

Hospital Harm Performance Measure Methodology Report: Opioid-Related Adverse Events

Submitted By:

Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation
(YNHHSC/CORE)

Prepared For:

Centers for Medicare & Medicaid Services (CMS)

November 2017

Table of Contents

1. EXECUTIVE SUMMARY.....	5
2. INTRODUCTION	7
2.1 Overview of Measure.....	7
2.2 Harm as Quality Indicator	8
3. MEASURE DEVELOPMENT APPROACH	11
3.1 Overview	11
3.2 Harm Prioritization.....	11
3.3 Stakeholder Engagement for Harm Prioritization.....	12
3.4 Data Sources	13
3.5 Cohort	14
3.6 Outcome	14
3.7 Risk Adjustment	14
3.8 Measure Calculation	15
3.9 Measure Testing.....	15
3.10 Measure Development Approach Summary	17
4. OPIOID-RELATED ADVERSE EVENTS	18
4.1 Opioid-Related Adverse Events Introduction	18
4.2 Approach to Opioid-Related Adverse Events Measure Development.....	20
5. SUMMARY	27
6. APPENDICES	28
Appendix A. Acknowledgement Details	28
Appendix B. Importance and Feasibility Prioritization Exclusion/Inclusion Criteria	32
Appendix C. Candidate List of Harms for Prioritization.....	33
Appendix D. Person and Family Engagement Network Survey Results	37

Appendix E.	Results from TEP and Technical Advisory Group Survey	38
Appendix F.	Summary of Patient Demographics	40
References		41

Center for Outcomes Research & Evaluation (CORE) Project Team

Susannah Bernheim, MD, MHS – Project Co-Lead and Director of Quality Measurement Programs
Kelly Kyanko, MD, MHS** – Project Co-Lead
Lisa Suter, MD* – New Measure Division Director and Associate Director of Quality Measurement Program
Shu-Xia Li, PhD – Lead Analyst and Associate Director for Data Management and Analytics
Kun Wang, PhD – Supporting Analyst
Zhenqiu Lin, PhD – Analytic Leadership and Director of Data Management and Analytics
Lynette Lines, MST, PMP – Project Manager
Kerry McCole, M.S.Ed, M.Phil.Ed – Project Coordinator
Jaymie Simoes, MPH – Supporting Project Coordinator
Nicole Cormier, MPH – Supporting Project Coordinator
Loralee Crowder, BS – Research Associate
Abby Alonso, BA – Research Assistant II
Becca Dendy, BA – Research Assistant II
Fior Rodriguez, BS – Research Assistant II
Elizabeth Triche, PhD – Research Scientist
Anita Arora, MD, MBA* – Research Scientist
Richard Andrew Taylor, MD, MHS* –Assistant Professor of Emergency Medicine; Director, ED Clinical Informatics and Analytics
Karen Dorsey, MD, PhD* – Associate Director of Quality Measurement Programs
Sarwat Chaudhry, MD* – Clinical Consultant
Harlan M. Krumholz, MD, SM* – Principal Investigator

*Yale School of Medicine

**New York University School of Medicine

Acknowledgements

This work is a collaborative effort and the authors gratefully acknowledge the individuals that provided feedback, experiences, and advice on the development of this patient-reported outcome-based performance measure; our Technical Expert Panel (TEP) members; our Technical Advisory Group members; Technical Working Group members and Consultants Dr. David Bates, Dr. David Classen, and Dr. Peter Pronovost; Subject matter experts Dr. Kasia Lipska, Dr. Francis Perry Wilson, and Dr. Lee Fleisher; The Person and Family Engagement Network; and Drs. Lein Han, Joseph Clift, and Pierre Yong at the Centers for Medicare & Medicaid Services (CMS) for their contributions to this work (please reference [Appendix A](#) for detailed acknowledgement list)

1. EXECUTIVE SUMMARY

Patient harm associated with delivery of care remains unacceptably high; in 2008 an estimated 13.5% of hospitalized Medicare beneficiaries experienced harm during their hospital stays¹. There is consensus among stakeholders on the need for new and improved measures of hospital harm to provide timely and accurate data and inform hospitals on their patient safety efforts². The broad availability of electronic health record (EHR) data presents an opportunity to address these measurement gaps.

The goal of the Hospital Harm Performance Measure is to assess multiple types of patient harm in the hospital setting using EHR data. We defined harm as any physical or psychological injury that occurs in the acute hospital setting as a result of active delivery of care or substandard care across all healthcare domains (diagnostic, treatment, preventive, and others). We aimed to develop a measure that would assess the rate at which specific harms occur in the hospital setting using a valid method that reliably allows comparison across hospitals. This measure is expected to inform and incentivize hospitals' harm reduction efforts and thus reduce their costs of medical interventions. Furthermore, hospital harm reduction will decrease death and disabilities and enhance patient satisfaction and quality-of-life.

While our ultimate goal is to develop a single summary measure that combines multiple harms to provide a broad view of patient safety and hospital care quality, we are committed to ensuring each component harm is measured in a rigorous fashion. For this reason, and to potentially expedite the availability of harm measures for implementation, we started measure development by creating multiple individual measures, each measuring one harm.

This report describes our approach to the selection of harms for initial development based on the harm's importance for measurement, feasibility of extracting data from the EHR, and input from multiple stakeholder groups including content experts and patients. Based on our environmental scan/literature review, a Technical Expert Panel (TEP), expert technical consultants, and the Person and Family Engagement Network input, we prioritized the following harms for initial measure development: hospital-acquired pressure injury, acute kidney injury, falls, opioid-related adverse events, and hypoglycemia. **In this report, we are only presenting the specifications and testing results of the opioid-related adverse events harm measure.** We are still conducting testing on the aforementioned harms and identifying future harms for measure development.

The report outlines the overall approach to measure development and testing, followed by detailed measure specifications for opioid-related adverse events. To develop preliminary specifications for each individual harm, we built on published harm specifications when available. Our aim was to develop specifications that will assess harms that are broadly defined when possible to capture as substantial a range of patients and outcomes. We also aimed to keep specifications as simple and straightforward as possible to ensure usability. Testing of the measures is being performed in two stages in hospital EHRs and includes clinical adjudication of harm and validation of individual data elements. Phase 1 testing confirmed the feasibility of measurement and validation of the harm outcomes. Phase 2 testing using the electronic specifications will further validate the measure and data elements.

We find in our initial testing that our specifications for the opioid-related adverse events harm measure are feasible to implement in a hospital EHRs and have high positive predictive value (94.6 % overall). Testing results, including clinical adjudication, demonstrated validity of capturing opioid-related adverse events using EHR data. Application of this measure will provide hospitals with reliable and timely measurement of their rates of this harm. Implementation will advance safe use of opioids in hospitals and prevent these serious and potentially lethal adverse drug events.

2. INTRODUCTION

2.1 Overview of Measure

The Centers for Medicare & Medicaid Services (CMS) contracted with Yale New Haven Health Services Corporation/Center for Outcomes Research and Evaluation (CORE) to develop a Hospital Harm Performance Measure based on Electronic Health Record (EHR) Data (hereinafter, Hospital Harm Performance Measure) for the adult all-payer population.

Hospital patient safety remains a critical focus of quality measurement and improvement efforts. The national effort to measure and improve performance on patient safety has resulted in quantifiable harm reductions and cost savings for hospitals and payers. The Department of Health and Human Services estimates 3.1 million occurrences of harmful hospital acquired conditions (HACs) were prevented between 2010 and 2015 with cost savings of \$28 billion^{3,4}. These harms include infections, pressure injury, falls, and other adverse events. Despite these improvements, as of 2015 harm rates among all patients remain high at 115 harm events per 1,000 discharges⁴.

Patient safety experts assert that achieving faster and greater reductions in hospital harm requires enhancements to our current approach. These include measuring more meaningful outcomes and using more valid and readily available data sources, like EHR data^{5,6}. An ideal measure of hospital harm provides a mechanism for identifying harms accurately, consistently, and rapidly across providers to ensure visibility to patients, providers, and policy makers and incentivize care improvement. It further results in a stronger patient safety and accountability culture while reducing healthcare costs across all settings².

The goal of the Hospital Harm Performance Measure is to assess multiple types of patient harm in the hospital setting for the adult all-payer population based on EHR data. We began by developing a number of measures of individual harms that could be extracted from the EHR, with a goal in future years to eventually incorporate them into a single measure. CORE developed these measures for hospital-level measurement, though we aim to make them flexible for application to other provider groups.

Activities of this first year of measure development included: 1) Engaging CMS and stakeholders to select and prioritize among harms and provide input on testing and specifications, 2) Performing review of literature and environmental scans to identify prior specifications on several harms, 3) Defining preliminary electronic specifications for feasibility testing of each harm outcome, 4) Testing EHR-based harm specifications and comparing with clinical adjudication, and 5) Reviewing testing results to finalize harm specifications. Fiscal Year 2018 (Year 2) will involve continued engagement with stakeholders on measure specifications, finalization of cohort and risk adjustment approach, e-specification of the measure, and completing testing of the Measure Authoring Tool (MAT) output in hospital EHRs. In addition, in the next phase of work, we will begin to consider additional harms to incorporate and define the means of rolling-up multiple harms into a single measure.

The persistently high hospital harm rate, its impact on health outcomes and cost, gaps in existing systematic and reliable measurement tools, and the increasing availability of reliable and timely EHR

data for measure development and implementation all support focus on development of an EHR-based hospital harm measure. This measure builds on prior work and will significantly advance patient safety.

2.2 Harm as Quality Indicator

2.2.1 Importance of Harm Measurement

Researchers ranked medical error as the third leading cause of death in 2013, confirming its severe impact on patient health outcomes⁷. The U.S. Department of Health and Human Services reported a decline in the rate of harm from 145 per 1,000 discharges in 2010 to 115 harms per 1,000 discharges in 2015⁴. Despite this, researchers estimated that over 400,000 deaths in 2013 were still due to medical error. Variation in harm rates across hospitals further signaled room for improvement. As an example, hospital fall rates ranged between 1.3⁸ to 8.9⁹ falls per 1,000 patient days¹⁰.

Hospital harm types vary by severity, incidence, and preventability. Looking at the most common harms, in 2013, the Agency for Healthcare Research and Quality (AHRQ) reported over one million pressure injuries, 760,000 adverse drug reactions to medicines that control blood sugar in diabetics, 290,000 catheter-associated urinary tract infections, and 240,000 falls. The Office of Inspector General (OIG) reports that 44% of hospital harm events are preventable¹¹.

