

**Center for Clinical Standards and Quality (CCSQ)
Centers for Medicare & Medicaid Services (CMS)**



Summary of Maintenance of Hospital Harms Technical Expert Panel (TEP) Evaluation of Measures

Patient Safety Measure Development and Maintenance

11/20/2020
Version # 2



SUBMITTED TO

Centers for Medicare & Medicaid Services (CMS)
Center for Clinical Standards and Quality (CCSQ)

ATTENTION

Annese Abdullah-Mclaughlin
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244-1850

SUBMITTED BY

IMPAQ International, LLC
10420 Little Patuxent Parkway
Suite 300
Columbia, MD 21044
(443)256-5500
www.impaqint.com

PROJECT

Patient Safety Measure Development and Maintenance
Contract Number: 75FCMC18D0027

TASK & DELIVERABLE

Chapter 4: Quality Measure Development and Reevaluation
Deliverable 4.3 Summary of TEP Evaluation of Measures

AUTHORS

Maggie Lohnes, IMPAQ International
Stacie Schilling, IMPAQ International
Jacqueline Stocking, UC Davis
Chana West, IMPAQ International
Bo Feng, IMPAQ International
Hannah Klein, IMPAQ International
Leah Dillard, IMPAQ International

TABLE OF CONTENTS

Background	3
Overview of the Technical Expert Panel	4
TEP Purpose & Objectives	4
Technical Expert Panel Meeting #1.....	5
Summary of Presentation	5
Summary of TEP Discussion	5
Summary Of TEP Decisions/ Voting Results.....	7
Conclusions and Next Steps	8
Technical Expert Panel Meeting #2.....	9
Summary of Presentation	9
Summary of TEP Discussion	9
Summary Of TEP Decisions/ Voting Results.....	14
Conclusions and Next Steps	14
Technical Expert Panel Meeting #3.....	16
Summary of Presentation	16
Conclusion and Next Steps.....	26
Appendix A: TEP Composition List	27
Appendix B: Project Staff	28

Background

The Centers for Medicare & Medicaid Services (CMS) has contracted with IMPAQ International to develop and maintain patient safety measures of hospital harm for implementation in CMS programs. The contract name is Measure & Instrument Development and Support (MIDS) Patient Safety Measure Development and Maintenance. The contract number is 75FCMC18D0027. As part of its measure development process, IMPAQ convenes groups of stakeholders and experts who contribute direction and thoughtful input to the measure developer during measure development and maintenance.

IMPAQ is obtaining expert and stakeholder input to inform improvements and changes for the measures. This report summarizes the feedback and recommendations made by the Technical Expert Panel (TEP) during the meetings to discuss the Hospital Harm electronic clinical quality measures (eCQMs). The report will be updated to include feedback and recommendations from future meetings as they occur.

Overview of the Technical Expert Panel

In alignment with the CMS Measures Management System Blueprint, the project team convened a TEP to provide guidance on the development of the Hospital Harm eQMs. The role of the TEP is to provide guidance to the measure developer on key methodological and clinical decisions. The Maintenance of Hospital Harms TEP is comprised of 11 individuals representing a variety of viewpoints and backgrounds, including experience with measures of hospital harm and expertise in performance measurement, quality and patient safety, and coding and informatics. One TEP member represents patient perspectives. The full TEP membership is listed below in Appendix A.

TEP PURPOSE & OBJECTIVES

The TEP shall be comprised of individuals with knowledge of measures of hospital harms and expertise in performance measurement, quality and patient safety, and coding and informatics. The overarching goals of the TEP are to provide feedback to the IMPAQ team regarding development of Hospital Harm eQMs. The primary areas of focus are clinical and methodological issues as well as broader issues related to the measurement cycle.

The TEP will:

- Review background materials provided by IMPAQ prior to each TEP meeting
- Participate in TEP conference calls
- Provide input on key clinical and methodological decisions
- Provide feedback to IMPAQ on key policy or other non-technical issues
- Review the TEP summary report prior to public release
- Be available to discuss recommendations following submission of the measures to CMS

The materials presented in this document do not represent final measure specifications for the Hospital Harm eQMs.

Technical Expert Panel Meeting #1

September 23, 2019 3:00 PM ET

SUMMARY OF PRESENTATION

IMPAQ convened the TEP to familiarize the TEP members with the new project team and vice versa, provide a status update on the Hospital Harm eCQMs, review considerations for potential refinement of measure specifications, and determine the TEP recommendations for hospital harm measure specifications. During the meeting, the TEP introduced themselves and announced any personal disclosures; IMPAQ gave a status update on the six Hospital Harm eCQMs, reviewed the public comments for the Opioid-Related Adverse Events (ORAE) measure and potential changes, discussed the Medication-Related Bleeding measure considerations and facilitated TEP discussion; and TEP members voted on two questions related to the hospital denominator exclusion. Prior to the meeting, IMPAQ provided TEP members with the presentation slide deck and the American College of Surgeons' Guidelines for the Perioperative Management of Antithrombotic Medication article for review and preparation for discussion.

Attendance:

TEP Members: Christine Norton, David Hopkins, Joseph Kunisch, Timothy Lowe, Karen Zimmer, Steven Jarrett

Not Present: David Baker, Cynthia Barnard, Lisa Freeman, Kevin Kavanagh, Amita Rastogi

CMS: Joseph Clift, Katrina Hoadley

IMPAQ: Kendall Hall, Mike Sacca, Anna Michie, Stacie Schilling, Maggie Lohnes, Chana West, Michelle Lefebvre, Bo Feng, Hannah Klein, Molly Mantus

SUMMARY OF TEP DISCUSSION

1. **Severe Hypoglycemia and Pressure Injury measures:** were recommended for NQF endorsement and were proposed for future rulemaking by CMS.
2. **Severe Hyperglycemia measure:** will be submitted for NQF endorsement this fall and has already been submitted to the Measures Under Consideration (MUC) List.

