NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 2723

Corresponding Measures:

De.2. Measure Title: Wrong-Patient Retract-and-Reorder (Wrong Patient-RAR) Measure

Co.1.1. Measure Steward: NewYork-Presbyterian Hospital

De.3. Brief Description of Measure: A Wrong-Patient Retract-and-Reorder (Wrong Patient-RAR) event occurs when an order is placed on a patient within an EHR, is retracted within 10 minutes, and then the same clinician places the same order on a different patient within the next 10 minutes. A Wrong-Patient Retract-and-Reorder rate is calculated by dividing Wrong Patient-RAR events by total orders examined.

1b.1. Developer Rationale: IMPORTANCE AND RATIONALE FOR THE WRONG-PATIENT RAR MEASURE

Wrong-patient errors (i.e., patient identification errors) have been identified as a serious and common health information technology (IT) safety hazard that requires prevention strategies, and the Wrong-Patient RAR measure has been endorsed by several leading health IT safety experts and organizations:

(1) The November 2014 final report from the Office of the National Coordinator for Health Information Technology (ONC), entitled "Health Information Technology Adverse Event Reporting: Analysis of Two Databases," reviewed 20,758 events that were reported to the ECRI Institute and United Healthcare PSOs and were identified as involving health IT. The researchers reported, "Despite national efforts, wrong-patient errors were among the most common human-computer interface issue—described in 15% of all health IT-related events associated with the human-computer interface."1

(2) In 2014, the Office of the National Coordinator (ONC) published nine SAFER Guides, each designed to help healthcare delivery organizations conduct self-assessments of recommended practices in areas known to be important to the safety and safe use of health IT. One of nine SAFER Guides is entitled "Patient Identification," and has 14 recommendations. Recommendation #14 recommends that "organizations regularly monitor their patient database for patient identification errors." Furthermore, the SAFER Guides recommend using the Wrong-Patient RAR measure, "to estimate the rate of erroneous orders due to patient ID errors," and cite the validation research published in JAMIA and presented in this application. The guide states "Monitoring reduces the likelihood that patients will be misidentified and harmed as a result." In addition, Recommendation #10 refers to verifying patient identification prior to placing orders, and references research using the Wrong- Patient RAR measure as the primary outcome. This research, which led to Recommendation #10, would not have been possible without the Wrong-Patient RAR measure.2

(3) The Office of the National Coordinator proposed the creation of a Health Information Technology (HIT) Safety Center. In 2014, three national HIT Safety leaders (Dean Sittig, David Classen, and Hardeep Singh) authored an article published in JAMIA entitled, “Patient safety goals for the proposed Federal Health Information Technology Safety Center.” In this article, the authors recommended that healthcare organizations report Wrong-Patient RAR rates to the HIT Safety Center on a quarterly basis.3

(4) In March 2015, The Joint Commission published Sentinel Event Alert 54 entitled "Safe use of health information technology." Two of the three examples provided in this Sentinel Event Alert were wrong-patient errors (i.e., a medication order for the wrong patient and an imaging test ordered for the wrong patient). The Joint Commission reiterated the importance of the ONC SAFER Guides, including, “using ongoing safety assessment tools for EHRs in operation to assure their safe performance.”4

(5) In February 2016, NQF issued the final report, “Identification and Prioritization of Health IT Patient Safety Measures,” by the HIT Safety Committee co-chaired by Hardeep Singh and Elisabeth Belmont. The committee agreed that patient identification was a priority area, and recognized the Wrong-Patient RAR measure as an example of a valuable HIT Safety Measure.5

(6) Accurate patient identification has been a Joint Commission National Patient Safety Goal (NPSG) since 2002. Effective as of 2019, The Joint Commission issued the first requirement that hospitals use distinct methods of patient identification for newborns (NPSG.01.01.01, EP 3).6 This requirement was informed by research conducted using the Wrong-Patient RAR measure.7

OPPORTUNITY FOR IMPROVEMENT

Nearly 96% of U.S. hospitals use computerized provider order entry (CPOE) systems.8 Although CPOE has been associated with a reduction in medical errors,9-12 when orders are placed electronically certain types of errors, including placing orders on the wrong patient, may occur more frequently.13-18

The danger of wrong-patient electronic orders was highlighted by one hospital’s report that after implementing CPOE, medications were prescribed for the wrong patient several times per month.19 In 2003, the United States Pharmacopeia analyzed 7,029 voluntarily reported medication errors over a 7-month period, and found a mean of 9 wrong-patient orders at each of 120 participating institutions using CPOE.20 In 2016, the ECRI Institute PSO analyzed 7,613 wrong-patient errors reported by 181 healthcare organizations.21 These reports likely underestimated the extent of wrong-patient orders, as voluntary reporting is known to be an unreliable method for identifying errors.22-23 At Montefiore Medical Center, the Wrong- Patient Retract-and Reorder (Wrong-Patient RAR) measure identified 5,246 orders placed on the wrong patient in 1 year, with a rate of 58 wrong-patient orders per 100,000 orders.24 In that year, 1 in 6 Montefiore clinicians placed an order on the wrong patient, and 1 in 37 hospitalized patients had an order placed for them in error. This was the first study using automated surveillance to identify wrong-patient orders, and it demonstrated the prevalence of wrong-patient orders to be significantly higher than previously thought.

Prior to the advent of the Wrong-Patient RAR report, wrong-patient error research required voluntary reported errors as the outcome measure. However, voluntary reporting of errors has proven unreliable and significantly underestimates adverse event rates. A 2011 study by Classen et al. found that among 393 adverse events, only 4 (1%) were identified through voluntary reporting.22 A 2010 study by the DHHS Office of the Inspector General found that only 8 (7%) adverse events were voluntarily reported of 120 total events discovered.23 As wrong-patient errors cannot be reliably measured through self-report, an unbiased mechanism was needed with which to test safety interventions. Wrong-Patient RAR events fill this gap, as it is a valid automated measure of near-miss wrong-patient errors.

The Wrong-Patient RAR measure has now been implemented at several different hospitals, using different EHR systems, including Montefiore Medical Center, NewYork-Presbyterian, Brigham and Women’s Hospital, and the Department of Veterans Affairs at the New York Harbor Healthcare System.24-26 Wrong-patient errors proved to be an HIT hazard for all four hospitals, with Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) rates varying from 72 to 163 errors per 100,000 orders. In addition, there have been eight peer-reviewed studies that have evaluated quality improvement and patient safety interventions to prevent wrong-patient electronic

orders using Wrong-Patient RAR events as the outcome measure.24,26-32 Finally, the Wrong-Patient RAR measure has been used as the outcome measure in four intervention studies to prevent wrong-patient order errors: three funded by Agency for Healthcare Research and Quality (R21HS023704, R01HS024713, R01HS024945) and one funded by National Institute for Child Health and Human Development (R01HD094793).

In sum, the Wrong-Patient RAR measure can be used as an ongoing safety assessment report for EHRs by monitoring the frequency of wrong-patient orders. Wrong-Patient RAR events may also be used as a primary outcome measure for intervention studies aimed at preventing wrong-patient errors. Finally, efforts are underway to implement the measure and compare data across healthcare systems. These uses for the Wrong-Patient RAR measure are consistent with recommendations from national experts in HIT Safety (see section on Usability and Use).

REFERENCES

1. Office of the National Coordinator for Health Information Technology. Health Information Technology Adverse Event Reporting: Analysis of Two Databases. 2014. http://healthit.gov/sites/default/files/Health_IT_PSO_Analysis_Final_Report_11-25-14.pdf

2. Sittig DF, Ash JS, Singh H. ONC issues guides for SAFER EHRs. J AHIMA 2014;85:50–2.

3. Sittig DF, Classen DC, Singh H. Patient safety goals for the proposed Federal Health Information Technology Safety Center. J Am Med Inform Assoc. 2015;22(2):472-478.

4. Joint Commission. Safe use of health information technology. Sentinel Event Alert, Issue 54 2015; http://www.jointcommission.org/assets/1/18/SEA_54.pdf.

5. Identification and Prioritization of Health IT Patient Safety Measures: Final Report. NQF HIT Safety Committee; 2016. http://www.qualityforum.org/Publications/2016/02/Identification_and_Prioritization_of_HIT_Patient_Safety_Measures.aspx

6. The Joint Commission. R3 Report: Distinct newborn identification requirement. Available at: https://www.jointcommission.org/standards/r3-report/r3-report-issue-17-distinct-newborn-identification-requirement/

7. Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc. 2013;20(2):305-310.

8. Pedersen CA, Schneider PJ, Scheckelhoff DJ. ASHP national survey of pharmacy practice in hospital settings: Prescribing and transcribing—2016. American Journal of Health-System Pharmacy. 2017;74(17):1336-1352.

9. Bates DW, Leape LL, Cullen DJ, Laird N, Petersen LA, Teich JM, Burdick E, Hickey M, Kleefield S, Shea B, Vliet MV, Seger, DL. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA.1998;280:1311-16.

10. Bates DW, Teich JM, Lee J, Seger D, Kuperman GJ, Ma´Luf N, Boyle D, Leape LL. The impact of computerized physician order entry on medication error prevention. J Am Med Inform Assoc 1999;6:313-21.

11. Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. Arch Intern Med 2003;163:1409-16.

12. Reckmann MH, Westbrook JI, Koh Y, Lo C, Day RO. Does computerized provider order entry reduce prescribing errors for hospital inpatients? A systematic review. J Am Med Inform Assoc 2009;16:613-23.

13. Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, Strom BL. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005;293:1197-203.

14. Harrison MI, Kopel R, Bar-Lev S. Unintended consequences of information technologies in health care: an interactive sociotechnical analysis. J Am Med Inform Assoc 2007;14:542-9.

15. Ash JS, Sittig DF, Poon EG, Guappone K, Campbell E, Dykstra RH. The extent and importance of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 2007;14:415-23.

16. Campbell EM, Sittig DF, Ash JS, Guappone KP, Dykstra RH. Types of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 2006;13:547-55.

17. Sittig DF, Ash JS. Clinical Information Systems: Overcoming Adverse Consequences. Boston, MA: Jones and Bartlett, 2010.

18. Ash J, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2003;11:104-12.

19. Cohen, M. (2006). Medication Errors (2nd ed.). Washington DC: American Pharmaceutical Association.

20. United States Pharmacopeia. MEDMARX 5th Anniversary Data Report: A Chartbook of 2003 Findings and Trends 1999-2003. Rockville, MD: United States Pharmacopeia, 2004.

21. ECRI Institute. ECRI Institute PSO Deep Dive: Patient Identification Full Report. 2016; https://www.ecri.org/patient-identification-deep-dive.

22. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, Whittington JC, Frankel A, Seger A, James BC. “Global trigger tool” shows that adverse events in hospitals may be ten times greater than previously measured. Health Aff (Millwood). 2011;30:581-589.

23. Levinson D. Adverse Events in Hospitals: Methods for Identifying Events. Washington, DC: United States Department of Health and Human Services, Office of Inspector General, 2010.

24. Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc. 2013;20(2):305-10.

25. Ingham M, Hughes D, Arslanian K, Ferdico R, Adelman J. Evaluating Wrong Patient Electronic Orders in a Department of Veterans Affairs Hospital. (Presented at the May 2014 National Patient Safety Foundation Annual Congress in Orlando, FL)

26. Green RA, Hripcsak G, Salmasian H, Lazar EJ, Bostwick SB, Bakken SR, Vawdrey DK. Intercepting wrong-patient orders in a computerized provider order entry system. Ann Emerg Med. 2015;65(6):679-686.

27. Adelman J, Aschner J, Schechter C, Angert R, Weiss J, Rai A, Berger M, Reissman S, Parakkattu V, Chacko B, Racine A, Southern W. Use of temporary names for newborns and associated risks. Pediatrics. 2015;136(2):327-333.

28. Adelman JS, Aschner JL, Schechter CB, Angert RM, Weiss JM, Rai A, Berger MA, Reissman SH, Yongue C, Chacko B, Dadlez NM, Applebaum JR, Racine AD, Southern WN. Evaluating serial strategies for preventing wrong-patient orders in the NICU. Pediatrics. 2017;139(5):pii: e20162863.

29. Kannampallil TG, Manning JD, Chestek DW, Adelman J, Salmasian H, Lambert BL, Galanter WL. Effect of number of open charts on intercepted wrong-patient medication orders in an emergency department. J Am Med Inform Assoc. 2018;25(6):739-743.

30. Adelman JS, Applebaum JR, Southern WN, Schechter CB, Aschner JL, Berger MA, Racine AD, Chacko B, Dadlez NM, Goffman D, Babineau J, Green RA, Vawdrey DK, Manzano W, Barchi D, Albanese C, Bates DW, Salmasian H. Risk of wrong-patient orders among multiple vs singleton births in the neonatal intensive care units of 2 integrated health care systems. JAMA Pediatr. 2019.

31. Adelman JS, Applebaum JR, Schechter CB, Berger MA, Reissman SH, Thota R, Racine AD, Vawdrey DK, Green RA, Salmasian H, Schiff GD, Wright A, Landman A, Bates DW, Koppel R, Galanter WL, Lambert BL,

Paparella S, Southern WN. Effect of restriction of the number of concurrently open records in an electronic health record on wrong-patient order errors: a randomized clinical trial. JAMA. 2019;321(18):1780-1787.

32. Lombardi D, Gaston-Kim J, Perlstein D, McDonald AM, McAuliffe K, Lazar J, Appelbaum EC, Last Z, Hulen RC. Preventing wrong-patient electronic orders in the emergency department. J Clin Outcomes Manage. 2016;23(12):550-554.

S.4. Numerator Statement: Total Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) events during a specified time period.

S.6. Denominator Statement: All electronic orders placed during a specified time period.

S.8. Denominator Exclusions: None

De.1. Measure Type: Process

S.17. Data Source: Electronic Health Data, Electronic Health Records, Other

S.20. Level of Analysis: Facility, Integrated Delivery System

IF Endorsement Maintenance – Original Endorsement Date: Dec 10, 2015 **Most Recent Endorsement Date:** Dec 10, 2015

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? This measure is not paired nor grouped.

Preliminary Analysis: Maintenance of Endorsement Measure

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence U*since the prior evaluation.*

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☐ **Yes** ☒ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☐ **Yes** ☒ **No**
- **Evidence graded?** ☐ **Yes** ☒ **No**

Summary of prior review in 2015

- This measure was initially endorsed in 2015
- Within the Montefiore Health System the developer identified 5,246 wrong-patient retract reorder errors. The developer states that this was aligned with on-going initiatives around Health Information Technology safety promulgated by the Office of the National Coordinator and allows for the monitoring of how systems are working and how hospitals are preventing wrong patient orders.
- The developer noted that there are healthcare actions that may reduce the incidence of Wrong Patient RAR, such as better system design (e.g., putting a patient's picture in the electronic health record to ensure that the orders are written on the right patient).