The impact of hospital harm varies from being temporary, as is the case with some adverse drug events, to long-term, such as lasting disability. Depending on their impact, harms may also influence other critical patient outcomes like length of stay, cost, readmissions, and mortality. One study reported that hospitalizations of patients harmed by medical care cost almost 20% more and lasted 15% longer than hospitalizations free of medical harm¹². Another study of Medicare patients with acute myocardial infarction reported that as a hospital's adverse events rate increased by 1%, its average 30-day risk-standardized mortality rate increased by 4.86% and the 30-day risk-standardized unplanned readmission rate increased by 3.44%¹³.

Hospital harm often leads to increased healthcare costs. In October 2008 alone, hospital care associated with adverse and temporary harm events cost Medicare approximately \$324 million¹. The Society of Actuaries Health Section documented that medical errors resulted in a cost of \$13,000 per error, producing a total cost to the United States economy of \$19.5 billion per year. This study estimated the 2,500 excess deaths cost \$1.4 billion. The authors also note \$1.1 billion due to loss of productivity as evident by short-term disability claims for the 10 million excess missed days from work due to medical errors¹². Others claim the cost of these deaths to be much higher (reaching \$1 trillion) when analyzing the data for quality-adjusted life years¹⁴.

2.2.2 Quality and Measurement Gaps

A number of national initiatives by CMS and others aim to address patient harm within hospitals. These incorporate a range of harms of varying scope and importance and utilize a variety of reporting mechanisms. The Hospital-Acquired Condition Reduction Program (HACRP) currently modifies hospital payment based on two domains: 1) AHRQ Patient Safety Indicator (PSI) 90, a composite of eight AHRQ PSIs (for example, pressure injury rates, accidental punctures or lacerations) and 2) Centers for Disease

Control and Prevention (CDC) National Healthcare Safety Network (NHSN) rates for healthcare-associated infections such as central line associated blood stream infections (CLABIs) and catheter-associated urinary tract infections (CAUTIs). Some harms, such as falls, hypoglycemia, objects left in during surgery, and pressure injuries receive non-payment through the HAC Present on Admission Program but are not publicly reported. While these initiatives provide an important surveillance function, they do not capture all harms and, in most cases, do not fully utilize the EHR. For now, this measure development project will not examine those harms captured by existing measures that currently utilize EHR data with clinical adjudication (i.e. the NHSN measures), but will consider harms included in other existing measures.

2.2.3 Preventability

National efforts to improve patient safety have demonstrated the success of focused measurement and improvement activities. The Department of Health and Human Services estimated that national efforts focused on adverse drug events, falls, infections and other forms of harm prevented the occurrence of 3.1 million harm events nationally between 2010 and 2015⁴. This effort to measure, track, and reduce hospital-acquired conditions, like bloodstream and urinary tract infections, pressure injuries, and adverse drug events, resulted in a 21% decline in HACs between 2010 and 2015, saving approximately 125,000 lives and \$28 billion⁴.

Despite these improvements, there is widespread agreement in an opportunity to reduce patient harm within hospitals^{4,7}. In particular, stakeholders have voiced concerns that the absence of a standard hospital harm measure is slowing down progress on patient safety improvement and reporting efforts^{5,7}. With a systematic EHR-based hospital harm measure in place, hospitals will be able to more reliably assess harm reduction efforts and modify improvement efforts in near real-time. Additionally, we can expect greater achievements in reducing hospital harms and enhancing hospital performance on patient safety outcomes.

2.2.4 Feasibility and Usability of an EHR-Based Harm Measure

Efforts to measure, track, and reduce hospital harm at the national level have resulted in encouraging results. For example, efforts targeting HACs between 2010 and 2015 resulted in a 21% decline in rate of HACs and saved nearly \$28 billion in healthcare costs⁴. The Hospital Harm Performance Measure under development will provide further comparative information across hospitals to aid patients in their decision-making regarding healthcare. As well, the measure will provide hospitals with benchmarking and actionable data to inform their quality improvement efforts. The use of EHR data also will provide a potential for real-time feedback to further a hospital safety culture and potentially mitigate harm.

Although reporting harm using EHR-based measures may require hospitals to initially invest resources, we anticipate that such investments are required to adopt a fully operational EHR system. The overall impact should produce reductions in healthcare costs. Results of the American Hospital Association national survey estimated that 84% of non-federal acute care hospitals had a basic EHR as of 2015, representing a 10% increase since 2014¹⁵. Also, 96% of these non-federal acute care hospitals reported having possession of a certified EHR through legal agreement¹⁵. As EHR adoption rate continues to

increase, we expect most U.S. hospitals will have the means of collecting and submitting the required data for the Hospital Harm Performance Measure by the time this measure is ready for implementation. Of note, basic EHR adoption rate among hospitals was quite high (80%) even among critical access hospitals and small rural hospitals¹⁵.

Having an EHR is only the first step to securing accurate and reliable data for measuring hospital harm. The quality of our measure results depends on the reliability of the data extracted from the structured fields in the EHR. We will test the feasibility of extracting the measure outcomes (types of harms) accurately from the EHR, thus ensuring the measure results will have high levels of validity and reliability. In order to reduce hospital burden, we aim to build measures based on extraction from structured fields in the EHR that are currently consistently completed during the course of clinical care and do not require natural language processing for analysis. We also aim to build a measure that will not require changes in clinical workflow to produce key data elements and will not require complex processing to capture measure results.

Stakeholders generally support the use of clinical data from the EHR for measuring hospital performance on patient safety metrics. Therefore, it is reasonable to expect general acceptance of an EHR data-based hospital harm measure. Using EHR data instead of administrative data allows for more credible, real-time measure results to support hospital quality improvement efforts for hospital harm reduction^{2,5,16}.

Based on our combined expertise in measure development, reevaluation, and implementation in public reporting and performance-based payment programs, we designed the measure specifications to enhance applicability to several CMS programs.

3. MEASURE DEVELOPMENT APPROACH

3.1 Overview

The goal of the Hospital Harm Performance Measure is to assess multiple types of patient harm in the hospital setting for the adult all-payer population based on EHR data. This measure will ultimately include multiple domains of harm across a broad cohort of hospitalizations. We began by selecting a limited set of individual harm measures based on their importance to stakeholders and feasibility of extraction from the EHR. We developed specifications and a testing plan for this first set of harms, with a goal in future years of measure development to eventually incorporate all harms into a single measure that can be expanded to include additional harms. To develop preliminary specifications for each individual harm, we built on published harm specifications when available. Our aim is to develop specifications that will assess harms that are broadly defined when possible to capture as substantial a range of patients and outcomes. We also aimed to keep specifications as simple and straightforward as possible to ensure usability. These preliminary specifications may be updated based on the results of measure testing.

Testing is being performed in two phases in hospital EHRs and includes clinical adjudication to confirm the validity of the outcome and individual data elements:

Phase 1: Testing of measure logic in multiple hospital EHRs with on-site clinical adjudication for the harms which will confirmed the feasibility of the measure specifications and provide validity testing of the outcome. Initial Phase 1 feasibility testing was done through a collaboration with a Patient Safety Organization (PSO). Building on the PSO's existing software infrastructure provided a patient-level, de-identified EHR data extract, along with clinical adjudication for both measure feasibility and validity testing through manual medical record abstraction. Testing determined whether the measure could be reliably calculated purely using EHR data available in structured fields.

Phase 2: Testing of data element and measure validity by confirming the accuracy of the extracted data elements and each instance of harm identified using the MAT output, EHR data query, and medical record abstraction. We will develop and test the harm measures' MAT output in several additional hospitals (outside of our collaboration with the PSO). This will ensure that we have tested both data element validity and measure feasibility in several hospitals and several EHR environments.

3.2 Harm Prioritization

There are numerous harms that a patient can experience while hospitalized, with impact spanning from mild, temporary discomfort to never events such as wrong site surgery. To identify initial harm outcomes to be included in the measure, we consulted with experts, conducted several literature reviews, and engaged with stakeholders, including patients. We envision this work as ongoing, adding new harm measures over time through the same development process.

As described in more detail below, we selected initial harms for inclusion based on importance and feasibility of extracting data from the EHR, with input from content experts and patients. After having

completed this Phase 1 testing, we are presenting the feasibility testing results for opioid-related adverse events.

3.2.1 Criteria for Harm Selection

During our feasibility assessment prior to beginning measure development, we identified a broad spectrum of harm outcomes that could be included within a hospital harm measure. In Fiscal Year 2017, we developed criteria for prioritizing and selecting a set of harms to incorporate into an initial measure. We reviewed these criteria with technical consultants, the Person and Family Engagement Network that supports CORE measure development, and CMS. In particular, we focused on identifying harms that are *important* and *feasible* to measure ([Appendix B](#)). *Importance* was based both on the severity of harm (i.e. potentially leads to prolonged hospital stay, permanent injury, life-sustaining intervention, or death) and on preventability (i.e., potentially avoidable with best practices). During this initial year of development, we selected only those harms that result in variable, wide ranging, or severe injury to patients for the severity of harm. For preventability, we considered only those harms that are partially preventable with best practices or fully preventable. *Feasibility* was assessed on whether all data needed to assess the harm outcome could be obtained from structured fields within EHRs, so minimal manual chart review or adjudication of harm would be required. For the feasibility criterion, we include harms with a moderate or easy degree of difficulty to extract from the EHR.

The two criteria, *importance* and *feasibility*, were applied to a comprehensive list of harms ([Appendix C](#)) compiled by our team. This initial prioritization allowed us to begin examining the literature on a few harms while we sought broader stakeholder input on harm selection. Potential harms were identified from: AHRQ's Medicare Patient Safety Monitoring System (MPSMS), PSI-90, the National Quality Forum (NQF) Serious Reportable Events, and Pascal/Institute for Healthcare Improvement (IHI) Electronic Global Trigger Tool (eGTT) tools, AHRQ Common Formats, World Health Organization International Classification for Patient Safety (WHO ICPS), as well as medical literature.

3.3 Stakeholder Engagement for Harm Prioritization

We engaged several stakeholder groups for their input on harm prioritization. These data will be used to guide ongoing selection of harms. Stakeholder groups included: *Person and Family Engagement Network*, *Technical Working Group*, *TEP*, and *Technical Consultants*. We solicited input from each of these groups to inform our selection of the first five harms for measurement. We plan to return to our summary of stakeholder input to determine additional harms to measure in subsequent phases of development.

3.3.1 Person and Family Engagement Network

To obtain patient and caregiver input on prioritizing harms, we surveyed the Person and Family Engagement Network that supports CORE measure development. This network consists of patients, family caregivers, consumers, and advocates. These individuals collaborate with CORE on many measure development projects and provide recommendations related to quality measurement to ensure our

efforts meet the needs of all health care stakeholders, including those constituent groups that make up the Network.