The materials presented in this document do not represent final measure specifications for the Hospital Harm eCQMs.

3. **Opioid-Related Adverse Events measure:** was not recommended by NQF for endorsement and was not finalized in rulemaking by CMS. IMPAQ received a substantial number of comments around the potential for unintended consequences of disincentivizing clinically appropriate treatment, naloxone often being used to treat conditions other than opioid-related overdoses, the broad denominator, the lack of risk adjustment, naloxone being considered a harm, and the potential for the measure to come in as a topped-out measure.
 - a. To address the public comments and NQF concerns, IMPAQ and CMS are considering limiting the denominator to those patients who were administered an opioid in the hospital setting, requiring that an opioid is administered prior to naloxone (and including a look-back time parameter), and using a new value set to represent opioid medications.

4. **Acute Kidney Injury measure:** will be submitted to the MUC list and for NQF endorsement in 2020. IMPAQ is continuing to conduct measure testing and plans to complete testing in Spring 2020.
 - a. The former contractors created two specifications for the measure, with the difference being how the creatinine levels are determined. The IMPAQ team will test these differences during the next measure testing cycle.

5. **Medication-Related Bleeding measure:** will be submitted to the MUC list and for NQF endorsement in 2020. Testing was completed under the prior contract, but the current development team is considering modifications that would require additional testing.
 - a. The IMPAQ eCQM team is making a few technical changes including replacing the current value set for blood transfusion to an expanded one that also includes ICD-10-PCS codes and expanding the current value set for general major bleeding events to also include SNOMED CT codes.
 - b. IMPAQ is considering adding a look-back timing parameter rule before excluding patients with a major surgery from the denominator. The eCQM team conducted a literature review to help inform the discussion around bleeding times following major surgery and found several key takeaways:
 - i. The management of perioperative bleeding is complicated and involves balancing the clinical consequences of potential deep vein thrombosis or other clotting along with the need to make sure patients don't experience any postoperative bleeding.

- ii. There is more attention regarding the diagnosis for which anticoagulation therapy has been prescribed than to the specific surgical procedure the patient underwent.
 - iii. Bleeding is highly variable across major surgical types.
 - iv. One perioperative bleeding classification framework includes three types: primary bleeding (occurs during the surgery), reactive bleeding (occurs within 24 hours of the surgery), and secondary bleeding (occurs 7 to 10 days postoperatively). The eCQM team was considering these categories as a way to define the timing parameter and limit the timeframe for postoperative surgery.
- c. Input from TEP members regarding potential timing parameter:
- i. Dr. Joseph Kunisch believes it will be extremely difficult to capture a timing parameter. He is uncertain if there is a single time period that would work across surgery types. He would not want to specify a time period unless there is research supporting it.
 - ii. Dr. David Hopkins suggests that IMPAQ separate the measure into two measures – one for medication-related bleeding and one for surgical-related bleeding.

SUMMARY OF TEP DECISIONS/ VOTING RESULTS

The TEP was asked to vote on two questions, with the option to either vote during the meeting or send their votes via email after the meeting. Post-meeting voting concluded on September 27, 2019. The results of the voting is as follows:

Exhibit 3: TEP Voting Results

Recommendation	TEP Voting Results
Do you agree with our recommendation to modify the denominator exclusion so that any bleeding event occurring prior to the major surgery should count as a harm?	89% Yes (8 votes) 11% Abstain (1 vote)
Should there be a timing parameter following major surgery to count as an exclusion? (e.g., if the bleeding event occurs more than 2 weeks after major surgery, then it is a harm event)	44% Yes (4 votes) 56% No (5 votes)

CONCLUSIONS AND NEXT STEPS

Based on the results of the TEP voting, IMPAQ plans to move forward with the recommendation to modify the denominator exclusion so that any bleeding event occurring prior to the major surgery counts as a harm. IMPAQ will not move forward with including a timing parameter following major surgery to count as an exclusion.

Following the TEP meeting, the MIDS Patient Safety team produced the meeting summary notes. IMPAQ will continue testing the Acute Kidney Injury and Medication-Related Bleeding measure, work through the NQF endorsement process for the Severe Hypoglycemia and Pressure Injury measures, go through the NQF review process for the Severe Hyperglycemia process, and keep the TEP posted on any further discussion needed around the ORAE measure. IMPAQ will schedule a subsequent TEP meeting to discuss testing results and further refinements needed for the Hospital Harm eCQMs.

Technical Expert Panel Meeting #2

November 22, 2019 2:00 PM ET

SUMMARY OF PRESENTATION

IMPAQ convened the TEP for the second meeting to provide a status update on the Opioid-Related Adverse Events (ORAE) eCQM, review considerations for proposed changes to ORAE measure specifications, and determine the TEP recommendations regarding the proposed changes to ORAE measure specifications. During the meeting, the TEP introduced themselves and announced any personal disclosures, IMPAQ discussed the ORAE considerations and facilitated TEP discussion, and TEP members voted on five recommendations. Prior to the meeting, IMPAQ provided TEP members with the presentation slide deck for review and preparation for discussion.