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided updated evidence by citing several ariticles in which the Wrong-Patient RAR measure has been used.
- The developer cites studies conducted at different healthcare settings (e.g., NICU, emergency department) showed reductions in wrong-patient order errors by displaying patient identification alerts when clinicians place orders in the electronic health record (EHR).
- The developer also cites that the Office of the National Coordinator for Health Information Technology recommends that healthcare organizations use the Wrong-Patient Retract-and-Reorder Measure to monitor for patient identification errors.

Questions for the Committee:

- *What is the relationship of this measure to patient outcomes?*
- *How strong is the evidence for this relationship?*
- *Is the evidence directly applicable to the process of care being measured?*

Guidance from the Evidence Algorithm

BOX 1: Outcome measure (No) → BOX 3: Systematic review and graded body of evidence (No) → BOX 7: Empirical evidence submitted by no systematic review (Yes) → BOX 8: Evidence submitted include all studies (Yes) → BOX 9: There is high certainty of the benefits of the measure (Yes) → MODERATE

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer cites an initial 2013 validation study, conducted at Montefiore Medical Center, which found a Wrong-Patient RAR rate of 58 wrong-patient orders per 100,000 orders within a single year.
- The developer cites a 2019 validation study conducted at New York-Presbyterian Hospital, but no Wrong-Patient RAR rate is reported.
- The developer does cite a 2015 sutdy in the "Disparities" section that was conducted at New York-Presbyterian Hospital in which a total of 3,457,342 electronic orders were recorded across five

emergency departments and a total of 5,637 Wrong-Patient RAR events were identified; 163 per 100,000 orders (95% CI 159 to 167) within a 2.5 year study period (Dec 2010 – June 2013).

- Developer provides a table in the reliability testing that reports testing at 6 hospitals and health systems resulting in a range of Wrong-Patient RAR performance between 64 – 163 events per 100,000 patients.

Disparities

- The developer cites a 2015 study conducted at New York-Presbyterian Hospital from 2010 – 2013 in which no statistically significant assosciations were found for gender and race.

Questions for the Committee:

- *Do the data provided by the developer indicate that there a gap in care that warrants a national performance measure?*
- *Does the SC know of any evidence that disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence

Comments:

**Evidence was provided and adequate

**The evidence is convincing that the mesure is capable of identifying near misses caused by patient misidentification.

**Unclear link of these errors to poor outcomes. By definition, these are the errors/events that are being caught and corrected. Its unclear the link of the performance of this measure to order entry errors overall that reach a patient.

**I know of nothing new that th authors have not already included

**Not aware of any new studies in this area. Studies cited are diretly related.

1b. Performance Gap

Comments:

**Gaps in care provided and adequate.

**Performance gap seems to be in the 2-fold range in the few hospitals where the measure has been tested.

**Unclear

**Gap was shown (2-3 fold in performance rate) but since this the entire stepped process result I would predict the variation is greater

**Limited gap based on information provided.

Disparities:

** No disparities.

**None

** Unclear. While there are clear disparities in patient safety literature, its not clear that there are disparities with this narrow target within patient safety

** Really only NICU

**n/a

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

2c. For composite measures: empirical analysis support composite approach

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

- There are no exclusions for this measure.
- There is no risk adjustment.

Reliability

- The Wrong-Patient Retract-and-Reorder (RAR) measure was been tested for reliability in six health care systems. Test-retest and signal-to-noise reliability was assessed. Test-retest reliability was performed by applying the Wrong-Patient RAR measure to NewYork-Presbyterian Hospital's EHR on three different instances (Epic, Allscripts, VistA). Signal-to-noise reliability was assessed by comparing rates of wrong-patient RAR events at all six institutions.
- Because the Wrong-Patient Retract-and-Reorder (RAR) measure captures actual errors in an EHR, and does not rely on documentation, chart abstraction, or voluntary reporting. The developer demonstrated that the data to be highly reliable and repeatable, producing the exact same results when assessing the same population in the same time period. In >12m orders, there were 7,128 wrong patient RAR events, with an event rates of 58 per 100,000 orders. Across three attempts (i.e. data pulls), the kappa score for inter-rater reliability was 1.0 compared to the first pull.

- Using the approach by Adams (2009), the developer calculated a signal-to-noise ratio as a function of the variance between hospitals. Reliability was estimated using a beta-binomial model. In each of the six hospitals tested, the reliability score was 0.99 (near perfect). In addition, measure-level reliability was 0.99.

Validity

- The developer used 2 approaches to validation
    - Comparing the measure to: "Another authoritative source of the same information"
    - Conducting testing such that: "Performance scores resulting from the measure as specified can be used to distinguish good from poor quality"
- Validation 1: Another authoritative source of the same information"
    - Clinicians who triggered the measure were contacted within 6-12 hours of the occurance to verify the event (in 3 health systems). The PPV range from 76.2% to 81.2% across published studies.
- Validation 2: "Performance scores resulting from the measure as specified can be used to distinguish good from poor quality"
    - The developer reported studies evaluating different interventions aimed at preventing wrong-patient errors using the Wrong-Patient RAR measure as the primary outcome of the study. Two studies described showed a significant decrease in wrong-patient RAR events when using an intervention aimed at preventing wrong-patient errors. Another study was presented that demonstrated the impact on different EHR configurations allowing clinicians to open varying numbers of workspaces at a time.

Questions for the Committee regarding reliability:

- *Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?*

Questions for the Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?*

Preliminary rating for reliability:	☒ **High**	☐ **Moderate**	☐ **Low**	☐ **Insufficient**
Preliminary rating for validity:	☐ **High**	☒ **Moderate**	☐ **Low**	☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability – Specifications
Comments:

**Reliability specs were clear and reliability is high

**The measure as pitched is well defined and readily implemented.

**Measurement does not seem to be an issue.

**As an electronic measure and query it has shown great reliability, chart reviews are also done with acceptable reliability

**Must be present in a EHR environment capable of pulling this data automatically

2a2. Reliability – Testing
Comments:

**No

**No

**None

**it is good but suggestions to consider would be to restudy the time till re-order filter to include variaous times so one can see if things are not being underestimated

**No

2b1. Validity – Testing

Comments:

**No

**No, as long as this is deemed a process measure. I would like to have it further developed into a genuine outcome measure in which the outcome of not catching the misidentification in the patients involved (the one with the wrong order and the one who would not have gotten the order)

**None

**It is empiric. I think the authors should consider a shortened measure version that tracks modified orders and that can be linked or correlated to safety outcomes

**No

2b4-7. Threats to Validity

Comments:

**None

**No

**No

**The data is complex (for mon EMR systems) to obtain. The threat to validity may be that it is too exacting a measure

**Yes

2b2-3. Other Threats to Validity

2b2. Exclusions

2b3. Risk Adjustment

Comments:

**No other threats noted

**none that I see

**Risk Adjustment not relevant here.

**I authors should consider the risk adjustment related to factors of clinical coverage - rather than patient racial/ethnic distinctions

**Not risk adjusted

Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 2723

Measure Title: Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) Measure

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☒ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

   - This is a measure that is inteneded to be abstracted from the EHR and has been tested in three EHRs successfully however it is not specified as an eMeasure.

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   - The reliability testing was appropriate.

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - The results demonstrated excellent reliability, 0.99-1.00 across multiple (6) hospitals.

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

Submission document: Testing attachment, section 2a2.2

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

Submission document: Testing attachment, section 2a2.2

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

☒ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☐ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

- Appropriate reliability testing was conducted across multiple EHRs (3) across multiple hospitals (6*).*

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

Submission document: Testing attachment, section 2b2.

- There are no exclusions

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

Submission document: Testing attachment, section 2b4.

- None

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
Submission document: Testing attachment, section 2b5.

- None

15. **Please describe any concerns you have regarding missing data.**

Submission document: Testing attachment, section 2b6.

- None

16. **Risk Adjustment**

16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☒ No ☐ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☒ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☒ No

16d. **Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

- There is no risk adjustment. It is appropriate that this measure is not risk adjusted as there is no conceptual reason for risk adjustment.

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☐ Data element ☐ Both**

18. **Method of establishing validity of the measure score:**

☐ Face validity

☒ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- The two approaches to test the validity of this measure were appropriate.

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

- The positive predictive value for clinician's stating they actually did retract an order on the wrong patient was 76%-81% across sites, which is high but not near perfect.
- Studies did demonstrate variation in the measure across appropriate covariates and the ability to improve performance through interventions.

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

NOTE that data element validation from the literature is acceptable.

Submission document*: Testing attachment, section 2b1.*

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

- Appropriate validity testing was conducted but the PPV was only 76-81% when compared to clinician report of order retraction on the wrong patient.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The data used for the Wrong-Patient RAR measures are electronic clinical data (i.e., EHR, Imaging/Diagnostic Study, Laboratory, Pharmacy, Registry) that are generated or collected by and used by healthcare personnel during the provision of care. The measure uses data that are routinely and automatically collected, and is readily available.
- All data elements are in defined fields in electronic health records (EHRs).

Questions for the Committee:

- *Does the SC have any concerns with respect to feasibility?*

Preliminary rating for feasibility: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility

Comments:

**May not be feasible in some smaller EMRs, but this is unlikely a significant number.

**On must always as what the root cause might be of a near miss, which is what this measure detects. How did the near miss happen and how can that be mitigated?

**Unclear, depending on the EHR, there may be data extraction issues.

**All are available electronically and with large health information-like systems may be easier to gather over time

**Would need to be pulled automatically from EHR.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☒ **Yes** ☒ **No**

Accountability program details

- The developer states that the measure is currently not being used within an accountability program. However, the developer further notes that at present, there is no regulatory body that oversees or mandates public reporting or benchmarking of health IT safety measures.
- Further, the developer states that the measure is currently being evaluated for use as part of a "Leapfrog CPOE Evaluation Tool."
- Additionally, the developer cites the the 21st Century Cures Act and the establishment of a new Electronic Health Record Reporting Programin, which the developer states that the Wrong-Patient Retract-and-Reorder measure has been discussed as a possible measure for this program.
- Lastly, developer provides several citations on recommendations for the use of Wrong-Patient RAR – namely The Joint Commission, ECRI Institute, and the ONC.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- The developer does provide information, since the measure is not being used.

Additional Feedback:

- The developer cites an NQF report in which an NQF-convened national Health IT Safety committee was convened in 2016 to discuss the need for Health IT Safety measures to help identify the nature, scope, and prevalence of HIT-related safety issues, and to assess how well providers, vendors, and others are preventing and/or mitigating HIT-related safety concerns. The developer states that the committee cited the Wrong Patient Retract-and-Reorder measure as an exemplary Health IT Safety measure.

***Questions for the Committee**:*

- *How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?*

- *How has the measure been vetted in real-world settings by those being measured or others?*

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- The developer notes that the Wrong-Patient RAR measure has been used in several studies evaluating the rate of order errors in EHR configurations from 2017 - 2020.
- The developer states that the Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) measure is being used to quantify and monitor wrong-patient order errors in different clinical settings and to evaluate quality and safety interventions to prevent wrong-patient order errors within and across health systems
- The developer reports measure rates from these studies ranging from 10/100,000 to 150/100,000.
- Within these studies, the developer states that the Wrong-Patient RAR measure has demonstrated the ability to identify statistically significant and clinically meaningful differences in rates of wrong-patient order errors across systems, settings, and populations.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- The developer does not provide information on any unexpected findings.
- The developer states that the benefits of the measure serve to identify areas where improvement is most needed and to design and test targeted interventions that address the underlying cause of these errors.

Potential harms

- The developer indicates “none”.

Additional Feedback:

- No additional feedback

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for Usability and Use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency

Comments:

**Accountable and transparent.

**Not reported

**No public reporting

**I did not see evidence of feedback or ways to make the EMR more full-proof. Studies looking at the system "why" would be bneeficial I am also not sure that the rate published on this measure is not withing the already withing the scope of usual data entry error rate

**n/a

4b1. Usability – Improvement

Comments:

**Finally and EMR-specific measure. No Potential harms noted.

**No unintended consequences but there should be a folloow up process to fully understand the near miss and its significance to patient outcomes.

**For the specific issue o wrong patient orders, this is useful for monitoring this specific problem. However each EHR instance may have different vulnerabilities and therefore comparisons across hospitals may not be relevant.

**Minimal harm

**I am not convinced this measure is of significant depth to warrant a national measure implementation strategy.

Criterion 5: Related and Competing Measures

Related or competing measures

- None

Harmonization

- Not applicable

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing

Comments:

**No

**No

**No

**None

**No

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 06/12/2020

- **No NQF Members have submitted support/non-support choices as of this date.**

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

NQF_evidence_attachment_New_Version_3-30-20.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

Yes

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 2723

Measure Title: Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) Measure

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 4/2/2020

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome: Click here to name the health outcome

- ☐ Patient-reported outcome (PRO): Click here to name the PRO

  PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: Wrong-patient order errors (orders placed for a patient that are intended for a different patient)

- ☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

The Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) Measure captures near-miss order errors that do not reach the patient and cause harm. The Wrong-Patient RAR measure is an electronic query that detects RAR events, defined as one or more orders placed for a patient that are retracted (cancelled) by the same clinician within 10 minutes, and then placed by the same clinician for a different patient within the next 10 minutes (Figure 1).[1]

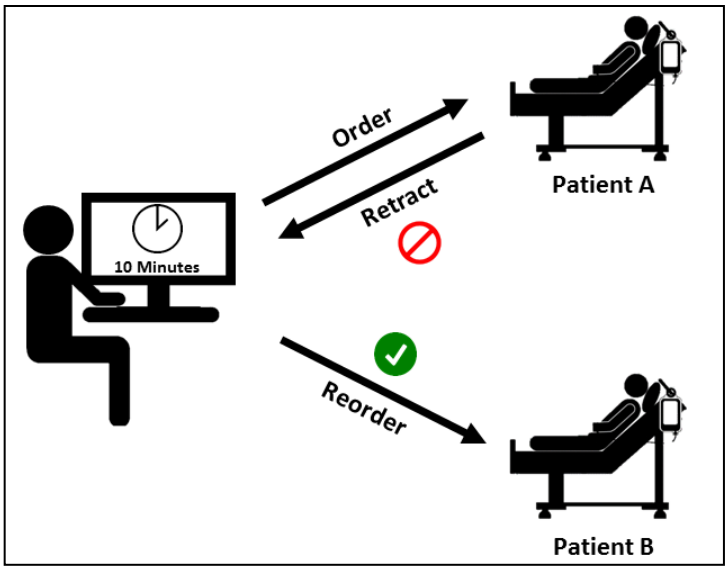

Figure 1. Wrong-Patient Retract-and-Reorder Measure.