The person and family engagement survey, fielded December 2016, asked individuals to categorize a list of 16 harms (identified and aggregated from the harm prioritization process) into higher, medium, and lower priority and list any additional harms they would like to see in future phases of measurement. The survey had a 71% response rate (n=30/42) and comprehensive results are available in [Appendix D](#). Patients and caregivers also corroborated CORE's criteria for harm selection: severity and preventability. They noted a priority for harms that are frequent and may be overlooked by hospital staff.

3.3.2 Technical Consultants

The technical consultant stakeholder group consists of clinicians and patient safety experts. This group recommended prioritizing: hypoglycemia and other glycemic control events, blood loss after surgery, delirium related to medication, falls without injury, and venous thromboembolism (VTE). The technical consultants also advised on initial specifications for our prioritized harms.

3.3.3 TEP and Technical Advisory Group Survey

The TEP is comprised of 13 experts in clinical medicine, performance measurement, coding and informatics, and representation from a consumer perspective. We engaged the TEP through surveys and a webinar. We also consulted a broader Technical Advisory Group comprised of 23 additional technical experts that provided input to supplement the TEP. These groups provided input on both harm prioritization, feasibility of extracting data for each harm from the EHR, and our overall measure concept and development and measure specifications.

In March 2017, we conducted an online survey with our TEP and Technical Advisory Group to prioritize harms by importance and feasibility of electronic specification. We had an 89% (n=32/36) response rate. 81% of the individuals self-identified as having EHR expertise to answer questions regarding the feasibility of electronically specifying harms. Individuals ranked a list of harms by importance and feasibility to extract from the EHR. Results from this survey are in [Appendix E](#). We also followed up with both groups during respective meetings to solicit any additional input on harm prioritization.

3.4 Data Sources

This measure is designed to be implemented solely with EHR data from structured fields. We do not plan to use free text data which would require the use of natural language processing (NLP), text mining, and/or machine learning in developing initial specification¹⁷. These methods are increasingly important tools in healthcare that transform narrative text, or unstructured data, such as progress notes or radiology reports into structured data that can be read and analyzed by machines. However, NLP is not routinely used across healthcare systems and is not usable for our purposes in this measure, though as hospitals' capabilities advance it may be useful in future measure iterations.

The data source for Phase 1 testing for feasibility and validity of the initial specifications was done in five hospitals within a PSO. This testing was completed using one year of data (June 1, 2016 to May 30, 2017)

for all eligible admissions, representing a total of 66,130 admissions. These hospitals use two different EHR vendors (Cerner and Epic). They represent a variety of bed sizes (between 100-199 beds and 300-399 beds), teaching and non-teaching hospitals, are located in two different states, and are all in an urban location. For a breakdown of patient demographics by hospital, reference [Table 4](#) in [Appendix F](#).

3.5 Cohort

3.5.1 General Approach and Overview

Our aim in defining individual measure cohorts is to assess a broad cohort of hospitalized patients at risk for a given harm. Feedback from our first TEP meeting also supports the use of measure that captured more “global” harm to advance a culture of patient safety for all hospitalized patients. Each harm will have its own unique cohort. The base cohort is completed admissions (as indicated by discharge date) for adult patients (18 years and older) within a defined period of time (preliminarily 12 months). For each individual harm measure, we defined specific exclusions relevant to the individual harm.

3.6 Outcome

3.6.1 General Approach to Harm Definition and Overview

We define harm as any physical or psychological injury that occurs in the acute hospital setting as a result from active delivery of care or substandard care across all healthcare domains (diagnostic, treatment, preventive, and others). For each individual harm, we created initial specifications that are as simple as possible and capture a wide range of severity of patient harm. We also do not aim to identify preventability of an individual harm instance or whether each instance of harm was an error, but rather assess the overall rate of the harm within a hospital incorporating a definition of harm that is likely to be overall decreased with hospital best practice. In addition, we utilize structured EHR fields that are currently populated in clinical care and do not require changes in clinician workflow. As well, these structured fields do not require NLP. Though we are currently developing measures for hospital-level measurement, we aim to make them flexible for application to other provider groups.

3.7 Risk Adjustment

3.7.1 General Approach to Risk Adjustment

Clinical characteristics, including a patient’s age, reason for hospitalization, or comorbid conditions all may influence the risk of hospital harm occurring. Therefore, if hospitals care for patients with different degree of risk it may be important for a performance measure of harm to adjust for patient risk factors in order to compare hospital rates. However, many harms should be avoidable, regardless of patient risk. Therefore, risk adjustment will be considered for each harm individually.

We anticipate that some harm measures will not require risk adjustment. In some cases, the rates of a specific harm should not be greatly influenced by patient characteristics (e.g. never events). Moreover, if patient risk can be largely ameliorated by best care practices then it may not necessary to adjust for

differences in hospital case mix. For example, if patients are at high risk of falling they may need closer supervision and modifications to their care that will then ameliorate that risk.

For harms where risk adjustment is necessary, we will consider adjusting the measures for factors present on admission, such as comorbid conditions or the Core Clinical Data Elements (CCDE). The CCDEs are the first captured data values for a basic set of vital signs and laboratory test results measured during the inpatient encounter. CCDEs provide clinical information about the patient's status at admission, and have been established for use in risk adjustment of hospital outcome performance measures. Examples of CCDEs include: demographic information (age, gender), vital signs (heart rate, systolic and diastolic blood pressure, respiratory rate, temperature, oxygen saturation, and weight), and laboratory results (hemoglobin, hematocrit, platelet, white blood cell (WBC) count, potassium, sodium, chloride, bicarbonate, anion gap, blood urea nitrogen (BUN), creatinine, glucose, and troponin).

3.8 Measure Calculation

Individual harm measures are defined as the proportion of patient admissions that experienced the specific harm. That is, the denominator is on a per admission basis rather than hospital day. The numerator is dichotomous as to whether the harm occurred. For example, if a patient experienced more than one opioid-related adverse event during an admission, the measure will count the patient once in the numerator. There are a number of rationales for taking this approach. First, it may be more useful and intuitive from the patient perspective to assess the rate of harm per hospitalization. Additionally, this approach avoids the risk of miscounting individual instances of a harm thus this approach is likely to provide greater measure accuracy. Finally, while for some harms multiple events per admission could be accurately captured using data from the EHR, other harms may not be as accurately captured and require reporting on a dichotomous basis. By reporting all harms in the same manner, it will allow for the combination of multiple harm measures into one composite measure. The means of creating a future composite measure that incorporates multiple individual EHR-based harm measures is still under development.

3.9 Measure Testing

3.9.1 Overview

There are three phases of testing planned for the overall hospital harm measure to assess measure feasibility and validity. Below, we describe the Phases 1 and 2 which are planned for the testing for the individual harm measures. In the future, if the measures are combined into a single composite measure, we will conduct an additional Phase 3 of testing measure.

Phase 1 testing aimed to establish the feasibility of the measure logic and validity of our harm outcome specifications (i.e. the definition of the harm), the measure numerator. In this phase of testing, we defined the data elements and logic used to identify the harms for each measure cohort and harm outcome. Phase 1 of testing was completed in collaboration with a PSO. This organization has as clients a set of hospitals with well-established mapping of multiple portions of their EHRs. These hospitals regularly export data into a database maintained by the PSO. These exports contain all data elements

needed to calculate the harm measures. Using this database, we extracted those instances of harm that met our measure specifications. Clinical adjudicators at each hospital reviewed medical records for each instance of harm and answered a standard set of questions to confirm whether harm criteria were met. This adjudication will provide the measure of validity of each harm outcome.

Phase 1 testing also established some components of data element and measure feasibility. The NQF has set forth a set of feasibility criteria that each EHR data element used to calculate quality measures must meet. These criteria include: 1) data “availability,” or whether the data are readily available in structured format; 2) data “accuracy,” or whether the information contained in the data elements is correct (defined as being from the most authoritative source); 3) data “standards,” or whether the data elements are coded using a nationally accepted terminology standard; and 4) “workflow,” or to what degree the data elements are captured during the course of care. By assessing the availability of each of the data elements required by the harm specification, we demonstrate that those data elements are currently captured for most patients in the measures’ target population under current clinical workflows. Additionally, we assessed data accuracy through the manual abstraction performed by adjudicators at each hospital providing data for Phase 1 testing.

Phase 2 testing will begin with the development and testing of a MAT output for each harm measure. The MAT output includes a human-readable version of the logic that hospitals will use to extract the data. This version is meant to be read and interpreted by hospitals’ health information technology (IT) staff. There is also a machine-readable version, which is meant to eventually interact directly with hospitals’ EHRs, eliminating the need for health IT personnel to interpret and separately code queries to extract data. In creating the MAT output, we will also define the value sets, or national terminology standards, for the data elements required for each harm measure. Next, we will test the MAT output in several hospitals’ EHRs. We will request Health IT staff apply the measure logic by querying their EHR to extract the data elements needed for each measure and calculating a measure result. For a subset of harms identified through each hospitals’ query, we will ask the Health IT staff to provide manually abstracted data to confirm the validity of the data elements extracted using the MAT output. We will also confirm the measure validity by ascertaining the harm outcome through standard clinical review. We will perform this testing in more than one hospital and more than one EHR vendor. This phase of testing will further confirm the data element and measure validity by confirming the accuracy of the extracted data elements and instance of harm identified using the MAT output and EHR data query.

3.9.2 Other Testing Considerations

Along with standard measure testing, a hospital harm measure warrants special considerations. For example, high-performing hospitals that are more diligent in surveying harms may be more apt to find and report them. This can result in surveillance bias, which could inappropriately penalize hospitals for vigilance in monitoring for harm, as has been shown for post-operative venous thromboembolism¹⁸. To address this, our selection process for initial harms targeted those harms that we thought may be less susceptible to such bias. Moreover, by addressing these issues in testing, we aim to examine any evidence that harm rates as determined by our measure are strongly correlated with testing rates.

Finally, we will seek opportunities throughout the measure development process to continue to assess measure validity. This will include systematic assessment to measure face validity with our TEP of national experts and stakeholder organizations.

3.10 Measure Development Approach Summary

In development of the Hospital Harm Performance Measure, we selected five initial harms based on importance and feasibility of extracting data from the EHR and broad stakeholder input. After initial feasibility testing, we are presenting the results for opioid-related adverse events. We aimed to develop specifications as simple as possible and based on existing published specifications when available to maximize usability. We did not aim to determine whether each instance of a harm is preventable. Rather, we aimed to select harms that with best practices, a hospital's overall rate of harm can be decreased. Validity and feasibility testing will be performed in two phases in hospital EHRs as noted above. Future work will aim to combine individual measures into a single summary measure. In the sections that follow, we present details on the opioid-related adverse event harm measure currently under development, including importance and research from our environmental scan and literature review, specifications, and Phase 1 test results.