Attendance:

TEP Members: David Baker, Cynthia Barnard, Christine Norton, David Hopkins, Kevin Kavanagh, Joseph Kunisch, Timothy Lowe, Amita Rastogi, Steven Jarrett

Not Present: Lisa Freeman, Karen Zimmer

IMPAQ: Kendall Hall, Mike Sacca, Anna Michie, Stacie Schilling, Maggie Lohnes, Chana West, Michelle Lefebvre, Bo Feng, Hannah Klein, Molly Mantus

SUMMARY OF TEP DISCUSSION

1. Status Update on Opioids Measure and Summary of CMS IPPS Rule Public

Comments: IMPAQ shared an update on the ORAE measure, noting that it was not recommended for NQF endorsement in Spring 2019 cycle as NQF had several concerns about the measure as specified, which arose again in the IPPS public comments. In consultation with CMS, IMPAQ presented a few options to address these concerns for the TEP to discuss.

- a. IMPAQ presented a brief summary of the three recommended changes to address the issues highlighted during the NQF review for the TEP's consideration:
 - i. Limit the denominator to those patients who were administered an opioid in the hospital setting.

- ii. Require that an opioid have been administered prior to naloxone and include a single lookback time period from naloxone administration (e.g., within 24 hours) to opioid administration to count as a harm.
 - iii. Use a new value set to represent opioid medications.
- b. The TEP raised a few clarifying questions and shared some comments to start off the discussion:
- i. Dr. Kavanagh asked to clarify if opioids administered with a naloxone reversal in the perioperative setting should not count for the measure, as this may be a planned procedure. Ms. Lohnes clarified that the measure does currently exclude perioperative opioid administration.
 - ii. Dr. Jarrett clarified that the denominator excludes patients in outpatient settings that are not part of the inpatient process. Ms. Lohnes confirmed the measure is only for the inpatient settings.
 - iii. Dr. Barnard shared her concern about using naloxone as a marker for harm, e.g., over-sedation, which could occur in the outpatient setting, such as with an endoscopy. Dr. Barnard disclosed that her organization, Northwestern is one of the potential test sites.
 - iv. Dr. Barnard also raised the question for discussion whether procedures that require naloxone for reversal should be excluded. Dr. Barnard asked to confirm the premise of the measure being avoiding over-sedation to prevent harms. Ms. West noted the current specifications are working through the testing process, and as part of that process we are looking at the frequency of these types of events to determine how best to inform the measure specification going forward; The results of the testing will be shared with the TEP once completed.
 - 1. Dr. Hall shared that the NQF panel raised the same philosophical question. Some believe administering naloxone is a good thing because it is reversing the harm.
 - 2. Dr. Barnard said she believes deep sedation is the harm and naloxone would not be a reliable marker for this harm. With instances of naloxone administration, the reason is not always clear, unlike other areas such as hyperglycemia. Additionally, instances of unplanned naloxone may not necessarily indicate there was an avoidable harm. Dr. Hall noted the challenge of working through

refining the inherited measure as best we can, based on what we heard from NQF and via public comment in the rule.

- v. Dr. Baker shared interest in understanding the false positive rate for true harms. With multiple procedures, it can be difficult to discern whether the harm occurred or if it was preventable. Dr. Rastogi suggested adding an exclusion for major inpatient procedures to the denominator. In response, Ms. West noted the testing process would allow for further refinements, depending on what it reveals and noted the testing process will look beyond the current specifications, including the frequency of procedural instances, for example.
- vi. The TEP requested that IMPAQ tests the effects of the proposed modifications and denominator exclusions and shares the results with the TEP.
- vii. Dr. Jarrett noted that his hospital has used naloxone as a trigger for improvement rather than a marker of harm and that this may be a better approach. He noted the use of a certain “scale” at his facility to help assess the use and intention of naloxone

2. Specification – Proposed Changes: To segue to the recommended changes, Ms. Lohnes summarized the public comment/NQF concerns overall and walked through the recommendations implemented in the updated version of the specification.

- a. **Recommendation #1:** Limit the denominator to those patients who were administered an opioid.
 - i. The TEP voted unanimous support to limit the denominator.
 - ii. Dr. Rastogi agreed but felt the exclusions should specify patients administered an opioid without surgery.
- b. **Recommendation #2:** Require that an opioid is administered prior to naloxone and include a single time period around opioid administration and naloxone administration (e.g., within 24 hours) to count as a harm.
 - i. Dr. Hopkins asked why the measure does not follow the AHRQ recommendations and include experiences such as respiratory depression and unresponsiveness following an opioid administration since these indicate a harm by opioid.

1. Consensus is that capturing detailed information regarding overdose symptoms in the EHR may require complicated data extraction such as natural language processing from narrative clinical notes.
- ii. A chart of medication half-life for a set of common opioid medications illustrated the variation in half-life; as opioid half-life varies in the context of co-morbidities and polypharmacy, both the minimum and maximum levels for each medication were presented. Review of opioid half-lives is intended to support identification of an appropriate 'look back' period. The half-life shows the point where the drug reaches its maximum effect; heroin is the shortest acting, and methadone is the longest acting. In consideration of the half-life for determining a lookback window, the TEP had several comments:
 1. Dr. Barnard would prefer to avoid including procedural medication in the same measure as IV medication and for the look back period to reflect the drug administration, so possibly less than 24 hours.
 2. Dr. Kavanagh highlighted that buprenorphine and methadone are used for treating opioid addiction while the rest of the medications listed are used to treat pain. He recommended using a lookback period based on the diagnosis; if addiction use 24 hours, if pain use 12-hour window. The option of creating a value set for each diagnosis, including only relevant drugs in each, is one approach to this, thought it could be difficult considering patients taking multiple opioids and other nuances, such as patient adherence to medical direction.
 3. Dr. Jarrett raised the fact that some opioids are administered in multiple doses and recommended looking at continued administration to identify patients who are getting too many doses and require a reversal agent. With this approach even long acting opioids would have more than one dose.
 4. Dr. Hopkins commented that the 24-hour look back period would be more appropriate for this measure and shared concerns that the 12-hour window would not be sufficient.
- iii. Two examples demonstrated how the 24-hour lookback period would operate. The first example demonstrated how an opioid given within 24 hours of naloxone is considered a harm, even with a half-life period of only a few hours. The second example demonstrated how an opioid with a half-life range of more than 24 hours could register as not a harm.

