:

- At the source location (i.e. a hospital), all data are sent combined, encrypted and transmitted via secure link to [redacted]
- Upon entrance to [redacted] the encrypted identifiers and any facial identifiers such as Name, Social Security Number, etc. are segregated from the rest of the record (record detail), checked to assure uniqueness within the system and placed in a secure transformation and loading zone. The record detail are given a sequence number, which will lookup the cluster ID to link this information into a persons record.
- The cluster ID is a number that allows information with the [redacted] system to be linked into coherent records without person specific information being present. The system storing the mapping unique person to cluster ID is kept in a high security system.
- The record detail is cleaned and normalized and placed in the operational zone of the [redacted] repository.

Inclusion of Women and Minorities

The inclusion of women and minorities will reflect the gender and racial composition of the patients who undergo an inpatient procedure at one of the partner IHM hospitals. Approximately 60% of patients are female and 80% are white. African Americans constitute 9% of the patients who undergo surgery at an [redacted] hospital, while 4% are Hispanic, 2% Asian American, 1% Native American and 4% are unknown.

Targeted/Planned Enrollment Table

This report format should NOT be used for data collection from study participants.

Study Title: [redacted]

Total Planned Enrollment: 407,097

TARGETED/PLANNED ENROLLMENT: Number of Subjects			
Ethnic Category	**Females**	**Males**	**Total**
Hispanic or Latino	9,933	6,351	16,284
Not Hispanic or Latino	238,396	152,417	390,813
Ethnic Category: Total of All Subjects *	248,329	158,768	407,097
Racial Categories			
American Indian/Alaska Native	2,483	1,588	4,071
Asian	4,967	3,175	8,142
Native Hawaiian or Other Pacific Islander	2,483	1,588	4,071
Black or African American	22,350	14,289	36,639
White	216,046	138,128	354,174
Racial Categories: Total of All Subjects *	248,329	158,768	407,097

* The "Ethnic Category: Total of All Subjects" must be equal to the "Racial Categories: Total of All Subjects."

Inclusion of Children

We will neither specifically include nor exclude children but do not expect them to be significantly represented in our cohort based on the procedure codes used to define our cohort. These are not procedures typically performed on children.