4. OPIOID-RELATED ADVERSE EVENTS

4.1 Opioid-Related Adverse Events Introduction

4.1.1 Importance of Opioid-Related Adverse Events

Opioids are often the foundation for sedation and pain relief. However, use of opioids can also lead to serious adverse events, including oversedation, delirium, respiratory depression, and death. Opioid-related adverse events have both patient-level and financial implications. Patients have been noted to have 55% longer lengths of stay, 47% higher costs, 36% higher risk of 30-day readmission, and 3.4 times higher payments than patients without these adverse events¹⁹. While noting that data are limited, the Joint Commission suggested that opioid-induced respiratory arrest may contribute substantially to the 350,000-750,000 in-hospital cardiac arrests annually²⁰.

Opioids are among the most frequently implicated medication in adverse drug events, with respiratory depression potentially leading to brain damage or, more seriously, death. A commonly cited estimate of opioid-related adverse respiratory events is 1%, but there appears to be considerable variability²¹. The incidence is difficult to accurately estimate because of differences in definitions of respiratory depression, routes of opioid administration, and patient populations (e.g., surgical vs. general patient population).

Administration of opioids varies widely by hospital, ranging from 5% in the lowest-use hospital to 72% in the highest-use hospital. Notably, hospitals that use opioids most frequently have increased adjusted risk of severe adverse events²². Among surgical patients in a large medical center, 98.6% received opioids and 13.6% experienced an opioid-related adverse drug event.

Most opioid-related adverse events are preventable. Of the adverse drug events reported to the Joint Commission's Sentinel Event database, 47% were due to a wrong medication dose, 29% to improper monitoring, and 11% to other causes (e.g., medication interactions, drug reactions). Additionally, in a closed-claims analysis, 97% of adverse events were judged preventable with better monitoring and response²³. While monitoring is key to prevention of opioid-related adverse events, there is considerable variability in monitoring practices. A 2013 study surveyed nurses from 90 institutions in the U.S. and found that pulse oximetry monitoring was more common than other monitoring methods²⁴. Nonetheless only about 58% reported using intermittent pulse oximetry and only 25% used continuous monitoring for Patient Controlled Analgesia (PCA). End-tidal carbon dioxide (ETCO₂) monitoring was only used for 2.2% of patients on epidural therapy and 1.5% for PCA patients. In another study, data from electronic medical records showed that, among 8 hospitals participating in a voluntary pilot to test a CMS e-quality measure, only 8.4% and 26.8% of patients received recommended assessments every 2.5 and 4.5 hours, respectively²⁵. The between-hospital range of the percentage of patients who were properly monitored was 0.0% to 43.6%.

Our environmental scan identified five adverse drug event measures related to naloxone use and/or opioid adverse events, two of which are EHR-specified measures. These measures are in various stages of development and use, but no fully specified EHR-based measure to identify opioid-related adverse

events exists. One of the EHR measures identified is an electronic clinical quality measure (eCQM) entitled “Patient Controlled Analgesia (PCA) Monitoring eCQM”. The PCA Monitoring eCQM is a process measure which examines PCA monitoring practices and surveillance of patients receiving PCA across hospitals²⁶. The other measure utilizing EHR data titled, “Naloxone Use for Reversal of Opioid Over Sedation per 1,000 Patient Days” from the Health Services Advisory Group (HSAG) examines naloxone use to reverse suspected opioid oversedation²⁷. This measure appears to be developed for quality improvement monitoring but was not fully developed, tested, or endorsed for accountability purposes. There are both EHR and process measures in development and under consideration for use in CMS programs. We will continue to conduct systematic environmental scans and literature reviews to learn more about related measures under development or in use.

4.1.2 Defining Opioid-Related Adverse Events

For this harm, we focused specifically on in-hospital opioid-related adverse events rather than opioid overdose events that happen in the community and may bring a patient into the emergency department. We also only focused on severe adverse events such as respiratory depression and oversedation, which can lead to brain damage and death, and do not include less severe adverse events such as constipation or vomiting that can be associated with opioid use.

Severe opioid-related adverse events are typically identified using either clinical signs and symptoms or naloxone administration. Defining opioid-related adverse events based on signs and symptoms can be challenging. Most studies have relied on intermittent monitoring of oxygen saturation (SpO₂), respiration rate, and/or observation of oversedation for defining adverse events. Cutoffs for defining hypoxia or respiratory depression vary, and supplemental oxygen can mask respiratory depression as identified by these signs. Capnography was found in several studies to be more sensitive than pulse oximetry for identifying respiratory depression²⁸⁻³³, but this and other advanced monitoring methods are not routinely used by most hospitals. Perhaps most challenging are the considerable differences in monitoring that may result in surveillance bias; only 8.4% and 26.8% of patients received recommended assessments every 2.5 and 4.5 hours, respectively²⁵. Hospitals with more stringent monitoring practices may actually show higher rates of opioid-related adverse events if measured by these signs and symptoms.

Naloxone is an opioid reversal agent typically used for severe opioid-related adverse events. Naloxone administration has been used in a number of studies as an indicator of severe opioid-related adverse events; many of these studies have used EHR data, suggesting feasibility of electronic specification. We found several studies assessing the accuracy of using a naloxone event as an indicator of respiratory depression by comparing to chart review, with 68%-91% confirmed by chart review^{34,35}. In some cases, naloxone, particularly low-dose naloxone, was administered for other clinical indications, including inflammation, nausea, and pruritus^{35,36}. In other cases, symptoms were not reversed with naloxone administration, so the naloxone events were deemed unrelated to opioid use (i.e., signs and symptoms due to another underlying cause)^{23,37}.

4.2 Approach to Opioid-Related Adverse Events Measure Development

4.2.1 Opioid-Related Adverse Events Cohort

4.2.1.1 Inclusion Criteria

The cohort for this measure is completed admissions (as indicated by discharge date) for adult patients (18 years and older) within a defined period of time (preliminarily 12 months).

4.2.1.2 Exclusion Criteria

There are no current denominator exclusion criteria for this measure.

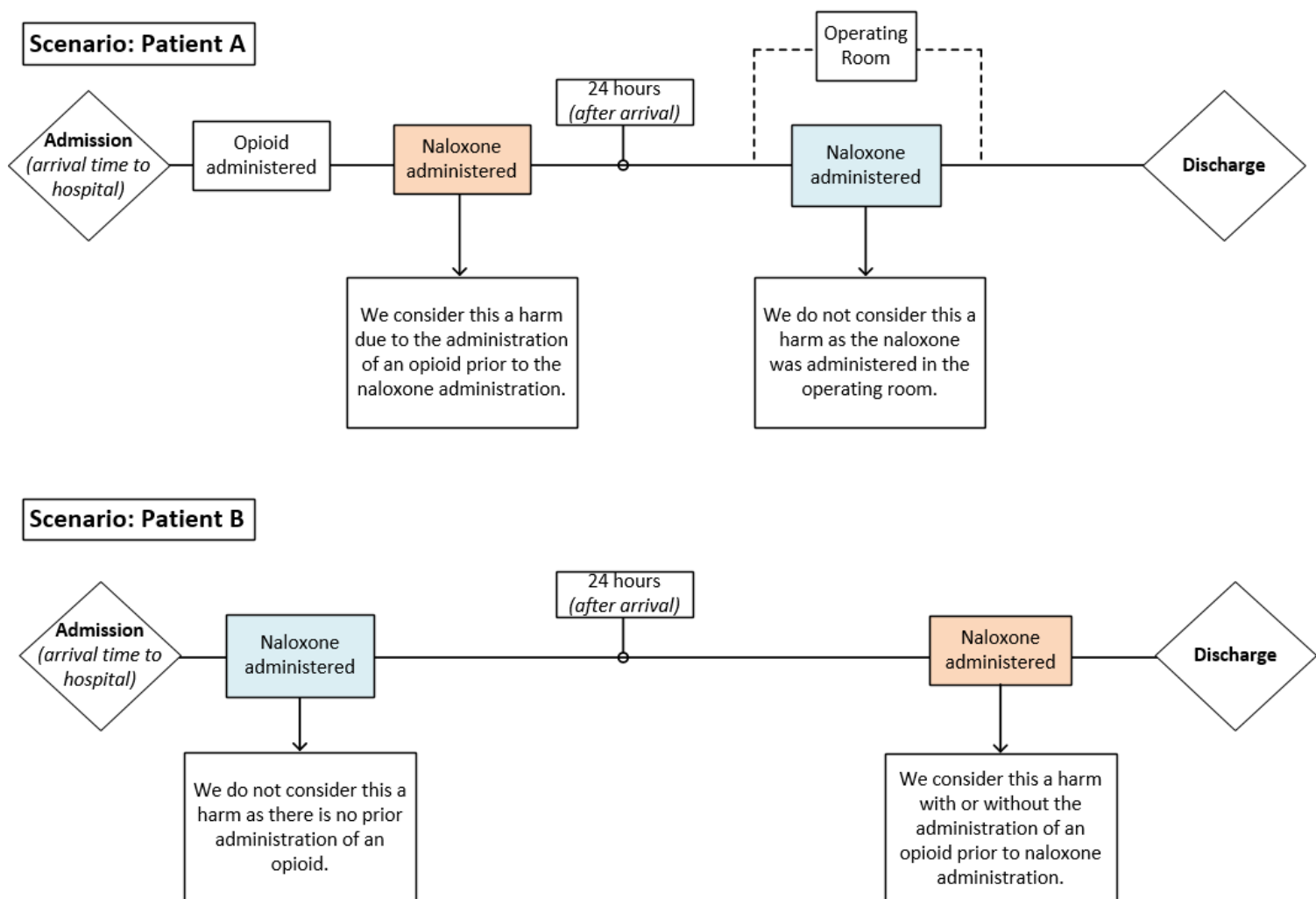
Rationale:

- We aim to capture a broad cohort of patients as most patients in whom opioids are administered in the hospital should not experience such extreme respiratory depression or oversedation to require naloxone.
- We do not limit the denominator to admissions where an opioid was administered (except in the first 24 hours of an admission). Phase 1 testing showed that opioids given in procedural areas (e.g. interventional radiology, bronchoscopy, and endoscopy) may not be captured in EHR data, yet may require naloxone for reversal of oversedation and respiratory depression. Limiting the cohort to only admissions where EHR data showed administration of an opioid could omit these cases. Furthermore, our testing demonstrates that the current specifications, which include all patients, successfully identify those events (i.e. use of naloxone) that are subsequent to over administration of opioids within the hospital, without restricting the measure denominator.

