



Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

Brief Measure Information

NQF #: 3136

Corresponding Measures:

Measure Title: GAPPS: Rate of preventable adverse events per 1,000 patient-days among pediatric inpatients

Measure Steward: Center of Excellence for Pediatric Quality Measurement

sp.02. Brief Description of Measure: GAPPS is a measure of the number of preventable adverse events per 1,000 patient-days among pediatric inpatients. It is designed to compare rates across institutions and over time. The GAPPS measure utilizes the GAPPS trigger tool to identify adverse events.

1b.01. Developer Rationale:

Patient safety is a core domain of healthcare quality and a major focus for quality improvement efforts.^{1,2} GAPPS is the first and only available global patient safety measure tailored for pediatric populations. By measuring preventable adverse event rates (i.e., harm) in inpatient pediatric populations, it provides important information to providers, hospital quality teams, and state health departments about outcomes of their patient care. Use of the measure will benefit patients, families, and providers because it enables stakeholders to identify and target areas of patient care that may benefit from quality improvement initiatives. Since GAPPS focuses on preventable adverse events, hospitals are able to assess and prioritize clinical areas with potential for immediate improvement.

The GAPPS Measure represents an opportunity to fill a notable void in safety measurement, specifically in pediatric preventable adverse events. National progress on quantifying and tracking AEs has stagnated because of the absence of an accepted national standard in all but a few defined areas. The GAPPS Measure can represent that national standard for pediatrics, and can be an important measure to evaluate and improve adverse event rates across different sites and time periods.

IMPORTANCE OF MEASURING ADVERSE EVENTS

Studies show that there is a high prevalence of medical errors and/or adverse events among the patient population. For example, in 1999, the Institute of Medicine (IOM) estimated that medical errors contribute up to 98,000 deaths and one million injuries each year.² In 2010, the Department of Health and Human Services' Office of the Inspector General estimated that 180,000 deaths due partly to adverse events occur among Medicare patients annually, making adverse events the third leading cause of death in the United States after heart disease and cancer.^{3,4} Hospitalized pediatric patients, who tend to have unique diseases and care distinct from adult patients, are also vulnerable to high adverse event rates: published studies report 11.1 adverse drug events per 100 inpatient pediatric patients, 74 adverse events per 100 neonatal intensive care unit (NICU) patients, and 203 adverse events per 100 pediatric intensive care unit (PICU) patients.⁵⁻⁹ Consequently, tracking adverse events in hospital settings is an important step towards understanding the current state of clinical care and creating initiatives aimed towards improving clinical quality.

Measuring preventable adverse event rates may also help hospitals better understand different aspects of their clinical quality. Some studies have found correlations between patient harm and other quality aspects such as performance on clinical processes of care and other health outcomes.^{10–14} For instance, studies suggest that patients who experience healthcare-related harms have greater odds of in-hospital and 30-day mortality, as well as 30-day readmission.^{11–14} As such, measuring preventable adverse event rates is an essential first step for hospitals to understand and improve their patient care.

IDENTIFICATION OF INPATIENT ADVERSE EVENTS

Various approaches exist for identifying adverse events. Voluntary passive reporting systems are commonly employed but recognized to have low sensitivity.¹⁵ A more reliable, sensitive methodology for capturing data on the safety of hospital care is thus essential.^{16,17}

In 2003, AHRQ released its Patient Safety Indicators (PSI), developed in response to a congressional mandate to reduce medical errors.¹⁸ PSIs are intended to identify events that most likely resulted from preventable medical errors.^{18–20} PSIs have been used with some success but have a number of limitations, in part due to their reliance on administrative data. They have also been found to have low sensitivity.^{17,18}

Use of trigger tools has been shown to be a faster, more sensitive, and more reliable method of adverse event detection than other approaches.^{5,9,17,21–23} “Triggers” are red flags in a medical record that may indicate the presence of an underlying adverse event and prompt further inspection to determine whether an adverse event occurred.^{21,22} An example trigger is the documented administration of an antidote-type medication (e.g., naloxone). Once a trigger is found, an in-depth review is undertaken to determine whether an adverse event occurred. In the case of naloxone, administration may indicate an adverse event occurred if the drug was given to counteract an overdose of opioids given in the hospital but may not if the overdose occurred due to voluntary recreational opioid use. Trigger tools detect adverse events in a high percentage of hospitalizations, ranging in published reports from 19% to 63%, and have evolved significantly over time.^{10–12}

The Global Trigger Tool for Measuring Adverse Events (GTT), developed by the Institute for Healthcare Improvement (IHI), has become widely accepted as an effective approach for identifying adverse events in hospitalized adult patients.^{2,9,15,23–26} The GTT approach identifies 10 times more adverse events than AHRQ’s PSIs and almost 100 times more events than voluntary reporting.^{1,17} However, the GTT has an exclusion of patients under age 18 so does not work for a pediatric population.

PEDIATRIC INPATIENT PATIENT SAFETY: LACK OF STANDARDIZED QUALITY MEASUREMENT

Although one study determined that a version of the GTT applied to the pediatric population could identify pediatric adverse events, the authors and other experts called for development of a standardized pediatric tool that focuses specifically on the problems of hospitalized children and that encompasses the breadth of inpatient pediatric care.^{12,18,25,29} The absence of a comprehensive pediatric trigger tool is a recognized limitation in quantifying the full scope of pediatric adverse events. An early effort to develop a pediatric-focused trigger tool led to the development of the Canadian Pediatric Trigger Tool.^{30,31}

We developed GAPPS to meet the need for a comprehensive, sensitive measure of pediatric patient safety. Our focus was on developing a global trigger tool for pediatric patients that could be more reliably applied across different hospital sites, both academic and community, than previous efforts. In addition, we sought to further refine the list of triggers to make a more robust global trigger list. We used methods similar to those used for GTT, including review of published tools and manual medical record review by experts in patient safety, which has been demonstrated to be a crucial component of developing patient safety measures.³⁰ We also utilized the RAND/UCLA Appropriateness Method, a 16-center field study, and post-analysis refinement of the trigger list to ensure GAPPS includes a more comprehensive trigger list than previous trigger tools.

GAPPS offers an enhancement in trigger tool methodology in that, unlike GTT, it requires that reviewers assess preventability. In the five years since IHI released the second edition of GTT, patient safety experts and national fiscal and quality improvement policies have increasingly focused on addressing preventable adverse events. The GAPPS measure uses the same approach to rate preventability as the North Carolina Patient Safety study, which was found to ascertain preventability with a high degree of reliability.³¹

DISPARITIES IN RISK OF HARM

Children with special healthcare needs experience elevated rates of medical errors.³² Among hospitalized pediatric patients, those with chronic conditions are at significantly higher risk for medical errors than those without chronic conditions.³²

What is known about racial/ethnic disparities in patient safety, particularly among children, is limited.³³ Black and Hispanic newborns are at higher risk of birth trauma.¹⁸ In addition, extrapolations from associations between race/ethnicity and known risk factors for harm suggest that Black and Hispanic children are likely at greater risk of harm than White children. Because severity and complexity of illness increase the risk of errors, and Black and Hispanic children are at higher risk for more complex conditions, these children are at greater risk for adverse events.³³ Emergency room visits are also associated with increased risk of adverse events, and Black and Hispanic children are known to visit the emergency department more frequently.^{28,33}

English language proficiency is also associated with adverse event prevalence. Hospitalized children of parents with limited English proficiency were shown to be significantly more likely to experience harm due to medical care.³⁶

COSTS OF GAPPS UPTAKE

As is the case with all new quality measures, we recognize that there may be concerns about the uptake of GAPPS due to hospital resource constraints. AE monitoring, like patient experience surveying, is not possible using administrative data and therefore requires additional resources. As previously discussed, intensive measurement of preventable AEs is one of the most important areas of quality measurement, one that has to this point been inadequately addressed in pediatric patient populations. The capacity of the GAPPS Measure to spur nationwide improvement in pediatric patient safety represents a significant potential return on investment, one that outweighs implementation resource concerns.

It is important to note that some of the resources and infrastructure to successfully implement the GAPPS measure are already in place, as states commit to systems of mandatory reporting of certain adverse events. About half of states currently have mandatory reporting systems in place, with many others reporting AEs on a voluntary basis.³⁴ GAPPS will require an augmentation of processes hospitals are already initiating to report AEs, but will yield far greater returns. Far from representing a duplicative burden, the GAPPS Measure can take advantage of recent prioritization of patient safety and the resources hospitals are putting in place to evaluate quality in this domain. The marginal resources used for GAPPS implementation provide a drastically more robust assessment of hospital safety than currently in place, as global trigger tools have been shown to capture up to ten times more AEs than alternative AE measurement methods.^{17,35}

REFERENCES

1. Levinson DR. Adverse Events in Hospitals: Methods for Identifying Events. Washington, DC: Office of the Inspector General, Department of Health and Human Services; 2010. Report No.: OEI-06-08-00221.
2. Kohn LT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System. Washington, DC: Institute of Medicine Committee on the Quality of Health Care in America; 1999.
3. James JT. A new, evidence-based estimate of patient harms associated with hospital care. *J Patient Saf.* 2013 Sep;9(3):122–8.
4. Allen M, ProPublica. How Many Die From Medical Mistakes In U.S. Hospitals? [Internet]. NPR.org. [cited 2014 Jun 14]. Available from: <http://www.npr.org/blogs/health/2013/09/20/224507654/how-many-die-from-medical-mistakes-in-u-s-hospitals>
5. Sharek PJ, Classen D. The incidence of adverse events and medical error in pediatrics. *Pediatr Clin North Am.* 2006 Dec;53(6):1067–77.
6. Pediatric Patient Safety Measurement Tool (PPSMT) Pilot Toolkit. Children's Hospital Association; 2012.
7. Stockwell DC, Slonim AD. Quality and safety in the intensive care unit. *J Intensive Care Med.* 2006 Aug;21(4):199–210.
8. Bates DW, Spell N, Cullen DJ, Burdick E, Laird N, Petersen LA, et al. The costs of adverse drug events in hospitalized patients. Adverse Drug Events Prevention Study Group. *JAMA J Am Med Assoc.* 1997 Jan 22;277(4):307–11.

9. Takata GS, Mason W, Taketomo C, Logsdon T, Sharek PJ. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children's hospitals. *Pediatrics*. 2008 Apr;121(4):e927-935.
10. Calder L, Tierney S, Jiang Y, Gagné A, Gee A, Hobden E, et al. Patient safety analysis of the ED care of patients with heart failure and COPD exacerbations: a multicenter prospective cohort study. *Am J Emerg Med*. 2014 Jan;32(1):29-35.
11. Joice GA, Deibert CM, Kates M, Spencer BA, McKiernan JM. "Never events": Centers for Medicare and Medicaid Services complications after radical cystectomy. *Urology*. 2013 Mar;81(3):527-32.
12. Rosen AK, Geraci JM, Ash AS, McNiff KJ, Moskowitz MA. Postoperative adverse events of common surgical procedures in the Medicare population. *Med Care*. 1992 Sep;30(9):753-65.
13. Rosen AK, Loveland S, Shin M, Shwartz M, Hanchate A, Chen Q, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Med Care*. 2013 Jan;51(1):37-44.
14. Sukumar S, Roghmann F, Trinh VQ, Sammon JD, Gervais M-K, Tan H-J, et al. National trends in hospital-acquired preventable adverse events after major cancer surgery in the USA. *BMJ Open*. 2013;3(6).
15. Sharek PJ, Horbar JD, Mason W, Bisarya H, Thurm CW, Suresh G, et al. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics*. 2006 Oct;118(4):1332-40.
16. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med*. 2010 Nov 25;363(22):2124-34.
17. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. "Global Trigger Tool" Shows That Adverse Events In Hospitals May Be Ten Times Greater Than Previously Measured. *Health Aff (Millwood)*. 2011 Apr 1;30(4):581-9.
18. Miller MR, Elixhauser A, Zhan C. Patient Safety Events During Pediatric Hospitalizations. *Pediatrics*. 2003 Jun 1;111(6):1358-66.
19. Dynan L, Goudie A, Smith RB, Fairbrother G, Simpson LA. Differences in Quality of Care Among Non-Safety-Net, Safety-Net, and Children's Hospitals. *Pediatrics*. 2013 Feb 1;131(2):304-11.
20. Miller MR, Zhan C. Pediatric Patient Safety in Hospitals: A National Picture in 2000. *Pediatrics*. 2004 Jun 1;113(6):1741-6.
21. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care*. 2003 Jun 1;12(3):194-200.
22. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care*. 2003 Dec 1;12(suppl 2):ii39-ii45.
23. Agarwal S, Classen D, Larsen G, Tofil NM, Hayes LW, Sullivan JE, et al. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc*. 2010 Sep;11(5):568-78.
24. Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care*. 2000 Mar;38(3):261-71.
25. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ Can Med Assoc J J Assoc Medicales Can*. 2004 May 25;170(11):1678-86.
26. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. *Med J Aust*. 1995 Nov 6;163(9):458-71.
27. Kirkendall ES, Kloppenborg E, Papp J, White D, Frese C, Hacker D, et al. Measuring adverse events and levels of harm in pediatric inpatients with the Global Trigger Tool. *Pediatrics*. 2012 Nov;130(5):e1206-1214.

28. Matlow AG, Baker GR, Flintoft V, Cochrane D, Coffey M, Cohen E, et al. Adverse events among children in Canadian hospitals: the Canadian Paediatric Adverse Events Study. CMAJ Can Med Assoc J J Assoc Med Can. 2012 Sep 18;184(13):E709-718.
29. Stockwell D, Bisarya H, Classen D, Kirkendall E, Landrigan C, Lemon V, et al. A trigger tool to detect harm in pediatric inpatient settings. Pediatrics. 2015;
30. Scanlon MC, Harris JM, Levy F, Sedman A. Evaluation of the Agency for Healthcare Research and Quality Pediatric Quality Indicators. Pediatrics. 2008 Jun 1;121(6):e1723-31.
31. Sharek PJ, Parry G, Goldmann D, Bones K, Hackbarth A, Resar R, et al. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. Health Serv Res. 2011 Apr;46(2):654-78.
32. Ahuja N, Zhao W, Xiang H. Medical Errors in US Pediatric Inpatients With Chronic Conditions. Pediatrics. 2012 Oct 1;130(4):e786-93.
33. Flores G, Ngui E. Racial/Ethnic Disparities and Patient Safety. Pediatr Clin North Am. 2006 Dec;53(6):1197-215.
34. QuPS.org - Medical Errors and Patient Safety - US States Comparison - Overview - [Internet]. [cited 2016 Nov 16]. Available from: http://www.qups.org/med_errors.php?c=internal&id=195
35. Office of the Inspector General. Adverse Events In Hospitals: National Incidence Among Medicare Beneficiaries. Department of Health and Human Services;; 2010.
36. Khan A, Yin HS, Brach C, et al. Association Between Parent Comfort With English and Adverse Events Among Hospitalized Children. JAMA Pediatr. 2020;174(12):e203215. doi:10.1001/jamapediatrics.2020.3215

sp.12. Numerator Statement: The number of preventable adverse events found in a patient sample.

sp.14. Denominator Statement: The denominator is 1,000 patient-days for all sampled pediatric patients who meet inclusion, but not exclusion, criteria.

sp.16. Denominator Exclusions:

We exclude patients who meet the above inclusion criteria but fall into the following exclusion categories:

- Patients discharged from the Emergency Department without admission to the hospital;
- Patients in newborn nurseries;
- Patients over age 18

Measure Type: Outcome

sp.28. Data Source:

Electronic Health Records: Electronic Health Records

Paper Medical Records

Electronic Health Records

sp.07. Level of Analysis:

Facility

IF Endorsement Maintenance – Original Endorsement Date: 2017-07-12 04:01 PM

Most Recent Endorsement Date: 7/12/2017 4:01:40 PM

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

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

sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?:

1. 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

1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

[Response Begins]

Yes

[Yes Please Explain]

New evidence that the target population values the measured outcome is included in response to 1a.02. The evidence concerns the clinical, familial, and societal burdens of patient harm, as well as the association of patient harm with future health.

[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

Current Submission:

Updated evidence information here.

Previous (Year) Submission:

Evidence from the previous submission here.

1a.01. Provide a logic model.