Figure 1 alt text: The Wrong-Patient Retract-and-Reorder measure uses an electronic query to detect orders placed for a patient that are cancelled within 10 minutes, and then reordered by the same clinician for a different patient within the next 10 minutes.

The link in the causal pathways is demonstrated graphically in the Incident Causation Model first described by industrial safety expert T.W. Van der Schaaf (Figure 2).[2] In this model, the key distinction between an adverse event and a near-miss error is that in the latter a "human recovery" occurs, just before the error reaches a patient and causes harm. This principle is the foundation for the Wrong-Patient RAR measure, which identifies self-caught errors. Quantifying and investigating near-miss errors enables identification of strengths and weaknesses of systems to prevent human error.

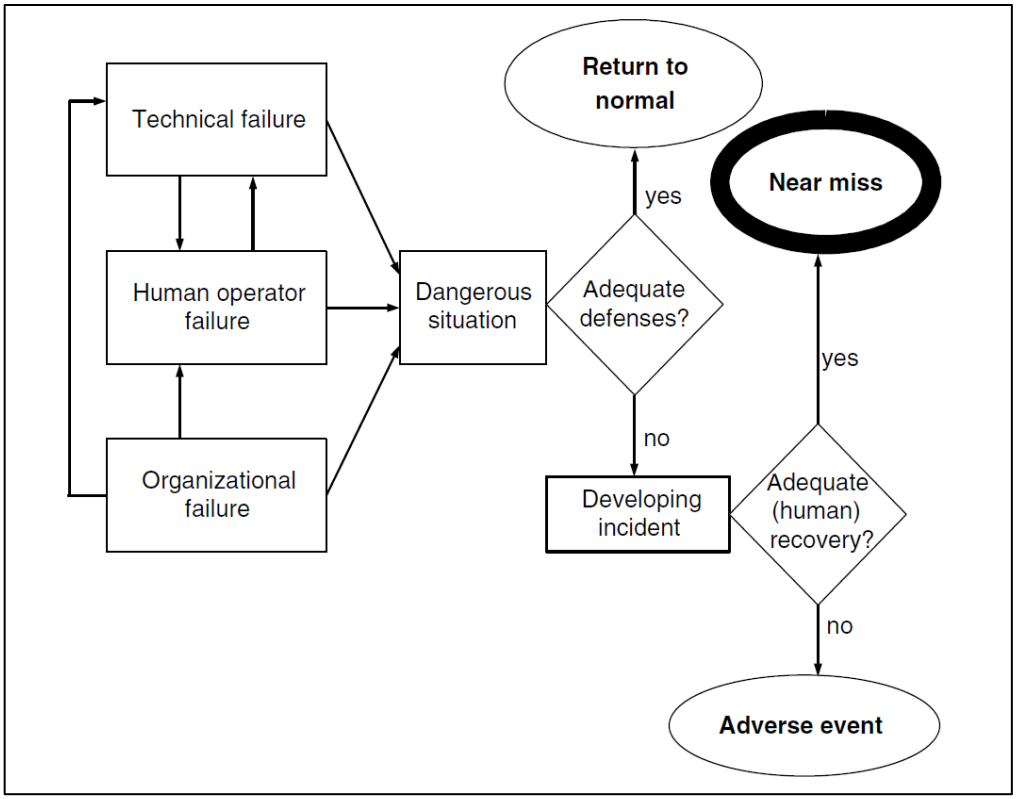

Figure 2. The Incident Causation Model.

Figure 2 alt text: The Incident Causation Model illustrates the pathway that begins with technical, human, and organizational failures and leads to potential adverse events. In this model, the key distinction between an adverse event and a near-miss error is that in a near-miss error a "human recovery" occurs just before the error reaches a patient and causes harm.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

Not applicable. This measure is not derived from patient report.

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses

explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) Measure was initially validated using real-time confirmatory telephone interviews with clinicians.[1] Results demonstrated that the measure correctly identified near-miss, wrong-patient orders in 170 of 223 cases, yielding a positive predictive value of 76.2% (95% CI 70.6% to 81.9%).

Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) events are near-miss order errors that do not reach the patient and cause harm. As in the model shown in Figure 2, near-miss errors are the immediate precursors to potential adverse events.[2] Near-miss errors are also referred to as "close calls" by the Department of Veterans Affairs,[3] "good catches" by the National Association for Healthcare Quality,[4] and "free lessons" by the safety expert James Reason.[5]

The use of near-miss errors to test safety improvements in healthcare is encouraged by every major patient safety organization including the Agency for Healthcare Research and Quality's (AHRQ), Institute of Medicine (IOM), World Health Organization (WHO), Institute for Healthcare Improvement (IHI), and The Joint Commission because they have been shown by safety experts to have the same causal pathway as errors that cause harm.[6-9]

This principle is supported by results of a study by Bates et al., demonstrating that among 4,031 randomly selected patient records 247 adverse drug events and 194 near-miss drug events had similar underlying causes.[10] Because near-miss and errors that reach the patient have similar proximate causes, interventions that reduce near-miss errors should also reduce errors that reach the patient and cause harm. In fact, in the seminal article that first demonstrated that CPOE systems prevent medication errors, Bates et al. demonstrated that CPOE systems decreased both serious errors and near-miss errors.[11]

AHRQ has supported the use of near-miss errors in Health Information Technology (HIT) patient safety initiatives. In fact, the AHRQ National Resource Center for Health Information Technology developed the **Health Information Technology Evaluation Toolkit**, and lists near-miss errors as a particularly useful outcome measure for the evaluation of the effectiveness and safety of Health Information Technology (HIT) projects.[12]

As of 2005, AHRQ and NQF have coordinated the development and maintenance of the "Common Formats" for national reporting of patient safety events to Patient Safety Organizations (PSOs), and have included near-miss errors as important patient safety data to be collected and analyzed.[13]

The Wrong-Patient RAR measure has been used as the outcome measure in eight peer-reviewed publications.[1,14-20] Using the Wrong-Patient RAR measure, studies conducted at three different healthcare systems showed reductions in wrong-patient order errors by displaying patient identification alerts when clinicians place orders in the electronic health record (EHR).[1,14,15] The Wrong-Patient RAR measure has been used in observational and before-after studies of wrong-patient order errors in neonatal intensive care units.[16-18] One study demonstrated a significantly higher rate of wrong-patient order errors in NICUs compared with general pediatric units[16]; a before-after study showed a significant reduction in wrong-patient order errors by using a distinct newborn naming convention that included the mother's first name (e.g., Judysboy/Judysgirl) compared with the typical nondistinct naming convention (e.g., Babyboy/Babygirl)[17]; and an observational study found a significantly higher rate of wrong-patient order errors among multiple-birth infants versus singletons despite using the distinct naming convention.[18] Finally, using the Wrong-Patient RAR measure, two studies assessed the rate of order errors comparing EHR configurations that allowed varying numbers of "charts" open at a time. In a randomized controlled trial, there was no difference in the rate of order errors between clinicians restricted to one chart open versus those allowed up to four charts open at once.[19]

Similarly, a time series analysis showed no difference in wrong-patient order errors when the EHR configuration was changed from allowing two charts to allowing four charts open.[20]

Citing this research, the *Patient Identification SAFER Guide* issued by the Office of the National Coordinator for Health Information Technology (ONC) recommends that healthcare organizations use the Wrong-Patient Retract-and-Reorder Measure to monitor for patient identification errors.[21]

As of 2019, The Joint Commission issued a requirement that all hospitals use distinct methods of identification for newborns as part of its National Patient Safety Goals,[22] citing the study showing a reduction in order errors in the NICU with a distinct naming convention, assessed using the Wrong-Patient RAR measure.[17]

1a.4.2 What process was used to identify the evidence?

Literature and web search.

1a.4.3. Provide the citation(s) for the evidence.

1. Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. *J Am Med Inform Assoc*. 2013;20(2):305-310.

   Relevant point from this reference: validation of the Wrong-Patient RAR measure demonstrated a positive predictive value on 76.2%, verified by real-time clinician interviews. The Wrong-Patient RAR measure demonstrated a significant reduction in order errors in a randomized controlled trial assessing patient identification alerts when placing orders.

2. Institute of Medicine (IOM). Patient Safety: Achieving a New Standard for Care. Washington, DC: National Academy Press, 2004.

   Relevant point from this reference: Near-miss errors are the immediate precursors to potential adverse events. Expresses the value IOM places on near-miss errors for patient safety.

3. Department of Veterans Affairs. VHA National Patient Safety Improvement Handbook. 2011 http://www.va.gov/vhapublications/ViewPublication.asp?pub_ID=2389

   cccc near-miss errors are referred to as “close calls” by the Department of Veterans Affairs.

4. National Association for Healthcare Quality. National Association for Healthcare Quality (NAHQ) Issues 'Call to Action' for Enhancing Health Care Quality and Patient Safety. 2012. http://www.prnewswire.com/news-releases/national-association-for-healthcare-quality-nahq-issues-call-to-action-for-enhancing-health-care-quality-and-patient-safety-174423331.html

   Relevant point from this reference: Near-miss errors are referred to as “good catches” by the National Association for Healthcare Quality.

5. Reason J. Human error: models and management. BMJ: British Medical Journal. 2000. 320(7237): 768.

   Relevant point from this reference: Near-miss errors are referred to as “free lessons” by the safety expert James Reason.

6. Agency for Healthcare Research and Quality: Doing What Counts for Patient Safety: Federal Actions to Reduce Medical Errors and Their Impact Report of the Quality Interagency Coordination Task Force (QuIC) To the President. 2000 http://archive.ahrq.gov/quic/Report/errors6.pdf

   Relevant point from this reference: Expresses the value AHRQ places on near-miss errors for patient safety.

7. Leape L., Abookire S. WHO Draft Guidelines for Adverse Event Reporting and Learning Systems: From Information to Action. Geneva, Switzerland: World Health Organization, 2005. http://www.who.int/patientsafety/events/05/Reporting_Guidelines.pdf .

Relevant point from this reference: Expresses the value the WHO places on near-miss errors for patient safety.

8. Institute for Healthcare Improvement. "Create a Reporting System" Apr. 2011. http://www.ihi.org/knowledge/Pages/Changes/CreateaReportingSystem.aspx

Relevant point from this reference: Expresses the value IHI places on near-miss errors for patient safety.

9. Wu, Albert W. The Value of Close Calls in Improving Patient Safety: Learning How to Avoid and Mitigate Patient Harm. Oakbrook Terrace, IL: Joint Commission Resources, 2011.

Relevant point from this reference: Expresses the value the Joint Commission places on near-miss errors for patient safety.

10. Bates DW, Cullen DJ, Laird N, et al, Incidence of adverse drug events and potential adverse drug events. JAMA. 1995. 274(1): 29-34.

Relevant point from this reference: In a study by Bates et al., of 4,031 randomly selected patient records, 247 adverse drug events and 194 near-miss drug events had similar underlying causes. Because near-miss and actual errors have similar proximate causes, interventions that reduce near-miss errors should also reduce actual errors.

11. Bates DW, Leape LL, Cullen DJ, Laird N, Petersen LA, Teich JM, Burdick E, Hickey M, Kleefield S, Shea B, Vliet MV, Seger, DL. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA.1998;280:1311-16.

Relevant point from this reference: In this seminal article that first demonstrated that CPOE systems prevent medication errors, Bates et al. demonstrated that CPOE systems decreased both serious errors and near-miss errors.

12. Cusack CM, Poon EG. Health Information Technology Evaluation Toolkit. Prepared for the AHRQ National Resource Center for Health Information Technology under contract No. 290-04- 0016. AHRQ Publication No. 08-0005-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2007.

Relevant point from this reference: The AHRQ National Resource Center for Health Information Technology developed the Health Information Technology Evaluation Toolkit, and lists near-miss errors as a particularly useful outcome measure for the evaluation of the effectiveness and safety of Health Information Technology (HIT) projects.

13. Agency for Healthcare Research and Quality. Concepts Underlying Common Formats Event Descriptions. http://www.pso.ahrq.gov/formats/eventdesc.htm

Relevant point from this reference: AHRQ and NQF have coordinated the development and maintenance of the "Common Formats" for national reporting of patient safety events to Patient Safety Organizations (PSOs), and has included near-miss errors as important patient safety data to be collected and analyzed.

14. Green RA, Hripcsak G, Salmasian H, Lazar EJ, Bostwick SB, Bakken SR, Vawdrey DK. Intercepting wrong-patient orders in a computerized provider order entry system. *Ann Emerg Med*. 2015;65(6):679-686.e671.

Relevant point from this reference: Reduction in wrong-patient order errors with patient identification alerts displayed in the EHR, assessed using the Wrong-Patient RAR measure.

15. Lombardi D, Gaston-Kim J, Perlstein D, MacDonald AM, McAuliffe K, Lazar J, Appelbaum EC, Last Z, Hulen RC. Preventing wrong-patient electronic orders in the emergency department. *JCOM*. 2016;23(12):550-554.

    Relevant point from this reference: Reduction in wrong-patient order errors with patient identification alerts displayed in the EHR, assessed using the Wrong-Patient RAR measure.

16. Adelman JS, Aschner JL, Schechter CB, Angert RM, Weiss JM, Rai A, Berger MA, Reissman SH, Yongue C, Chacko B, Dadlez NM, Applebaum JR, Racine AD, Southern WN. Evaluating serial Strategies for preventing wrong-patient orders in the NICU. *Pediatrics*. 2017;139(5).

    Relevant point from this reference: Higher rate of wrong-patient order errors in NICUs compared with general pediatric units, assessed using the Wrong-Patient RAR measure.

17. Adelman J, Aschner J, Schechter C, Angert R, Weiss J, Rai A, Berger M, Reissman S, Parakkattu V, Chacko B, Racine A, Southern W. Use of temporary names for newborns and associated risks. *Pediatrics*. 2015;136(2):327-333.

    Relevant point from this reference: Reduction in wrong-patient order errors in NICUs using a distinct newborn naming convention, assessed using the Wrong-Patient RAR measure.