4.2.2 Opioid-Related Adverse Events Outcome

Admissions with administration of a narcotic antagonist (i.e., naloxone) except if the administration a) occurred in in the operation room or b) occurred in the first 24 hours of admission without documentation of hospital administration of an opioid preceding the naloxone (reference [Figure 1](#) for a more detailed description of this outcome definition).

Figure 1. Description of Opioid-Related Adverse Events Outcome Definition



Rationale:

- The numerator definition was adapted from the AHRQ Common Formats Surveillance Beta Release October 2015³⁸ which defined opioid adverse drug events as: *Patient receiving opioids (e.g., morphine, fentanyl, meperidine) during hospital stay and experiencing any of the following within 24 hours: 1) respiratory depression/reduction in adequacy of ventilation (as indicated by any of the following: respiratory rate < 9 breaths per minute, pulse oximetry SpO2 < 85%, arterial blood gas SaO2 < 85%); 2) unresponsiveness or responsive only to noxious stimulation; or 3) administration of either of the following: narcotic antagonist (i.e., IV naloxone), unless administered during or within 2 hours following a procedure, or a Respiratory stimulant (i.e., doxapram).*
- We narrowed the Common Formats harm definition to focus only on naloxone administration. Criteria for respiratory depression as assessed by patient monitoring of respiratory rate and oxygenation status were excluded as they may not be reliably captured in EHR data and could be subject to substantial surveillance bias given the variability in hospital monitoring. Phase 1

testing demonstrated that for patients receiving naloxone, respiratory rate, pulse oximetry SpO₂, and arterial blood gas SaO₂ +/-2 hours of naloxone administration were frequently missing, ranging from 15.4% of events missing oxygen saturation values 2 hours before the naloxone event to 59.8% of events missing SaO₂ two hours before the naloxone event.

- We also did not include the criterion: “unresponsiveness or responsive only to noxious stimulation” because it is not easily captured in EHR data.
- We also we did not include administration of a respiratory stimulant (i.e., doxapram) as this is not standard procedure for opioid overdose.
- The Common Formats criteria for naloxone included the following exclusion: “unless administered during or within two hours following a procedure.” We narrowed this to only naloxone use in the operating room where it would be part of the sedation plan as administered by an anesthesiologist. Use of naloxone for procedures outside of the operating room (for example, bone marrow biopsy) would be inappropriate because it would indicate the patient was over sedated.
- Matching the Common Format definition, we limit the numerator to naloxone administration within 24 hours of opioid administration for events within the first 24 hours of admission. The intended goal is to obtain evidence of hospital administration of opioids to distinguish naloxone use for patients who present to the emergency room with a community opioid overdose from those who have an adverse event related to in-hospital administration of an opioid.
- For naloxone administered >24 hours after admission, we do not require documentation of opioid administration. Phase 1 testing showed that opioids given in procedural areas (e.g. interventional radiology, cardiac catheterization laboratory, and endoscopy) may not be captured in EHR data, yet may require naloxone for reversal of over sedation and respiratory depression. Clinical adjudication showed that with the current specifications, 92.2% of patients were given an opioid medication given within 24 hours prior to the naloxone administration. Also, sensitivity analyses showed that the requirement of opioid administration for naloxone events >24 hours did not significantly change the positive predictive value of the measure (94.6% vs. 96.4%)

4.2.3 Opioid-Related Adverse Events Risk Model Development and Candidate Variables

We do not anticipate requiring risk adjustment for this measure as most instances of opioid-related adverse events should be preventable. There are several risk factors that affect sensitivity to opioids, including age, sex, weight, concurrent medications (particularly central nervous system (CNS) depressants/other sedating medications), chronic opioid use, and comorbid cardiac, renal, and/or respiratory disease. However, physicians should take these factors into account when dosing opioids. Risk adjustment is only needed if certain hospitals have distinctly different risk profiles for their patients that cannot be addressed by best care. In a July 2017 survey of TEP and Technical Advisory Group Members, 61% (14/23) of survey respondents did not think this measure required risk adjustment, 22% (5/23) thought it should, and 17% (4/23) were unsure.

4.2.4 Opioid-Related Adverse Events Measure Results

Below are results from Phase 1 testing for feasibility and validity of the opioid-related adverse events measure specifications. Testing was completed in five hospitals within a PSO with two different EHRs systems using one year of data (June 1, 2016 to May 30, 2017) for all eligible admissions, representing a total of 66,130 admissions, as described in [Section 3.4](#).

Table 1. Opioid-Related Adverse Events Admission-Level Positive Predictive Value

	Number of admissions in which harm occurred identified through EHR	Number of admissions verified by clinical adjudication	Positive Predictive Value (PPV)
Hospital 1	18	16	88.9%
Hospital 2	38	35	92.1%
Hospital 3	85	83	97.7%
Hospital 4	75	75	100.0%
Hospital 5	62	54	87.1%
Total	278	263	94.6%

Table 2. Opioid-Related Adverse Events Overall and Subgroup Harm Rate, % of Admissions in which Harm Occurred, by Hospitals

	Total	Hospital 1	Hospital 2	Hospital 3	Hospital 4	Hospital 5
Total number of admissions	66127	5529	10448	14467	18070	17613
Harm rate: percentage of patient admissions in which harm occurred	0.4%	0.3%	0.4%	0.6%	0.4%	0.4%
Harm rate by age group						
18 - 54	0.2%	0.3%	0.1%	0.4%	0.2%	0.2%
55 - 64	0.5%	0.3%	0.4%	0.7%	0.4%	0.4%
65 +	0.6%	0.3%	0.6%	0.7%	0.6%	0.5%
Harm rate by gender						
Female	0.4%	0.4%	0.4%	0.6%	0.5%	0.3%
Male	0.4%	0.3%	0.3%	0.6%	0.4%	0.5%
Unknown	0.0%	-	0.0%	-	0.0%	0.0%
Harm rate by race						
Missing Race	0.0%	-	-	-	-	0.0%
American Indian or Alaska Native	0.7%	0.0%	0.0%	1.9%	0.0%	0.0%
Asian	0.5%	11.1%	0.5%	0.5%	0.2%	0.5%
Black or African American	0.5%	0.4%	0.5%	0.6%	0.5%	0.0%
Decline	0.0%	-	0.0%	-	0.0%	0.0%
Native Hawaiian or Other Pacific Islander	0.5%	0.0%	0.7%	0.6%	0.0%	0.0%
Other Races	0.3%	0.0%	0.2%	0.6%	0.2%	0.3%
Unknown	0.2%	-	0.0%	0.0%	0.0%	0.9%
White	0.4%	0.3%	0.3%	0.6%	0.4%	0.4%

Table 3. Summary of Additional Opioid-Related Adverse Events Clinical Adjudication Questions

	Total (N=423)	Hospital 1 (N=23)	Hospital 2 (N=49)	Hospital 3 (N=125)	Hospital 4 (N=117)	Hospital 5 (N=109)
Percentage of patients with naloxone administered because of excessive opioid medication administration	93.9%	95.7%	89.8%	96.8%	100.0%	85.3%
Percentage of patients where symptoms reversed with naloxone	81.8%	69.6%	79.6%	77.6%	86.3%	85.3%
Percentage of patients given naloxone in the post-acute care unit to reverse anesthesia	3.8%	8.7%	0.0%	2.4%	0.9%	9.2%
Percentage of patients with an opioid medication given within 24 hours prior to the naloxone administration	92.2%	95.7%	89.8%	94.4%	94.9%	87.2%

*Charts were reviewed to determine if naloxone administration represented a valid opioid-related adverse event.

4.2.5 Opioid-Related Adverse Events Measure Results Summary

Overall admission-level positive predictive value for the opioid-related adverse events measure was 94.6%, with a range across the five hospitals of 87.1%-100.0%. We are unable to assess negative predictive value from this data; this will be calculated in the next phase of testing. However, we believe that it is unlikely that a hospital would have a number of severe respiratory events related to opioid administration that would not be followed by naloxone administration. The rate of the harm was 0.4% of all admissions (range across hospitals: 0.3%-0.6%). Rates among studies using naloxone administration as a surrogate measure of respiratory depression were 0.3% (0.1-1.3 %)²¹, supporting the validity of this measure.

Overall, the measure captured the intended harm: In 93.9% of events, adjudicators noted that naloxone was administered because of excessive opioid medication administration, and for 81.8% of events patients' symptoms were reversed with naloxone. On review of clinical adjudication data, in only two of the events was naloxone given for pruritus. In only 3.8% of events the naloxone was given in the post-anesthesia care unit to reverse anesthesia. The measure also captured iatrogenic or hospital harm, rather than events of opioid overdose that occurred in the community and required naloxone administration in the emergency room. An opioid medication was given to the patient in the hospital within 24 hours of the naloxone in 92.2% of events. In two events, the principal diagnosis in claims data for the admission was for opioid poisoning (poisoning by heroin, undetermined, initial encounter; poisoning by other opioids, accidental (unintentional), initial encounter), however clinical adjudicators still found that the opioid-related adverse events for these admissions were iatrogenic harm.

As with all potential harms, it will be important to consider unintended consequences of measurement and how to best minimize them. One of the main concerns for this measure is the unintended consequence of potentially discouraging clinicians from giving naloxone for altered mental status and respiratory depression. Unintended consequences for this measure will require consideration by CMS and stakeholders in preparation for implementation.

In summary, testing results demonstrated the feasibility and validity of capturing opioid-related adverse events using EHR data. Application of this measure will provide hospitals with reliable and timely measurement of their rates of this harm. Implementation will advance safe use of opioids in hospitals and prevent these serious and potentially lethal adverse drug events.

5. SUMMARY

Hospital harm continues to occur despite a national focus on hospital quality and patient safety. To address this, stakeholders recommend strategies to achieve greater reductions in hospital harm and prioritize the development of a systematic and reliable measurement tool to accurately capture hospital harm events^{2,16}. Evidence supports that measurement of harms has produced reductions in harm events⁷, however, there is more work to do to eliminate preventable hospital harm.

While a number of hospital harm measures are currently in use, stakeholders criticize their validity, reliability, and reliance on retrospective administrative data². CMS can strengthen its harm measurement strategy by comprehensively reassessing the types of hospital harms that are most important, employing more reliable methods for comparing hospital performance, and using clinical data available in the hospital EHR. With 84% of hospitals having a basic EHR system (and 96% having legal agreement for a certified EHR system)⁵, it is now possible to readily obtain the necessary clinical data for development and subsequent implementation of a novel hospital-level harm measure.

6. APPENDICES

Appendix A. Acknowledgement Details

We would like to acknowledge the ongoing support from our technical consultants. These individuals gave generously of their time, providing guidance on key clinical and methodological decisions.