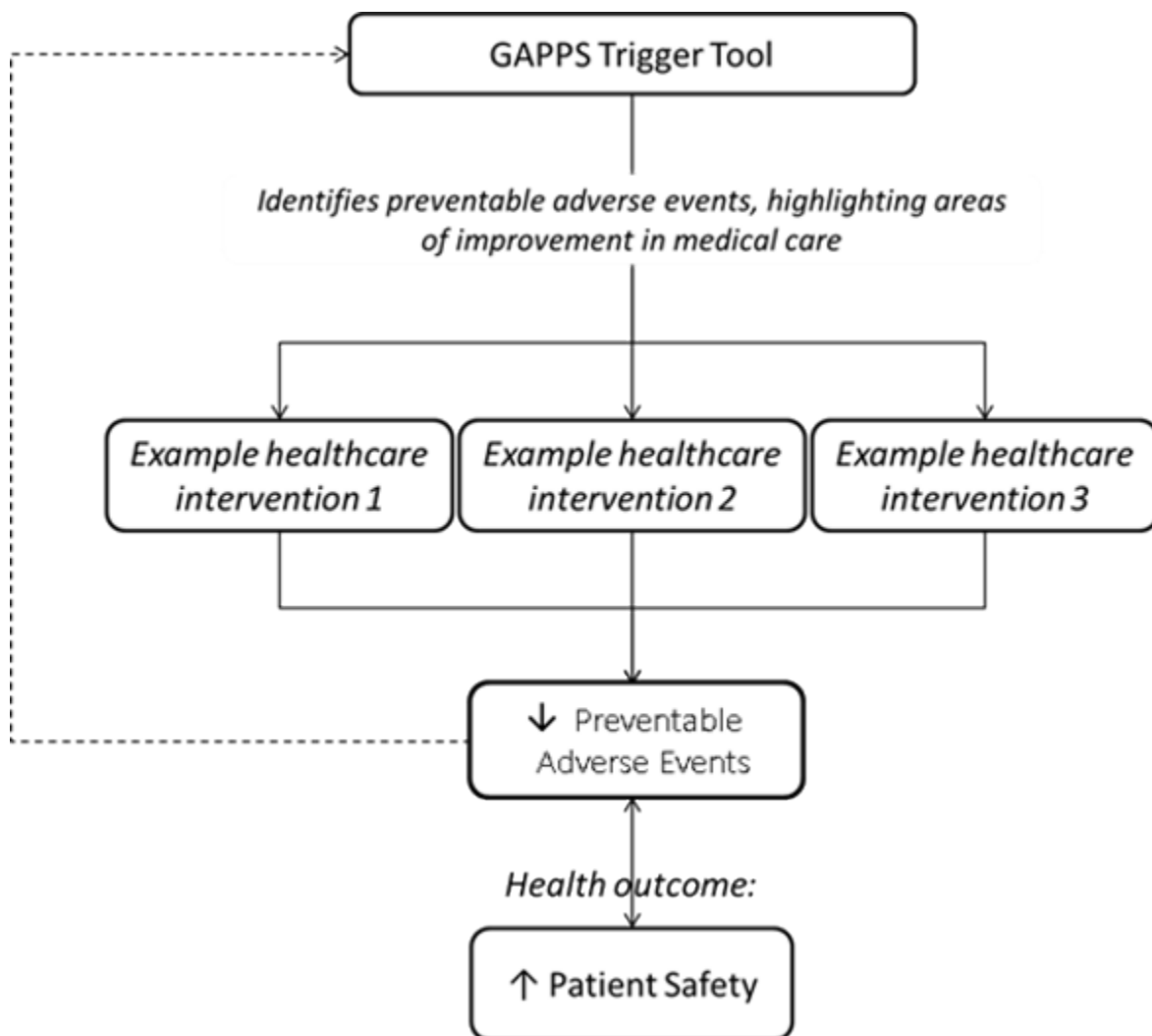
Briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

[Response Begins]

Previous (2017) Submission:

GAPPS is a measure of global patient safety that specifically focuses on identifying preventable adverse events. All the AEs that GAPPS identifies represent instances in which existing patient safety interventions could be applied to prevent patient harm. GAPPS improves the capacity of hospitals to identify preventable AEs, target resources, and implement methods for decreasing risk of AEs in these prioritized domains. For example, GAPPS may assist hospitals in identifying preventable AEs resulting from errors. A hospital may implement new protocols to decrease the use of narcotics in an attempt to address preventable AEs related to the Naloxone (Narcan) administration trigger (see medication administration records [MARs] section of the included trigger list). Another example might be the initiation of rapid response teams to decrease AEs related to the trigger transfer to higher level of care (see physician orders section of trigger list). The advantage of the GAPPS Measure is that it allows hospitals to make notable improvements in patient safety while complementing voluntary event reporting. Using GAPPS, hospitals can implement evidence-based approaches to improve patient safety and measure the

effectiveness of these efforts, allowing for ongoing improvement. Additionally, the GAPPs Measure can be used for analyses across institutions that may serve as the foundation for development of new evidence-based interventions to reduce preventable AEs.



[Response Ends]

1a.02. Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.

Describe how and from whom input was obtained.

[Response Begins]

Current Submission:

CLINICAL, FAMILIAL, AND SOCIETAL BURDENS OF PATIENT HARM

Failures in patient safety lead to substantial morbidity and mortality and thus have grave ramifications for patients and families.¹ Deaths due to adverse events or medical errors cause enormous suffering for families. Furthermore, temporary or long-term injury, increased hospital length of stay, as well as additional medical or surgical

interventions necessitated by adverse events, place psychological and financial burdens on patients and families.^{2,3,4}

ASSOCIATION OF PATIENT HARM WITH FUTURE HEALTH

Adverse events due to healthcare can have lasting negative effects on children's health. Although adverse events most often result in transient harms, some lead to permanent injury and death. Research is underway on the long-term outcomes of children who suffer from healthcare-associated harm. GAPPS is designed to identify permanent as well as transient harms and thus facilitates the identification of events that affect the future health of hospitalized children.

REFERENCES

1. Zhan C, Miller MR. EXcess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA*. 2003;290(14):1868-1874. doi:10.1001/jama.290.14.1868.
2. National Quality Forum. Serious Reportable Events in Healthcare--2006 Update. Washington, DC; 2007.
3. Tessier L, Guilcher SJT, Bai YQ, Ng R, Wodchis WP. The impact of hospital harm on length of stay, costs of care and length of person-centred episodes of care: a retrospective cohort study. *CMAJ*. 2019;191(32):E879-E885. doi:10.1503/cmaj.181621
4. Hoogervorst-Schilp J, Langelaan M, Spreeuwenberg P, de Bruijne MC, Wagner C. Excess length of stay and economic consequences of adverse events in Dutch hospital patients. *BMC Health Serv Res*. 2015;15:531. Published 2015 Dec 1. doi:10.1186/s12913-015-1205-5

[Response Ends]

1a.03. Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.

[Response Begins]

USE OF TRIGGER TOOLS TO IDENTIFY PATIENT HARM

The GAPPS measure uses a trigger tool to detect preventable AEs in hospitalized children. Multiple studies suggest that trigger tools are valid and reliable for tracking the incidence of patient harms in hospital settings, including within pediatric populations; trigger tools, including automated approaches to trigger tools, are better able to detect AEs than other methods (e.g., traditional voluntary incident reporting and detection tools used with administrative databases).¹⁻¹² Trigger tool studies have identified specific triggers that have high yields for AE detection, such as "return to surgery," "positive blood culture," or "abrupt medication stop."^{8,13,14} These findings were applied when developing the triggers used in the GAPPS measure.

In addition, nearly all studies evaluating mean medical record review times reported times under 30 minutes,^{2,15-17} indicating that trigger tools can be used to adequately detect AEs with a reasonably small time burden. The GAPPS methodology likewise employs a maximum 30-minute time frame. Since GAPPS can be applied using either a manual or automated approach (the difference is whether triggers are initially identified by a primary reviewer or by an algorithm programmed into an electronic health record [EHR] system), the automated approach to identifying preventable AEs may decrease the time burden even further as shown in several studies.¹⁸⁻²¹ As previously indicated, the automation of the trigger identification system has no impact on the measure beyond changing the means by which triggers are identified. The remaining pieces of the measure process following trigger identification are exactly the same for the automated and manual trigger approaches.

RELATIONSHIP BETWEEN PATIENT SAFETY AND OTHER ASPECTS OF QUALITY

Some studies in adult populations have shown that patient harm rates are associated with other aspects of clinical quality. These studies demonstrate that rates of patient harm directly correlate with other quality metrics, including performance on clinical processes of care and other health outcomes.²²⁻²⁶ For example, evidence shows that patients who experience healthcare-related harms have greater odds of in-hospital and 30-day mortality, as well as 30-day readmission.²³⁻²⁶ It is likely that the same general associations hold true for pediatric populations.

MEASURING PATIENT SAFETY TO DRIVE QUALITY IMPROVEMENT

Patient safety is a core domain of healthcare quality and a major focus for quality improvement efforts.^{10,27} Hospitals have been able to demonstrate that having a reliable means to track AEs leads to improvements in patient safety and associated clinical outcomes.^{28–32} For example, hospitals that institute real-time adverse drug event surveillance systems are able to intervene before AEs become severe, or are able to prevent future AEs altogether.^{28,30} In addition, the ability to track AEs allows for the design, implementation, and evaluation of targeted interventions, resulting in fewer AEs and decreased mortality.^{29,30,32}

REFERENCES

1. Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. *JAMA J Am Med Assoc.* 1991 Nov 27;266(20):2847–51.
2. Agarwal S, Classen D, Larsen G, Tofil NM, Hayes LW, Sullivan JE, et al. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc.* 2010 Sep;11(5):568–78.
3. Call RJ, Burlison JD, Robertson JJ, Scott JR, Baker DK, Rossi MG, et al. Adverse Drug Event Detection in Pediatric Oncology and Hematology Patients: Using Medication Triggers to Identify Patient Harm in a Specialized Pediatric Patient Population. *J Pediatr.* 2014 Apr 24;
4. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. “Global Trigger Tool” Shows That Adverse Events In Hospitals May Be Ten Times Greater Than Previously Measured. *Health Aff (Millwood).* 2011 Apr 1;30(4):581–9.
5. Griffin FA, Classen DC. Detection of adverse events in surgical patients using the Trigger Tool approach. *Qual Saf Health Care.* 2008 Aug 1;17(4):253–8.
6. Hooper AJ, Tibballs J. Comparison of a Trigger Tool and voluntary reporting to identify adverse events in a paediatric intensive care unit. *Anaesth Intensive Care.* 2014 Mar;42(2):199–206.
7. Kennerly DA, Kudyakov R, da Graca B, Saldaña M, Compton J, Nicewander D, et al. Characterization of Adverse Events Detected in a Large Health Care Delivery System Using an Enhanced Global Trigger Tool over a Five-Year Interval. *Health Serv Res.* 2014 Mar 13;
8. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care.* 2003 Jun 1;12(3):194–200.
9. Takata GS, Mason W, Taketomo C, Logsdon T, Sharek PJ. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children’s hospitals. *Pediatrics.* 2008 Apr;121(4):e927–935.
10. Levinson DR. Adverse Events in Hospitals: Methods for Identifying Events. Washington, DC: Office of the Inspector General, Department of Health and Human Services; 2010. Report No.: OEI-06-08-00221.
11. Stockwell DC, Kirkendall E, Muething SE, Kloppenborg E, Vinodrao H, Jacobs BR. Automated adverse event detection collaborative: electronic adverse event identification, classification, and corrective actions across academic pediatric institutions. *J Patient Saf.* 2013 Dec;9(4):203–10.
12. Ferranti J, Horvath MM, Cozart H, Whitehurst J, Eckstrand J, Pietrobon R, et al. A Multifaceted Approach to Safety: The Synergistic Detection of Adverse Drug Events in Adult Inpatients. *J Patient Saf.* 2008 Sep;4(3):184–90.
13. Naessens JM, O’Byrne TJ, Johnson MG, Vansuch MB, McGlone CM, Huddleston JM. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care J Int Soc Qual Health Care ISQua.* 2010 Aug;22(4):266–74.
14. Resar RK, Rozich JD, Simmonds T, Haraden CR. A trigger tool to identify adverse events in the intensive care unit. *Jt Comm J Qual Patient Saf Jt Comm Resour.* 2006 Oct;32(10):585–90.
15. Lander L, Roberson DW, Plummer KM, Forbes PW, Healy GB, Shah RK. A trigger tool fails to identify serious errors and adverse events in pediatric otolaryngology. *Otolaryngol–Head Neck Surg Off J Am Acad Otolaryngol–Head Neck Surg.* 2010 Oct;143(4):480–6.

16. Sharek PJ, Horbar JD, Mason W, Bisarya H, Thurm CW, Suresh G, et al. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics*. 2006 Oct;118(4):1332–40.
17. Unbeck M, Schildmeijer K, Henriksson P, Jürgensen U, Muren O, Nilsson L, et al. Is detection of adverse events affected by record review methodology? an evaluation of the “Harvard Medical Practice Study” method and the “Global Trigger Tool.” *Patient Saf Surg*. 2013;7(1):10.
18. Kaiser AM, de Jong E, Evelein-Brugman SF, Peppink JM, Vandenbroucke-Grauls CM, Girbes AR. Development of trigger-based semi-automated surveillance of ventilator-associated pneumonia and central line-associated bloodstream infections in a Dutch intensive care. *Ann Intensive Care*. 2014 Dec 21;4:40.
19. Jha AK, Kuperman GJ, Teich JM, Leape L, Shea B, Rittenberg E, et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc JAMIA*. 1998 Jun;5(3):305–14.
20. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting Adverse Events Using Information Technology. *J Am Med Inform Assoc*. 2003 Mar 1;10(2):115–28.
21. Kirkendall ES, Spires WL, Mottes TA, Schaffzin JK, Barclay C, Goldstein SL. Development and Performance of Electronic Acute Kidney Injury Triggers to Identify Pediatric Patients at Risk for Nephrotoxic Medication-Associated Harm. *Appl Clin Inform*. 2014 Apr 2;5(2):313–33.
22. Calder L, Tierney S, Jiang Y, Gagné A, Gee A, Hobden E, et al. Patient safety analysis of the ED care of patients with heart failure and COPD exacerbations: a multicenter prospective cohort study. *Am J Emerg Med*. 2014 Jan;32(1):29–35.
23. Joice GA, Deibert CM, Kates M, Spencer BA, McKiernan JM. “Never events”: Centers for Medicare and Medicaid Services complications after radical cystectomy. *Urology*. 2013 Mar;81(3):527–32.
24. Rosen AK, Geraci JM, Ash AS, McNiff KJ, Moskowitz MA. Postoperative adverse events of common surgical procedures in the Medicare population. *Med Care*. 1992 Sep;30(9):753–65.
25. Rosen AK, Loveland S, Shin M, Shwartz M, Hanchate A, Chen Q, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Med Care*. 2013 Jan;51(1):37–44.
26. Sukumar S, Roghmann F, Trinh VQ, Sammon JD, Gervais M-K, Tan H-J, et al. National trends in hospital-acquired preventable adverse events after major cancer surgery in the USA. *BMJ Open*. 2013;3(6).
27. Kohn LT, Corrigan JM, Donaldson MS. *To Err Is Human: Building a Safer Health System*. Washington, DC: Institute of Medicine Committee on the Quality of Health Care in America; 1999.
28. Evans RS, Pestotnik SL, Classen DC, Bass SB, Burke JP. Prevention of adverse drug events through computerized surveillance. *Proc Annu Symp Comput Appl Sic Med Care Symp Comput Appl Med Care*. 1992;437–41.
29. Kaufman J, Rannie M, Kahn MG, Vitaska M, Wathen B, Peyton C, et al. An interdisciplinary initiative to reduce unplanned extubations in pediatric critical care units. *Pediatrics*. 2012 Jun;129(6):e1594-1600.
30. Piontek F, Kohli R, Conlon P, Ellis JJ, Jablonski J, Kini N. Effects of an adverse-drug-event alert system on cost and quality outcomes in community hospitals. *Am J Health-Syst Pharm AJHP Off J Am Soc Health-Syst Pharm*. 2010 Apr 15;67(8):613–20.
31. Cohen MM, Kimmel NL, Benage MK, Cox MJ, Sanders N, Spence D, et al. Medication safety program reduces adverse drug events in a community hospital. *Qual Saf Health Care*. 2005 Jun;14(3):169–74.
32. Garrett PR Jr, Sammer C, Nelson A, Paisley KA, Jones C, Shapiro E, et al. Developing and implementing a standardized process for global trigger tool application across a large health system. *Jt Comm J Qual Patient Saf Jt Comm Resour*. 2013 Jul;39(7):292–7.

[Response Ends]

1b.01. Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.

[Response Begins]

Patient safety is a core domain of healthcare quality and a major focus for quality improvement efforts.^{1,2} GAPPs is the first and only available global patient safety measure tailored for pediatric populations. By measuring preventable adverse event rates (i.e., harm) in inpatient pediatric populations, it provides important information to providers, hospital quality teams, and state health departments about outcomes of their patient care. Use of the measure will benefit patients, families, and providers because it enables stakeholders to identify and target areas of patient care that may benefit from quality improvement initiatives. Since GAPPs focuses on preventable adverse events, hospitals are able to assess and prioritize clinical areas with potential for immediate improvement.

The GAPPs Measure represents an opportunity to fill a notable void in safety measurement, specifically in pediatric preventable adverse events. National progress on quantifying and tracking AEs has stagnated because of the absence of an accepted national standard in all but a few defined areas. The GAPPs Measure can represent that national standard for pediatrics, and can be an important measure to evaluate and improve adverse event rates across different sites and time periods.

IMPORTANCE OF MEASURING ADVERSE EVENTS

Studies show that there is a high prevalence of medical errors and/or adverse events among the patient population. For example, in 1999, the Institute of Medicine (IOM) estimated that medical errors contribute up to 98,000 deaths and one million injuries each year.² In 2010, the Department of Health and Human Services' Office of the Inspector General estimated that 180,000 deaths due partly to adverse events occur among Medicare patients annually, making adverse events the third leading cause of death in the United States after heart disease and cancer.^{3,4} Hospitalized pediatric patients, who tend to have unique diseases and care distinct from adult patients, are also vulnerable to high adverse event rates: published studies report 11.1 adverse drug events per 100 inpatient pediatric patients, 74 adverse events per 100 neonatal intensive care unit (NICU) patients, and 203 adverse events per 100 pediatric intensive care unit (PICU) patients.⁵⁻⁹ Consequently, tracking adverse events in hospital settings is an important step towards understanding the current state of clinical care and creating initiatives aimed towards improving clinical quality.

Measuring preventable adverse event rates may also help hospitals better understand different aspects of their clinical quality. Some studies have found correlations between patient harm and other quality aspects such as performance on clinical processes of care and other health outcomes.¹⁰⁻¹⁴ For instance, studies suggest that patients who experience healthcare-related harms have greater odds of in-hospital and 30-day mortality, as well as 30-day readmission.¹¹⁻¹⁴ As such, measuring preventable adverse event rates is an essential first step for hospitals to understand and improve their patient care.

IDENTIFICATION OF INPATIENT ADVERSE EVENTS

Various approaches exist for identifying adverse events. Voluntary passive reporting systems are commonly employed but recognized to have low sensitivity.¹⁵ A more reliable, sensitive methodology for capturing data on the safety of hospital care is thus essential.^{16,17}

In 2003, AHRQ released its Patient Safety Indicators (PSI), developed in response to a congressional mandate to reduce medical errors.¹⁸ PSIs are intended to identify events that most likely resulted from preventable medical errors.¹⁸⁻²⁰ PSIs have been used with some success but have a number of limitations, in part due to their reliance on administrative data. They have also been found to have low sensitivity.^{17,18}

Use of trigger tools has been shown to be a faster, more sensitive, and more reliable method of adverse event detection than other approaches.^{5,9,17,21-23} "Triggers" are red flags in a medical record that may indicate the presence of an underlying adverse event and prompt further inspection to determine whether an adverse event occurred.^{21,22} An example trigger is the documented administration of an antidote-type medication (e.g.,

naloxone). Once a trigger is found, an in-depth review is undertaken to determine whether an adverse event occurred. In the case of naloxone, administration may indicate an adverse event occurred if the drug was given to counteract an overdose of opioids given in the hospital but may not if the overdose occurred due to voluntary recreational opioid use. Trigger tools detect adverse events in a high percentage of hospitalizations, ranging in published reports from 19% to 63%, and have evolved significantly over time.¹⁰⁻¹²

The Global Trigger Tool for Measuring Adverse Events (GTT), developed by the Institute for Healthcare Improvement (IHI), has become widely accepted as an effective approach for identifying adverse events in hospitalized adult patients.^{2,9,15,23-26} The GTT approach identifies 10 times more adverse events than AHRQ's PSIs and almost 100 times more events than voluntary reporting.^{1,17} However, the GTT has an exclusion of patients under age 18 so does not work for a pediatric population.