18. Adelman JS, Applebaum JR, Southern WN, Schechter CB, Aschner JL, Berger MA, Racine AD, Chacko B, Dadlez NM, Goffman D, Babineau J, Green RA, Vawdrey DK, Manzano W, Barchi D, Albanese C, Bates DW, Salmasian H. Risk of wrong-patient orders among multiple vs singleton births in the neonatal intensive care units of 2 integrated health care systems. *JAMA Pediatr*. 2019.

    Relevant point from this reference: Higher rate of wrong-patient order errors among multiples versus singletons receiving care in NICUs, assessed using the Wrong-Patient RAR measure.

19. Adelman JS, Applebaum JR, Schechter CB, et al. Effect of Restriction of the Number of Concurrently Open Records in an Electronic Health Record on Wrong-Patient Order Errors: A Randomized Clinical Trial. *JAMA.* 2019;321(18):1780-1787.

    Relevant point from this reference: Randomized controlled trial comparing two EHR configurations showed no difference in wrong-patient order errors, assessed using the Wrong-Patient RAR measure.

20. Kannampallil TG, Manning JD, Chestek DW, Adelman J, Salmasian H, Lambert BL, Galanter WL. Effect of number of open charts on intercepted wrong-patient medication orders in an emergency department. *J Am Med Inform Assoc*. 2017;25(6):739-743.

    Relevant point from this reference: Time series analysis of two EHR configurations showed no difference in wrong-patient order errors, assessed using the Wrong-Patient RAR measure.

21. Office of the National Coordinator for Health Information Technology. Patient Identification SAFER Guide January 2014. Available at: https://www.healthit.gov/sites/default/files/safer/guides/safer_patient_identification.pdf

    Relevant point from this reference: ONC recommends that healthcare organizations use the Wrong-Patient RAR measure to monitor patient identification errors.

22. The Joint Commission. R3 Report: Distinct newborn identification requirement. Available at: https://www.jointcommission.org/standards/r3-report/r3-report-issue-17-distinct-newborn-identification-requirement/

    Relevant point from this reference: The Joint Commission issued a requirement that all hospitals use distinct methods of newborn identification, citing a study using the Wrong-Patient RAR measure.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

IMPORTANCE AND RATIONALE FOR THE WRONG-PATIENT RAR MEASURE

Wrong-patient errors (i.e., patient identification errors) have been identified as a serious and common health information technology (IT) safety hazard that requires prevention strategies, and the Wrong-Patient RAR measure has been endorsed by several leading health IT safety experts and organizations:

(1) The November 2014 final report from the Office of the National Coordinator for Health Information Technology (ONC), entitled "Health Information Technology Adverse Event Reporting: Analysis of Two Databases," reviewed 20,758 events that were reported to the ECRI Institute and United Healthcare PSOs and were identified as involving health IT. The researchers reported, "Despite national efforts, wrong-patient errors were among the most common human-computer interface issue—described in 15% of all health IT-related events associated with the human-computer interface."1

(2) In 2014, the Office of the National Coordinator (ONC) published nine SAFER Guides, each designed to help healthcare delivery organizations conduct self-assessments of recommended practices in areas known to be important to the safety and safe use of health IT. One of nine SAFER Guides is entitled "Patient Identification," and has 14 recommendations. Recommendation #14 recommends that "organizations regularly monitor their patient database for patient identification errors." Furthermore, the SAFER Guides recommend using the Wrong-Patient RAR measure, "to estimate the rate of erroneous orders due to patient ID errors," and cite the validation research published in JAMIA and presented in this application. The guide states "Monitoring reduces the likelihood that patients will be misidentified and harmed as a result." In addition, Recommendation #10 refers to verifying patient identification prior to placing orders, and references research using the Wrong- Patient RAR measure as the primary outcome. This research, which led to Recommendation #10, would not have been possible without the Wrong-Patient RAR measure.2

(3) The Office of the National Coordinator proposed the creation of a Health Information Technology (HIT) Safety Center. In 2014, three national HIT Safety leaders (Dean Sittig, David Classen, and Hardeep Singh) authored an article published in JAMIA entitled, "Patient safety goals for the proposed Federal Health Information Technology Safety Center." In this article, the authors recommended that healthcare organizations report Wrong-Patient RAR rates to the HIT Safety Center on a quarterly basis.3

(4) In March 2015, The Joint Commission published Sentinel Event Alert 54 entitled "Safe use of health information technology." Two of the three examples provided in this Sentinel Event Alert were wrong-patient errors (i.e., a medication order for the wrong patient and an imaging test ordered for the wrong patient). The Joint Commission reiterated the importance of the ONC SAFER Guides, including, "using ongoing safety assessment tools for EHRs in operation to assure their safe performance."4

(5) In February 2016, NQF issued the final report, "Identification and Prioritization of Health IT Patient Safety Measures," by the HIT Safety Committee co-chaired by Hardeep Singh and Elisabeth Belmont. The committee agreed that patient identification was a priority area, and recognized the Wrong-Patient RAR measure as an example of a valuable HIT Safety Measure.5

(6) Accurate patient identification has been a Joint Commission National Patient Safety Goal (NPSG) since 2002. Effective as of 2019, The Joint Commission issued the first requirement that hospitals use distinct

methods of patient identification for newborns (NPSG.01.01.01, EP 3).6 This requirement was informed by research conducted using the Wrong-Patient RAR measure.7

OPPORTUNITY FOR IMPROVEMENT

Nearly 96% of U.S. hospitals use computerized provider order entry (CPOE) systems.8 Although CPOE has been associated with a reduction in medical errors,9-12 when orders are placed electronically certain types of errors, including placing orders on the wrong patient, may occur more frequently.13-18

The danger of wrong-patient electronic orders was highlighted by one hospital's report that after implementing CPOE, medications were prescribed for the wrong patient several times per month.19 In 2003, the United States Pharmacopeia analyzed 7,029 voluntarily reported medication errors over a 7-month period, and found a mean of 9 wrong-patient orders at each of 120 participating institutions using CPOE.20 In 2016, the ECRI Institute PSO analyzed 7,613 wrong-patient errors reported by 181 healthcare organizations.21 These reports likely underestimated the extent of wrong-patient orders, as voluntary reporting is known to be an unreliable method for identifying errors.22-23 At Montefiore Medical Center, the Wrong- Patient Retract-and Reorder (Wrong-Patient RAR) measure identified 5,246 orders placed on the wrong patient in 1 year, with a rate of 58 wrong-patient orders per 100,000 orders.24 In that year, 1 in 6 Montefiore clinicians placed an order on the wrong patient, and 1 in 37 hospitalized patients had an order placed for them in error. This was the first study using automated surveillance to identify wrong-patient orders, and it demonstrated the prevalence of wrong-patient orders to be significantly higher than previously thought.

Prior to the advent of the Wrong-Patient RAR report, wrong-patient error research required voluntary reported errors as the outcome measure. However, voluntary reporting of errors has proven unreliable and significantly underestimates adverse event rates. A 2011 study by Classen et al. found that among 393 adverse events, only 4 (1%) were identified through voluntary reporting.22 A 2010 study by the DHHS Office of the Inspector General found that only 8 (7%) adverse events were voluntarily reported of 120 total events discovered.23 As wrong-patient errors cannot be reliably measured through self-report, an unbiased mechanism was needed with which to test safety interventions. Wrong-Patient RAR events fill this gap, as it is a valid automated measure of near-miss wrong-patient errors.

The Wrong-Patient RAR measure has now been implemented at several different hospitals, using different EHR systems, including Montefiore Medical Center, NewYork-Presbyterian, Brigham and Women's Hospital, and the Department of Veterans Affairs at the New York Harbor Healthcare System.24-26 Wrong-patient errors proved to be an HIT hazard for all four hospitals, with Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) rates varying from 72 to 163 errors per 100,000 orders. In addition, there have been eight peer-reviewed studies that have evaluated quality improvement and patient safety interventions to prevent wrong-patient electronic orders using Wrong-Patient RAR events as the outcome measure.24,26-32 Finally, the Wrong-Patient RAR measure has been used as the outcome measure in four intervention studies to prevent wrong-patient order errors: three funded by Agency for Healthcare Research and Quality (R21HS023704, R01HS024713, R01HS024945) and one funded by National Institute for Child Health and Human Development (R01HD094793).

In sum, the Wrong-Patient RAR measure can be used as an ongoing safety assessment report for EHRs by monitoring the frequency of wrong-patient orders. Wrong-Patient RAR events may also be used as a primary outcome measure for intervention studies aimed at preventing wrong-patient errors. Finally, efforts are underway to implement the measure and compare data across healthcare systems. These uses for the Wrong-Patient RAR measure are consistent with recommendations from national experts in HIT Safety (see section on Usability and Use).

REFERENCES

1. Office of the National Coordinator for Health Information Technology. Health Information Technology Adverse Event Reporting: Analysis of Two Databases. 2014. http://healthit.gov/sites/default/files/Health_IT_PSO_Analysis_Final_Report_11-25-14.pdf

2. Sittig DF, Ash JS, Singh H. ONC issues guides for SAFER EHRs. J AHIMA 2014;85:50–2.

3. Sittig DF, Classen DC, Singh H. Patient safety goals for the proposed Federal Health Information Technology Safety Center. J Am Med Inform Assoc. 2015;22(2):472-478.

4. Joint Commission. Safe use of health information technology. Sentinel Event Alert, Issue 54 2015; http://www.jointcommission.org/assets/1/18/SEA_54.pdf.

5. Identification and Prioritization of Health IT Patient Safety Measures: Final Report. NQF HIT Safety Committee; 2016. http://www.qualityforum.org/Publications/2016/02/Identification_and_Prioritization_of_HIT_Patient_Safety_Measures.aspx

6. The Joint Commission. R3 Report: Distinct newborn identification requirement. Available at: https://www.jointcommission.org/standards/r3-report/r3-report-issue-17-distinct-newborn-identification-requirement/

7. Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc. 2013;20(2):305-310.

8. Pedersen CA, Schneider PJ, Scheckelhoff DJ. ASHP national survey of pharmacy practice in hospital settings: Prescribing and transcribing—2016. American Journal of Health-System Pharmacy. 2017;74(17):1336-1352.

9. Bates DW, Leape LL, Cullen DJ, Laird N, Petersen LA, Teich JM, Burdick E, Hickey M, Kleefield S, Shea B, Vliet MV, Seger, DL. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA.1998;280:1311-16.

10. Bates DW, Teich JM, Lee J, Seger D, Kuperman GJ, Ma´Luf N, Boyle D, Leape LL. The impact of computerized physician order entry on medication error prevention. J Am Med Inform Assoc 1999;6:313-21.

11. Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. Arch Intern Med 2003;163:1409-16.

12. Reckmann MH, Westbrook JI, Koh Y, Lo C, Day RO. Does computerized provider order entry reduce prescribing errors for hospital inpatients? A systematic review. J Am Med Inform Assoc 2009;16:613-23.

13. Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, Strom BL. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005;293:1197-203.

14. Harrison MI, Kopel R, Bar-Lev S. Unintended consequences of information technologies in health care: an interactive sociotechnical analysis. J Am Med Inform Assoc 2007;14:542-9.

15. Ash JS, Sittig DF, Poon EG, Guappone K, Campbell E, Dykstra RH. The extent and importance of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 2007;14:415-23.

16. Campbell EM, Sittig DF, Ash JS, Guappone KP, Dykstra RH. Types of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc 2006;13:547-55.

17. Sittig DF, Ash JS. Clinical Information Systems: Overcoming Adverse Consequences. Boston, MA: Jones and Bartlett, 2010.

18. Ash J, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2003;11:104-12.

19. Cohen, M. (2006). Medication Errors (2nd ed.). Washington DC: American Pharmaceutical Association.

20. United States Pharmacopeia. MEDMARX 5th Anniversary Data Report: A Chartbook of 2003 Findings and Trends 1999-2003. Rockville, MD: United States Pharmacopeia, 2004.

21. ECRI Institute. ECRI Institute PSO Deep Dive: Patient Identification Full Report. 2016; https://www.ecri.org/patient-identification-deep-dive.

22. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, Whittington JC, Frankel A, Seger A, James BC. "Global trigger tool" shows that adverse events in hospitals may be ten times greater than previously measured. Health Aff (Millwood). 2011;30:581-589.

23. Levinson D. Adverse Events in Hospitals: Methods for Identifying Events. Washington, DC: United States Department of Health and Human Services, Office of Inspector General, 2010.

24. Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc. 2013;20(2):305-10.

25. Ingham M, Hughes D, Arslanian K, Ferdico R, Adelman J. Evaluating Wrong Patient Electronic Orders in a Department of Veterans Affairs Hospital. (Presented at the May 2014 National Patient Safety Foundation Annual Congress in Orlando, FL)

26. Green RA, Hripcsak G, Salmasian H, Lazar EJ, Bostwick SB, Bakken SR, Vawdrey DK. Intercepting wrong-patient orders in a computerized provider order entry system. Ann Emerg Med. 2015;65(6):679-686.

27. Adelman J, Aschner J, Schechter C, Angert R, Weiss J, Rai A, Berger M, Reissman S, Parakkattu V, Chacko B, Racine A, Southern W. Use of temporary names for newborns and associated risks. Pediatrics. 2015;136(2):327-333.

28. Adelman JS, Aschner JL, Schechter CB, Angert RM, Weiss JM, Rai A, Berger MA, Reissman SH, Yongue C, Chacko B, Dadlez NM, Applebaum JR, Racine AD, Southern WN. Evaluating serial strategies for preventing wrong-patient orders in the NICU. Pediatrics. 2017;139(5):pii: e20162863.

29. Kannampallil TG, Manning JD, Chestek DW, Adelman J, Salmasian H, Lambert BL, Galanter WL. Effect of number of open charts on intercepted wrong-patient medication orders in an emergency department. J Am Med Inform Assoc. 2018;25(6):739-743.

30. Adelman JS, Applebaum JR, Southern WN, Schechter CB, Aschner JL, Berger MA, Racine AD, Chacko B, Dadlez NM, Goffman D, Babineau J, Green RA, Vawdrey DK, Manzano W, Barchi D, Albanese C, Bates DW, Salmasian H. Risk of wrong-patient orders among multiple vs singleton births in the neonatal intensive care units of 2 integrated health care systems. JAMA Pediatr. 2019.

31. Adelman JS, Applebaum JR, Schechter CB, Berger MA, Reissman SH, Thota R, Racine AD, Vawdrey DK, Green RA, Salmasian H, Schiff GD, Wright A, Landman A, Bates DW, Koppel R, Galanter WL, Lambert BL, Paparella S, Southern WN. Effect of restriction of the number of concurrently open records in an electronic health record on wrong-patient order errors: a randomized clinical trial. JAMA. 2019;321(18):1780-1787.