1. Dr. Baker commented that if a patient goes into respiratory arrest 24 hours after receiving codeine, there are large problems going on. Example 1 is likely showing a false positive and a false negative in example 2.
 - iv. The TEP voted in support of the recommendation to require documentation of the provision of an opioid by hospital staff prior to any naloxone administration during a hospital stay.
 - v. The TEP voted did not support the recommendation to have a single time period for opioid administration and subsequent naloxone administration to qualify as a harm.
 - vi. The TEP votes did not support the recommendation for a 24- hour time period between hospital opioid administration and subsequent naloxone administration to qualify as a harm.
 - vii. Dr. Rastogi commented that from a measure development perspective it could be very difficult to reconcile the various time periods associated with the different opioids for patients might be on more than one opioid. If the period is shortened, the measure may miss some of the longer acting opioids. However usually naloxone would not be administered to a patient that received a short acting opioid a few days prior.
- c. **Recommendation #3:** Replace the current value set with a new value set that includes an updated list based on current medications in the market.
- i. Preliminary feasibility results show results for four hospitals that indicate improved capture of opioid administrations.
 - ii. The TEP voted in support of adopting the new value set.

3. Final Comments

- a. Dr. Jarrett requested IMPAQ test the measure in smaller facilities such as critical access hospitals, because they see very few opioid events and it would be interesting to see those results.
- b. Dr. Barnard reiterated a few recommendations for testing including testing the timing between opioid and naloxone administration, the reason the opioid and the naloxone was given, and finally look at procedural doses vs opioids given for treatment at intervals through IV or PCA. It would also would be useful to know the clinical

standpoint on how to treat chronic opioid abusers and how this measure captures chronic opioid reversals.

- c. Dr. Kavanagh shared that in his experience, chronic opioid users have a higher pain threshold and then require a higher opioid dosage, but the result is a lower therapeutic margin. With these patients, the risk is higher for respiratory arrest and the need to administer naloxone when trying to relieve pain.
- d. Dr. Rastogi added that buprenorphine is now recommended for treatment therapy for opioid overdose and it could be interesting to look into the use of such combinations of drugs.

SUMMARY OF TEP DECISIONS/ VOTING RESULTS

The TEP was asked to vote on five questions, with the option to either vote during the meeting or send their votes via email after the meeting. Post-meeting voting concluded on November 25, 2019. The results of the voting are as follows:

Exhibit 3: TEP Voting Results

Recommendation	TEP Voting Results
For the denominator, do you agree with the recommendation to only include patients who have explicit documentation of an opioid administration during the inpatient hospitalization?	100% Yes (9 votes)
Do you agree with the recommendation to require documentation of the provision of an opioid by hospital staff prior to any naloxone administration during stay?	100% Yes (8 votes)
Do you agree with the recommendation to include a single time period around opioid administration and subsequent naloxone administration qualify as a harm?	33% Yes (3 votes) 63 % No (6 votes)
Do you agree with the recommendation to use a 24-hour time period between hospital opioid administration and subsequent naloxone administration in order to qualify as a harm?	22% Yes (2 votes) 78% No (7 votes)
Do you agree with the recommendation to replace the current value set "Opioids for pain control" with a new "Opioids, All" medication value set?	100% Yes (9 votes)

CONCLUSIONS AND NEXT STEPS

Following the TEP meeting, the MIDS Patient Safety team produced the meeting summary notes. IMPAQ plans to complete the final analysis of the recommended value set query results from four test sites; finalize the specifications, incorporating TEP recommendations; complete

The materials presented in this document do not represent final measure specifications for the Hospital Harm eQMs.

full measure testing with a minimum of two test sites using different electronic health record vendor systems; analyze reliability and validity testing results; and provide the TEP with the final test results in a subsequent meeting.

Technical Expert Panel Meeting #3

October 29, 2020 12:00 PM ET

SUMMARY OF PRESENTATION

IMPAQ convened the TEP for the third meeting to review measure testing results for the Opioid-Related Adverse Events (ORAE) eCQM, and to provide a status update on the Acute Kidney Injury (AKI) and Pressure Injury (PI) eCQMs. During the meeting, the TEP introduced themselves and announced any personal disclosures, IMPAQ presented the ORAE testing results, presented the status updates for AKI and PI, and facilitated TEP discussion. Prior to the meeting, IMPAQ provided TEP members with the presentation slide deck for review and preparation for discussion.

Attendance:

TEP Members: David Baker, Cynthia Barnard, Christine Norton, Lisa Freeman, David Hopkins, Steven Jarrett, Kevin Kavanagh, Joseph Kunisch, Timothy Lowe, Amita Rastogi,

Not Present: Kevin Kavanagh, Karen Zimmer

IMPAQ: Kendall Hall, Jensen Chiu, Anna Michie, Stacie Schilling, Maggie Lohnes, Chana West, Michelle Lefebvre, Bo Feng, Hannah Klein, Leah Dillard

UC Davis: Jacqueline Stocking, Patrick Romano, Meghan Weyrich

SUMMARY OF TEP DISCUSSION

1. **Opioid-Related Adverse Events eCQM Testing Results:** Chana West and Bo Feng reviewed the alpha and beta testing results that showed the measure meets expectations for feasibility, reliability, and validity.
 - a. Ms. West reviewed the measure history and primary changes to the measure since it was not recommended for NQF endorsement. The primary concern was the use of naloxone as an indicator for quality and potential unintended consequences. In response, the IMPAQ team: updated the measure value sets to ensure the most current codes were used; limited the denominator to include only those encounters where patients received an opioid during the hospitalization; and added a timeframe between the opioid administration and subsequent naloxone administration to better ensure the opioid was the cause for the naloxone administration.