PEDIATRIC INPATIENT PATIENT SAFETY: LACK OF STANDARDIZED QUALITY MEASUREMENT

Although one study determined that a version of the GTT applied to the pediatric population could identify pediatric adverse events, the authors and other experts called for development of a standardized pediatric tool that focuses specifically on the problems of hospitalized children and that encompasses the breadth of inpatient pediatric care.^{12,18,25,29} The absence of a comprehensive pediatric trigger tool is a recognized limitation in quantifying the full scope of pediatric adverse events. An early effort to develop a pediatric-focused trigger tool led to the development of the Canadian Pediatric Trigger Tool.^{30,31}

We developed GAPPS to meet the need for a comprehensive, sensitive measure of pediatric patient safety. Our focus was on developing a global trigger tool for pediatric patients that could be more reliably applied across different hospital sites, both academic and community, than previous efforts. In addition, we sought to further refine the list of triggers to make a more robust global trigger list. We used methods similar to those used for GTT, including review of published tools and manual medical record review by experts in patient safety, which has been demonstrated to be a crucial component of developing patient safety measures.³⁰ We also utilized the RAND/UCLA Appropriateness Method, a 16-center field study, and post-analysis refinement of the trigger list to ensure GAPPS includes a more comprehensive trigger list than previous trigger tools.

GAPPS offers an enhancement in trigger tool methodology in that, unlike GTT, it requires that reviewers assess preventability. In the five years since IHI released the second edition of GTT, patient safety experts and national fiscal and quality improvement policies have increasingly focused on addressing preventable adverse events. The GAPPS measure uses the same approach to rate preventability as the North Carolina Patient Safety study, which was found to ascertain preventability with a high degree of reliability.³¹

DISPARITIES IN RISK OF HARM

Children with special healthcare needs experience elevated rates of medical errors.³² Among hospitalized pediatric patients, those with chronic conditions are at significantly higher risk for medical errors than those without chronic conditions.³²

What is known about racial/ethnic disparities in patient safety, particularly among children, is limited.³³ Black and Hispanic newborns are at higher risk of birth trauma.¹⁸ In addition, extrapolations from associations between race/ethnicity and known risk factors for harm suggest that Black and Hispanic children are likely at greater risk of harm than White children. Because severity and complexity of illness increase the risk of errors, and Black and Hispanic children are at higher risk for more complex conditions, these children are at greater risk for adverse events.³³ Emergency room visits are also associated with increased risk of adverse events, and Black and Hispanic children are known to visit the emergency department more frequently.^{28,33}

English language proficiency is also associated with adverse event prevalence. Hospitalized children of parents with limited English proficiency were shown to be significantly more likely to experience harm due to medical care.³⁶

COSTS OF GAPPS UPTAKE

As is the case with all new quality measures, we recognize that there may be concerns about the uptake of GAPPS due to hospital resource constraints. AE monitoring, like patient experience surveying, is not possible using administrative data and therefore requires additional resources. As previously discussed, intensive measurement of preventable AEs is one of the most important areas of quality measurement, one that has to this point been inadequately addressed in pediatric patient populations. The capacity of the GAPPS Measure to spur nationwide

improvement in pediatric patient safety represents a significant potential return on investment, one that outweighs implementation resource concerns.

It is important to note that some of the resources and infrastructure to successfully implement the GAPPs measure are already in place, as states commit to systems of mandatory reporting of certain adverse events. About half of states currently have mandatory reporting systems in place, with many others reporting AEs on a voluntary basis.³⁴ GAPPs will require an augmentation of processes hospitals are already initiating to report AEs, but will yield far greater returns. Far from representing a duplicative burden, the GAPPs Measure can take advantage of recent prioritization of patient safety and the resources hospitals are putting in place to evaluate quality in this domain. The marginal resources used for GAPPs implementation provide a drastically more robust assessment of hospital safety than currently in place, as global trigger tools have been shown to capture up to ten times more AEs than alternative AE measurement methods.^{17,35}

REFERENCES

1. Levinson DR. Adverse Events in Hospitals: Methods for Identifying Events. Washington, DC: Office of the Inspector General, Department of Health and Human Services; 2010. Report No.: OEI-06-08-00221.
2. Kohn LT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System. Washington, DC: Institute of Medicine Committee on the Quality of Health Care in America; 1999.
3. James JT. A new, evidence-based estimate of patient harms associated with hospital care. *J Patient Saf.* 2013 Sep;9(3):122–8.
4. Allen M, ProPublica. How Many Die From Medical Mistakes In U.S. Hospitals? [Internet]. NPR.org. [cited 2014 Jun 14]. Available from: <http://www.npr.org/blogs/health/2013/09/20/224507654/how-many-die-from-medical-mistakes-in-u-s-hospitals>
5. Sharek PJ, Classen D. The incidence of adverse events and medical error in pediatrics. *Pediatr Clin North Am.* 2006 Dec;53(6):1067–77.
6. Pediatric Patient Safety Measurement Tool (PPSMT) Pilot Toolkit. Children’s Hospital Association; 2012.
7. Stockwell DC, Slonim AD. Quality and safety in the intensive care unit. *J Intensive Care Med.* 2006 Aug;21(4):199–210.
8. Bates DW, Spell N, Cullen DJ, Burdick E, Laird N, Petersen LA, et al. The costs of adverse drug events in hospitalized patients. Adverse Drug Events Prevention Study Group. *JAMA J Am Med Assoc.* 1997 Jan 22;277(4):307–11.
9. Takata GS, Mason W, Taketomo C, Logsdon T, Sharek PJ. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children’s hospitals. *Pediatrics.* 2008 Apr;121(4):e927–935.
10. Calder L, Tierney S, Jiang Y, Gagné A, Gee A, Hobden E, et al. Patient safety analysis of the ED care of patients with heart failure and COPD exacerbations: a multicenter prospective cohort study. *Am J Emerg Med.* 2014 Jan;32(1):29–35.
11. Joice GA, Deibert CM, Kates M, Spencer BA, McKiernan JM. “Never events”: Centers for Medicare and Medicaid Services complications after radical cystectomy. *Urology.* 2013 Mar;81(3):527–32.
12. Rosen AK, Geraci JM, Ash AS, McNiff KJ, Moskowitz MA. Postoperative adverse events of common surgical procedures in the Medicare population. *Med Care.* 1992 Sep;30(9):753–65.
13. Rosen AK, Loveland S, Shin M, Shwartz M, Hanchate A, Chen Q, et al. Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: the case of readmissions. *Med Care.* 2013 Jan;51(1):37–44.
14. Sukumar S, Roghmann F, Trinh VQ, Sammon JD, Gervais M-K, Tan H-J, et al. National trends in hospital-acquired preventable adverse events after major cancer surgery in the USA. *BMJ Open.* 2013;3(6).
15. Sharek PJ, Horbar JD, Mason W, Bisarya H, Thurm CW, Suresh G, et al. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics.* 2006 Oct;118(4):1332–40.

16. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med*. 2010 Nov 25;363(22):2124–34.
17. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. “Global Trigger Tool” Shows That Adverse Events In Hospitals May Be Ten Times Greater Than Previously Measured. *Health Aff (Millwood)*. 2011 Apr 1;30(4):581–9.
18. Miller MR, Elixhauser A, Zhan C. Patient Safety Events During Pediatric Hospitalizations. *Pediatrics*. 2003 Jun 1;111(6):1358–66.
19. Dynan L, Goudie A, Smith RB, Fairbrother G, Simpson LA. Differences in Quality of Care Among Non–Safety-Net, Safety-Net, and Children’s Hospitals. *Pediatrics*. 2013 Feb 1;131(2):304–11.
20. Miller MR, Zhan C. Pediatric Patient Safety in Hospitals: A National Picture in 2000. *Pediatrics*. 2004 Jun 1;113(6):1741–6.
21. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care*. 2003 Jun 1;12(3):194–200.
22. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care*. 2003 Dec 1;12(suppl 2):ii39–ii45.
23. Agarwal S, Classen D, Larsen G, Tofil NM, Hayes LW, Sullivan JE, et al. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc*. 2010 Sep;11(5):568–78.
24. Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care*. 2000 Mar;38(3):261–71.
25. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ Can Med Assoc J J Assoc Medicale Can*. 2004 May 25;170(11):1678–86.
26. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. *Med J Aust*. 1995 Nov 6;163(9):458–71.
27. Kirkendall ES, Kloppenborg E, Papp J, White D, Frese C, Hacker D, et al. Measuring adverse events and levels of harm in pediatric inpatients with the Global Trigger Tool. *Pediatrics*. 2012 Nov;130(5):e1206–1214.
28. Matlow AG, Baker GR, Flintoft V, Cochrane D, Coffey M, Cohen E, et al. Adverse events among children in Canadian hospitals: the Canadian Paediatric Adverse Events Study. *CMAJ Can Med Assoc J J Assoc Medicale Can*. 2012 Sep 18;184(13):E709–718.
29. Stockwell D, Bisarya H, Classen D, Kirkendall E, Landrigan C, Lemon V, et al. A trigger tool to detect harm in pediatric inpatient settings. *Pediatrics*. 2015;
30. Scanlon MC, Harris JM, Levy F, Sedman A. Evaluation of the Agency for Healthcare Research and Quality Pediatric Quality Indicators. *Pediatrics*. 2008 Jun 1;121(6):e1723–31.
31. Sharek PJ, Parry G, Goldmann D, Bones K, Hackbarth A, Resar R, et al. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Serv Res*. 2011 Apr;46(2):654–78.
32. Ahuja N, Zhao W, Xiang H. Medical Errors in US Pediatric Inpatients With Chronic Conditions. *Pediatrics*. 2012 Oct 1;130(4):e786–93.
33. Flores G, Ngui E. Racial/Ethnic Disparities and Patient Safety. *Pediatr Clin North Am*. 2006 Dec;53(6):1197–215.
34. QuPS.org - Medical Errors and Patient Safety - US States Comparison - Overview - [Internet]. [cited 2016 Nov 16]. Available from: http://www.qups.org/med_errors.php?c=internal&id=195
35. Office of the Inspector General. Adverse Events In Hospitals: National Incidence Among Medicare Beneficiaries. Department of Health and Human Services;; 2010.

36. Khan A, Yin HS, Brach C, et al. Association Between Parent Comfort With English and Adverse Events Among Hospitalized Children. JAMA Pediatr. 2020;174(12):e203215.
doi:10.1001/jamapediatrics.2020.3215

[Response Ends]

1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and 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 (4b) under Usability and Use.

[Response Begins]

We conducted a National Field Test at 16 hospitals from across the United States that represented diverse geographic regions (four hospitals in each of the four US census regions: Northeast, South, West, and Midwest) and included eight teaching and eight non-teaching hospitals (teaching status was based on categorization set by the American Hospital Association). 3,790 hospitalizations occurring between 2007 and 2012 were included in this analysis.

Current rate of preventable AEs. Of the 414 AEs identified, 210 (50.7%) AEs were preventable, representing 9.5 preventable AEs [CI 8.2-10.8]/1,000 patient days. Compared to community hospitals, academic hospitals had higher preventable harm rates (13.1 [CI 11.4-15.2] vs. 2.4 [CI 1.5-3.8] AEs/1,000 patient days, $p < 0.001$). GAPPs is a measure of preventable adverse events. Because the identified adverse events are preventable, the ideal would be to have no preventable AEs. The GAPPs Measure specifically focuses on preventable adverse events because it outlines areas for immediate improvement. As we will discuss in more detail later in the application, the GAPPs Measure indicates there is significant room for improvement across most institutions and patient demographics. In addition to widespread incidence of preventable AEs, there is dramatic variation across institutions (preventable AEs/1000 patient days ranged from 0-20.4) and significant subpopulation disparities for race, medical complexity, and insurance type. In sum, GAPPs demonstrates areas to target to reduce disparities for subpopulations that may be at higher risk of encountering a preventable AE.

Changes in the rate of preventable AEs over time. Multivariate analyses controlling for demographic characteristics and chronic conditions showed no significant changes in preventable AE rates over time. Poisson regression accounting for hospital-level clustering and changes over time found no significant changes over time in preventable AEs (risk factor=1.00/1,000 patient days [CI 0.98-1.02]). When stratified by hospital type, neither academic nor community hospitals experienced significant temporal trends in preventable AEs/1,000 patient days.

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, 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. Include citations.

[Response Begins]

N/A

[Response Ends]

1b.04. 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.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. 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 (4b) under Usability and Use.

[Response Begins]

RACE/ETHNICITY

We assessed differences in pediatric patient safety associated with race/ethnicity by evaluating whether the rate of preventable adverse events (AEs) identified by reviewers varied among racial/ethnic groups. Race/ethnicity was recorded in our National Field Test using the categories Alaska Native, American Indian, Asian, Black, Hispanic, Native Hawaiian or other Pacific Islander, Other, and White. For our analysis, we combined Alaska Native, American Indian, Asian, Native Hawaiian or other Pacific Islander, and non-White, non-Hispanic Other patients into a single “Other” category because each of the categories was represented by a very small number of hospitalizations.

Table 1 – Preventable AEs per 1,000 patient days by race/ethnicity (n = 3,231)

Race/Ethnicity	n (%)	prev AE rate	P-Value
White	2,152 (56.8%)	8.9	Reference
Black	476 (12.6%)	5.9	0.13
Hispanic	419 (11.1%)	15.9	0.002
Other	184 (4.9%)	11.1	0.47

Table displaying the rate of preventable AEs per 1,000 patient days by race/ethnicity

Across all sites evaluated, we found that Hispanic patients had a higher unadjusted preventable AE rate at 15.9, compared to White patients at 8.9 (p=0.002).

PATIENTS WITH CHRONIC CONDITIONS

We assessed differences in pediatric patient safety associated with presence of chronic conditions by evaluating whether the rate of preventable AEs identified by reviewers varied based on the chronic conditions present among patients (as classified in AHRQ’s CCI system), controlling for length of hospitalization.¹

Table 2–Preventable AEs per 1,000 patient days by numbers of Chronic Condition Indicators per patient (n= 3,524)

Chronic Condition Indicators	n (%)	prev AE rate	P-Value
0 body system	1,990 (52.5%)	6.5	Reference
1 body system	1,085 (28.6%)	9.5	0.04
2 body systems	321 (8.5%)	17.9	<0.001
3 or more body systems	128 (3.4%)	19.8	<0.001

Table displaying the rate of preventable AEs per 1,000 patient days by numbers of Chronic Condition Indicators per patient

Overall, we found that patients with a body system affected by a chronic condition had higher unadjusted preventable AE rates than those without any body system affected by a chronic condition. Particularly, patients with 3 or more body systems affected by a chronic condition had the highest unadjusted preventable AE rate at 19.8 (p<0.001) as shown in Table 2.

INSURANCE STATUS

We assessed differences in pediatric patient safety associated with socioeconomic status (SES) by using insurance status as a proxy for SES and examining whether the rate of preventable AEs identified by reviewers varied with

insurance status. Insurance status was captured in our National Field Test using six non-mutually exclusive categories: Medicaid, Medicare, Private Insurance, Self-Pay, No Insurance, and Not Recorded. These sorted the cohort into eight unique categories (some of which indicate that a patient had multiple insurance types listed during the hospitalization included in our field test): no insurance; private insurance; public insurance; private insurance and self-pay; public and private insurance; public and no insurance; public, self-pay and private insurance; and insurance not recorded.

We chose to exclude hospitalizations for patients covered by Medicare from the analysis because pediatric eligibility for Medicare is based on having specific medical conditions rather than being based solely on family income.² Therefore, we evaluated patients with public insurance (Medicaid), private insurance, and no insurance. Patients who were recorded to have both private insurance and public insurance were categorized as patients with private insurance.

Table 3 shows the distribution of the insurance types included in our analysis. 37.5% of the patients had public insurance, 59.5% had private insurance, and 3.0% did not have insurance. 8.5% were missing insurance information and not included in this analysis. Patients with private insurance (preventable AE rate 8.5) and no insurance (preventable AE rate 3.9) had lower crude preventable AE rates across sites than patients with public insurance (AE rate 12.1).

Table 3 – Preventable AEs per 1,000 patient days by insurance type (n = 3,468)

Insurance	n (%)	prev AE rate	P-Value
Public	1,300 (37.5%)	12.1	Reference
Private	2,064 (59.5%)	8.5	0.02
No Insurance	104 (3.0%)	3.9	0.11

Table displaying the rate of preventable AEs per 1,000 patient days by insurance type

REFERENCES

1. Chronic Condition Indicator (CCI) for ICD-10-CM [Internet]. Healthcare Cost and Utilization Project (HCUP); Available from: https://www.hcup-us.ahrq.gov/toolssoftware/chronic_icd10/chronic_icd10.jsp
2. Medicare C for, Baltimore MS 7500 SB, Usa M. Medicare [Internet]. 2014 [cited 2014 Sep 2]. Available from: <http://www.cms.gov/Medicare/Medicare.html>

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, 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 above.

[Response Begins]

N/A

[Response Ends]

2. 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.

spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.

[Response Begins]

No

[Response Ends]

spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

For example, specifications may have been updated based on suggestions from a previous NQF CDP review.

[Response Begins]

N/A

[Response Ends]

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see [What Good Looks Like](#)).

[Response Begins]

GAPPs: Rate of preventable adverse events per 1,000 patient-days among pediatric inpatients

[Response Ends]

sp.02. Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]

GAPPs is a measure of the number of preventable adverse events per 1,000 patient-days among pediatric inpatients. It is designed to compare rates across institutions and over time. The GAPPs measure utilizes the GAPPs trigger tool to identify adverse events.

[Response Ends]

sp.04. Check all the clinical condition/topic areas that apply to your measure, below.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Surgery: General*

[Response Begins]

Other (specify)

[Other (specify) Please Explain]

Include except for patients admitted for behavioral health and acute events requiring medical intervention.

[Response Ends]

sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.

[Response Begins]

Safety

Safety: Complications

Safety: Healthcare Associated Infections

Safety: Medication

Safety: Overuse

[Response Ends]

sp.06. Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Children (Age < 18)

[Response Ends]

sp.07. Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Facility

[Response Ends]

sp.08. Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.

[Response Begins]

Inpatient/Hospital

[Response Ends]

sp.09. 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. If no URL is available, indicate "none available".

[Response Begins]

<https://www.childrenshospital.org/research/centers/center-excellence-pediatric-quality-measurement-cepqm-research/cepqm-measures/global-tool-patient-safety>

[Response Ends]

sp.12. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, [contact staff](#). Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

[Response Begins]

Available in attached Excel or csv file

[Response Ends]

Attachment: 3136_3136_S.2b_Data_Dictionary_Code_Table_Manual_-_Automated_Trigger_Lists-Dev rev-508.xlsx

For the question below: state the outcome being measured. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.13. State the numerator.