CORE Technical Consultants:

David Bates, MD, MSc, FACMI

Chief, Division of General Medicine at Brigham and Women's Hospital, and Medical Director, Clinical and Quality Analysis at Partners Healthcare System, Inc., Boston, MA

David Classen, MD

Chief Medical Officer, Pascal Metrics, Co-Chair, of National Quality Forum Common Formats Committee, Salt Lake City, UT

Peter Pronovost, MD, PhD, FCCM

Senior Vice President for Patient Safety and Quality and Director of the Armstrong Institute for Patient Safety and Quality at Johns Hopkins Medicine, and Co-Director of Cardiac Surgical Intensive Care Unit at Johns Hopkins University School of Medicine, Baltimore, MD

Lee Fleisher, MD

Robert D. Dripps Professor and Chair of Anesthesiology and Critical Care and Professor of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

Kasia Lipska, MD, MHS, BS

Assistant Professor of Medicine, Endocrinology, Yale School of Medicine, New Haven, CT

Francis Perry Wilson, MD, MS

Assistant Professor of Medicine, Nephrology, Yale School of Medicine, New Haven, CT

We would also like to thank the members of a combined TEP called by CMS. The TEP members provided important insight and feedback on key measure decisions for the development of the hospital-level patient-reported outcome-based performance measure.

TEP members:

David Baker, MD, MPH

Executive Vice President for Health Care Quality Evaluation, The Joint Commission, Oakbrook Terrace, IL

Cynthia Barnard, PhD, MBA, MSJS

Vice President, Quality, Northwestern Memorial HealthCare, Chicago, IL

Lisa Freeman

Patient Advocate

Executive Director Connecticut Center for Patient Safety, Fairfield, CT

Patrick Guffey, MD

Associate Professor, Department of Anesthesiology, University of Colorado Department of Anesthesiology, Aurora, CO

Julia Hallisy

Patient Advocate

Founder, The Empowered Patient Coalition, San Francisco, CA

David Hopkins, MS, PhD

Performance Measurement, Stanford University, Stanford, CA

Kevin Kavanagh, MD, MS

Board Chairman, Health Watch USA, Somerset, KY

Joseph Kunisch, PhD, RN-BC, CPHQ

Enterprise Director of Clinical Quality Informatics, Memorial Hermann Hospital System, Houston, TX

Timothy Lowe, PhD

Director of Research Services, Premier Inc., Charlotte, NC

Jennifer Meddings, MD, MSc

Assistant Professor in Internal Medicine, Division of General Medicine, University of Michigan Health System, Ann Arbor, MI

Amita Rastogi, MD, MHA, CHE, MS

Chief Medical Officer, HCI3/Altarum Institute, Newtown, CT

Eric Thomas, MD, MPH

Professor of Medicine, McGovern Medical School at University of Texas Health, Houston, TX

Karen Zimmer, MD, MPH

Instructor and Associate Professor, Department of Pediatrics, Jefferson School of Population Health, Philadelphia, PA

We would also like to acknowledge the contributions of our Technical Advisory Group. The Technical Advisory Group provided important insight and feedback on key measure decisions.

Technical Advisory Group Members:

Ruric (Andy) Anderson, MD, MBA, FACP

Chief Medical Officer, Aurora Health Care, Milwaukee, WI

J. Matthew Austin, PhD

Assistant Professor, Armstrong Institute for Patient Safety and Quality, Johns Hopkins Medicine, Baltimore, MD

Ann Borzecki, MD, MPH

Physician-Researcher, Department of Veteran's Affairs, Bedford, MA

John Bott, MSSW, MBA

Healthcare Performance Measurement, Consumer Reports, Yonkers, NY

Kyle Bruce, DPM, MPH

Podiatrist, Riverbend Medical Group, Springfield, MA

David Chang, PhD, MPH, MBA

Associate Professor of Surgery, Massachusetts General Hospital, Boston, MA

Hazel Crews, MHS, MHA, CPHQ

Regional Director of Quality; Patient Safety Officer and Director of Clinical Research, Indiana University Health, Bloomington, Indiana

Melissa Danforth, BA

Vice President for Hospital Ratings, The Leapfrog Group, Washington DC

Richard P. Dutton, MD, MBA

Chief Quality Officer, United States Anesthesia Partners, Dallas, TX

Marybeth Foglia, RN, PhD, MA

Health Care Ethicist, Veterans Health Administration, Seattle, WA

Jeffrey Giullian, MD, MBA

Vice President of Medical Affairs, Chief Medical Officer and National Medical Director for Hospital Services Group, DaVita Kidney Care, Denver, CO

Maryellen Guinan, JD

Policy Analyst, America's Essential Hospitals, Washington, DC

Kate Kovich, MS, OTL, CPPS

Vice President, Patient Safety, Advocate Health Care, Downers Grove, IL

David Levine, MD, FACEP

Senior Vice President Advanced Analytics and Medical Director, Vizient, Chicago, IL

Karen Lynch, BSN

Director of Analytics Applied Informatics, Massachusetts General hospital, Boston, MA

Milisa Manojlovich, PhD, RN

Associate Professor of Nursing, University of Michigan, Ann Arbor, MI

Barbara Pelletreau, RN, MPH

Senior Vice President of Patient Safety, Dignity Health, San Francisco, CA

Marc Philip Pimentel, MD, MPH

Anesthesiologist and Assistant Medical Director for Quality and Safety, Brigham and Women's Hospital, Boston, MA

Christine Sammer, DrPH, RN, CPPS, FACHE

Director of Corporate Patient Safety, Adventist Health System, Altamonte Springs, FL

Brett Stauffer, MD, MHS

Associate Chief Quality Officer, Baylor Scott and White Health, Dallas, TX

Scott Trenhaile, MD

Orthopedic Surgeon, Orthoillinois, Rockford, IL

Brooks Udelsman, MD, MHS

General Surgery Resident, Post Graduate Year 3, Massachusetts General Hospital, Boston, MA

Boback Ziaieian, MD, PhD

Clinical Instructor, UCLA Division of Cardiology, Los Angeles, CA

Appendix B. Importance and Feasibility Prioritization Exclusion/Inclusion Criteria

- Importance
 - Degree of Harm:
 - **Excluded:** 1) minimal;
 - **Included:** 2) variable or wide range – potentially severe; 3) Severe – leading to prolonged hospital stay, permanent harm, life-sustaining intervention, or death.
 - Preventability:
 - **Excluded:** 1) not preventable;
 - **Included:** 2) partially preventable; 3) fully preventable.
- Feasibility
 - Difficulty of electronic extraction:
 - **Excluded:** 1) Difficult;
 - **Included:** 2) Moderate; 3) Easy.
 - Reliability/validity/risk for surveillance bias:
 - Will be determined by future testing

Appendix C. Candidate List of Harms for Prioritization

- Adverse Drug Events
 - Blood clots and other occlusions related to medication (deep vein thrombosis (DVT)/pulmonary embolism (PE))
 - Chemotherapy-related adverse drug events
 - Contrast dye-related acute renal injury
 - Medication-related electrolyte imbalance
 - IV overload/pulmonary edema
 - Medication-related acute renal insufficiency/renal injury
 - Medication-related allergic reaction
 - Medication-related bleeding (IV heparin, low molecular weight heparin, and warfarin adverse drug events)
 - Medication-related cardiac event/arrhythmia
 - Medication-related coagulopathy
 - Medication-related constipation
 - Medication-related dehydration
 - Medication-related delirium, confusion, or over sedation
 - Medication-related diarrhea
 - Medication-related glycemic events (insulin adverse drug events/hypoglycemia)
 - Medication-related headache
 - Medication-related hypotension
 - Medication-related nausea and vomiting
 - Medication-related skin-mucosal reaction/itching
 - Medication-related neurological complication (e.g. seizure)
 - Medication-related respiratory complication
 - Transfusion-related event/blood incompatibility/death or disability from unsafe administration of blood products
 - Digoxin adverse drug events
 - Patient death or serious disability associated with a medication error
 - Opioid-related adverse events/naloxone administration
- Healthcare-Associated Infections
 - Candidiasis infection related to antibiotic use or other medication
 - Catheter-associated urinary tract infection
 - Central line-associated bloodstream infections
 - Healthcare-associated clostridium difficile infection
 - Peripheral or central line related non-bloodstream infections
 - Postoperative/post-procedure fever/infection (non-wound associated)
 - Respiratory infection (non-ventilator associated)
 - Septicemia/bloodstream infections (non-central line related)
 - Ventilator-associated pneumonia
 - Methicillin-resistant Staphylococcus aureus bacteria/infection in a sterile site
 - Vancomycin-resistant Enterococci infection in a sterile site
- Patient Care/Management Events
 - Altered nutrition

- Cardiac arrest
- Catheter-related urinary retention
- Constipation/obstipation
- Death
- Electrolyte imbalance
- Fall (with and without injury)
- Intensive care unit psychosis/hospitalization-related delirium
- IV infiltration
- Intubation-related event
- Mental status change
- Pressure injury
- Poor glycemic control/diabetic ketoacidosis/secondary diabetes
- Readmission
- Respiratory complications/aspiration
- Accidental puncture or laceration/skin tear, abrasion, or other breakdown
- Stroke or other neurological complication
- VTE
- Diagnostic errors/death or serious injury resulting from failure to follow up or communicate laboratory, pathology, or radiology test results
- Failure to rescue from a complication of care or underlying illness
- Iatrogenic pneumothorax
- Artificial insemination with the wrong donor sperm or egg
- Patient death or serious injury from irretrievable loss of an irreplaceable biological specimen
- Exacerbation of preexisting medical condition
- Prolonged weakness or dizziness
- Immobility
- Perinatal Care/Labor and Delivery Events
 - 3rd or 4th degree lacerations after vaginal delivery
 - Administration of general anesthesia during childbirth
 - Maternal hypotension requiring intrauterine resuscitation
 - Maternal tissue trauma, bruising and/or laceration due to instrument delivery
 - Postpartum hematoma, cervical laceration or other bleeding events
 - Postpartum hemorrhage resulting in blood product administration
 - Retained placenta or tissue
 - Tachysystole with non-reassuring fetal heart rate due to oxytocin use
 - Neonate death or serious injury associated with labor and delivery in a low-risk pregnancy
 - Maternal death or serious injury associated with labor and delivery in low-risk pregnancy
- Surgical/Procedure-Related Events
 - Retention/removal of retained foreign body
 - Removal, injury, or repair of organ
 - Complications related to peripheral venous or arterial puncture
 - Epidural/spinal-related event
 - Hypotension related to blood loss after procedure
 - Neurological complication related to surgery or procedure
 - Postoperative/post-procedure fever/infection (non-wound associated)