32. Lombardi D, Gaston-Kim J, Perlstein D, McDonald AM, McAuliffe K, Lazar J, Appelbaum EC, Last Z, Hulen RC. Preventing wrong-patient electronic orders in the emergency department. J Clin Outcomes Manage. 2016;23(12):550-554.

1b.2. Provide performance scores on the measure as specified (<u>current and over time</u>) at the specified level of analysis. *(<u>This is required for maintenance of endorsement</u>. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

Wrong-patient errors are considered "Never Events"—mistakes that should never happen in healthcare. Multiple intervention studies have demonstrated wrong patient errors are preventable by verification alerts, intrusive interventions, patient photographs and safer names for newborns, and all of these interventions are either recommend or required by the Office of the National Coridnator's SAFER Guides, the Joint Commission or both. **Despite these recommendations, multiple studies and reports have demonstrated that hospitals are still struggling with wrong-patient errors. Below is a list of recently reported Wrong-Patient RAR rates (references at end of this section):**

	Year	RAR rate per 100,000 orders
Wrong-Patient Orders across the Veterans Health Administration Health System (over 500 facilities)	2020 (manuscript in development)	70/100,000
Reducing wrong-patient order entry errors by displaying patient photos in the Electronic Health Records. 2020	2020 (manuscript under review)	150/100,000
Wrong-Patient Orders in Obstetrics: An Unrecognized Patient Safety Risk	2020 (manuscript under review)	80/100,000
Effect of an Alternative Newborn Naming Strategy on Wrong-Patient Errors	2020	150/100,000
ECRI White Paper: Gathering, Using and Sharing Systems Data to Drive Safety Efforts	2019	10 – 30/100,000
Effect of Restriction of the Number of Concurrently Open Records in an Electronic Health Record on Wrong-Patient Order Errors	2019	50/100,000
A Comparison of One- and Four-Open-Chart Access: No Change in Computerized Provider Order Entry Error Rates	2019	35/100,000
Risk of Wrong-Patient Orders Among Multiple vs Singleton Births in the Neonatal Intensive Care Units of 2 Integrated Health Care Systems	2019	42 – 66/100,000
Effect of number of open charts on intercepted wrong-patient medication orders in an emergency department	2018	85/100,000
Evaluating Serial Strategies for Preventing Wrong-Patient Orders in the NICU	2017	50-110/100,000

Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) events consist of orders placed for a patient that are retracted within 10 minutes, and then placed by the same clinician for a different patient within the next 10 minutes. In a study at Montefiore Medical Center, real-time confirmatory telephone interviews with Montefiore clinicians who triggered the measure, demonstrated that among 223 events identified by the RAR

report, 170 were confirmed as near-miss wrong-patient order errors resulting in a positive predictive value of 76.2% (95% CI 70.6% to 81.9%). In the initial validation study, the Wrong-Patient RAR measure identified 5,246 orders placed on the wrong patient in 1 year, resulting in a rate of 58 wrong-patient orders per 100,000 orders (applying the positive predictive value of 76.2%). In 2019, validity testing conducted at NewYork-Presbyterian Hospital using similar methods yielded consistent results. In a sample of 245 providers interviewed, 188 were confirmed wrong-patient order errors, resulting in a positive predictive value of 76.7% (95% CI 71.4% to 82.0%).

To better understand if our results were dependent on the specific definition of retract-and-reorder events, we performed a sensitivity analysis using several combinations of time-to-retraction and time-to-reorder intervals up to a maximum of 30 minutes. In the initial validity testing of time intervals up to 10 minutes to retraction and 10 minutes to reorder, all intervals had similar positive predictive values, but the 10-minute interval identified the most errors. In 2019, we conducted similar sensitivity analysis for the interval 11-30 minutes to retract and 11-30 minutes to reorder. Of 84 interviews with clinicians who triggered the measure, 24 were confirmed wrong-patient order errors for a positive predictive value of 28.7% (95% CI 18.9 to 38.2%). In addition, examining all Wrong-Patient RAR events up to 30 minutes to retract and 30 minutes to reorder over a 1-year period (2017), we found that 73% of events occurred within the 10 minutes to retract and 10 minutes to reorder time interval. Therefore, we confirmed that the 10/10 time interval maximizes both the positive predictive value and the number of Wrong-Patient RAR events.

(REF) Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc. 2013;20(2):305-10.

(REF) Kneifati-Hayek J, Salmasian H, Fernandes Y, Robles P, Redman C, Green R, Vawdrey D, Applebaum J, Southern W, Adelman J. Validation and enhancement of a National Quality Forum measure for wrong-patient orders. 25th Annual AHRQ NRSA Trainees Research Conference; June 1, 2019; Washington DC.

(REF) Salmasian H., Blanchfield B.B., Bates D.W., Adelman J.S., Landman A.B. Reducing wrong-patient order entry errors by displaying patient photos in the Electronic Health Records. 2020 (Under Review JAMA)

(REF) Kern-Goldberger AR, Kneifati-Hayek J, Fernandes Y, Applebaum JR, Adelman JS, Goffman D. Wrong-Patient Orders in Obstetrics: An Unrecognized Patient Safety Risk. Presented at The Society for Maternal Fetal Medicine (SMFM) 40th Annual Meeting. Grapevine, TX. (2020, February 8). (Under review Am J Obstet Gynecol).

(REF) Effect of an Alternative Newborn Naming Strategy on Wrong-Patient Errors: A Quasi-Experimental Study. Ethan Pfeifer, Margaret Lozovatsky, Joanna Abraham, Thomas KannampallilAppl Clin Inform. 2020 Mar; 11(2): 235–241. Published online 2020 Apr 1. doi: 10.1055/s-0040-1705175

(REF) Gathering, Using, and Sharing Systems Data to Drive Safety Efforts. The Partnership for Health IT Patient Safety. (2019). Retrieved from https://assets.ecri.org/PDF/HIT-Partnership/ECRI-Systems-Data-Drive-Safety-Efforts-2020-Partnership.pdf

(REF) Adelman JS, Applebaum JR, Schechter CB, et al. Effect of restricting of the number of concurrently open records in an electronic health record on wrong-patient order errors: a randomized clinical trial. JAMA. 2019;321(18):1780-1787.

(REF) A Comparison of One- and Four-Open-Chart Access: No Change in Computerized Provider Order Entry Error Rates Paula Scariati, Herschel Knapp, Stuart Gray. Appl Clin Inform. 2019 Oct; 10(5): 804–809. Published online 2019 Oct 23. doi: 10.1055/s-0039-1697599

(REF) Adelman JS, Applebaum JR, Southern WN, Schechter CB, Aschner JL, Berger MA, Racine AD, Chacko B, Dadlez NM, Goffman D, Babineau J, Green RA, Vawdrey DK, Manzano W, Barchi D, Albanese C, Bates DW, Salmasian H. Risk of wrong-patient orders among multiple vs singleton births in the neonatal intensive care units of 2 integrated health care systems. JAMA Pediatr. 2019.

(REF) Kannampallil TG, Manning JD, Chestek DW, Adelman J, Salmasian H, Lambert BL, Galanter WL. Effect of number of open charts on intercepted wrong-patient medication orders in an emergency department. J Am Med Inform Assoc. 2018;25(6):739-743.

(REF) Adelman JS, Aschner JL, Schechter CB, Angert RM, Weiss JM, Rai A, Berger MA, Reissman SH, Yongue C, Chacko B, Dadlez NM, Applebaum JR, Racine AD, Southern WN. (2017). Evaluating serial strategies for preventing wrong-patient orders in the NICU. Pediatrics.139(5). pii: e20162863. PMID: 28557730.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

In a study conducted at New York-Presbyterian, during the entire study period (December 2010 through June 2013) a total of 3,457,342 electronic orders were recorded across five emergency departments. In the same period, a total of 5,637 Wrong-Patient Retract-and-Reorder events were identified, indicating an estimated average rate for wrong-patient orders of 163 per 100,000 orders (95% CI 159 to 167). Of all orders, 40.6% were for diagnostic procedures (of which 15% were for imaging modalities and 85% for laboratory tests), 21.1% were for medications, and 38.2% were nursing and miscellaneous orders. The majority of orders were placed by resident physicians (50.7%), followed by attending physicians (34.1%), physician assistants (12.1%), and others (3.1%).

Regression analysis indicated that no patient demographic potential confounder variables had a statistically significant association with the rate of near-miss wrong-patient orders.

PATIENT SEX

Male: Odds Ratio 0.99 (0.89–1.10)

Female NA

PATIENT RACE

Black: Odds Ratio 1.15 (0.71–1.84)

Hispanic: Odds Ratio 1.24 (0.69–2.21)

White: Odds Ratio 1.17 (0.74–1.85)

Other: Odds Ratio 0. 1.22 (0.77–1.93)

Unknown: Odds Ratio 1.00 (0.99–1.01)

Asian: NA

(REF) Green RA, Hripcsak G, Salmasian H, Lazar EJ, Bostwick SB, Bakken SR, Vawdrey DK. Intercepting wrong-patient orders in a computerized provider order entry system. Ann Emerg Med. 2015;65(6):679-686.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

De.6. Non-Condition Specific*(check all the areas that apply):*

Safety, Safety : Medication

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Children, Elderly, Populations at Risk, Populations at Risk : Dual eligible beneficiaries, Populations at Risk : Individuals with multiple chronic conditions, Populations at Risk : Veterans, Women

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

The Wrong-Patient Retract-and-Reorder (RAR) Measure Web Page is currently in process.

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

No data dictionary **Attachment:**

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

Since the initial endorsement of the measure, we have conducted additional validity testing of the Wrong-Patient Retract-and-Reorder (RAR) measure. The “24-hour rule,” defined as exclusion of orders as Wrong-Patient RAR events if they were reordered on the initial patient by any clinician within 24 hours of retraction, was eliminated from the specifications. Applying this rule excluded 20%–25% of cases without an appreciable increase in the positive predictive value (true positives/(true positives + false positives)).

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Total Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) events during a specified time period.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

A Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) event occurs when an electronic order, including medications, lab tests, imaging, procedures and general care orders, is placed on a patient, is retracted within 10 minutes by the same provider, and then the same clinician places the same order on a different patient within the next 10 minutes.

Note 1: Definition of a Retracted Order – an order that is discontinued and never acted upon. For EMRs that do not support the "retraction" function, retracted orders can be defined as orders that are "discontinued" or "cancelled", excluding those in which an action has been charted prior to being discontinued or cancelled.

Note 2: Definition of an Ordering Clinician - for this measure, the ordering clinician is the person who enters the order into the computer. Example 1: if a nurse takes a verbal order from a physician and enters the order into the computer, it is the nurse who may select the wrong patient and is considered the ordering clinician. Example 2: if a medical student enters an order for a patient that is co-signed by a supervising resident, it is the medical student who may select the wrong patient and is considered the ordering clinician.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

All electronic orders placed during a specified time period.

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

All electronic orders including medications, lab tests, imaging, procedures and general care orders placed by ordering clinicians during a specified time period.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

None

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

None

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate –*

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)

Results may be stratified by provider type (e.g. MD, RN, PA, Pharmacist, etc.), patient type (e.g. age group, gender, race, ethnicity, etc.), order type (e.g. medications, lab tests, imaging, etc.), or location (e.g. ED, Inpatient, Outpatient, etc.).

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

Stratification by risk category/subgroup

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Lower score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

Measure Logic for Wrong-Patient Retract-and-Reorder (Wrong Patient-RAR) Events

Numerator

1. Obtain all orders and retraction of orders for a specified time period. For each order and retraction of an order, capture patient and provider demographic characteristics of interest, as well as order information including date and time of order or retraction, and type of order with order details (e.g. Tylenol 325 mg orally three times per day for seven days).

2. Identify the First Order of a potential Wrong-Patient RAR event (orders that are retracted within 10 minutes of being placed by the same clinician).

3. Identify the Second Order of a potential Wrong-Patient RAR event. Get the next non-retracted order that was placed within 10 minutes of the above retracted order by the same clinician on a different patient, where the order is the same as the retracted order. The order should be the same general order, but the underlying details do not need to be an exact match (e.g. dose can change as computer may adjust dose based on patient weight).

4. Any order that meets the above criteria is a Wrong-Patient RAR event. Each RAR event involves a single ordering clinician and two different patients.

Denominator

1. Obtain all orders examined in the specified time period. For each order, capture patient and provider demographic characteristics of interest, as well as order information including date and time of order and type of order with order details.

Rate Calculation (per 100,000 orders)

1. For a specified time period, the Wrong-Patient RAR Rate is calculated as Total Wrong-Patient RAR Events divided by Total Orders multiplied by 100,000.

   (Total Wrong-Patient RAR Events/Total Orders) ? 100,000

2. The Wrong-Patient RAR Rate can be stratified by subgroups of interest.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

The Wrong-Patient RAR measure is not based on a sample. The measure uses all electronic orders placed during a specified time period.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not applicable.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Electronic Health Data, Electronic Health Records, Other

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

The data source for the Wrong-Patient RAR measure is a replicate EHR or data warehouse. The Wrong-Patient RAR measure uses an electronic query to retrospectively extract information on all electronic orders placed during a specified time period.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility, Integrated Delivery System

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Emergency Department and Services, Inpatient/Hospital, Outpatient Services

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

2. Validity – See attached Measure Testing Submission Form

WP-RAR_MeasSubm_MeasTesting_6_30_15B-637067518098861334-637067518289422573.docx,NQF_current_testing_attachment_v7.1_REVISED_3-26-20_CLEAN.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include

information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 2723
Measure Title: Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) Measure
Date of Submission: Click here to enter a date

Type of Measure:

☐ **Outcome (*including PRO-PM*)**	☐ **Composite – *STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.17***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ claims	☐ claims
☐ registry	☐ registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

The Wrong-Patient Retract-and-Reorder (RAR) measure uses existing EHR data for all orders placed within the specified timeframe.

1.3. What are the dates of the data used in testing? 2009 - 2019

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (***must be consistent with levels entered in item S.20***)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

The Wrong-Patient Retract-and-Reorder (RAR) measure has been tested for reliability in six health care systems shown in the table below (Table 1). Test-retest and signal-to-noise reliability was assessed. Test-retest reliability was performed by applying the Wrong-Patient RAR measure to NewYork-Presbyterian Hospital's EHR on three different instances. Signal-to-noise reliability was assessed by comparing rates of wrong-patient RAR events at all six institutions.

Validity testing was conducted at four institutions in three different EHR systems by using the Wrong-Patient RAR measure to identify RAR events then verifying if the event was a near-miss, wrong-patient order based on telephone interviews with ordering clinicians or chart review.