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.

[Response Begins]

The number of preventable adverse events found in a patient sample.

[Response Ends]

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.14. Provide details needed to calculate the numerator.

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 sp.11.

[Response Begins]

Adverse events are defined as “unintended physical injuries resulting from or contributed to by medical care that require additional monitoring, treatments, or hospitalizations, or that result in death.”(1,2) This matches the Institute for Healthcare Improvement’s adult Global Trigger Tool’s (IHI GTT’s) definition of harm since “harm” and “adverse event” are used synonymously in the context of patient safety.(1) GAPPs includes assessments of preventability to facilitate the identification of clinical areas with potential for immediate improvement.

The GAPPs measure requires two physicians to review and independently rate the preventability of each adverse event case they review. When physicians disagree on an event’s preventability, they discuss the rationale for their ratings with one another until both agree on whether an adverse event is preventable or not. A third physician is consulted in the rare occasion that the two physicians continue to disagree on an event’s preventability after discussing with one another.

REFERENCES

1. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). Institute for Healthcare Improvement; 2009. (IHI Innovation Series white paper).
2. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. “Global Trigger Tool” Shows That Adverse Events In Hospitals May Be Ten Times Greater Than Previously Measured. Health Aff (Millwood). 2011 Apr 1;30(4):581–9.

Below is a list of example triggers from the GAPPs Measure that are often found by reviewers in various sections of the medical record. For a full list of GAPPs triggers and a description of each, see appendix A.1.

Discharge summary

- All inpatient deaths
- Mechanical ventilation >48 hours
- Hospital readmission within 30 days
- Return to surgery

Laboratory reports

- Valproic acid >170 mcg/ml
- Carbamazepine >20 mcg/ml
- Serum creatinine doubling
- Nephrotoxin use (e.g., aminoglycosides, cyclosporine, tacrolimus, vancomycin) and rising creatinine (Cr)

- Hepatotoxic medications and elevated liver enzymes (AST, ALT)
- Drop of hemoglobin (Hgb) or hematocrit (Hct) of >25% in less than 24 hours

Radiology results

- Patient fall

Physician orders

- Abrupt medication stop
- Transfer to higher level of care

Medication administration records (MARs)

- Vitamin K administration after warfarin
- Naloxone administration
- Hypoglycemia (<2 mmol/L or 40 mg/dL)

Nursing flow sheets

- Surgical site infection
- Infiltration/phlebitis documentation
- Embolus/thrombus documentation
- Pressure ulcer documentation (= stage 2)

Procedure notes (diagnostic, surgical)

- Any code or arrest, or rapid response team activation
- Mechanical ventilation greater than 48 hours post-operative

Nursing/Physician/Multi-disciplinary progress notes

- Opiate-related constipation with intermittent laxative use
- Healthcare-associated infections: positive C. difficile test
- Healthcare-associated infections: positive blood culture (only after 48 hours from admission)
- Healthcare-associated infections: positive urine culture (only after 48 hours from admission)
- Healthcare-associated infections: positive respiratory or GI viral test (only after 48 hours from admission)
- Racemic epinephrine administration (patients mechanically ventilated within the last 24 hours)

[Response Ends]

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.15. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

The denominator is 1,000 patient-days for all sampled pediatric patients who meet inclusion, but not exclusion, criteria.

[Response Ends]

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.16. Provide details needed to calculate the denominator.

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 sp.11.

[Response Begins]

The denominator includes all patients who meet the following criteria:

1. Patients <18 years of age at admission;
2. Patients with length of stay (LOS) greater than or equal to 24 hours;
3. Patients admitted for acute care. Acute care does not include patients discharged from the Emergency Department without admission to the hospital; or patients in rehabilitation and residential units, non-acute inpatient psychiatric units, newborn nurseries, and day treatment areas. If a patient is initially admitted acutely but subsequently transferred to inpatient psychiatric care, the acute portion of the hospitalization should be included; and
4. Patients who were discharged from, who were transferred out of, or who died during the inpatient or observation hospital stay.

[Response Ends]

sp.17. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

We exclude patients who meet the above inclusion criteria but fall into the following exclusion categories:

- Patients discharged from the Emergency Department without admission to the hospital;
- Patients in newborn nurseries;
- Patients over age 18

[Response Ends]

sp.18. Provide details needed to calculate the denominator exclusions.

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 sp.11.

[Response Begins]

Stratification is not required within institutions. However, if desired, quality improvement teams may choose to stratify preventable adverse event rates. Variables commonly used to stratify outcome measures include service (e.g., medical versus surgical), department (e.g., cardiology, neurology, etc.), and patient safety focus area (e.g., healthcare-associated infections).

For comparisons between institutions, preventable adverse event rates should be stratified by teaching versus community hospitals due to differences in types (e.g., complexity) of patient populations.

[Response Ends]

sp.19. Provide all information required to stratify the measure results, if necessary.

Include 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 in the Data Dictionary field.

[Response Begins]

Stratification is not required within institutions. However, if desired, quality improvement teams may choose to stratify preventable adverse event rates. Variables commonly used to stratify outcome measures include service (e.g., medical versus surgical), department (e.g., cardiology, neurology, etc.), and patient safety focus area (e.g., healthcare-associated infections).

For comparisons between institutions, preventable adverse event rates should be stratified by teaching versus community hospitals due to differences in types (e.g., complexity) of patient populations

[Response Ends]

sp.20. Is this measure adjusted for socioeconomic status (SES)?

[Response Begins]

No

[Response Ends]

sp.21. Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.

[Response Begins]

Statistical risk model

[Response Ends]

sp.22. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

sp.23. Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score

[Response Begins]

Better quality = Lower score

[Response Ends]

sp.24. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

[Response Begins]

GAPPs allows quality improvement teams to measure preventable adverse event rates over time among pediatric inpatients. GAPPs can be applied within entire hospitals, individual divisions or services, or specific programs. The original candidate trigger list (n=54 triggers) was developed through literature searches and expert panel determination. After the national field test, we selected the final manual triggers (n=27 triggers) based on incidence and positivity rates (i.e., the frequency with which a trigger identifies an AE). To form our automated trigger list, we compiled all of the manual triggers that could be automated in an academic tertiary care hospital's EHR system as well as candidate triggers that had a low frequency in the national field test that could feasibly be automated and had a positivity rate $\geq 10\%$ when further tested at the academic tertiary care hospital (n=30 triggers), and recommended inclusion of all manual triggers in a final automated trigger list. As compared with our final manual list (n=27 triggers), the final automated list added triggers that are relatively rare, but when present have a high positivity rate for identifying AEs (there is a lower bar for including triggers in the automated tool because it does not involve manual effort).

The main advantages of using the automated, rather than manual, GAPPs approach are speed (it eliminates the need to find triggers manually in medical records and allows primary reviewers to avoid looking at non-flagged records) and consistency of trigger detection (it reduces human error during review). Whether an institution uses the manual or automated trigger list, the implementation of the measure to identify adverse events is the same. For more detailed instructions on how to find preventable adverse events using either GAPPs' manual or automated approach, refer to Appendix A.

Step 1 – Assemble a review team

The GAPPs review team should consist of:

- Two primary reviewers who are responsible for reviewing and identifying adverse events in medical records. The second primary reviewer will only review a subset of the first primary reviewer's charts for a reliability check. It is recommended that each primary reviewer have extensive clinical experience, have familiarity with multiple clinical settings and interventions (including diagnostic tests, medications, and procedures), and be well-acquainted with the hospital's medical record system and typical delivery of care. The primary reviewer in trigger tool applications has historically been a nurse, but physicians, physician assistants, and pharmacists – among others – may also be good candidates.
- Two secondary reviewers who are responsible for reviewing any suspected adverse event identified by a primary reviewer. The secondary reviewers verify the occurrence of adverse events, as well as the ratings of severity and preventability for the events. They do not review medical records directly; instead, they listen to the primary reviewer's description of the adverse events he or she identified and ask questions as needed for clarification. Some secondary reviewers may choose to read the primary reviewer's written assessment in addition to listening to the reviewer's description of the hospitalization. Secondary reviewers should be physicians.

Step 2 – Select relevant hospitalizations

We recommend that the main primary reviewer selects a random sample of at least 20 inpatient hospitalizations each month from a list of all inpatient hospitalizations with discharge dates that fall within the month being reviewed; the hospitalizations may be drawn from an entire hospital or from a specific division, service, or program. The hospitalizations should meet eligibility criteria (noted below) for a minimum of 60 hospitalizations per quarter. For institutions with high pediatric patient volume, records for 60 unique patients typically will be reviewed. However, patients who have multiple discharges that fall within a given quarter may have their records reviewed multiple times.

A two-stage process is used to determine which pediatric medical records should be included in the GAPPs sample frame. The first stage determines whether patients meet the inclusion criteria listed below. For patients who meet inclusion criteria, certain exclusion criteria – also described below, are then applied.

Inclusion Criteria:

GAPPs is intended for broadly reviewing the medical records of pediatric patients who meet the following criteria:

- Patients <18 years of age at admission;
- Patients with length of stay (LOS) ≥24 hours;
- Patients admitted for acute care. Acute care does not include patients discharged from the Emergency Department without admission to the hospital; or patients in rehabilitation and residential units, non-acute inpatient psychiatric units, newborn nurseries, and day treatment areas. If a patient is initially admitted acutely but subsequently transferred to inpatient psychiatric care, the acute portion of the hospitalization should be included; and
- Patients who were discharged from, who were transferred out of, or who died during the inpatient or observation hospital stay.

Exclusion Criteria:

Patients with inpatient LOS <24 hours are excluded because patients with brief hospital stays are less likely to have received the amount of medical intervention necessary to evaluate the quality of care.

Patients ≥18 years of age at admission are excluded because the Center of Excellence for Pediatric Quality Measurement's (CEPQM) task was to create a tool for measuring patient safety in the pediatric age group (i.e., <18 years of age). With this in mind, GAPPs is designed to perform exclusively in pediatric patients.

Step 3 – Review of patient records by primary reviewers and secondary reviewers

Primary reviewers should spend up to 30 minutes reviewing each hospitalization in a medical record. They should focus on identifying and recording triggers and adverse events (for lists of the GAPPs manual and automated triggers, see Appendix A).

- Identifying triggers: When a trigger is discovered in the record (either manually or automatically via an electronic health record (EHR) system that flags hospitalizations), primary reviewers should look for information relevant to that trigger to investigate whether an adverse event occurred. Reviewers typically identify many more triggers than adverse events. If no adverse event is found, continue reviewing the remainder of the record for additional triggers. The manner in which the trigger is identified (manually or automatically) has no impact on the rest of the GAPPs measure process. The automated trigger list removes the arduous human identification factor from the process, but the measure remains exactly the same following trigger identification.
- Some adverse events will be found without the identification of a related trigger. These events should still be recorded in the Primary Review Forms and Suspected Adverse Event Forms.
- Identifying adverse events: Whether discovered due to a positive trigger or encountered while searching for triggers, adverse events and their corresponding information should be recorded by the primary reviewer. We recommend that reviewers consider the following items when determining whether an adverse event has occurred:
- Harm likely occurred through event(s) in which people experiencing the event would be unhappy the event occurred (e.g., IV infiltrate, even if minor).

- Adverse events are, by definition, the result of medical treatment. If an incident was part of the natural progression of a patient's disease process, it is unlikely to be an adverse event (e.g., patient admitted for respiratory failure due to pneumonia worsens despite appropriate management and consequently needs to be intubated), unless medical care somehow contributed to the incident.
- Incidents that are the intended results of medical care are not considered adverse events (e.g., neutropenia with chemotherapy).
- Psychological harm alone is not generally considered an adverse event (e.g., stress).

All identified adverse events should be recorded, regardless of location. The Primary Review Forms and Suspected Adverse Event Forms allow reviewers to specify where harms occurred, so harms occurring outside the hospital can be analyzed separately or removed from assessments of unit/hospital care quality as needed.

- Determining severity
- Severity: Reviewers should assign severity to an adverse event using the five-point severity scale below, which is a modified version of the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Index for Categorizing Errors. Since the categories are not mutually exclusive, reviewers should assign the highest severity category that applies to the adverse event. It is important to note that adverse events in high-severity categories do not have to meet all of the requirements of lower-harm-level categories. For example, an adverse event can be categorized in harm level H (i.e., insulin bolus) but not qualify as a G-level harm (i.e., permanent injury).

Category E: Temporary harm to the patient and required intervention

Category F: Temporary harm to the patient and required initial or prolonged hospitalization

Category G: Permanent patient harm

Category H: Intervention required to sustain life

Category I: Patient death

Step 4- Determine preventability of adverse event

Primary reviewers (nurses) record preventability for data collection and internal validity assessment purposes. However, the final determination of preventability is made by the secondary reviewers (physicians). All reviewers should rely on the category definitions provided below and their own clinical experience when determining preventability. Training sessions, discussions with the review team, and experience with reviews will be crucial in developing consistent preventability ratings.

Categories of Preventability [1]

Definitely not preventable: Events in which no obvious error occurred; necessary precautions were taken; no alteration in method or care exists to prevent the event.

Drug-associated rash (no prior exposure or history): A 9-year-old male with no known allergies presented to the emergency department for a sore throat, cough, and fever. When the patient was given ibuprofen for his fever, he developed hives and itching. The patient was then given diphenhydramine and responded well to the drug with no respiratory distress. Ibuprofen was discontinued and listed as an allergy on the patient's medical record.

Probably not preventable: Events that do not appear preventable but would require further investigation to assess certainty.

Procedural complications (with skilled proceduralist and no errors): Despite nursing standards being followed, a 7-year-old female developed an IV infiltrate.

Probably preventable: Events that appear preventable but would require further investigation to assess certainty.

Hospital-acquired infections: A male infant born at 35 weeks estimated gestation age had an umbilical catheter placed. An inflamed wound developed at the catheter site, and he was started on antibiotics. An abscess formed at the site over the next few days, so the wound was drained, and cultures were obtained that were positive for MRSA and Enterobacter spp.

Definitely preventable: Events where error was identified; necessary precautions were not taken; event was preventable by modification of behavior, technique, or care.

Medication overdose: A 13-year-old female was given an overdose of insulin during treatment for diabetic ketoacidosis. Her blood glucose dropped precipitously, and she required a D50 bolus.

Category ^[1]	Description	Specific Case Example
Definitely not preventable	Events in which no obvious error occurred; necessary precautions were taken; no alteration in method or care exists to prevent the event.	<i>Drug-associated rash (no prior exposure or history):</i> A 9-year-old male with no known allergies presented to the emergency department for a sore throat, cough, and fever. When the patient was given ibuprofen for his fever, he developed hives and itching. The patient was then given diphenhydramine and responded well to the drug with no respiratory distress. Ibuprofen was discontinued and listed as an allergy on the patient's medical record.
Probably not preventable	Events that do not appear preventable but would require further investigation to assess certainty.	<i>Procedural complications (with skilled proceduralist and no errors):</i> Despite nursing standards being followed, a 7-year-old female developed an IV infiltrate.
Probably preventable	Events that appear preventable but would require further investigation to assess certainty.	<i>Hospital-acquired infections:</i> A male infant born at 35 weeks estimated gestation age had an umbilical catheter placed. An inflamed wound developed at the catheter site, and he was started on antibiotics. An abscess formed at the site over the next few days, so the wound was drained, and cultures were obtained that were positive for MRSA and Enterobacter spp.
Definitely preventable	Events where error was identified; necessary precautions were not taken; event was preventable by modification of behavior, technique, or care.	<i>Medication overdose:</i> A 13-year-old female was given an overdose of insulin during treatment for diabetic ketoacidosis. Her blood glucose dropped precipitously, and she required a D50 bolus.

This table displays description and examples of adverse events by degrees of preventability.

[1] While secondary reviewers can select one of four preventability rankings for each adverse event, preventability rankings are categorized into two groups when assessing secondary reviewer agreement and during data analysis. Specifically, adverse events ranked as “definitely not preventable” and “probably not preventable” are considered “nonpreventable,” and adverse events ranked as “definitely preventable” and “probably preventable” are considered “preventable.”

Step 5 – Record data in appropriate forms

Primary reviewers

Primary reviewers should complete the Primary Review Form for each hospitalization. For each adverse event, they should also complete the Suspected Adverse Event Form.

Secondary reviewers

Secondary Reviewer A should complete the Secondary Review Form A for each suspected adverse event identified by a primary reviewer, either confirming or denying that an adverse event occurred. Secondary Reviewer B should

complete the Secondary Review Form B for each suspected adverse event identified by the primary reviewers, either confirming or denying that an adverse event occurred.

In cases in which Secondary Reviewers A and B disagree about whether an adverse event occurred or do not independently rate an adverse event with the same severity and preventability (note: preventability agreement is determined dichotomously, i.e., definitely/probably preventable vs. definitely/probably not preventable), the secondary reviewers must discuss the issues and reach consensus on all rankings. If the two secondary reviewers are unable to reach a consensus after discussing the case, a third physician should be consulted. Once reviewers agree on all rankings, one of the reviewers should complete the Consensus Form.

Step 6 – Check reliability

To assess the reliability with which institutions use GAPPs to identify triggers and adverse events, a second primary reviewer should perform a completely independent review of a random 10% sample of the medical records reviewed by the main primary reviewer from each sampling time frame (i.e., 6 records per quarter). This second review should occur at the end of each year on a total of 24 records annually. During this check, the second primary reviewer completes the same forms as the first primary reviewer: the Primary Review Form and, for each adverse event identified in a medical record, the Suspected Adverse Event Form. Knowing the rates at which primary reviewers identify and agree about adverse events will allow institutions to assess the reliability of their adverse event detection and to improve training efforts for reviewers as needed.

Step 7 – Analyze data

After the primary and secondary reviewers complete their reviews in each collection period, the data should be analyzed by computing preventable adverse events per 1,000 patient-days using the following equation: $[(\text{Total number of preventable adverse events identified in all the medical records in the sampling frame}) / (\text{Sum of the total number of inpatient days for all of the medical records reviewed in the sampling period})] * 1,000$. When comparing across institutions, the unit of time should be annual.

Case-mix adjustment for inter-hospital comparisons:

We recommend groups use mixed effects negative binomial regression to adjust preventable adverse event rates based on patient characteristics and type of service. Specifically, the outcome is the number of preventable adverse events for an admission (exposure time equal to length of stay), case-mix variables are fixed effects, and a hospital-level random intercept represents the variation between hospitals. Case-mix models should be stratified by hospital type (teaching vs. community). The case-mix data are obtained from the Primary Review Forms.