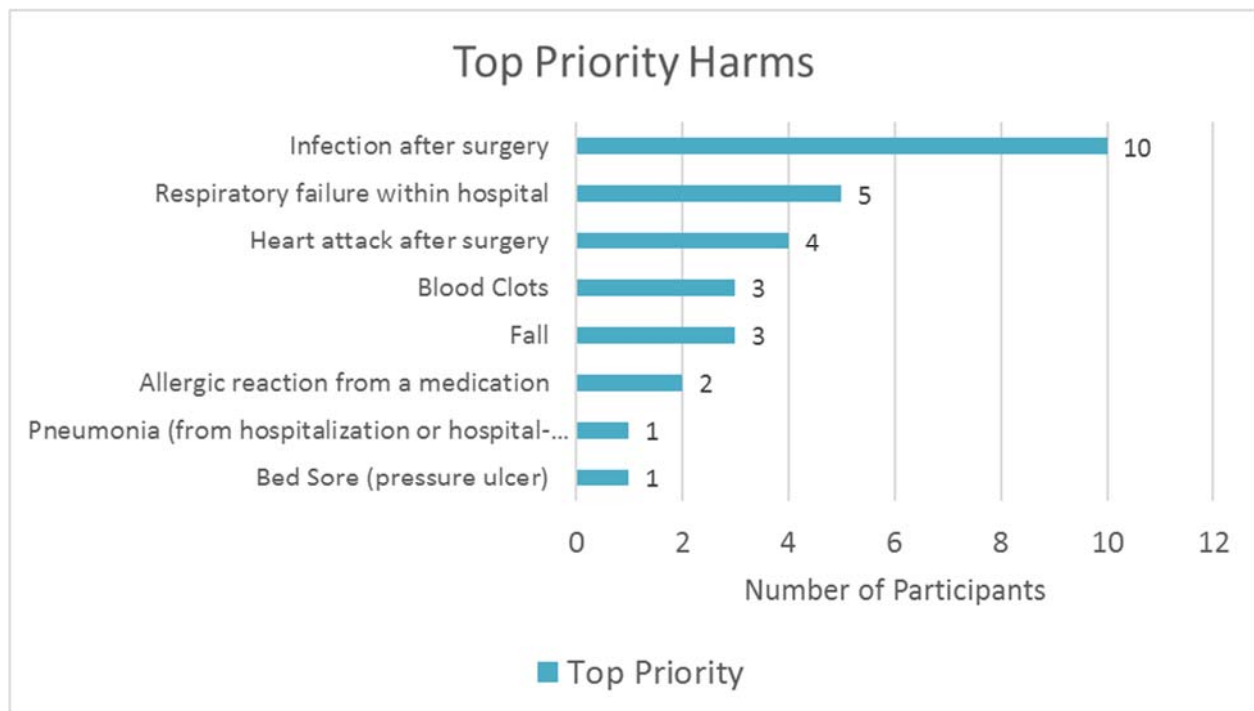
- Postoperative/post-procedure acute renal failure
- Postoperative/post-procedure ileus
- Postoperative/post-procedure nausea and vomiting
- Postoperative/post-procedure pain
- Postoperative/post-procedure urinary retention
- Post-spinal tap headache
- Premature extubation causing respiratory failure/intubation/reintubation/Bilevel Positive Airway Pressure in Post-Anesthesia Care Unit
- Radiation-related injury
- Abnormal bleeding/blood loss/hematoma following surgery or procedure
- Cardiac complications related to surgery or procedure
- Respiratory failure/complications related to surgery or procedure
- Wound dehiscence
- Adverse events after femoral artery puncture for angiography
- Adverse events after hip replacement
- Adverse events after knee replacement
- Intraoperative or postoperative death in American Society of Anesthesiologists (ASA) Class I (i.e. low risk) patient
- Mechanical adverse events associated with central line placement
- Perioperative DVT or PE
- Postoperative pneumonia
- Postoperative hip fracture
- Postoperative physiologic and metabolic derangement
- Postoperative sepsis
- Surgical site/wound infection
- Surgery on wrong body part
- Surgery on wrong patient
- Wrong surgical procedure
- Urinary catheter associated trauma
- Potential Criminal Events
 - Abduction of a patient
 - Any instance of care ordered by or provided by someone impersonating a physician, nurse, pharmacist, or other licensed health care provider
 - Death or significant injury of a patient or staff member resulting from a physical assault (i.e., battery)
 - Sexual assault/abuse on a patient or staff
- Environmental and Occupational Events
 - Death or serious injury of a patient or staff associated with the introduction of a metallic object into the magnetic resonance imaging area
 - Any incident in which a line designated for oxygen or other gas to be delivered to a patient contains the wrong gas or is contaminated by toxic substances
 - Patient or staff death or serious disability associated with a burn
 - Patient or staff death or serious disability associated with electric shock
 - Patient death or serious disability associated with the use of restraints or bedrails
- Patient Protection Events

- Discharge of patient unable to make decisions to unauthorized person
 - Patient death or serious disability associated with patient elopement for more than four hours
 - Patient suicide, attempted suicide, or self-harm resulting in serious disability while being cared for in a health care facility
- Product/Device Events
 - Patient death/disability associated with contaminated drugs, devices, or biologics provided by the health care facility
 - Patient death/disability associated with use or function of a device in patient care, when device is used for function other than as intended
 - Patient death/disability associated with intravascular air embolism that occurs while being cared for in a health care facility
 - Delay in surgery due to equipment malfunction
 - Equipment failure/malfunction
 - Equipment-related event

Appendix D. Person and Family Engagement Network Survey Results

We engaged with the National Partnership for Women and Families to survey the Person and Family Engagement Network to obtain patient and caregiver input on prioritizing harms. The Person and Family Engagement Network consists of patients, family caregivers, and patient advocates. These individuals provide periodic recommendations related to quality measurement and collaborate with CORE on many projects to ensure their efforts meet the needs of all health care stakeholders, including those constituent groups that make up the Network. This survey asked individuals to categorize a list of 16 harms (identified and aggregated from the process mentioned previously) into higher, medium, and lower priority and list any additional harms they would like to see in future measurement. We had a 71% response rate, with the majority of respondents prioritizing infection after surgery ([Figure 2](#)). Patients and caregivers corroborated CORE's list of criteria for harm selection: severity, preventability, frequency. They also noted a priority for harms that may be overlooked by hospital staff.

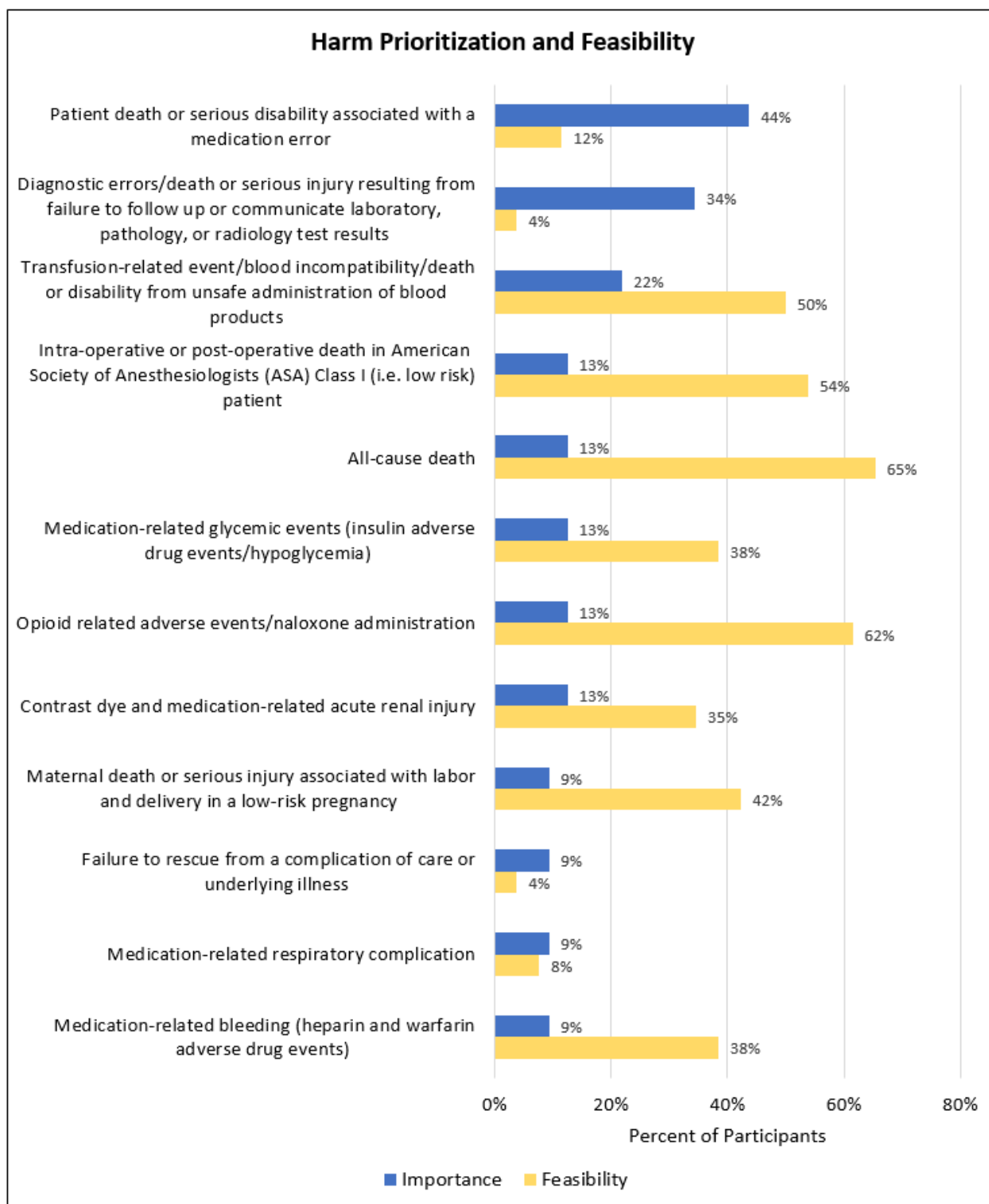
Figure 2. Survey Results – Person and Family Engagement Network



Appendix E. Results from TEP and Technical Advisory Group Survey

In March 2017, we conducted an online survey with our TEP and Technical Advisory Group members to prioritize harms by importance and feasibility of electronic specification. 32 of 36 participants (89%) completed the survey. Of the 32 participants, 26 (81%) self-identified as having EHR expertise to answer questions regarding the feasibility of electronically specifying harms. Data can be found in [Figure 3](#). In [Figure 3](#) the importance value (blue, or top/first bar) is derived from the responses to the question asking all participants to rank their top three harms from those they had prioritized as “High Importance”. The bar represents the percent of participants who ranked each harm in their top three most important harms to measure. The feasibility value (yellow, or bottom/second bar) is derived from the responses to the survey question asking participants who self-identified as having EHR expertise to identify the level of feasibility (easy, moderate, difficult, or unsure) of all harms ([Appendix C](#)). The bar represents the percent of participants with EHR expertise who selected each harm as “easy” to specify in the EHR.