The materials presented in this document do not represent final measure specifications for the Hospital Harm eCQMs.

- b. Ms. West reviewed the changes to the denominator and numerator and noted that the team added a 12 hour look back to tie the opioid administration to the event. In testing this change, the team found very few patients were removed from the measure when lowering the timeframe from 24 hours to 12 hours.
- c. Ms. West reviewed the testing process used by the team, including the alpha and beta testing processes. The alpha testing results were reviewed with the TEP during the second TEP meeting. The alpha testing included quantitative data queries to determine the impact of changing the opioid value set within the measure. Beta testing involved feasibility assessment, implementation testing using synthetic test data prior to running the measure programming against actual patient data, and parallel forms validity testing to determine the level of agreement between actual patient data reports from each facility and the results from clinical abstraction.
- d. Ms. West reviewed the results from the feasibility assessments, showing that the test sites had varying qualities to ensure the test results would show variability. The results highlighted an issue with the ability to retrieve data from anesthesia records in facilities where paper documentation is used. These elements that are likely to be impacted by this documentation practice are: opioid administration, naloxone administration and facility location operating room suite. Feasibility testing determined that this was not a vendor specific issue, but rather a hospital implementation that could be remedied with technical and clinical workflow modifications.
- e. Dr. Feng then provided an overview of the beta testing results. The measure performance rate in 2019 ranged from 0.11% to 0.45% across the six implementation test sites. Larger hospitals (based on bed size) did not necessarily have higher measure performance rates, though they had more qualified admissions (measure denominator counts). The beta testing focused on the data element level because the small number of implementation sites that participated in testing can yield unreliable measure score level reliability test results.
- f. The team used two metrics to assess the data element reliability: 1) the rate of missing or erroneous values for all the critical data elements required for measure implementation, and 2) Cohen's Kappa coefficients to determine concordance between the test site's quality reporting engine and the clinical abstractor. Test results showed a rate of 0 for missing or erroneous data for all six hospitals, suggesting that the critical data elements are reliably and consistently captured in the EHR. The Kappa coefficients are either 0.98 or 1 for the measure's critical data elements across six hospitals, suggesting near perfect inter-rater agreement.

- g. Although beta testing focused on the data element level, the team assessed measure score level reliability using the signal-to-noise ratios (SNRs) based on the Adams' beta-binomial method and the intraclass correlation coefficient (ICC) via the split-half sample approach. Testing results showed that the SNRs approximated zero for all six hospitals. The estimated ICCs via the split-half sample approach and across the 300 simulation runs showed an average score of 0.78. However, the estimated ICCs exhibit a wide confidence interval, suggesting uncertainty associated with the small sample size. The team further conducted a hypothesis testing on the relationship between the size of SNR and the number of hospitals included for estimation. For this test, the team found a general tendency of rising SNRs when the number of hospitals included for estimation increases. The hypothesis testing result reinforced the team's rationale of focusing on the data element level reliability test and suggested the measure's ability to differentiate real performance across accountable entities from noises.
- h. The data element level validity testing results showed a near perfect match between the test site's quality reporting engine and the clinical abstractor for the randomly selected patient cases captured in the measure. In one test site, however, the clinical abstractor identified a case that amounted to a false negative. Specifically, for the encounter of interest, the test site's quality reporting engine deemed it to be denominator only, but the clinical abstraction found it to be a numerator event. In light of the finding, the team assessed the prevalence of false negatives the measure may suffer and found that the rate is 0.006%. The team believed that the concern of false negative is not grave.
- i. The measure score level validity testing, based on the positive predictive value, sensitivity, negative predictive value, and specificity, showed the values of 98% or higher for the four statistics. To assess the magnitude of false positives, the team categorized nurse notes on patient responses to the naloxone administration and found that nearly four-fifths of the reviewed numerator cases showed that patients demonstrated clear signs of reaction, such as be more awake and responsive. This suggested that naloxone was used to address the excessive use of opioids. The remaining proportion of the reviewed numerator cases had nurse notes indicating that patients showed modest signs of reaction after the naloxone administration. The team caution that they do not necessarily suggest that naloxone was used for alternative purposes other than opioid reversal, as in some cases patients became responsive only after the second naloxone was given. Therefore, it is still possible that naloxone was used to reverse the opioid, but its dosage was inadequate the first time. Overall, the validity testing results suggested the measure as currently specified can correctly predict the true positives.

- 2. TEP Discussion of ORAE Testing Results:** The TEP then asked questions for the project team about the Opioid eCQM testing results and several TEP members voiced concern about the low measure performance rate and the importance of measuring this event.
- a. Amita Rastogi asked how many of the events received the opioid inside the operating room (OR) versus outside the OR.
 - i. Chana West clarified that the measure denominator looks at whether the patient received the opioid during the hospitalization. An event will count for the numerator if they received an opioid in the OR and then received the naloxone outside the OR. Nearly all of the cases received the opioid and the naloxone outside of the OR.
 - b. David Baker- clarified that patients who receive the opioid and naloxone inside of the OR are then excluded from the measure.
 - i. Ms. West confirmed that is correct.
 - c. Dr. Tim Lowe asked anesthesiology has their own system in this example hospital, so their information does not go into the general EHR?
 - i. Ms. West clarified that the information from the OR is in the EHR, but it is scanned in and not entered directly, so it cannot be pulled from a structured field.
 - ii. Dr. Lowe followed up to confirm that this data is therefore entered as a PDF and the data would have to be pulled in a different way.
 - iii. Ms. West noted that sites would have to modify workflows to enter this data since these EHR vendors offer an anesthesia module, but the facilities are not using them.
 - iv. Dr. Lowe followed up and commented about the frequency of this issue and asked if it's an issue at one or two hospitals?
 - v. Ms. West responded that of the 23 facilities the team looked at during feasibility testing, there were 5 facilities with this issue, 1 Allscripts and 4 Cerner, but the Cerner sites were all from the same hospital system. However, there were some hospitals that do use the anesthesia module within both Allscripts and Cerner systems, so it is not an issue with availability within the EHR, but rather the hospital's implementation of that system.