[Response Ends]

sp.27. If measure testing is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

Examples of samples used for testing:

- *Testing may be conducted on a sample of the accountable entities (e.g., hospital, physician). The analytic unit specified for the particular measure (e.g., physician, hospital, home health agency) determines the sampling strategy for scientific acceptability testing.*
- *The sample should represent the variety of entities whose performance will be measured. The [2010 Measure Testing Task Force](#) recognized that the samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.*
- *The sample should include adequate numbers of units of measurement and adequate numbers of patients to answer the specific reliability or validity question with the chosen statistical method.*
- *When possible, units of measurement and patients within units should be randomly selected.*

[Response Begins]

Hospitals and departments using GAPPs to measure preventable adverse event rates are responsible for generating complete, accurate, and valid lists of all pediatric inpatient hospitalizations (<18 years old) discharged between the first and last days of each month (e.g., for January, any qualifying discharges between and including the 1st and 31st days). The hospitalizations may be drawn from an entire hospital or from a specific division, service, or program.

We recommend that primary reviewers select a random sample of at least 20 inpatient hospitalizations each month from a list of all inpatient hospitalizations with discharge dates that fall within the month being reviewed. The hospitalizations should meet eligibility criteria (see Appendix A) for a minimum of 60 hospitalizations per quarter. We recommend a minimum sample size of 60 records per quarter in order for institutions to achieve adequate reliability for estimates of hospital-level preventable adverse event incidence. This sample size is based on the assumption that the trigger tool will be used in an improvement setting, for which the aim is to detect trends in the data showing meaningful change over time. According to Perla and colleagues, to plot the data quarterly, the appropriate sample size of medical records is given by $9/R$, where R is the average number of adverse events per person.⁽³⁾ Assuming an adverse event rate of at least 0.15, the recommended sample size computes to $9/0.15=60$.

For institutions with high pediatric patient volume, records for 60 unique patients will typically be reviewed. However, patients who have multiple discharges that fall within a given quarter may have their records reviewed multiple times.

Records should be selected through a random process to eliminate any potential bias. A variety of selection methods can be used to ensure a random sample, such as:

- For a given month, number all discharge records (including those ending with deaths) sequentially starting with one. Using random-number generating software, generate 25 numbers between one and the total number of records. Select the hospitalizations labeled with the random number. Keep the first 20 hospitalizations that meet eligibility criteria. If 25 records are insufficient to yield 20 eligible hospitalizations, select more hospitalizations using the random-number generating software until you obtain the 20 that are needed.

For medical facilities, departments, or programs that have fewer than 20 pediatric inpatients per month, review records for all hospitalizations during the month.

[Response Ends]

sp.30. Select only the data sources for which the measure is specified.

[Response Begins]

Electronic Health Records

Paper Medical Records

[Response Ends]

sp.31. Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.

[Response Begins]

Primary Review Form, Suspected Adverse Event Form, Secondary Review Form A Secondary Review Form B, Consensus Form

[Response Ends]

sp.32. Provide the data collection instrument.

[Response Begins]

Available in attached appendix in Question 1 of the Additional Section

[Response Ends]

Attachment: 3136_3136_AAA_MOO_-_Appendices_FINAL_20170201-508.pdf

2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

No

[Response Ends]

2ma.02. Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

No

[Response Ends]

2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

[Response Begins]

Yes

[Response Ends]

2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

[Response Begins]

No additional risk adjustment analysis included

[Response Ends]

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the [Submitting Standards webpage](#).
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the [2021 Measure Evaluation Criteria and Guidance](#).

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a. 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 and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;

AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the

information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration
- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses 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 non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions

Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v.

75 percent) is clinically meaningful; or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 v.\$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous (Year) Submission:

Testing from the previous submission here.

2a.01. Select only the data sources for which the measure is tested.

[Response Begins]

Electronic Health Records

Paper Medical Records

[Response Ends]

2a.02. 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).

[Response Begins]

N/A

[Response Ends]

2a.03. Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]

2007-2012

[Response Ends]

2a.04. Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- Clinician: Clinician

- *Population: Population*

[Response Begins]

Facility

[Response Ends]

2a.05. List the measured entities 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.

[Response Begins]

During the National Field Test, GAPPs was used at 16 hospital study sites. The hospitals were situated across the United States, representing diverse geographic regions (four hospitals in each of the four US census regions: Northeast, South, West, Midwest) and included eight teaching and eight non-teaching hospitals (teaching status was based on standards set by the American Hospital Association). Hospitals were identified to participate through the Pediatric Research in Inpatient Settings (PRIS) network. During the range of years in which we were sampling patient records (2007-2012) each hospital site was at a different stage in the evolution from paper to electronic records, with varying use of electronic health records.

Table 1 - Hospital Characteristics:

Hospital Characteristics	Hospitals (Total N = 16)
<i>Teaching Status</i>	*
Teaching	8
Non-Teaching	8
<i>Regions</i>	*
Northeast	4
Midwest	4
South	4
West	4
<i>Type</i>	*
Free-Standing	5
General Population	11

This table displays hospital characteristics for the 16 hospitals included in the GAPPs National Field Test.

*Intentionally left blank

[Response Ends]

2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

[Response Begins]

We reviewed 3,814 medical records for pediatric patients with discharge dates from January 1, 2007 to December 31, 2012 across the 16 hospital sites (≈240 records/hospital). Sampled patients who met the following criteria were included in the study:

- Patients <18 years of age at admission;
- Patients with length of stay (LOS) ≥24 hours;
- Patients admitted for acute care. Acute care does not include patients discharged from the Emergency Department without admission to the hospital; or patients in rehabilitation and residential units, non-acute inpatient psychiatric units, newborn nurseries, and day treatment areas. If a patient is initially admitted acutely but subsequently transferred to inpatient psychiatric care, the acute portion of the hospitalization should be included; and
- Patients who were discharged from, who were transferred out of, or who died during the inpatient or observation hospital stay.

Exclusion Criteria:

Patients with inpatient LOS <24 hours are excluded because patients with brief hospital stays are less likely to have received the amount of medical intervention necessary to evaluate the quality of care.

Patients ≥18 years of age at admission are excluded because the Center of Excellence for Pediatric Quality Measurement's (CEPQM) task was to create a tool for measuring patient safety in the pediatric age group (i.e., <18 years of age). With this in mind, GAPPs is designed to perform exclusively in pediatric patients.

Table 2 - National Field Test Patient Characteristics:

Characteristics	*	
Gender, No. (%)	*	*
Female	1,698	(44.5%)
Male	2,058	(54.0%)
*	*	*
Age, No. (%)	*	*
0 to <3 months	1,031	(27.0%)
3 months to < 3 years	917	(24.0%)
3 years to < 10 years	784	(20.6%)
10 years to < 18 years	1,052	(27.6%)
*	*	*
Race/Ethnicity, No. (%)	*	*
Black	477	(12.5%)
Hispanic	421	(11.0%)
White	2,165	(56.8%)
Other	184	(4.8%)
Missing	567	(14.9%)
*	*	*
Chronic Condition Indicators**, No. (%)	*	*

Characteristics	*	
0 body system	2,003	(52.5%)
1 body system	1,090	(28.6%)
2 body systems	324	(8.5%)
more than 2 body systems	129	(3.4%)
*	*	*
Insurance, No. (%)	*	*
Private	2,073	(54.4%)
Public	1,306	(34.2%)
Other	106	(2.8%)
*	*	*
Length of Stay in day, Median (IQR)	3	(2,5)

This table displays patient characteristics from the GAPPs National Field Test.

*Intentionally left blank

** Chronic Condition Indicators allow researchers to determine whether a diagnosis is a chronic condition based on ICD-9-CM codes, and if so, what specific body system(s) are affected by a chronic condition.

[Response Ends]

2a.07. 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.

[Response Begins]

N/A

[Response Ends]

2a.08. List 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.

[Response Begins]

Race/Ethnicity; Presence of chronic conditions as measured by AHRQ's Chronic Condition Indicator (CCI) (38); Insurance Status.

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter “see validity testing section of data elements”; and enter “N/A” for 2a.11 and 2a.12.

2a.09. Select the level of reliability testing conducted.

Choose one or both levels.

[Response Begins]

Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)

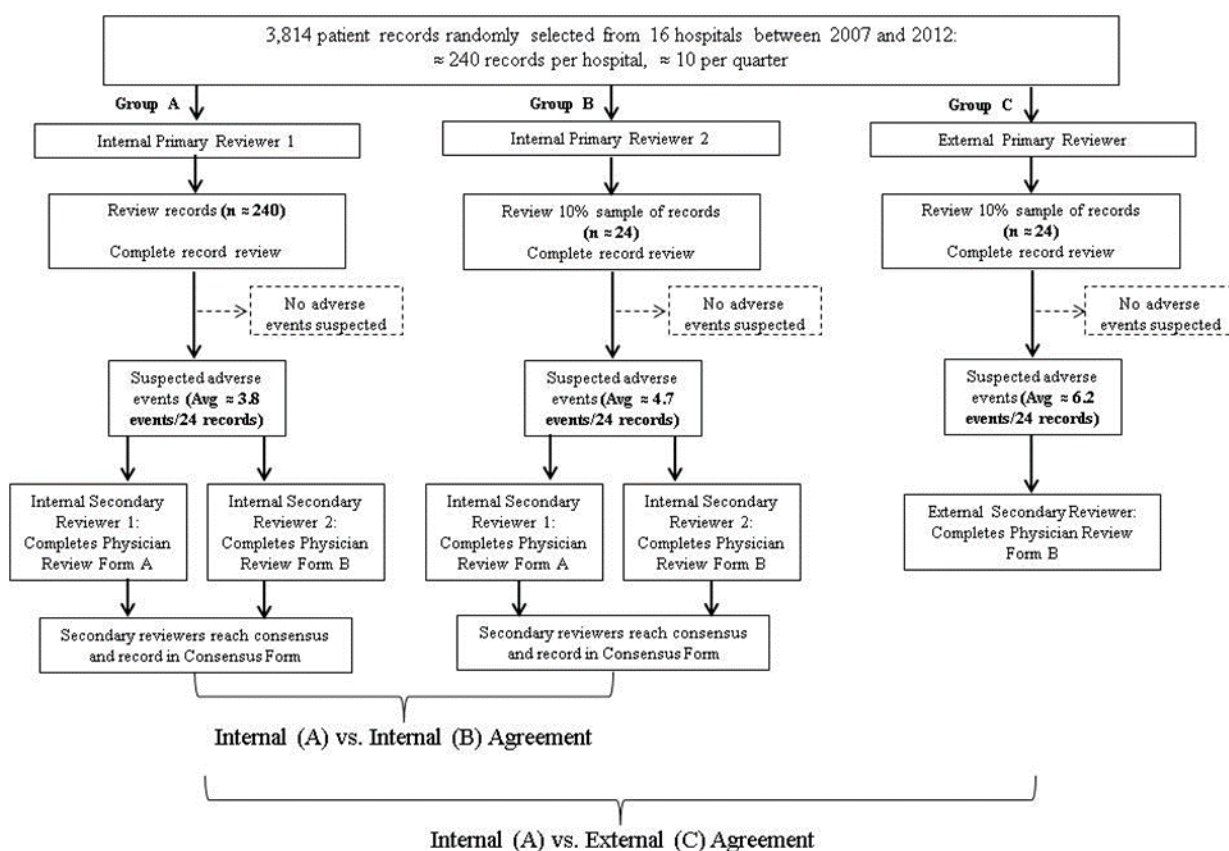
[Response Ends]

2a.10. For each level of reliability testing 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.

[Response Begins]

Figure 1 – Flowchart of the medical record review process in the GAPPs study:



The candidate triggers were evaluated using three groups of medical records in each hospital (see Figure 1). Group A consisted of the total sample of medical records from each hospital that were reviewed by primary (nurse) and secondary (physician) reviewers internal to the hospital. Group B consisted of a random sub-sample of the records in Group A (i.e., 24 records per site) that were reviewed again by additional primary and secondary reviewers internal to each hospital. Group C consisted of a separate random sub-sample (i.e., 24 records per site) of the records reviewed in Group A that were reviewed again by primary and secondary reviewers external to each hospital.

We used Groups A and B to evaluate the reliability of the measure. We compared ratings from pairs of independent secondary reviewers within Group A. We also compared primary reviewer findings (Group A) to a second primary reviewer's findings (Group B) for the same medical records. To assess reliability, we used a Kappa

statistic for variables with only two possible outcomes and a weighted Kappa computed with Fleiss-Cohen weights for variables with more than two possible ordinal outcomes.(1,2) We used the categorization of Landis and Koch to interpret reliability for ranges of Kappa scores (k <0: poor, k = 0.00-0.20: slight, k = 0.21-0.40: fair, k = 0.41-0.60: moderate, k = 0.61-0.80: substantial, k = 0.81-1.00: almost perfect).

REFERENCES

- Fleiss J, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas.* 1973;33:613–9.
- Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics.* 1977 Mar;33(1):159.

[Response Ends]

2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, [NQF Measure Evaluation Criteria](#)).

[Response Begins]

Table 3 - Internal Primary Reviewer (Nurse) Reliability:

Primary Review (Nurse)	Kappa	Agreement
# Triggers	0.68	83.1%
Any Trigger	0.67	87.9%
# Adverse Events	0.73	88.4%
Any Adverse Events	0.69	91.8%

This table displays internal primary reviewer reliability for nurses.

Table 4 - Internal Secondary Reviewer (Physician) Reliability:

Secondary review (Physician)	Kappa	Agreement
Verification of Same Adverse Events	0.81	91.6%
AE Severity	0.86	89.6%
AE Preventability	0.72	77.4%

This table displays internal primary reviewer reliability for physicians.

[Response Ends]

2a.12. Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

[Response Begins]

The reliability for internal primary reviewers (Group A versus Group B, [n = 379]) was “substantial” for both determination of the total number of suspected AEs (Kappa = 0.73, 95% CI 0.62 – 0.85) and identification of at least one suspected AE (Kappa = 0.69, 95% CI 0.59 – 0.79). Both primary reviewers agreed on the total number of AEs 88% of the time and agreed that a record did or did not contain at least one AE 92% of the time. In some cases, however, the AEs identified in the medical records differed. From the total sample of records reviewed, primary reviewers identified the same AEs 62% of the time.

The two internal secondary reviewers in Group A independently determined the presence or absence of an AE among suspected AEs identified by the primary reviewer in Group A (n = 617). Internal secondary reviewers verified the same suspected AEs 92% of the time, with “almost perfect” reliability (Kappa = 0.81, 95% CI 0.76 – 0.86).

The reliability for internal secondary reviewers in Group A versus Group B (n = 379) was “substantial” for both determination of the total number of suspected AEs (Kappa = 0.73, 95% CI 0.57 – 0.89) and verification of at least one suspected AE (Kappa = 0.70, 95% CI 0.59 – 0.81). Internal secondary reviewers in Group A and Group B agreed on the total number of AEs 92% of the time and agreed that a record did or did not contain at least one AE 94% of the time. In some cases, however, the AEs identified in the medical records differed.

As our team conducted reliability testing with patient-level data elements (AEs identified in individual medical records) using appropriate methodology (Kappa statistic for variables with only two possible outcomes and a weighted Kappa computed with Fleiss-Cohen weights for variables with more than two possible ordinal outcomes), and as there was substantial or moderate agreement between primary and secondary reviewers in both determination of the total number of suspected AEs and identification of at least one suspected AE, we deem that there is moderate confidence in the measure’s reliability as per the NQF Algorithm #2: Guidance for Evaluating Reliability.

[Response Ends]

2b.01. Select the level of validity testing that was conducted.

[Response Begins]

Patient or Encounter-Level (data element validity must address ALL critical data elements)

[Response Ends]

2b.02. 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.

[Response Begins]

We developed the draft trigger tool used in the GAPPs measure through the RAND/UCLA Appropriateness Method, which is a modified Delphi process.(1–3) We first compiled a set of 78 candidate triggers from a literature review of existing pediatric and adult trigger tools and input from trigger tool experts.(4–6) We then recruited nine panelists from national pediatric and patient safety organizations and asked them to rate separately the validity and feasibility of the candidate triggers on a nine-point scale (where 1 is the least valid/feasible and 9 is the most valid/feasible). A trigger was considered valid if it was judged to be reasonably likely to identify an underlying AE, indicating that harm potentially occurred. A trigger was considered feasible if it was judged likely to be accurately and consistently documented in either paper or electronic medical records as part of patient care at a wide range of hospitals, from smaller community sites to larger tertiary care centers. Applying the RAND/UCLA Appropriateness Method, we accepted triggers that had both median validity and feasibility ratings greater than or equal to seven. This approach resulted in inclusion of 54 of the initial 78 candidate triggers in the draft GAPPs trigger list.

It is not possible to assess the performance of the GAPPS measure against a true “gold standard” for detection of preventable AEs because such a gold standard does not yet exist. We therefore focused our validity testing on evaluation of how accurately and completely “typical reviewers” (i.e., clinicians who are trained in GAPPS methodology but not necessarily trigger tool experts) were able to identify preventable AEs using the measure as compared to expert reviewers. The expert reviewers had extensive experience with using trigger tools for preventable AE identification and consequently were most likely to identify preventable AEs accurately and completely. To evaluate the validity of the GAPPS measure, we assessed the performance of the National Field Test hospitals' internal reviewers relative to the performance of external expert reviewers in applying the measure (i.e., we compared findings of reviewers in Group A versus Group C, as shown in Figure 1). For this comparison, we calculated the specificity and sensitivity between reviewer groups.

REFERENCES

1. Fitch K, Bernstein S, Aguilar MD, Burnand B, LaCalle JR, Lázaro P, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND; 2001.
2. Brown B. DELPHI PROCESS: A Methodology Used for the Elicitation of Opinions of Experts. Rand Corp. 1968 Sep;1–14.
3. Sweidan M, Williamson M, Reeve JF, Harvey K, O'Neill JA, Schattner P, et al. Identification of features of electronic prescribing systems to support quality and safety in primary care using a modified Delphi process. BMC Med Inform Decis Mak. 2010 Apr 15;10(1):21.
4. Stockwell D, Bisarya H, Classen D, Kirkendall E, Landrigan C, Lemon V, et al. A trigger tool to detect harm in pediatric inpatient settings. Pediatrics. 2015;
5. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). Institute for Healthcare Improvement; 2009. (IHI Innovation Series white paper).
6. Kirkendall ES, Kloppeborg E, Papp J, White D, Frese C, Hacker D, et al. Measuring adverse events and levels of harm in pediatric inpatients with the Global Trigger Tool. Pediatrics. 2012 Nov;130(5):e1206-1214.