Table 1: Measured Entities for Reliability Testing

Measured Entity	**Type of Testing**	**EHR**
NewYork–Presbyterian Hospital	Reliability, Validity	Allscripts
Henry Ford Health System	Reliability	EPIC
VA New York Harbor Health Care System	Reliability, Validity	VistA
Montefiore Medical Center	Reliability, Validity	EPIC

Brigham and Women's Hospital	Reliability	EPIC
Children's Hospital of Pittsburgh of UPMC	Reliability, Validity	EPIC

1.6. How many and which <u>patients</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

All patients for whom an order was placed during the 1-year data collection period were included in the analysis and any clinician identified by the Wrong-Patient RAR measure during the validation period was eligible to be included in the analysis. No specific demographic characteristics (age, sex, race, etc.) were targeted.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

The same data were used to conduct all aspects of testing.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

No social risk factors were analyzed. The measure uses order information available in the EHR and electronic data systems.

2a2. RELIABILITY TESTING

***<u>Note</u>**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Several types of reliability were examined for the Wrong-Patient RAR measure. Test-retest reliability was carried out by applying Wrong-Patient RAR measure to one institution's data set at multiple instances, yielding similar results. Signal-to-noise analysis was also conducted to determine reliability across institutions by using the measure at multiple institutions, which again yielded similar results. Because wrong-patient RAR events occur at a rate ranging from 58 to 163 per 100,000 orders, we used a minimum sample size of approximately 1 million orders to conduct reliability testing (as shown in Tables 1 and 2).

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

See below.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Reliability

We used several approaches suggested by NQF to test reliability:

1) The measure should be well defined and precisely specified.
2) Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing include:
    a. Test-retest reliability;
    b. Inter-rater/abstractor or intra-rater/abstractor studies;
    c. Signal-to-noise.

1) Reliability: The measure is well defined and precisely specified:

1. The Wrong-Patient Retract-and-Reorder measure identifies actual errors in real time.
2. The Wrong-Patient Retract-and-Reorder measure uses EHR data with computer generated time stamps.
3. The Wrong-Patient Retract-and-Reorder measure does not require human interpretation for numerator or denominator data [i.e. no coders (e.g. AHRQ PSI 90), no chart abstraction (e.g. NHSN CAUTIs), no voluntary reporting (e.g. NDNQI Falls), and no documentation (e.g. Medication Reconciliation).

2) Reliability: Test-Retest and Inter-Rater Reliability Evaluations

The Wrong-Patient Retract-and-Reorder (RAR) measure captures actual errors in an EHR, and does not rely on documentation, chart abstraction, or voluntary reporting. As such, the data are highly reliable and repeatable, producing the exact same results when assessing the same population in the same time period (Table 2).

Table 2: Test-Retest and Inter-Rater Reliability Evaluations

	Total Orders	**Wrong-Patient Retract-and-Reorder Events**	**Wrong-Patient Retract-and-Reorder Rate (per 100,000 orders)**	**Inter-Rater Reliability (Kappa Score)**
Attempt #1	12,262,023	7,128	58	*Reference
Attempt #2	12,262,023	7,128	58	1.0
Attempt #3	12,262,023	7,128	58	1.0

3) Signal-to-Noise Reliability Evaluation

We used the NQF recommended Adams (2009) approach to assess the signal-to-noise measure reliability of the Wrong-Patient Retract-and-Reorder measure in the context of the observed variability across 6 measurement units (i.e., 6 hospital facilities as per Table 3). The signal-to-noise ratio was calculated as a function of the variance between hospitals (signal) and the variance within a hospital (noise). Reliability was estimated using a beta-binomial model.

Table 3. Reliability Statistics by Hospital

	WP-RAR Events	Total orders	WP-RAR events per 100,000 orders	Reliability score
	(1)	(2)	(3)	(4)
NewYork–Presbyterian Hospital	5,637	3,457,342	163	0.99
Henry Ford Health System	2,572	2,280,945	113	0.99
VA New York Harbor Health Care System	2,306	3,202,728	72	0.99
Montefiore Medical Center	6,885	9,024,723	76	0.99
Brigham and Women's Hospital	875	921,644	95	0.99
Children's Hospital of Pittsburgh of UPMC	644	1,002,901	64	0.99

The reliability estimate for each hospital describes how well we can apply this hospital's measure result to compare with other hospitals. Table 2 reports the reliability scores for all six hospitals were 0.99. In addition to the hospital-level reliability, we report the measure-level reliability, which characterizes overall how well one can confidently use this measure to distinguish the performance of one hospital from another. The measure-level reliability estimate is 0.99, which indicates excellent measure-level reliability for the Wrong-Patient Retract-and-Reorder measure.

Following the methodology described in Adams (2009), the hospital reliability scores in column 4 were estimated using the following method:

Reliability score for hospital i = $\frac{\sigma^2_{between-hospital}}{\sigma^2_{between-hospital}+\sigma^2_{hospital\ i\ error}}$

where

$\sigma^2_{hospital\ i\ error} = \frac{p_i(1-p_i)}{n_i}$

$\sigma^2_{between-hospital} = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$; where $E(p_i) = \frac{\alpha}{\alpha+\beta}$

The estimates of α and β were obtained by fitting a beta-binomial model using the xtnbreg command in Stata:

$\alpha = e^{(\gamma_0+\gamma_1)}$

$\beta = \gamma_0$; where γ_0 and γ_1 are model parameters

The measure-level reliability score is estimated as:

Measure-level reliability score = $\frac{\sigma^2_{between-hospital}}{\sigma^2_{between-hospital}+\sigma^2_{within-hospital}}$, where

$\sigma^2_{within-hospital} = \frac{\sum p_i(1-p_i)}{N}$

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☒ **Empirical validity testing**

☐ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

See below.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

See below.

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Validity

We used 2 separate approaches suggested by NQF for validation:

1) "Another authoritative source of the same information"
2) "Performance scores resulting from the measure as specified can be used to distinguish good from poor quality"

1) Validation: "Another authoritative source of the same information"

The Wrong-Patient Retract-and-Reorder (RAR) measure has been tested using the same methodology in three large health systems and yielded similar results. Clinicians who triggered the measure were contacted and interviewed within 6-12 hours of the occurrence of the event to verify if the event was a true wrong-patient order. The positive predictive value (percentage of true positives) ranged from 76.2% to 81.0% in the three validation studies (described in detail below).

Montefiore Medical Center

Wrong-Patient Retract-and-Reorder (RAR) events consist of orders placed for a patient that are retracted within 10 minutes, and then placed by the same clinician for a different patient within the next 10 minutes. In a study at Montefiore Medical Center in 2010, we ran the Wrong-Patient RAR measure every 12 hours, and contacted a convenience sample of Montefiore clinicians who triggered the measure within 12 hours of the event. These near real-time confirmatory telephone interviews with clinicians demonstrated that among 223 events identified by the Wrong-Patient RAR measure, 170 were confirmed as near-miss wrong-patient order errors, resulting in a positive predictive value of 76.2%, (95% CI 70.6% to 81.9%).

To understand if our results were dependent on the specific definition of retract-and-reorder events, we performed a sensitivity analysis using several combinations of time-to-retract and time-to-reorder intervals up to a maximum of 10 minutes. Of the time intervals we tested, all had similar positive predictive values, but the

10-minute interval identified the most errors. The mean time to retraction of the 5,246 orders placed on the wrong patient in 2009 using the 10-minute time interval was 1 minute and 18 seconds, suggesting that on average errors are caught very quickly. An evaluation of Retract-and-Reorder events that extended beyond a 10-minute interval revealed very few additional events.

(REF) Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. *J Am Med Inform Assoc*. 2013;20(2):305-310.

VA New York Harbor Health Care System

In 2015, the VA New York Harbor Health Care System implemented the Wrong-Patient RAR measure in its VistA (Veterans Health Information Systems and Technology Architecture) EHR system, and followed the Montefiore protocol to validate the measure. The measure ran every 12 hours and VA clinicians who triggered the measure were contacted by phone within 12 hours of the event. In a convenience sample of 100 clinicians interviewed, 81 confirmed that the event was a wrong-patient order error, resulting in a positive predictive value of 81.0% (95% CI 73.3% to 88.7%).

NewYork-Presbyterian Hospital

The Wrong-Patient RAR measure was validated in 2019 in the NewYork-Presbyterian Hospital health system. Validation was conducted using similar methods as described above. However, the query ran every 30 minutes and clinicians were contacted within 6 hours of triggering the event. In a convenience sample of 245 providers interviewed, 188 were confirmed wrong-patient order errors, resulting in a positive predictive value of 76.7% (95% CI 71.4% to 82.0%).

To examine further if our results were dependent on the definition of retract-and-reorder events, we performed a sensitivity analysis using time-to-retract and time-to-reorder intervals greater than 10 minutes and up to a maximum of 30 minutes. For the retract-and-reorder interval of 11-30 minutes, we conducted 84 interviews with clinicians and confirmed 24 wrong-patient order errors, for a positive predictive value of 28.7% (95% CI 18.9 to 38.2%). In addition, examining all wrong-patient RAR events up to 30 minutes to retract and 30 minutes to reorder over a 1-year period (2017), we found that 73% of events occurred within the 10 minutes to retract and 10 minutes to reorder (10/10) time interval. Therefore, we confirmed that the 10/10 time interval maximizes the positive predictive value and the number of wrong-patient RAR events. We also examined the effect of the "24-hour rule," defined as exclusion of orders as wrong-patient RAR events if they were reordered on the initial patient by any clinician within 24 hours of retraction. We found that applying this rule excluded 20%–25% of cases without an appreciable increase in positive predictive value. Therefore, we have eliminated the 24-hour rule in the revised specifications.

(REF) Kneifati-Hayek J, Salmasian H, Fernandes Y, Robles P, Redman C, Green R, Vawdrey D, Applebaum J, Southern W, Adelman J. Validation and enhancement of a National Quality Forum measure for wrong-patient orders. 25th Annual AHRQ NRSA Trainees Research Conference; June 1, 2019; Washington DC.

2) Validation: "Performance scores resulting from the measure as specified can be used to distinguish good from poor quality"

Multiple studies evaluated different interventions aimed at preventing wrong-patient errors using the Wrong-Patient RAR measure as the primary outcome of the study. The studies described in a) and b) below showed a significant decrease in wrong-patient RAR events when using an intervention aimed at preventing wrong-patient errors. Studies described in c) below used the Wrong-Patient RAR measure to test EHR configurations allowing clinicians to open varying numbers of workspaces at a time, and yielded consistent results. Taken together, these findings further support the validity of the Wrong-Patient RAR measure.

a) **Interventions to Prevent Wrong-Patient Orders in Computerized Provider Order Entry Systems**

Understanding and Preventing Wrong-Patient Electronic Orders: The Wrong-Patient RAR measure was used in a randomized controlled trial to test two interventions designed to prevent wrong-patient

orders. One intervention used an ID-verify alert, an alert that displayed the patient's name, gender, and age for the clinician to verify with one click, and the other used an ID-reentry function that blocked access to the order entry screen until the clinician entered the patient's initials, gender, and age. Over 4,000 clinicians who placed orders on inpatients were randomly assigned to always receive the ID-verify alert, the ID-reentry function, or neither from December 2009 to June 2010. Over 1 million orders were placed in each arm of the study. Compared with the control population, the ID-verify alert reduced the odds of a wrong-patient RAR event by 16% (P=0.03) and the ID-reentry function reduced the odds of a wrong-patient RAR event by 41% (P<0.001). The study found that wrong-patient orders identified by the Wrong-Patient RAR measure are common, and that interventions like the ID-verify alert and the ID-reentry function can lower the frequency of wrong-patient orders.

(REF) Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. *J Am Med Inform Assoc*. 2013;20(2):305-310.

Intercepting Wrong-Patient Orders in a Computerized Provider Order Entry System:

STUDY OBJECTIVE: A before-after study evaluating the short and long-term effect of a computerized provider order entry–based patient verification intervention to reduce wrong-patient orders in 5 emergency departments. METHODS: A patient verification dialog appeared at the beginning of each ordering session, requiring providers to confirm the patient's identity after a mandatory 2.5-second delay. Using the retract-and-reorder technique, the researchers estimated the rate of wrong-patient orders before and after the implementation of the intervention to intercept these errors. This research consists of a short and long-term quasi-experimental study with both historical and parallel controls. RESULTS: Wrong-patient RAR events were reduced by 30% immediately after implementation of the verification dialog intervention. This reduction persisted when inpatients were used as a parallel control. After 2 years, the rate of wrong-patient RAR events remained at 24.8% less than before intervention.

CONCLUSION: A computerized provider order entry–based patient verification system led to a moderate reduction in wrong-patient orders that was sustained over time. Interception of wrong-patient orders at data entry is an important step in reducing these errors.

(REF) Green RA, Hripcsak G, Salmasian H, Lazar EJ, Bostwick SB, Bakken SR, Vawdrey DK. Intercepting wrong-patient orders in a computerized provider order entry system. *Ann Emerg Med*. 2015;65(6):679-686.

b) **Risk of Wrong-Patient Orders in the Neonatal Intensive Care Unit (NICU)**

Assessing the Risk for Wrong Patient Errors in the NICU: Since there can be no delay in registering newborns and giving them identification wristbands, newborns are commonly assigned temporary first names such as Babyboy or Babygirl, or BB or BG. This convention for assigning temporary, nondistinct, first names in the NICU is thought to increase the risk of patient misidentification. To assess this potential risk, we compared the wrong-patient order rate in the NICU to that of all non-neonates, using the automated Wrong-Patient RAR measure of near miss wrong patient errors. In neonates from 1/2007-6/2010 there were 382 wrong-patient RAR events identified at the individual order level among 341,408 orders for a rate of 112/100,000 orders. This rate was 1.6 times higher than the rate for all non-neonates of 72/100,000 (13,319 wrong-patient RAR events among 18,439,819 individual orders). This evidence is consistent with previous research that suggests assigning temporary, nondistinct, first names to neonates increases the risk of wrong patient errors in the NICU.

(REF) Adelman JS, Southern W, Schechter C, Aschner J, Racine A, Weiss J, Rai A, Berger M, Reissman S, Chacko B, Angert R. Assessing the risk for wrong patient errors in the NICU. American Academy of Pediatrics National Conference and Exhibition, October 2014, San Diego, CA.