- d. Christine Norton asked if the team is aware of the reason that sites have not yet purchased the anesthesia module; is it because it is cumbersome or is it because it has a high cost?
 - i. Ms. West noted that the team did not explore those questions specifically, instead they focused on the current clinical and technical workflow processes and the timing it would take to implement these changes. Clearly it would be a heavy lift to implement these changes to both the system and the workflow, but the team does not have specific numbers.
 - ii. Dr. Steve Jarrett added that Atrium Health does not currently use an anesthesia module in the OR due to both complexity and cost concerns. They are also moving from Cerner to Epic, so they will probably never get the anesthesia module from Cerner. He added that the anesthesia modules are improving and will probably become more widely implemented, but it does involve a complex workflow so it will take a while.
- e. Dr. Cindy Barnard asked about the current standard for reliable data across multiple facilities, what is the general rule of thumb for adopting an eCQM when there are variations among facilities and whether they collect data?
 - i. Ms. West replied that she is not sure if there is a reliable standard, and even though eCQMs have been around for several years, there is still an emphasis on making improvements along the way. The challenge comes with any measure that is being developed to either work with the current documentation practices or push the envelope for improvements. As we move to more outcome-based measures, it is going to require more changes along the way, but there is not necessarily a standard. The goal is to use this measure because of its value and that may require facilities to make changes to do so.
 - ii. Dr. Barnard remarked that it depends on how the measure is used. If it's used for internal improvement, that is one thing. But if we want to compare facilities and one facility scans their information while another enters their information in discrete, there are going to be challenges.
 - iii. Ms. West clarified that in the case where a facility is entering their data by scanning, they would be excluded from the measure.
 - iv. Dr. Barnard raised several concerns with this scenario, mainly that there would not be an ability to compare facilities, and you wouldn't know which facilities lack the data since a lack of data would be the same as no event.

- v. Ms. West clarified that to date not all eCQMs are required for reporting.
 - vi. Dr. Lowe raised similar concerns about publicly reporting this measure but would be comfortable with it as an optional measure. He shared a concern that not having sufficient data for some hospitals could be a form of cheating.
- f. Dr. Steve Jarrett asked if this is widespread and there is a definite distinguisher, is there a way to have 2 separate eCQMs, so that there is a way to participate and you would be able to compare like groups. One of the issues with measuring adverse events in general is that there is not a national standard to compare to.
- i. Ms. West replied that the team would probably not develop separate specifications to support this, but for example we have three data elements that are pertinent to the OR. We are looking at the facility location to determine the time that the patient was in the OR, which could be captured from various places to capture that time period. The issue with that data is that you would only be able to capture the medication and doses that are captured in that part of the EMR. So, if you are looking at data within your system for internal performance improvement, then you would be able to look at it with the caveat that you could have some false negatives or positives because you are not capturing medications given within the operating room.
 - ii. Dr. Lowe remarked that it seems that there will be some level of noise in the data or a couple of missed cases. His concern is that around the differences between hospitals- if they are slight then this kind of thing could cause a problem, but we don't know that yet.
 - iii. Dr. Bo Feng acknowledged Dr. Lowe's point.
- g. Dr. Rastogi acknowledged that when the team made the updated the value set used within the measure, the number of codes went up. What other changes may have impacted the measure? When NQF rejected the team made some changes but did not fully understand the changes that were made.
- i. Ms. West clarified that originally the measure included any adult patient that was admitted to the hospital in the denominator. So, the team narrowed the denominator by changing it to be any adult patient that was admitted and administered an opioid during that stay, to only include patients that were at risk for the harm. Another change was to make it more explicit that an opioid was administered by requiring that an opioid be administered within 12 hours of the naloxone administration to count as a harm. Before there was no timeframe for the numerator.

- ii. Dr. Rastogi noted that the second piece gave her some concern. Would it have an unintended consequence if a provider set it out but did not give the naloxone- for example waiting to give naloxone at 13 hours? Perhaps anesthesiologists know more about this and can say whether naloxone must be given within 12 hours to have a reversal effect. is there a reason for the 12 hours?
 - iii. Ms. West agreed that there previously was some concern that some providers would hold back on using naloxone to improve their scores. To alleviate these concerns, the team made an effort to explicitly link the naloxone to an opioid. Additionally, the team looked at data and the bulk of events occurred well before that 12-hour mark.
 - iv. Dr. Jarrett added that the measure of 12 hours has been used manually at lots of hospitals and other places where the naloxone would be used for opioid over-sedation, so many clinicians would be used to this metric. He agreed that 12 hours would give a good window for determining if the naloxone is being used in that way.
- h. Dr. David Hopkins thanked the project team for the thorough testing protocol that was followed. From the results it looks like the measure is well specified and valid, but he is concerned that the rate was less than .1% and what that means for the importance of the measure. Especially with all the patient safety activities going on and a scarcity of resources, he worries that this is not an issue worth measuring.
- i. Dr. Feng acknowledged that the concern about the rate being too low is not enough to say the event is not worth measuring and monitoring. For example, some Patient Safety Indicator (PSI) measures have low rates but are important to measure. Additionally, Dr. Feng pointed to the variability in measure performance rates, which suggest that there is room for improvement. There are two ways to look at importance, the rate itself and the rate in relative terms.
 - ii. Dr. Baker agreed that it is easy to say that these low rates would mean that it is not worth measuring. He added that there are other measures that have even lower rates but there is still an emphasis on those areas because the importance is still high. For events where the event is both preventable and fatal, there is importance. It is important to consider the severity of the adverse event.
 - iii. Dr. Jarrett added that many hospitals are tracking this information manually so to have this measure available electronically would improve the ability to