[Response Ends]

2b.03. Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

[Response Begins]

Table 5:

Identifying a record with 1 or more AEs	Specificity	Sensitivity
Internal primary reviewers versus External expert primary reviewers	0.91	0.40
Internal secondary reviewers versus External expert secondary reviewers	0.95	0.33

This table displays the statistical results from validity testing, including measures of sensitivity and specificity.

[Response Ends]

2b.04. Provide 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?)

[Response Begins]

As summarized in Table 5, using the findings of the external reviewers as the standard of comparison, the specificity for identifying a record with one or more AEs was 0.91 for primary reviewers and 0.95 after taking secondary reviewer verification into account. The sensitivity was 0.40 for primary reviewers and 0.33 after taking secondary reviewer verification into account. The lower sensitivity is likely due in part to the novice reviewers' lack of experience with the tool and their inability to make up for their inexperience by increasing the amount of time they took to perform their review, given that there was a 30-minute time limit per record.

[Response Ends]

2b.05. 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 in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]

Since GAPPs facilitates preventable AE identification among pediatric inpatients, teams can conduct intra-hospital comparisons of preventable AE rates between different hospital divisions. Hospitals can also track their preventable AE rates over time to evaluate their state of patient safety, and can compare preventable AEs based on preventability and severity.

Comparison of preventable AE rates across hospitals would require reviewers at each institution to receive adequate training in the trigger tool methodology to ensure standardization of the preventable AE detection process. We anticipate that such training would contribute to increasing the measure's reliability, which would be necessary to reach a level appropriate for inter-hospital comparisons. Using the automated approach would further increase reliability because it removes the human error involved in finding triggers. However, using an automated trigger identification system instead of the manual system has no impact on the rest of the measure calculation, and we anticipate there are no obvious differences for either approach. A national, state, or other multi-hospital database would be ideal for inter-hospital comparisons but does not yet exist. If GAPPs is used for inter-hospital comparisons, preventable AE rates can be compared based on hospital type (academic vs. community, see figure 3).

[Response Ends]

2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include 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.

[Response Begins]

Figure 3 - Distribution of preventable AE rates by hospital:

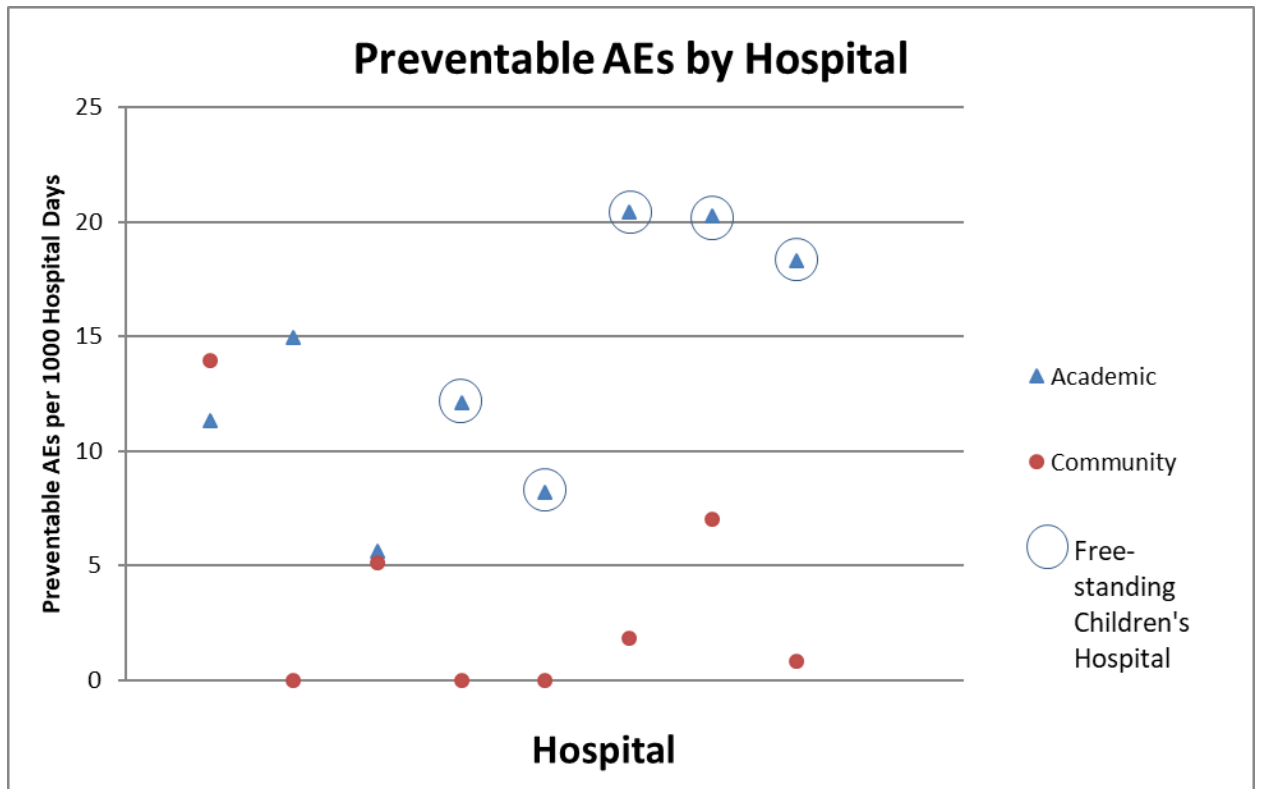
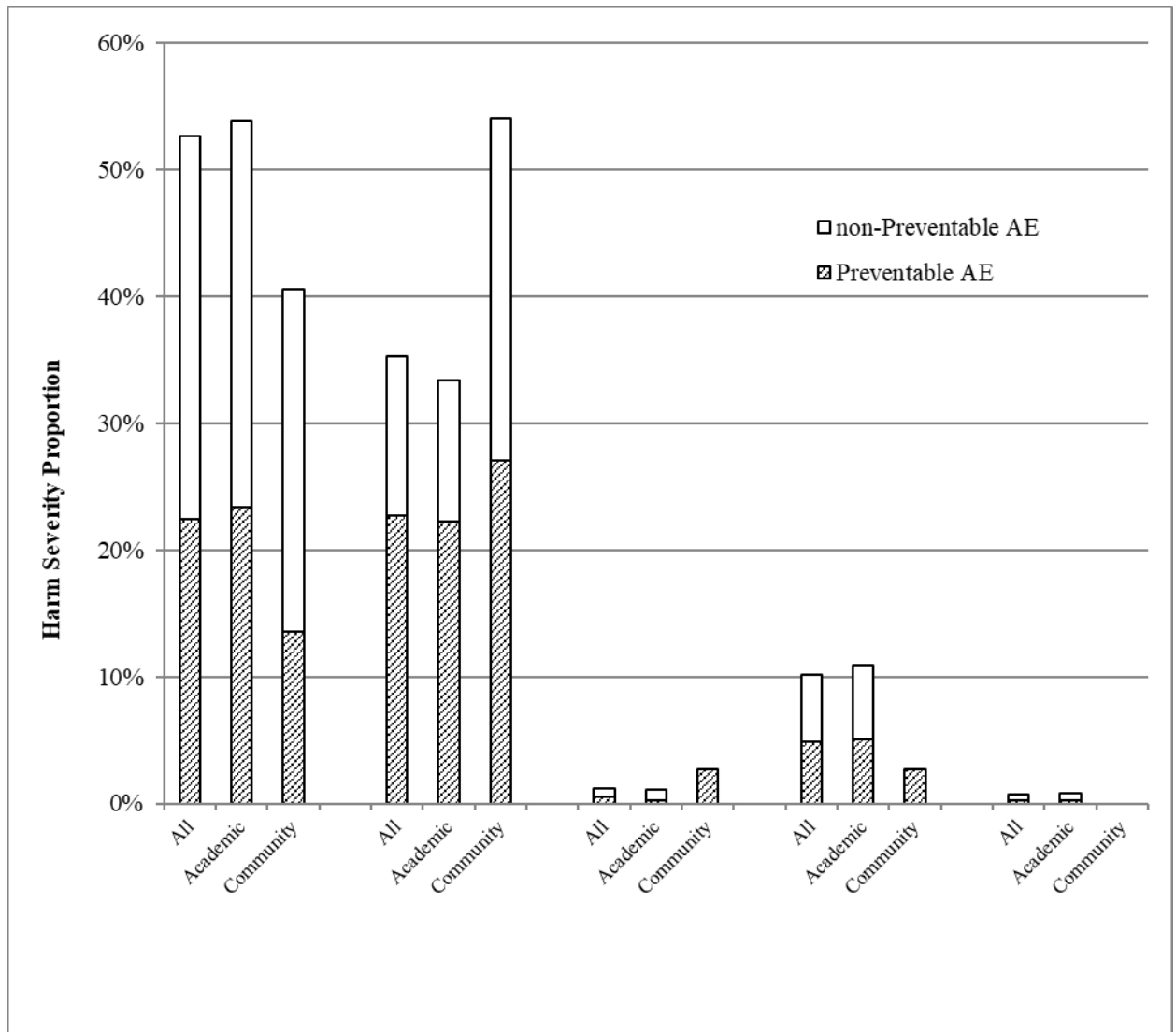


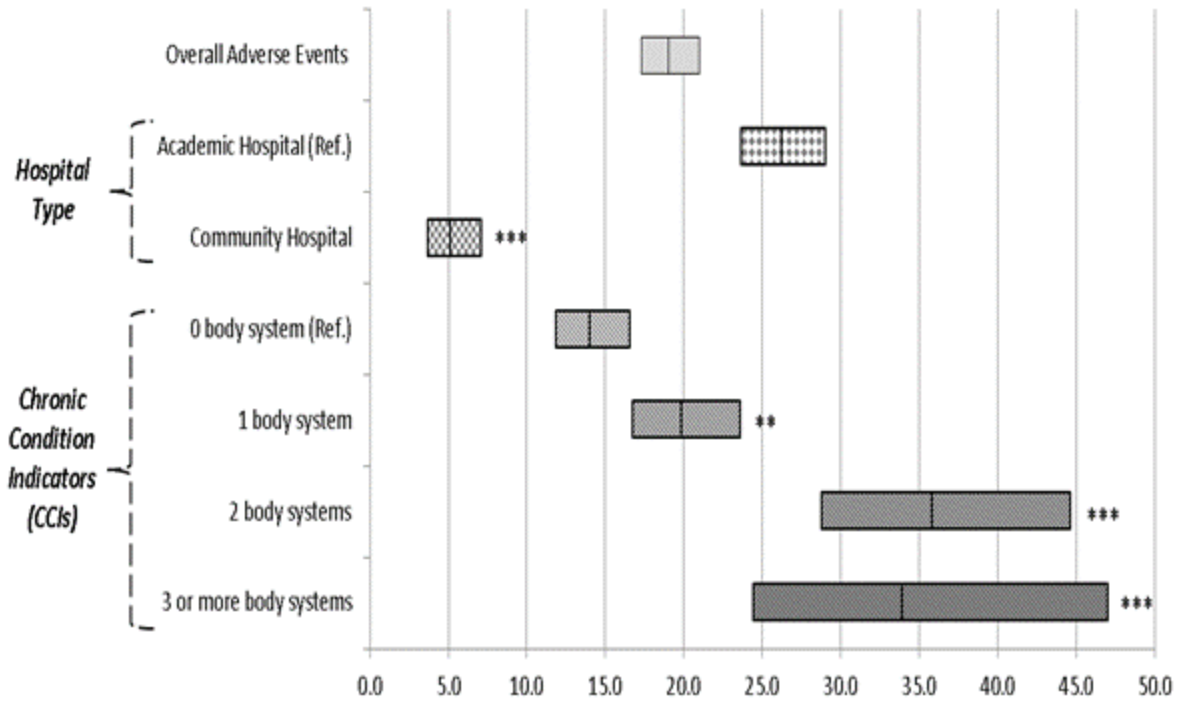
Figure 4 - Severity of all harms and preventable harms:



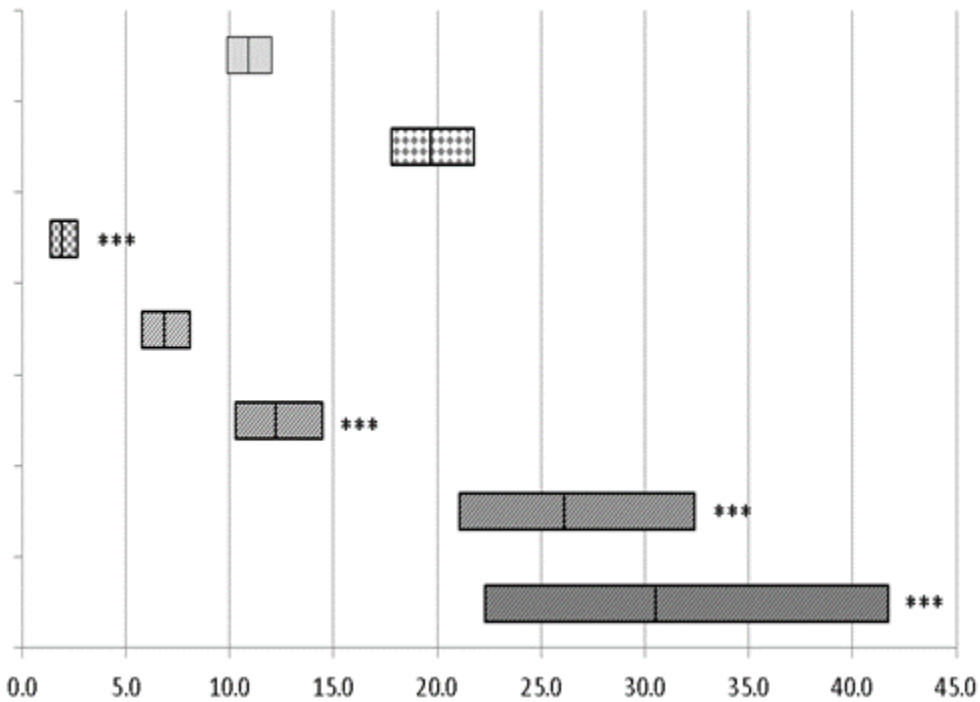
* NCC MERP Categories: E: contributed to or resulted in temporary harm to the patient and required intervention, F: contributed to or resulted in temporary harm to the patient and required initial or prolonged hospitalization, G: contributed to or resulted in permanent patient harm, H: required intervention to sustain life, I: contributed to or resulted in the patient's death.

Figure 5 - Distribution of Adverse Events (AEs) by Hospital and Clinical Characteristics:

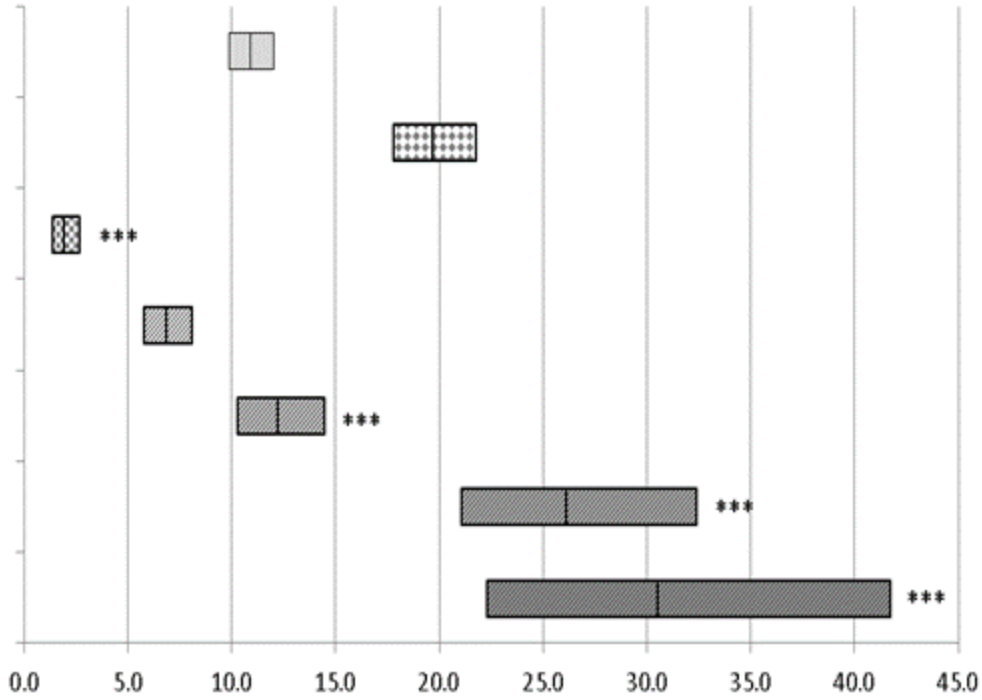
A. All AEs per 1,000 Patient Days



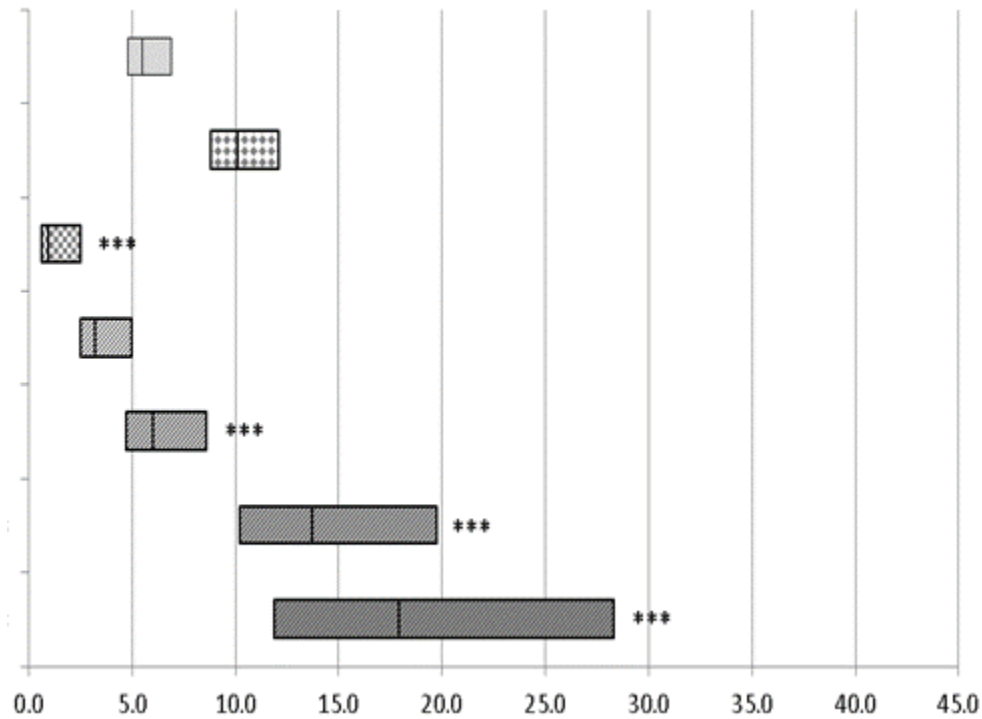
B. All AEs per 100 Admissions



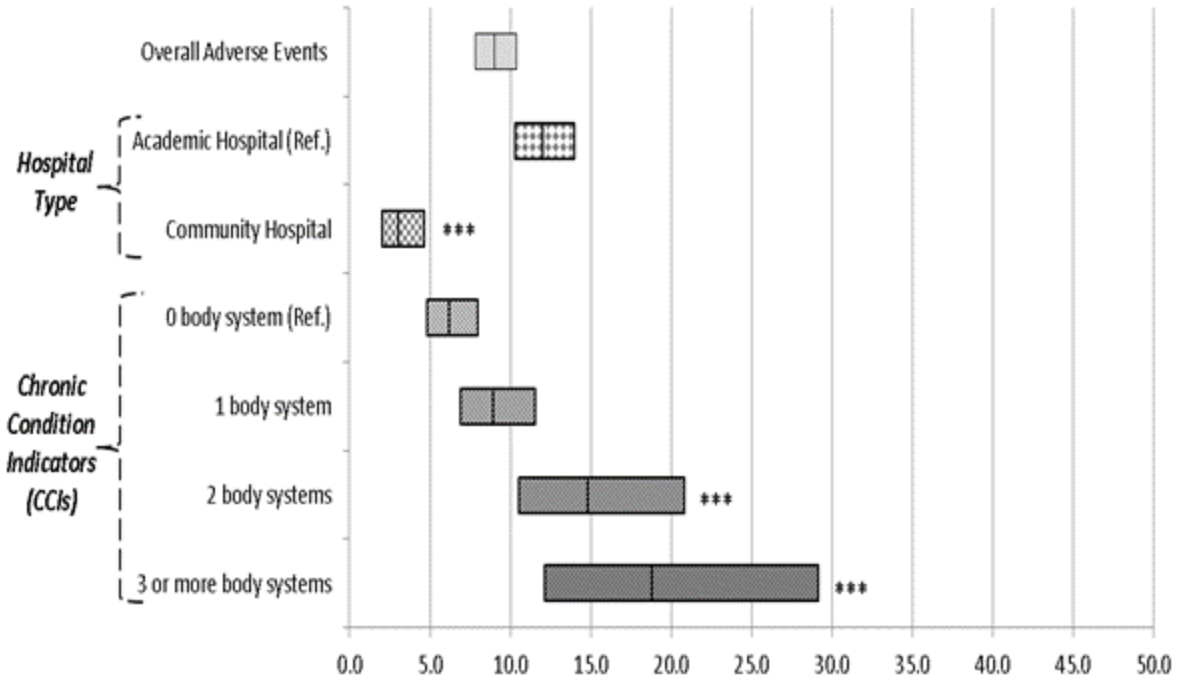
C. Preventable AEs per 1,000 Patient Days



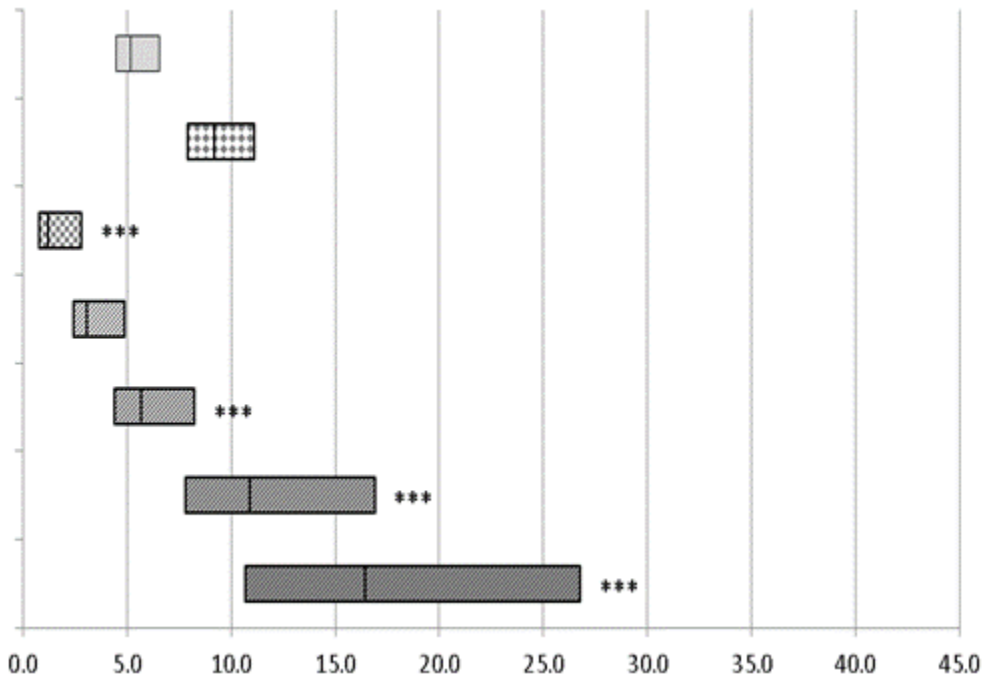
D. Preventable AEs per 100 Admissions



E. High-Severity AEs^A per 1,000 Patient Days



F. High-Severity AEs^A per 100 Admissions

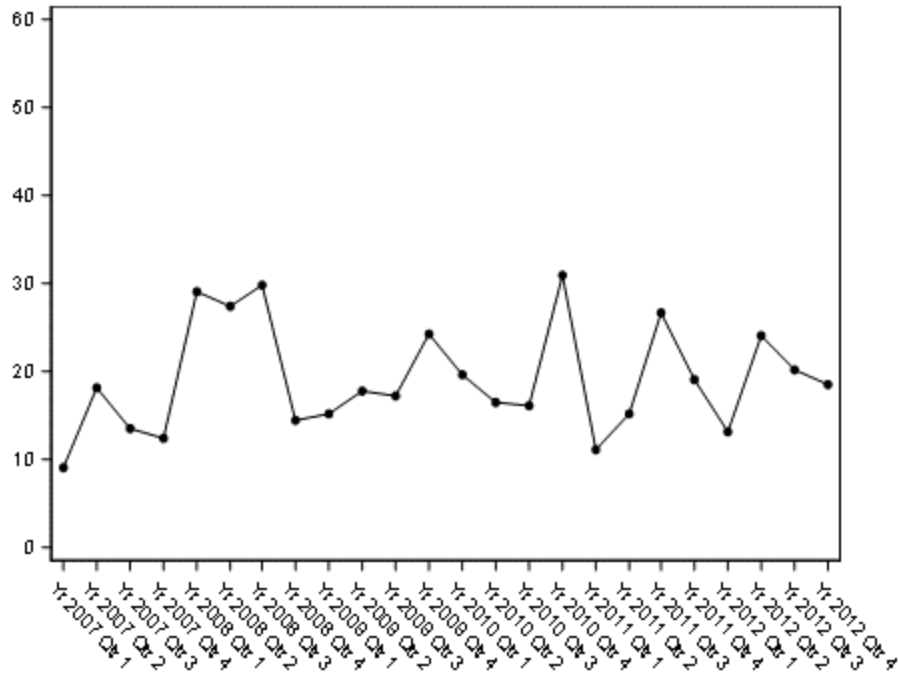


A: High-Severity AEs are defined as NCC MERP categories F to I

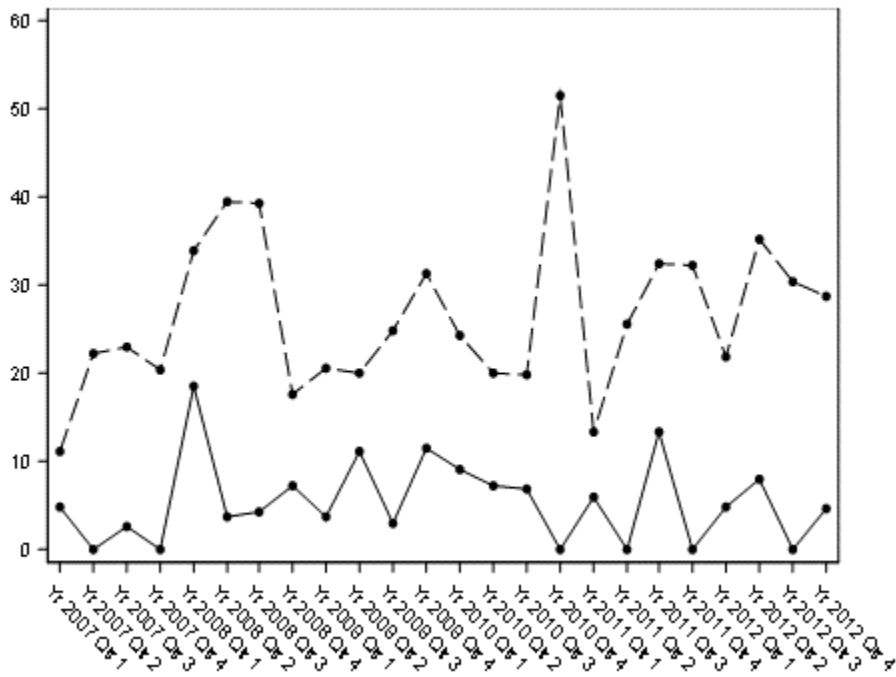
: p<0.01; *: p<0.001

Figure 6 – Rates of all harms, preventable harms, and high-severity harms per 1,000 patient-days, according to quarter:

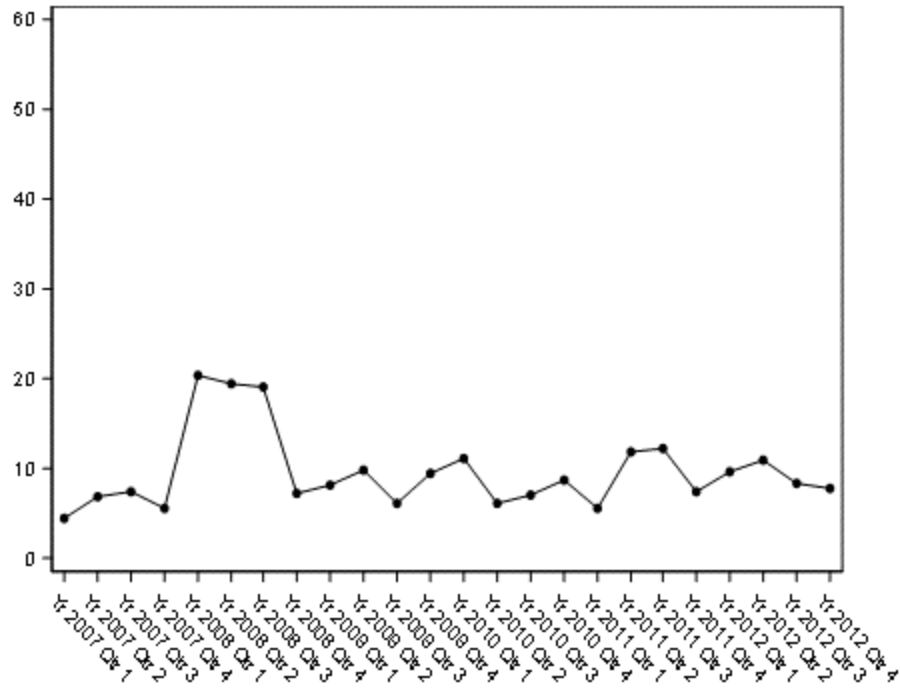
A. All AEs (per 1,000 Patient-Days)



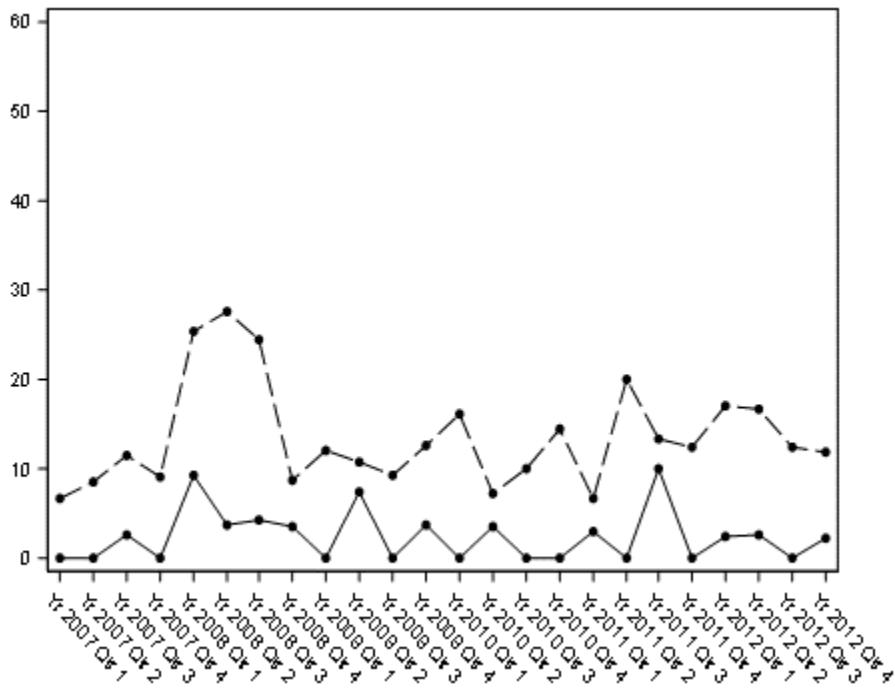
B. By Hospital Type^A, All AEs (per 1,000 Patient-Days)



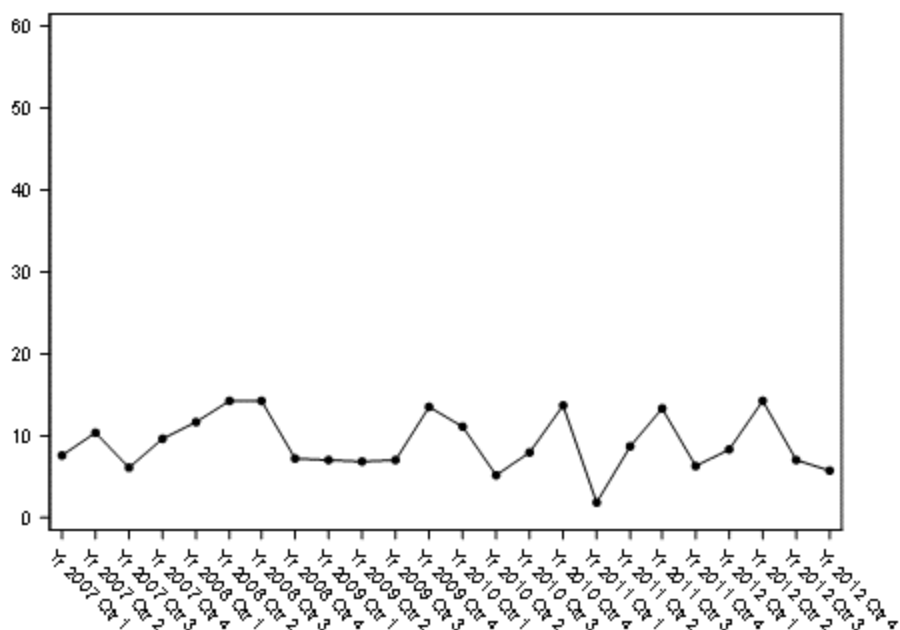
C. Preventable AEs (per 1,000 Patient-Days)



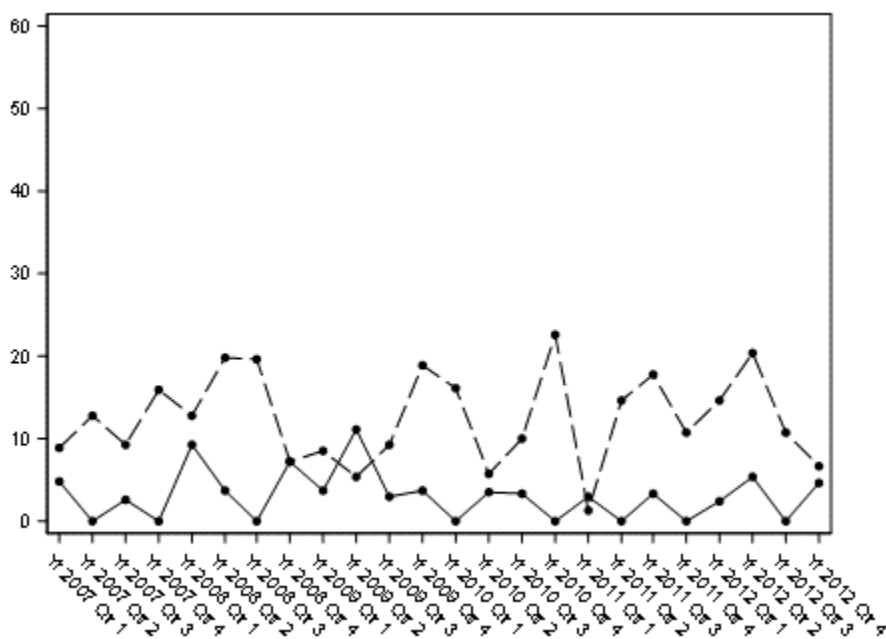
D. By Hospital Type^A, Preventable AEs (per 1,000 Patient-Days)