Changing the First Name of NICU Patients to Reduce Wrong Patient Orders:

BACKGROUND: To determine the level of risk associated with nondistinct naming conventions such as babyboy or babygirl, we performed an intervention study to evaluate if assigning distinct first names at birth would result in a reduction in wrong-patient errors. METHODS: We conducted a two year before-after implementation study to examine the effect of a distinct naming convention (that incorporated the mother's first name into the newborn's first name (e.g., Wendysgirl) on the incidence of wrong-patient errors. RESULTS: The reduction in wrong-patient RAR events post-vs-pre intervention was 36.3%. After accounting for clusters of orders within order sessions, the odds ratio of a wrong-patient RAR event post-vs-pre intervention was 0.64 (95% CI: 0.42-0.97; p = 0.04). CONCLUSIONS: The study results suggest that nondistinct naming conventions are associated with an increased risk for wrong-patient errors, and that this risk can be mitigated by changing to a more distinct naming convention.

(REF) Adelman JS, Southern W, Schechter C, Aschner J, Angert R, Weiss J, Rai A, Berger M, Reissman S, Chacko B, Racine A. Changing the first name of NICU patients to reduce wrong patient orders. American Academy of Pediatrics National Conference and Exhibition, October 2014, San Diego, CA.

(REF) Adelman JS, Schechter C, Aschner J, Angert R, Weiss J, Rai A, Berger M, Reissman S, Chacko B, Racine A, Southern W. The "Babyboy/Babygirl" problem: evaluating the risk of nondistinct, temporary first names for newborns and measuring the effect of changing the paradigm to reduce wrong patient orders. Vermont Oxford Network Conference, November 2014, Chicago, IL.

Use of Temporary Names for Newborns and Associated Risks: We conducted a 2-year before-after intervention study to examine the effect of a distinct naming convention that incorporates the mother's first name into the newborn's first name (eg, Wendysgirl) on the incidence of wrong-patient errors. We used the Wrong-Patient RAR measure for detecting the outcome of wrong-patient electronic orders. The reduction in wrong-patient RAR events pre- versus post-intervention was 36.3%. After accounting for clusters of orders within order sessions, the odds ratio of an RAR event post-intervention was 0.64 (95% CI 0.42 to 0.97). Using the Wrong-Patient RAR measure, results suggest that nondistinct naming conventions are associated with an increased risk of wrong-patient order errors and that this risk can be mitigated by changing to a more distinct naming convention.

(REF) Adelman JS, Aschner J, Schechter C, Angert R, Weiss J, Rai A, Berger M, Reissman S, Parakkattu V, Chacko B, Racine A, Southern W. Use of temporary names for newborns and associated risks. *Pediatrics*. 2015;136(2):327-333.

Evaluating Serial Strategies for Preventing Wrong-Patient Orders in the NICU:

BACKGROUND: We conducted a before-after intervention study to assess the effectiveness of an ID reentry intervention and a distinct naming convention (eg, "Wendysgirl") for reducing wrong-patient order errors, using non-NICU pediatric units as a comparator. METHODS: Using the Wrong-Patient RAR measure, we examined the rate of wrong-patient orders in NICU and non-NICU pediatric units during 3 periods: baseline (before implementing interventions), ID reentry intervention (reentry of patient identifiers before placing orders), and combined intervention (addition of a distinct naming convention for newborns). RESULTS: We reviewed >850,000 NICU orders and >3.5 million non-NICU pediatric

orders during the 7-year study period. At baseline, wrong-patient orders were more frequent in NICU than in non-NICU pediatric units (117.2 vs 74.9 per 100,000 orders, respectively; odds ratio 1.56; 95% confidence interval, 1.34–1.82). The ID reentry intervention reduced the frequency of errors in the NICU to 60.2 per 100,000 (48.7% reduction; *P* < .001). The combined ID reentry and distinct naming interventions yielded an additional decrease to 45.6 per 100,000 (61.1% reduction from baseline; *P* < .001). CONCLUSIONS: The risk of wrong-patient orders in the NICU was significantly higher than in non-NICU pediatric units. Implementation of a combined ID reentry intervention and distinct naming convention greatly reduced this risk.

(REF) Adelman JS, Aschner JL, Schechter CB, Angert RM, Weiss JM, Rai A, Berger MA, Reissman SH, Yongue C, Chacko B, Dadlez NM, Applebaum JR, Racine AD, Southern WN. Evaluating serial strategies for preventing wrong-patient orders in the NICU. *Pediatrics*. 2017;139(5).

Risk of Wrong-Patient Orders Among Multiple vs Singleton Births in the Neonatal Intensive Care Units of Two Integrated Health Care Systems: Using the Wrong-Patient RAR measure, we assessed the risk of wrong-patient orders among multiple-birth infants and singletons receiving care in the NICU, and examined the proportion of wrong-patient orders that occurred between multiples and their siblings and between multiples and nonsiblings. The data were collected from six NICUs of two large, integrated health care systems in New York City that used newborn naming conventions for newborns per the requirements of The Joint Commission. A total of 10,819 infants were included: 85.5% were singleton-birth infants and 14.5% were multiple-birth infants. The overall wrong-patient order rate was significantly higher among multiple-birth infants than among singletons (66.0 vs 41.7 RAR events per 100,000 orders, respectively; adjusted odds ratio, 1.75; 95% CI, 1.39-2.20; P < .001). The rate of RAR events between multiple-birth infants and nonsiblings (36.1 per 100,000 orders) was similar to that of singleton-birth infants (41.7 per 100,000 orders). The excess risk among multiple-birth infants (29.9 per 100,000 orders) appears to be owing to errors between siblings. The risk increased as the number of siblings receiving care in the NICU increased; a wrong-patient order error occurred in 1 in 7 sets of twin births and in 1 in 3 sets of higher-order multiple births.

(REF) Adelman JS, Applebaum JR, Southern WN, et al. Risk of wrong-patient orders among multiple vs singleton births in the neonatal intensive care units of 2 integrated health care systems. *JAMA Pediatr*. 2019.

c) **Risk of Wrong-Patient Orders Associated with the Number of Workspaces Open in the EHR**

Effect of Restriction of the Number of Concurrently Open Records in an Electronic Health Record on Wrong-Patient Order Errors: A Randomized Clinical Trial:

To assess the risk of wrong-patient orders in an EHR configuration, we compared the proportion of wrong-patient order sessions between clinicians able to open 1 record at a time and those able to open up to 4 records concurrently. Clinicians in a large health system in New York were randomly assigned to 1 record (restricted; n = 1669) or up to 4 records open (unrestricted; n = 1687). The primary outcome was order sessions that included one or more wrong-patient orders identified by the Wrong-Patient RAR measure. The study included 12,140,298 orders under 4,486,631 order sessions and placed for 543,490 patients. There was no significant difference in wrong-patient order sessions per 100 000 in the restricted vs unrestricted group, respectively, overall (90.7 vs 88.0; odds ratio [OR], 1.03 [95% CI, 0.90-1.20]; P = .60) or in any setting (ED: 157.8 vs 161.3, OR, 1.00 [95% CI, 0.83-1.20], P = .96; inpatient: 185.6 vs 185.1, OR, 0.99 [95% CI, 0.89-1.11]; P = .86; or outpatient: 7.9 vs 8.2, OR, 0.94 [95% CI, 0.70-1.28], P = .71). The effect did not differ among settings (P for interaction = .99). In the unrestricted group overall, 66.2% of the order sessions were completed with 1 record open,

including 34.5% of ED, 53.7% of inpatient, and 83.4% of outpatient order sessions. A strategy that limited clinicians to 1 patient record compared with allowing up to 4 records open concurrently did not reduce the rate of wrong-patient order errors.

(REF) Adelman JS, Applebaum JR, Schechter CB, et al. Effect of restricting of the number of concurrently open records in an electronic health record on wrong-patient order errors: a randomized clinical trial. *JAMA*. 2019;321(18):1780-1787.

Effect of Number of Open Charts on Intercepted Wrong-Patient Medication Orders in an Emergency Department: To reduce the risk of wrong-patient errors, safety experts recommend allowing only one patient chart to be open at a time. The number of patient records allowed open in the EHR has been based on anecdotal evidence or institutional preference, and hence varies across institutions. This study was an interrupted time series analysis of wrong-patient medication orders in an emergency department during 2010–2016, using the Wrong-Patient RAR measure as the outcome. Overall, there were 83.6 intercepted wrong-patient events per 100,000 orders, and there was no significant decrease in the number of wrong-patient medication orders during the transition from a maximum of 4 open records to a maximum of 2 ($b = -0.19$, $P = .33$) and no significant increase during the transition from a maximum of 2 open records to a maximum of 4 ($b = 0.08$, $P = .67$). These results have implications regarding decisions about the numbeopen charts in the emergency department in relation to the impact on workflow and efficiency.

(REF) Kannampallil TG, Manning JD, Chestek DW, Adelman JS, Salmasian H, Lambert BL, Galanter WL. Effect of number of open charts on intercepted wrong-patient medication orders in an emergency department. *J Am Med Inform Assoc.* 2018;25(6):739-743.

_____

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b4*

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

Not applicable.

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Not applicable.

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)
Not applicable.

_____

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other**

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Not applicable.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

Not applicable.

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

Not applicable.

2b3.4a. What were the statistical results of the analyses used to select risk factors?

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

Not applicable.

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable.

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**
Not applicable.

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):
Not applicable.

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:
Not applicable.

2b3.9. Results of Risk Stratification Analysis:

Not applicable.

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)
Not applicable.

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not applicable.

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

Generally, to identify clinically significant differences in rates of wrong-patient RAR events across different patient populations and clinical settings, logistic regression models are constructed to calculate odds ratios with their 95% confidence intervals. Depending on the analysis, the *order* or *order session* may be used as the unit of analysis. Table 4 below reports results at the order level. The primary outcome was dichotomous, indicating whether each order was a wrong-patient RAR event and reported as the number of wrong-patient RAR events per 100,000 orders. Logistic regression analyses were performed, using patient as the random intercept to account for nesting of orders within patients, for the overall study population and for each study site. To test the difference in rates of RAR events between sites, we used a logistic regression model with study site as a covariate and patient-level random intercept. In Table 5, we examined the rate of RAR events among singletons, twins, and higher-order multiples at the level of the order, infant, and set of multiples. Each RAR event was attributed to the patient for whom the *initial* order was placed, i.e., a wrong-patient order between a set of twins was counted once.

In Tables 6 and 7, the unit of analysis was the *order session*, defined as a series of orders placed consecutively by a single clinician for a single patient that began with opening the patient's order file and terminated when an order was placed for another patient or after 60 minutes of inactivity, whichever occurred first. If a clinician places orders in the wrong patient's record, several orders may be entered and subsequently

retracted together. Thus, the order session, rather than each order, represents an independent opportunity for a wrong-patient error to occur. Logistic regression models were constructed as described above, using clinician as the random intercept to account for nesting of order sessions within clinicians. The primary outcome was dichotomous, indicating whether each order session contained a wrong-patient RAR event and reported as the number of RAR order sessions per 100,000 order sessions. In the randomized trial shown in Table 7, the primary analysis included all order sessions performed by clinicians according to their assigned randomization group. To determine the effect of group on wrong-patient order sessions, random effects logistic regression models were used with RAR order sessions as the outcome, randomization group as the independent variable, and clinician as the random intercept. The effect was estimated using the odds ratio and 95% confidence interval.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

As shown in Table 4, results of an observational study demonstrated significantly greater odds of wrong-patient RAR events among multiples compared with singletons receiving care in neonatal intensive care units. Rates were remarkably similar at both study sites. In Table 5, although statistical tests were not performed, clinically meaningful differences in rates of RAR events were evident, with increasing order error rates as the number of siblings receiving care in the NICU increased.

Table 4: Observational Study of the Rate of Wrong-Patient Orders in Neonatal Intensive Care Units: Singletons vs Multiples

	Orders		
	Singletons	**Multiples**	
Overall			**Odds Ratio (95% CI)**
RAR rate, per 100,000	41.7	66.0	1.72 (1.37-2.15)
No. RAR events	511	205	
Total No. Orders	1,225,632	310,528	
Site 1			
RAR rate, %	41.6	65.3	1.72 (1.35–2.20)
No. RAR events	405	181	
Total No. Orders	974,022	277,110	
Site 2			
RAR rate, %	42.1	71.8	1.83 (1.01–3.29)
No. RAR events	106	24	
Total No. Orders	251,610	33,418	

Table 5: Observational Study of the Rate of Wrong-Patient Orders Among Singletons, Twins, and Higher-Order Multiples in Neonatal Intensive Care Units

	Singletons	**Twins**	**Higher-order multiples**
Orders			
RAR rate, per 100,000	**41.7**	**61.2**	**98.2**

No. RAR events	511	165	40
Total No. Orders	1,225,632	269,808	40,720
Infants			
RAR rate**, %**	**3.8%**	**7.9%**	**14.5%**
No. RAR events	352	112	23
Total No. Orders	9250	1410	159
Sets of Multiples			
RAR rate**, %**	—	**14.3%**	**38.5%**
No. RAR events	—	101	20
Total No. Orders	—	705	52

(REF) Adelman JS, Applebaum JR, Southern WN, et al. Risk of Wrong-Patient Orders Among Multiple vs Singleton Births in the Neonatal Intensive Care Units of 2 Integrated Health Care Systems. *JAMA Pediatr.* 2019; Aug 26.

As shown in Table 6, we compared the rates of wrong-patient order sessions among women in obstetric units compared with women of child-bearing age receiving care in medical/surgical units over a 3-year period from 2017-2019 at a large academic medical center. Rates of wrong-patient order sessions were significantly greater in obstetric units versus medical/surgical units overall, among attendings and housestaff, and during day and night shifts.

Table 6: Observational Study of the Rate of Wrong-Patient Order Sessions in Obstetric vs Medical/Surgical Units

	Order Sessions		
	Obstetric Units	**Medical/Surgical Units**	**Odds Ratio (95% CI)**
Overall			
RAR rate, per 100,000	**79.5**	**42.3**	1.98 (1.64, 2.39)
No. RAR order sessions	538	276	
Total no. order sessions	676,643	652,820	
Clinician Type			
Attending			
RAR rate, per 100,000	**127.0**	**41.4**	3.70 (2.27, 6.04)
No. RAR order sessions	81	30	
Total no. order sessions	63,769	72,411	
House staff			

RAR rate, per 100,000	**119.9**	**56.0**	2.81 (2.15, 3.67)
No. RAR order sessions	276	146	
Total no. order sessions	230,205	260,552	
Nurse Pract/Physician Asst			
RAR rate, per 100,000	**47.3**	**31.2**	1.29 (0.96, 1.74)
No. RAR order sessions	181	100	
Total no. order sessions	382,669	319,857	
Order Timing			
Day			
RAR rate, per 100,000	**87.1**	**43.2**	2.12 (1.71, 2.64)
No. RAR order sessions	347	186	
Total no. order sessions	398,383	430,749	
Night			
RAR rate, per 100,000	**68.6**	**40.5**	1.78 (1.34, 2.36)
No. RAR order sessions	191	90	
Total no. order sessions	278,260	222,071	

(REF) Goffman D, Kern-Goldberger A, Kneifati-Hayek J, Fernandes Y, Applebaum J, Adelman J. Wrong patient orders in obstetrics: an unrecognized patient safety risk. SMFM 40th Annual Pregnancy Meeting, February 3-8, 2020, Grapevine, TX. *Am J Obstet Gynecology.* 2020;222:S39.