track this information. So, tracking this with is an eCQM would be more effective and efficiently to see where a hospital lands in the variability.

- iv. Dr. Rastogi countered that the potential adverse events, such as brain damage or death, would not happen if the naloxone is administered.
- v. Dr. Baker acknowledged that the reversal would prevent such harm, but if the individual is not rescued with the naloxone then they face a more severe event (i.e. fatality). This could be a good measure for quality improvement because it could show a systems issue. For example, if you have 3 times the national average for rescues with naloxone, the odds are that you may have a fatal event.
- i. Dr. Lowe commented that he is happy to see that the numbers are low, because that indicates that the systems are working. Also, a lot of the low hanging fruit are gone so now we are working on the smaller measures. It is probably a good thing to differentiate which resources are needed. Often asked where to put the hospital efforts and usually they have to do some data monitoring to find the answer and can indicate other areas of concern that were previously unknown. Would be also useful to see where resources are needed to make improvements.
 - i. Dr. Joe Kunisch added that when he looked at this measure the naloxone was administered not to completely reverse the opioid, but to reverse the adverse effects such as vomiting, etc. versus respiratory depression. Did you do anything about that patient population? Did you do any testing to look at these patients?
 - ii. Ms. West explained that the team reviewed nursing notes and physician orders for all the sampled numerator cases to determine both the indication for administration and patient response and did not find any instances in the measure where the patient received the naloxone for the alternative reasons presented.
 - iii. Dr. Kunisch's organization, Memorial Hermann, has a policy to administer the naloxone in these cases so they have some events in this category, though it is possible that they may be outliers.
 - iv. Dr. Kendall Hall asked Dr. Kunisch if there is a concern when the naloxone is given in the post-op that there would be a pain killer reversal?
 - 1. Dr. Kunisch clarified that no, the dosage is so low that it is really just to treat a side effect of the opioids.

- v. Ms. West added that the value set only includes dosages that are high enough for a reversal, so any dosage that is only to treat a side effect (provided they are lower) would not be captured by this measure.
- j. Dr. Barnard raised concern about small numbers problem and the statement that the measure has the ability to distinguish performance across hospitals. She doesn't see that yet, as she has seen measures with a more robust picture.
 - i. Dr. Feng acknowledged her concern and added that the eCQM testing with six hospitals is not enough to statistically distinguish hospital-level variance from within-hospital variance. It would require more data (i.e., more hospitals) to separate the signal from the noise. To test a hypothesis, the team used a simulation process to calculate the signal to noise ratio by varying the number of hospitals included in estimation. Those simulations showed a general tendency for the SNR to rise as the sample size (number of hospitals included for estimation) increases. If we have a national test dataset to run the measure on, then we could calculate a new signal to noise ratio and see whether the tendency seen in the simulations holds in real life.
 - ii. Dr. Baker followed up on Dr. Barnard's point and asked if you have 5,000 hospitals in the US and the range across those hospitals varies. For the hospitals that are on the high range, are the results valid?
 - iii. Dr. Feng noted that it would be possible to run a simulation to test this question using fictitious data. From the tests we have done we see the general trend upward.
 - iv. Dr. Barnard acknowledged the trend, but still felt that if this is a very low volume event, the meaningfulness and usefulness of a rate is not there. But if you use the events for quality improvement, that would be useful as a trigger tool. This could add noise to the reporting because you could have a really poor rate with one event.
- k. Dr. Jarrett commented in the chat that based on the work we have done, he believed the naloxone measure is low. Many of our efforts have been in the case review setting. The conversation and concern expressed make sense to me as well.
- l. Dr. Hopkins commented in the chat that he remained concerned about the naloxone measure, the very low positive rate (which is good news) and questioned whether it is able to differentiate performance between hospitals if that is the purpose.

3. Measure Status Update and Overview- Acute Kidney Injury eCQM: Dr. Kendall Hall reviewed the measure changes since IMPAQ received the measure. In May 2020, IMPAQ

The materials presented in this document do not represent final measure specifications for the Hospital Harm eQMs.

reached out to the American Society of Nephrology and the Renal Physicians Association to review the measure.

- a. The main feedback supported incorporation of Kidney Disease Improving Global Outcomes (KDIGO) guidelines into the measure.
- b. In response to the feedback, the team revised the measure to refine the initial population, the denominator, and the numerator.
 - i. To narrow the initial population, the team added a length of stay requirement of greater than 48 hours.
 - ii. The team revised the denominator to require at least one serum creatinine result after 48 hours of arrival to the hospital and refined the denominator exclusions to ensure the measure does not include patients that have an AKI upon arrival.
 - iii. The team increased the numerator serum creatinine value requirement from a 1.5 times increase to a 2.0 times increase from the lowest serum creatinine value during the encounter.
- c. Dr. Hall presented the current measure specifications, noting that they may change as the measure will undergo testing and the results may require tweaks to the measure.
- d. Dr. Jarrett asked if the measure will be risk adjusted.
 - i. Dr. Hall confirmed that the measure at this time will not be risk adjusted.
 - ii. Dr. Jarrett commented that there are well known kidney injury risks, specifically for people with diabetes and recommended they be considered for risk adjustment.
 - iii. Dr. Hall confirmed that the specialty groups mentioned this as well in their feedback.
- e. Dr. Lowe asked if the team is looking at certain specific injuries that could occur within the first 48 hours in the hospital that might trigger a harm.
 - i. Dr. Hall clarified that we would expect another Scr value at 49 hours or later to capture if someone was given too much contrast, etc. that would happen within the first 48 hours. Dr. Hall added that the team will review this in detail during testing and make refinements as needed to ensure the measure is capturing events appropriately.