Figure 3. Priority Harms for Measurement Based on Importance and Feasibility Ranked by Technical Expert Panel and Technical Advisory Group



Appendix F. Summary of Patient Demographics

Table 4. Summary of Patient Demographics, Overall and By Hospitals

	Total (N=66127)	Hospital 1 (N=5529)	Hospital 2 (N=10448)	Hospital 3 (N=14467)	Hospital 4 (N=18070)	Hospital 5 (N=17613)
Mean age at admission, years, (SD)	58.7 (20.4)	65.5 (19.0)	57.4 (20.5)	61.3 (18.3)	59.2 (20.0)	54.7 (21.9)
Male, N, (%)	27647 (41.8%)	2565 (46.4%)	3773 (36.1%)	6739 (46.6%)	7865 (43.5%)	6705 (38.1%)
Female, N, (%)	38475 (58.2%)	2964 (53.6%)	6673 (63.9%)	7728 (53.4%)	10204 (56.5%)	10906 (61.9%)
White, N, (%)	42613 (64.5%)	5087 (92.0%)	4929 (47.2%)	9715 (67.2%)	14622 (80.9%)	8260 (46.9%)
Asian, N, (%)	5258 (8.0%)	9 (0.2%)	1898 (18.2%)	1604 (11.1%)	1007 (5.6%)	740 (4.2%)
Black or African American, N, (%)	6421 (9.7%)	246 (4.5%)	1958 (18.7%)	1958 (13.5%)	1575 (8.7%)	684 (3.9%)
American Indian or Alaska Native, N, (%)	155 (0.2%)	3 (0.1%)	31 (0.3%)	54 (0.4%)	47 (0.3%)	20 (0.1%)
Other Races, N, (%)	10380 (15.7%)	181 (3.3%)	1304 (12.5%)	786 (5.4%)	483 (2.7%)	7626 (43.3%)
Decline, N, (%)	46 (0.1%)	0 (0.0%)	2 (0.0%)	0 (0.0%)	5 (0.0%)	39 (0.2%)
Unknown, N, (%)	571 (0.9%)	0 (0.0%)	56 (0.5%)	185 (1.3%)	215 (1.2%)	115 (1.0%)

* Numbers for categories might not add to total number due to the missing data.

References

1. Office of the Inspector General (OIG), US Department of Health and Human Services. *Adverse Events in Hospitals: National Incidence Among Medicare Beneficiaries*. 2010.
2. National Patient Safety Foundation. *Free from Harm: Accelerating Patient Safety Improvement Fifteen Years after To Err Is Human*. Boston, MA: National Patient Safety Foundation;2015.
3. Whitman E. Patient safety efforts saved \$28 billion over five years. *Modern Healthcare*. 2016.
4. Agency for Healthcare Research and Quality. National Scorecard on Rates of Hospital-Acquired Conditions 2010 to 2015: Interim Data From National Efforts to Make Health Care Safer. 2016; https://www.ahrq.gov/professionals/quality-patient-safety/pfp/2015-interim.html?utm_source=AHQ&utm_medium=PSLS&utm_term=&utm_content=14&utm_campaign=AHQ_NSOHAC_2016. Accessed January 13, 2017.
5. Jha A, Pronovost P. Toward a Safer Health Care System: The Critical Need to Improve Measurement. *JAMA*. 2016;315(17):1831-1832.
6. Winters BD, Bharmal A, Wilson RF, et al. Validity of the Agency for Health Care Research and Quality Patient Safety Indicators and the Centers for Medicare and Medicaid Hospital-acquired Conditions: A Systematic Review and Meta-Analysis. *Med Care*. 2016;54(12):1105-1111.
7. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. *BMJ*. 2016;353:i2139.
8. Tan KM, Austin B, Shaughnassy M, et al. Falls in an acute hospital and their relationship to restraint use. *Ir J Med Sci*. 2005;174(3):28-31.
9. Schwendimann R, Buhler H, De Geest S, Milisen K. Falls and consequent injuries in hospitalized patients: effects of an interdisciplinary falls prevention program. *BMC Health Serv Res*. 2006;6:69.
10. *National Quality Measure Clearinghouse. Acute care prevention of falls: rate of inpatient falls per 1,000 patient days. NQMC measure summary. April 2012; Available at: <https://www.qualitymeasures.ahrq.gov/summaries/summary/36944>.*
11. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (2013). National Action Plan for Adverse Drug Event Prevention (draft). Washington, DC: Author. Retrieved November 4, 2013, from <http://www.hhs.gov/ash/initiatives/ade/ade-action-plan.pdf>.
12. Shreve J., Van Den Bos J., Gray T., Halford M., Rustagi K., Ziernkiewicz E. The economic measurement of medical errors. *Milliman*. 2010:1-274.
13. Wang Y, Eldridge N, Metersky ML, et al. Association Between Hospital Performance on Patient Safety and 30-Day Mortality and Unplanned Readmission for Medicare Fee-for-Service Patients With Acute Myocardial Infarction. *J Am Heart Assoc*. 2016;5(7).

14. Andel C, Davidow SL, Hollander M, Moreno DA. The economics of health care quality and medical errors. *J Health Care Finance*. 2012;39(1):39-50.
15. Henry J, Pylypchuk Y, Searcy T, V. P. *Adoption of electronic health record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2015. ONC Data Brief 35. May 2016*. 2016.
16. Gandhi TK, Berwick DM, Shojania KG. Patient Safety at the Crossroads. *JAMA*. 2016;315(17):1829-1830.
17. Toyabe S. Detecting inpatient falls by using natural language processing of electronic medical records. *BMC Health Serv Res*. 2012;12:448.
18. Bilimoria KY, Chung J, Ju MH, et al. Evaluation of surveillance bias and the validity of the venous thromboembolism quality measure. *JAMA*. 2013;310(14):1482-1489.
19. Kessler ER, Shah M, Gruschus SK, Raju A. Cost and quality implications of opioid-based postsurgical pain control using administrative claims data from a large health system: opioid-related adverse events and their impact on clinical and economic outcomes. *Pharmacotherapy*. 2013;33(4):383-391.
20. Overdyk FJ. Postoperative respiratory depression and opioids. *Initiatives in Safe Patient Care*. 2009.
21. Cashman JN, Dolin SJ. Respiratory and haemodynamic effects of acute postoperative pain management: evidence from published data. *Br J Anaesth*. 2004;93(2):212-223.
22. Herzig SJ, Rothberg MB, Cheung M, Ngo LH, Marcantonio ER. Opioid utilization and opioid-related adverse events in nonsurgical patients in US hospitals. *J Hosp Med*. 2014;9(2):73-81.
23. Lee LA, Caplan RA, Stephens LS, et al. Postoperative opioid-induced respiratory depression: a closed claims analysis. *Anesthesiology*. 2015;122(3):659-665.
24. Willens JS, Jungquist CR, Cohen A, Polomano R. ASPMN survey--nurses' practice patterns related to monitoring and preventing respiratory depression. *Pain Manag Nurs*. 2013;14(1):60-65.
25. Jungquist CR, Correll DJ, Fleisher LA, et al. Avoiding Adverse Events Secondary to Opioid-Induced Respiratory Depression: Implications for Nurse Executives and Patient Safety. *Journal of Nursing Administration*. 2016;46(2):87-94.
26. Development, Maintenance and Support of Hospital Clinical Quality Measures for ARRA HITECH: PCA Monitoring eCQM: Abt Associates; 2013.
27. Health Services Advisory Group, Hospital Quality Institute. *Compendium of Measures (CoM)*. 2017.
28. Adams L, Butas S, Spurlock D, Jr. Capnography (ETCO₂), respiratory depression, and nursing interventions in moderately sedated adults undergoing transesophageal echocardiography (TEE). *Journal of PeriAnesthesia Nursing*. 2015;30(1):14-22.

29. Cacho G, Perez-Calle JL, Barbado A, Lledo JL, Ojea R, Fernandez-Rodriguez CM. Capnography is superior to pulse oximetry for the detection of respiratory depression during colonoscopy. *Revista Espanola de Enfermedades Digestivas*. 2010;102(2):86-89.
30. Deitch K, Chudnofsky CR, Dominici P. The utility of supplemental oxygen during emergency department procedural sedation and analgesia with midazolam and fentanyl: a randomized, controlled trial. *Annals of Emergency Medicine*. 2007;49(1):1-8.
31. Deitch K, Chudnofsky CR, Dominici P. The utility of supplemental oxygen during emergency department procedural sedation with propofol: a randomized, controlled trial. *Annals of Emergency Medicine*. 2008;52(1):1-8.
32. Deitch K, Miner J, Chudnofsky CR, Dominici P, Latta D. Does end tidal CO2 monitoring during emergency department procedural sedation and analgesia with propofol decrease the incidence of hypoxic events? A randomized, controlled trial. *Annals of Emergency Medicine*. 2010;55(3):258-264.
33. Hutchison R, Rodriguez L. Capnography and respiratory depression. *American Journal of Nursing*. 2008;108(2):35-39.
34. Nwulu U, Nirantharakumar K, Odesanya R, McDowell SE, Coleman JJ. Improvement in the detection of adverse drug events by the use of electronic health and prescription records: an evaluation of two trigger tools. *Eur J Clin Pharmacol*. 2013;69(2):255-259.
35. Eckstrand JA, Habib AS, Williamson A, et al. Computerized surveillance of opioid-related adverse drug events in perioperative care: a cross-sectional study. *Patient Saf Surg*. 2009;3(1):18.
36. Ramachandran SK, Haider N, Saran KA, et al. Life-threatening critical respiratory events: a retrospective study of postoperative patients found unresponsive during analgesic therapy. *Journal of Clinical Anesthesia*. 2011;23(3):207-213.
37. Lin RJ, Reid MC, Chused AE, Evans AT. Quality Assessment of Acute Inpatient Pain Management in an Academic Health Center. *Am J Hosp Palliat Care*. 2016;33(1):16-19.
38. Agency for Healthcare Research and Quality. *Common Formats for Surveillance Event Description Medication*; https://www.psoppc.org/psoppc_web/publicpages/surveillancecommonformats. 2015.