In a randomized controlled trial, we randomly assigned more than 3300 clinicians to one of two EHR configurations: the Restricted group was limited to open one patient record at a time; the Unrestricted group was allowed to open up to four patient records concurrently. Results showed no significant difference between groups in any clinical setting examined. However, there was considerable variation across clinical settings, with higher rates in the emergency department and inpatient units and low rates in the outpatient setting.

le 7: Randomized Controlled Trial Assessing the Rate of Wrong-Patient Order Sessions in Two Different EHR Configurations

	Order Sessions		
	Restricted (1 record)	**Unrestricted (up to 4 records)**	**Odds Ratio (95% CI)**
Overall			
RAR rate, per 100,000	**90.7**	**88.0**	1.03 (0.90 to 1.20)
No. RAR order sessions	1980	2026	
Total no. order sessions	2,183,365	2,303,266	

Emergency department			
rate, per 100,000	**157.8**	**161.3**	1.00 (0.83 to 1.20)
RAR order sessions	560	576	
al no. order sessions	354,882	357,047	
Inpatient			
rate, per 100,000	**185.6**	**185.1**	0.99 (0.89 to 1.11)
RAR order sessions	1324	1340	
al no. order sessions	713,417	723,746	
Outpatient			
rate, per 100,000	**7.9**	**8.2**	0.94 (0.70 to 1.28)
RAR order sessions	86	97	
al no. order sessions	1,082,855	1,176,344	

(REF) Adelman JS, Applebaum JR, Schechter CB, et al. Effect of restricting of the number of concurrently open records in an electronic health record on wrong-patient order errors: a randomized clinical trial. *JAMA*. 2019;321(18):1780-1787.

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

Taken together, these results demonstrate the ability of the Wrong-Patient RAR measure to identify statistically significant and clinically meaningful differences in rates of wrong-patient order errors across systems, settings, and populations. The differences observed are consistent with expected variation, for example, higher rates among multiples vs singletons in neonatal intensive care units, women in obstetric vs medical/surgical units, and patients receiving care in emergency and inpatient vs outpatient settings. The ability of the measure to detect these differences supports the use of the measure to evaluate interventions to reduce wrong-patient order errors.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note*: This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

_____

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

There is no missing data. The measure is run on all orders placed within a specified timeframe.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

As above.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

As above.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in electronic health records (EHRs)

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment: Feasibility_Scorecard_v1_0_4-6-20_SUBMIT.xlsx

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

The Wrong-Patient Retract-and-Reorder (RAR) Measure retrospectively extracts all orders placed for all patients in electronic health record (EHR) systems over a specified time period. The measure captures orders placed in all clinical settings within the health system, including inpatient, emergency department, and outpatient settings. The Wrong-Patient RAR measure identifies orders placed for a patient that are retracted within 10 minutes by the same clinician, then reordered by the same clinician within the next 10 minutes. Required data elements (at the order, clinician, and patient level) are all collected as part of routine care and extracted from standard fields in the health system's data warehouse or replica server. Because the Wrong-Patient RAR measure extracts data for all orders, there are no issues of missing data. The Wrong-Patient RAR measure is an electronic query that has been programmed successfully in several EHR systems, with minimal modifications to the query, including in Allscripts, Epic, Cerner, VistA, and others. Once programmed into the system, the query can be run for specified time periods (monthly, quarterly, annually) with no manual data collection. Taken together, the automated Wrong-Patient RAR measure is highly feasible, can be implemented in a range of EHRs, extracts required data for the universe of orders placed in electronic systems, and therefore is not subject to missing data or need for any manual data collection.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

None

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Regulatory and Accreditation Programs Quality Improvement (Internal to the specific organization)	

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

The Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) measure is being used to quantify and monitor wrong-patient order errors in different clinical settings, and to evaluate quality and safety interventions to prevent wrong-patient order errors within and across health systems. The Wrong-Patient RAR measure has demonstrated the ability to identify statistically significant and clinically meaningful differences in rates of wrong-patient order errors across systems, settings, and populations (see Validity and Evidence sections). The query and technical assistance have been provided to several healthcare systems and the query is publicly available to Epic users through the Epic UserWeb. There are no policies or actions restricting access to implementation or results of the Wrong-Patient RAR measure.

Results demonstrated by the Wrong-Patient RAR measure have led to national patient safety requirements and recommendations by The Joint Commission, the Office of the National Coordinator for Health Information Technology (ONC), and other patient safety experts (described in 4a1.3.). At present, there is no regulatory body that oversees or mandates public reporting or benchmarking of health IT safety measures; however, the Wrong-Patient RAR measure has been proposed as a such a measure to be used for that purpose.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

In 2016 NQF convened a committee of national Health IT Safety experts to discuss the need for Health IT Safety measures to help identify the nature, scope, and prevalence of HIT-related safety issues, and to assess how well providers, vendors, and others are preventing and/or mitigating HIT-related safety concerns. The group concluded that among the patient safety concerns that have emerged with increasing use of EHRs is the issue of accurate and reliable patient identification. The NQF report titled, *Identification and Prioritization of Health IT Patient Safety Measure*, states "Because of the safety risks in this area, the Committee agreed that patient identification is an important area for measurement." The committee then cited the Wrong Patient Retract-and-Reorder measure as an exemplary Health IT Safety measure. Following this report, AHRQ funded our research team to develop additional Retract-and-Reorder measures that will identify a series of medication errors included wrong-drug, wrong-dose, wrong-route and wrong frequency.1

The Wrong-Patient Retract-and-Reorder (Wrong-Patient RAR) measure can be used as an ongoing safety assessment tool for EHRs by monitoring the frequency of wrong-patient orders. National experts have recommended that organizations submit their Wrong-Patient RAR rates to the proposed Federal Health Information Technology Safety Center. **This measure is currently being evaluated for use along with CPOE simulator tests to be used as part of the "Leapfrog CPOE Evaluation Tool." This work is being funded by Gordon and Betty Moore Foundation Grant #6925 (David Classen PI). The Leapfrog Group is an accountability organization that reports data on almost 2,000 hospitals.**

In addition, the 21st Century Cures Act (H.R. 34) directed the US Department of Health and Human Services (HHS) to establish a new Electronic Health Record (EHR) Reporting Programin. The Act states in section 4002(c), "HHS must support the convening of stakeholders to develop reporting criteria" and in section 4005(c), "HHS must report on best practices and current trends provided by patient safety organizations to improve the integration of health IT into clinical practice." Development of the program is underway, with a plan to begin collecting data and to publicly release EHR comparison information in late 2022. 5 The Wrong-Patient Retract-and-Reorder Measure has been discussed as a possible measure for this program.2

Finally, Wrong-Patient RAR events may continute to be used as the primary outcome measure for intervention studies aimed at preventing wrong-patient errors. These proposed uses for the Wrong-Patient RAR measure are consistent with the recommendations below from national experts in HIT Safety.

(1) In Patient Identification SAFER Guides issued by ONC, Recommendation #2.4 suggests that to avoid wrong-patient errors clinicians be required to verify patient identity by typing in selected identifiers prior to placing orders. Recommendation #3.1 endorses that "organizations regularly monitor their patient database for patient identification errors." Furthermore, the SAFER Guides recommend using the Wrong-Patient RAR measure "to estimate the rate of erroneous orders due to patient ID errors." These recommendations cite the validation research on the Wrong-Patient RAR measure published in JAMIA and presented in this application.3,4

(2) HIT Safety leaders (Dean Sittig, David Classen, and Hardeep Singh) authored an article published in JAMIA entitled, "Patient safety goals for the proposed Federal Health Information Technology Safety Center." In this article, the authors recommend that healthcare organizations report Wrong-Patient RAR rates to the HIT Safety Center on a quarterly basis. 5

(3) In March 2015, The Joint Commission published Sentinel Event Alert 54: Safe Use of Health Information Technology." In this report, The Joint Commission recommends, "using ongoing safety assessment tools for EHRs in operation to assure their safe performance."6

(4) Passage of The 21st Century Cures Act (H.R. 34) in 2016 created the HIT Advisory Committee (HITAC) to develop policies, standards, implementation standards, and certification criteria for health IT. HITAC's priority target areas include the use of HIT to improve the quality of health care, to support data for use in quality and public reporting programs, and to address other priority target areas that affect patient safety. In addition, Act states in section 4002(c), "HHS must support the convening of stakeholders to develop reporting criteria" and

in section 4005(c), “HHS must report on best practices and current trends provided by patient safety organizations to improve the integration of health IT into clinical practice.”7

(5) Toward these goals, ECRI Institute and its Partnership for Health IT Patient Safety led a group of 13 healthcare organizations and 4 EHR vendors in a collaboration to collect and aggregate EHR-generated patient safety data using the Wrong-Patient RAR measure. As proof of principle, the collaborative was successful in supporting the programming of the measure and collecting, aggregating, and reporting of wrong-patient order data from multiple healthcare organizations.8

(6) Effective 2019, The Joint Commission issued a requirement as part of its National Patient Safety Goals that all hospitals must now use distinct methods of newborn identification, such as including the mother’s first name in the newborn’s name. This requirement cited studies using the Wrong-Patient RAR measure that showed a reduction in wrong-patient order errors in neonatal intensive care units after implementation of a more distinct newborn naming convention.9

REFERENCES

(1) EHR Reporting Program. The Office of the National Coordinator for Health Information Technology. (2019, July 29). Retrieved from https://www.healthit.gov/sites/default/files/page/2019 07/EHRReportingProgram072519v1.pdf

(2) Identification and Prioritization of Health IT Patient Safety Measures. National Quality Forum. (2016, February 11). Retrieved from http://www.qualityforum.org/Publications/2016/02/Identification_and_Prioritization_of_HIT_Patient_Safety_Measures.aspx

(3) Sittig DF, Ash JS, Singh H. ONC issues guides for SAFER EHRs. J AHIMA. 2014;85:50–52.

(4) Adelman JS, Kalkut GE, Schechter CB, Weiss JM, Berger MA, Reissman SH, Cohen HW, Lorenzen SJ, Burack DA, Southern WN. Understanding and preventing wrong-patient electronic orders: a randomized controlled trial. J Am Med Inform Assoc. 2013;20(2):305-310.

(5) Sittig DF, Classen DC, Singh H. Patient safety goals for the proposed Federal Health Information Technology Safety Center. J Am Med Inform Assoc. 2015;22(2):472-478.

(6) The Joint Commission. Safe use of health information technology. Sentinel Event Alert, Issue 54, March 31, 2015. http://www.jointcommission.org/assets/1/18/SEA_54.pdf.

(7) The 21st Century Cures Act, H.R. 34, Public Law 114–255. December 13, 2016.

https://www.congress.gov/114/plaws/publ255/PLAW-114publ255.pdf

(8) ECRI Institute. Partnership for Health IT Patient Safety. Gathering, Using, and Sharing Systems Data to Drive Safety Efforts. 2019. https://assets.ecri.org/PDF/HIT-Partnership/ECRI-Systems-Data-Drive-Safety-Efforts-2020-Partnership.pdf.

(9) The Joint Commission. R3 report: distinct newborn identification requirement. Available at: https://www.jointcommission.org/standards/r3-report/r3-report-issue-17-distinct-newborn-identification-requirement/.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

Not applicable.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

Not applicable.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

Not applicable.

4a2.2.2. Summarize the feedback obtained from those being measured.

Not applicable.

4a2.2.3. Summarize the feedback obtained from other users

Not applicable.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Not applicable.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

As described in detail in the Testing and Evidence sections, the Wrong-Patient RAR measure has demonstrated statistically significant and clinically meaningful reductions in wrong-patient order errors in evaluating quality improvement and patient safety interventions in varied settings, health systems, populations, and EHRs. Patient identification alerts were shown to significantly reduce wrong-patient order errors in two health systems using two different EHRs, including in the emergency department where rates of these errors are among the highest. Using a more distinct newborn naming convention was shown to significantly reduce wrong-patient order errors among infants receiving care in neonatal intensive care units. These findings informed national recommendations and regulations to improve patient safety. The Wrong-Patient RAR measure has been used in several studies evaluating the rate of order errors in EHR configurations allowing clinicians to open 1, 2, or 4 patient records at a time, which has implications for patient safety as well as for clinician efficiency and satisfaction. The finding of no statistically significant difference in order rates in any study in any clinical setting has informed decision-making about EHR configuration in several large health systems, and national recommendations are being reconsidered. Use of the Wrong-Patient RAR measure has advanced evidence-based patient safety, enabling the use of rigorous methodology to test safety interventions and inform best practices.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

None.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

The Wrong-Patient RAR measure has demonstrated that near-miss wrong-patient order errors occur frequently across healthcare systems and particularly in emergency and critical care settings, i.e., among the most vulnerable patient populations. In addition, because the Wrong-Patient RAR measure is an automated electronic query that accesses EHR data in real time, clinicians can be contacted within minutes of the occurrence of an event to investigate the circumstances and root causes of the errors. These "benefits" of the measure serve to identify areas where improvement is most needed and to design and test targeted interventions that address the underlying cause of these errors.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

No

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment **Attachment:** Final_Binder_7_1_15-637067518101431376-637067518290122687.pdf

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): NewYork-Presbyterian Hospital

Co.2 Point of Contact: Jason, Adelman, jaa9106@nyp.org, 646-317-4803-

Co.3 Measure Developer if different from Measure Steward: NewYork-Presbyterian Hospital

Co.4 Point of Contact: Jason, Adelman, jaa9106@nyp.org, 646-317-4803-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2012

Ad.3 Month and Year of most recent revision: 01, 2020

Ad.4 What is your frequency for review/update of this measure? Every 3 Years

Ad.5 When is the next scheduled review/update for this measure? 01, 2020

Ad.6 Copyright statement: None

Ad.7 Disclaimers: None

Ad.8 Additional Information/Comments: None