- f. Dr. Barnard commented that it is an interesting question whether a lab value constitutes harm. If there is a need to initiate dialysis or another treatment that may involve risk, then that would seem to be clear evidence of harm. But is a change in lab values, in itself, harm? Also, to the discussion about contrast, is it the assumption that all events measured in this metric are preventable?

4. Measure Status Update and Overview- Pressure Injury eCQM: Dr. Jacqueline Stocking reviewed the Pressure Injury measure updates since the last TEP meeting and provided an overview of the current measure specifications. The team met with additional stakeholders to review potential concerns and recommendations for consideration.

- a. In response to concerns about the lack of availability of certified wound care clinicians and the low inter-rater reliability of staging, the team removed stage 2 injuries from the measure.
- b. Lengthened deep tissue injury (DTI) timeframe to 72 hours as it can take this long to manifest and become visible to the clinician.
- c. The team is currently assessing the feasibility of using precoordinated codes to differentiate between a new, worsening, or same present on arrival pressure injury.
- d. The current measure specifications assess the proportion of adult inpatient encounters who suffer the harm of developing a new stage 3, stage 4, deep tissue, or unstageable pressure injury.
- e. Dr. Kunisch commented that the information on present on admission indicator, which coders enter to the documentation, will be available in the quality data model in the next year in the updated version. This should help with documentation once it is available.

CONCLUSION AND NEXT STEPS

Following the conclusion of the TEP meeting, the Project Team summarized the discussion in the updated TEP Summary Report for sending to the TEP and CMS. The team will continue with the testing of the Acute Kidney Injury, Medication- Related Bleeding, and Pressure Injury eCQMs as the team refines and finalizes the measures. The team will also submit the Opioid Related Adverse Events eCQM to the National Quality Forum (NQF) for consideration for endorsement in the NQF Spring 2021 review cycle.

Appendix A: TEP Composition List

Name, Credentials, and Professional Role	Organizational Affiliation, City, State	Conflict of Interest Disclosure
Christine Norton, MA; Patient Representative	Minnesota	None
David Baker, MD, MPH; Executive Vice President for Health Care Quality Evaluation	The Joint Commission, Oakbrook Terrace, IL	None
Cynthia Barnard, PhD, MBA, MSJS; Vice President, Quality	Northwestern Memorial Healthcare, Chicago, IL	None
Lisa Freeman, Patient Advocate; Executive Director	Connecticut Center for Patient Safety, Fairfield, CT	None
David Hopkins, PhD, MS; Performance Measurement	Stanford University, Stanford, CA	None
Kevin Kavanagh, MD, MS; Board Chairman	Health Watch USA, Somerset, KY	None
Joseph Kunisch, PhD, RN-BC, CPHQ; Enterprise Director of Clinical Quality Informatics	Memorial Hermann Hospital System, Houston, TX	None
Timothy Lowe, PhD; Director, Research Services	Memorial Hermann Hospital System, Houston, TX	None
Amita Rastogi, MD, MHA, CHE, MS; Vice President Commercial Lines Bundled Payments, Medical Director	Signify Health, Norwalk, CT	None
Karen Zimmer, MD, MPH; Instructor and Associate Professor, Department of Pediatrics	Jefferson School of Population Health, Philadelphia, PA	None
Steven Jarrett, PharmD; Medication Safety Officer, Quality Division	Atrium Health, Charlotte, NC	CMS/CDC development of measures

The materials presented in this document do not represent final measure specifications for the Hospital Harm eQMs.

Appendix B: Project Staff

IMPAQ Team	
Name	Role
Kendall Hall, MD, MS	Project Director
Jensen Chiu, MHA	Senior Oversight
Anna Michie, MHS, PMP	Project Manager
Stacie Schilling, MPH	NQF Lead
Bo Feng, PhD	NQF SME
Maggie Lohnes, RN, MSN	eCQM Lead
Michelle Lefebvre, RN, BSN	eCQM Authoring and Maintenance Lead
Chana West, RN, MSN	eCQM Testing Lead
Katie Magoulick, MPH, MSW, LGSW	eCQM Project Manager
Hannah Klein	TEP Lead
Leah Dillard	TEP Meeting Coordination & Support
Kennell Team	
Name	Role
Allison Russo, DrPH, MPH	Information Gathering Lead
Christina Superina, MPP	Project Manager
Courtney Colahan	Team Member
UC Davis Team	
Name	Role
Patrick Romano, MD, MPH	PSI Measure Development Lead
Richard White, MD	Clinical SME
Christian Sandrock, MD, MPH	Clinical SME
Jacqueline Stocking, PhD, MSN, RN	Clinical SME
Garth Utter, MD, MSc	Clinical SME
Daniel Tancredi, PhD	Statistical SME
Guibo Xing, PhD	Measure Testing Lead
Monika Ray, PhD	Computer Science SME
Meghan Weyrich, MPH	Project Manager

The materials presented in this document do not represent final measure specifications for the Hospital Harm eQMs.