E. High-Severity AEs^B (per 1,000 Patient-Days)



F. By Hospital Type^A, High-Severity AEs (per 1,000 Patient-Days)

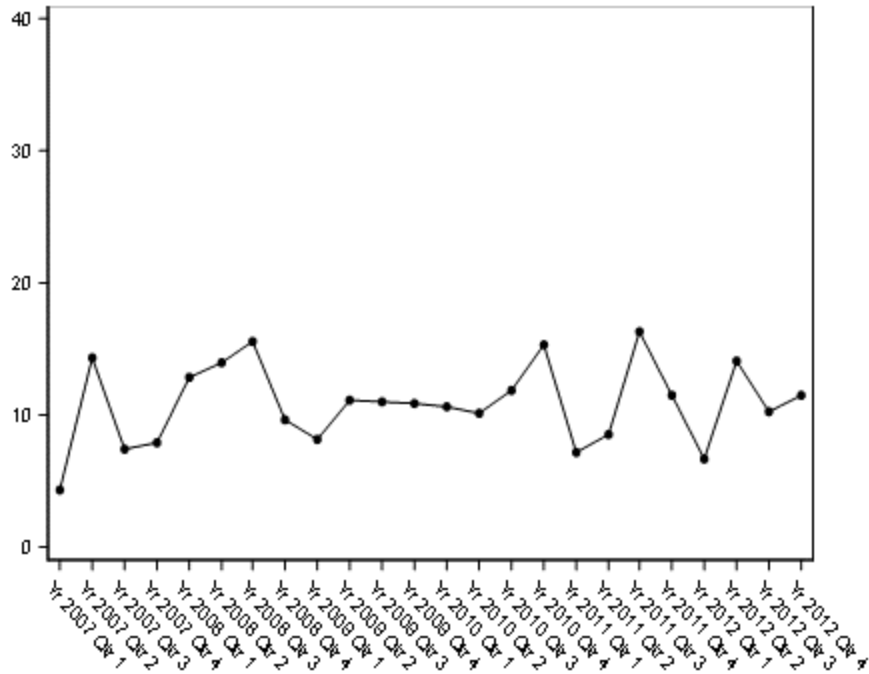


A: ----- Academic Hospital, — Community Hospital

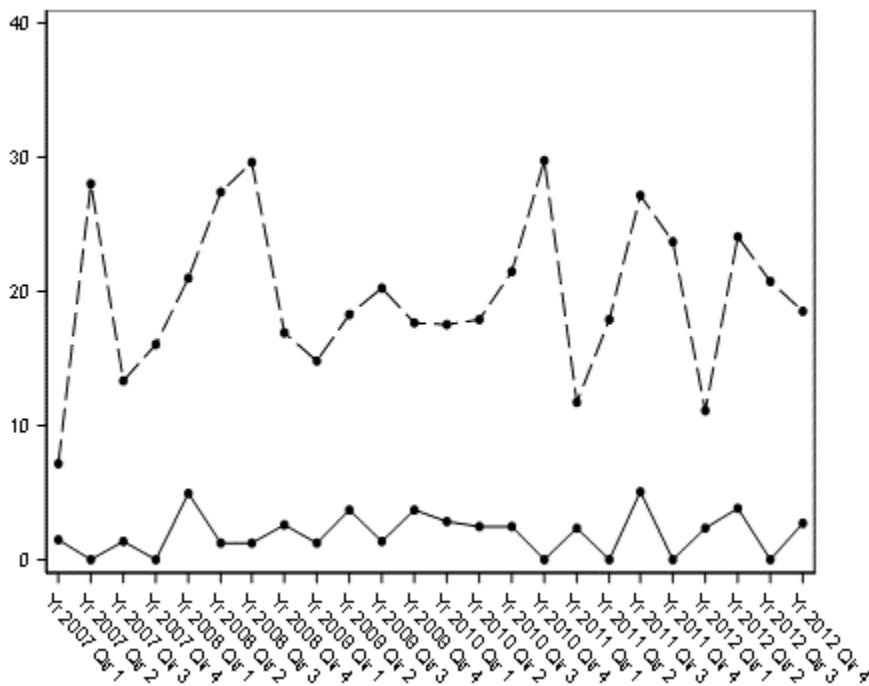
B: High-severity AEs are defined as NCC MERP categories F to I

Note: The values presented are unadjusted.

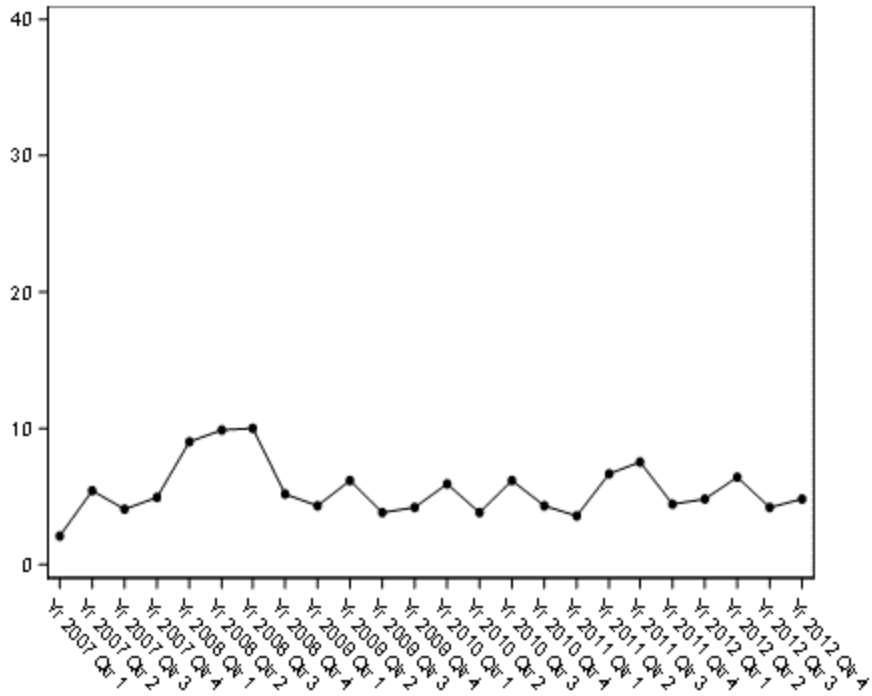
A. All AEs (per 100 Admissions)



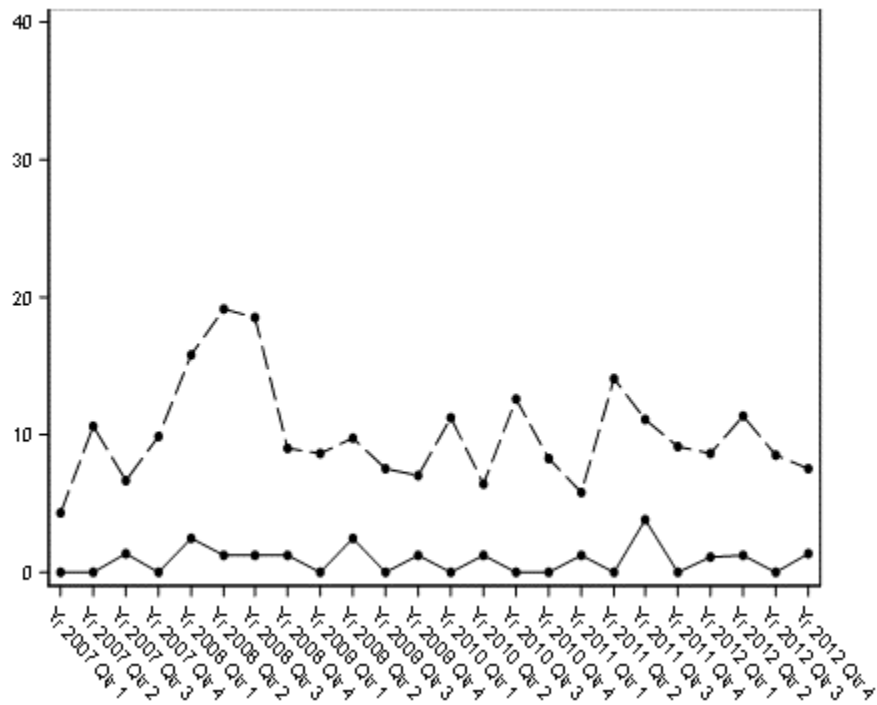
B. By Hospital Type^A, All AEs (per 100 Admissions)



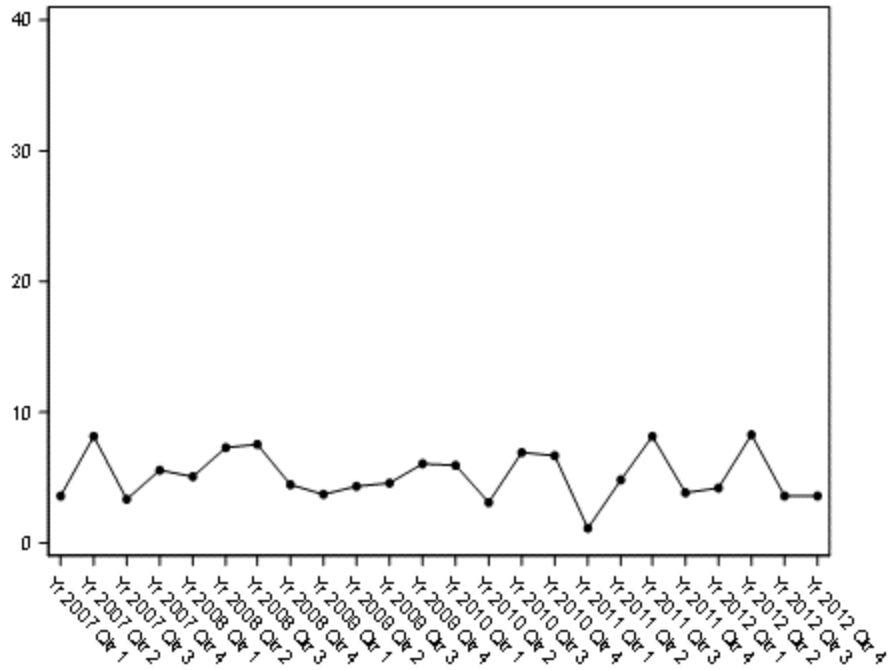
C. Preventable AEs (per 100 Admissions)



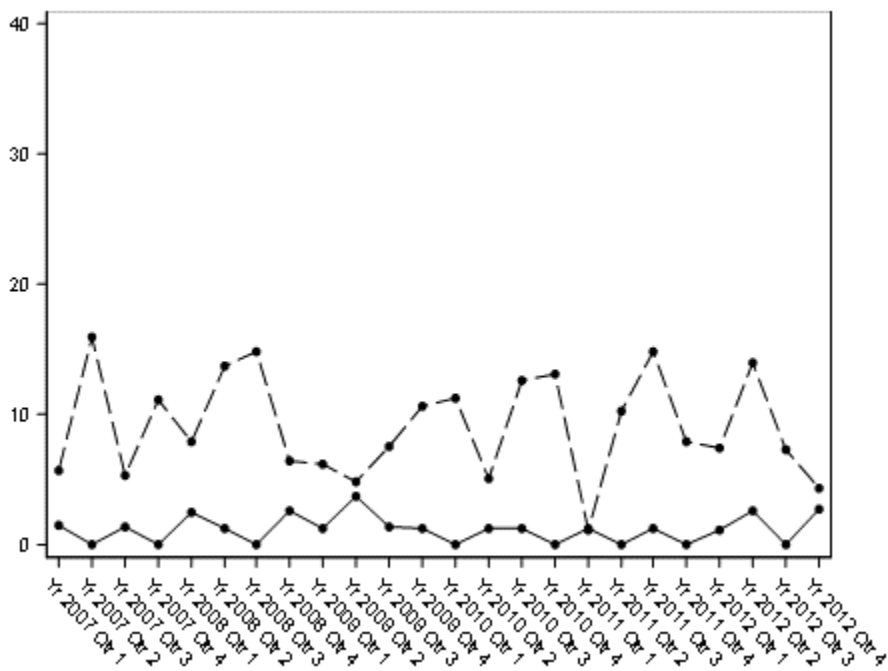
D. By Hospital Type^A, Preventable AEs (per 100 Admissions)



E. High-Severity AEs^B (per 100 Admissions)



F. By Hospital Type^A, High-Severity AEs (per 100 Admissions)



A: ----- Academic Hospital, — Community Hospital

B: High-severity AEs are defined as NCC MERP categories F to I

Note: The values presented are unadjusted.

[Response Ends]

2b.07. Provide 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.

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]

In a study of 16 academic and community hospitals that care for children across the US, we found that harm due to medical care remained common from 2007 to 2012 and did not decrease significantly over time. Approximately half of all harms are preventable, indicating that these harms can be immediately targeted by quality improvement teams. AEs were most commonly severity levels of E and F.

We found wide disparities in the rates of harm in academic and community hospitals. The reasons for this difference are unclear, but major differences in the frequency of complex chronic conditions as well as in the types and severity of illness seen in the two types of hospitals likely explain much of the difference. Neither community nor academic centers experienced improvements over the six-year span studied, suggesting that effectively controlling pediatric patient safety problems has proven similarly difficult in both settings, despite their baseline differences in populations and harm epidemiology.

[Response Ends]

2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

[Response Begins]

The only source of missing data for this measure is missing documents for a selected hospitalization. In the rare circumstance that a medical record was missing all documents, the selected hospitalization was discarded and replaced by another hospitalization. We only have record of this happening with two medical records in our study (n=3,814). We have no reason to believe that missing a record is anything more than a random and rare, non-systematic event. With the increasing uptake of electronic medical records and transition away from paper charts, missing documentation for hospitalizations will become vanishingly rare. For these reasons, we believe missing data did not represent a significant source of systematic bias. Given the infrequency of missing data, we were unable to conduct additional testing on the cause or impact of missing documentation.

[Response Ends]

2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

[Response Begins]

Given the rarity of missing documents, especially in settings which have adopted electronic health records, no overall frequency of missing data was calculated. As a result, no additional testing was conducted related to the missing data. In addition, we do not believe missing data could have biased results based on the infrequency of missing records.

[Response Ends]

2b.10. Provide 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 non-responders), and how the specified handling of missing data minimizes bias.

In other words, 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 was conducted, justify the selected approach for missing data.

[Response Begins]

The rarity of missing data made empirical analysis on the cause and impact of missing data unfeasible. We have no reason to believe missing data systematically biased performance results in any way.

[Response Ends]

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 eQMs). 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.

2b.11. Indicate whether there is more than one set of specifications for this measure.

[Response Begins]

No, there is only one set of specifications for this measure

[Response Ends]

2b.12. 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. Indicate what statistical analysis was used.

[Response Begins]

[Response Ends]

2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]

[Response Ends]

2b.15. Indicate whether the measure uses exclusions.

[Response Begins]

N/A or no exclusions

[Response Ends]

2b.16. Describe the method of testing exclusions and what was tested.

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?

[Response Begins]

N/A

[Response Ends]

2b.17. Provide 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.

[Response Begins]

N/A

[Response Ends]

2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, 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.

[Response Begins]

N/A

[Response Ends]

2b.19. Check all methods used to address risk factors.

[Response Begins]

Statistical risk model with risk factors (specify number of risk factors)

[Statistical risk model with risk factors (specify number of risk factors) Please Explain]

3

Stratification by risk category (specify number of categories)

[Stratification by risk category (specify number of categories) Please Explain]

2

[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]

For inter-hospital comparisons of preventable adverse event rates, case-mix adjustment models should be fit with mixed effects negative binomial regression. In each model, the dependent variable is the number of preventable adverse events with exposure time equal to length of stay (a random intercept at the hospital level) and the fixed effects are the three case-mix adjusters: patient age group, number of chronic conditions, and service type (medical vs. surgery).

[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]

[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]

Published literature

[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of $p < 0.10$ or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

[Response Begins]

One of the methodological issues associated with making comparisons across institutions is the need to adjust appropriately for case-mix differences. Case-mix refers to patient characteristics, such as demographic characteristics and health status, that may affect measures of outcomes or processes. Systematic effects of this sort create the potential for a population's rates to be higher or lower because of its characteristics, rather than

because of the quality of care provided, making comparisons of unadjusted rates potentially misleading. The basic goal of adjusting for case-mix is to estimate how different institutions would be rated if they all provided care to comparable groups of patients in terms of case-mix variables.

To evaluate potential variables for case-mix adjustment of GAPPs rates, we evaluated: patient age, sex, number of chronic conditions as determined using the Chronic Condition Indicator (CCI), and service type based on data collected by the reviewers from the patient record.⁽¹⁾ We included a variable in our multivariate case-mix models if its bivariate association with preventable adverse events was $p < 0.10$.

REFERENCES

1. Chronic Condition Indicator (CCI) for ICD-10-CM [Internet]. Healthcare Cost and Utilization Project (HCUP); Available from: https://www.hcup-us.ahrq.gov/toolssoftware/chronic_icd10/chronic_icd10.jsp

[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]

Gender was not associated with preventable adverse events (PAEs) in both the teaching and community hospitals ($p=0.69$ and $p=0.29$, respectively). Age group (< 3 years, 3 years to < 10 years, 10 years to < 18 years) was associated with PAEs in the community ($p < 0.001$) and teaching ($p=0.01$) hospitals. Surgery services (versus medical) were also associated with PAEs in both the community ($p=0.002$) and the teaching ($p < 0.001$) hospitals. We evaluated the association of the number of CCIs (0, 1, 2, and 3+) and PAEs and found that in community and teaching hospitals the risk associated with 0 and 1 CCIs was not different ($p=0.49$ and $p=0.90$, respectively), and the risk associated with 2 and 3+ CCIs was not different ($p=0.91$ and $p=0.87$, respectively). After collapsing the CCI variable, patients with 2+ CCIs were significantly more likely to have PAEs compared to patients with 0 or 1 CCIs in both community ($p=0.009$) and teaching ($p=0.008$) hospitals. Therefore, we included CCIs (2+ CCIs versus 0-1 CCI), age (< 3 years, 3 years to < 10 years, 10 years to < 18 years), and type of service (surgery versus medical) as adjusters in our case-mix model.

[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]

GAPPs is a measure of preventable adverse events in hospitals. It is measuring in-hospital processes of care (e.g., the safety of medication ordering and delivery, procedural performance, care coordination) that should be equally applied to all, regardless of SDS. Unlike many other common measures (e.g., readmissions, where social factors beyond the purview of the healthcare system are an important factor in readmission rates), GAPPs is focused on in-hospital care quality that is within the control of the healthcare system largely irrespective of patient SDS. Therefore we have chosen not to include SDS characteristics in our risk adjustment model because we do not believe that SDS characteristics should be inherently related to rates of preventable adverse events at a substantial level, except insofar as they may be associated with true differences in care quality. That said, we do recommend the reporting of stratified analyses by SDS groups, to facilitate identification of disparities in care.

[Response Ends]

2b.26. 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). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter “N/A” for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]

Pseudo R-squared values for the teaching hospital and community hospital models were 0.02 and 0.11, respectively.

[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]

Our models showed good discrimination between any PAEs and no PAEs (community hospitals c-statistic= 0.80, 95% confidence interval [CI] 0.69-0.91; teaching hospitals c-statistic= 0.76, 95% CI 0.71-0.81).

[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]

The teaching and community hospital models had good calibration across all twelve risk groups represented in our model (goodness-of-fit tests $p=0.31$ and $p=0.99$, respectively).

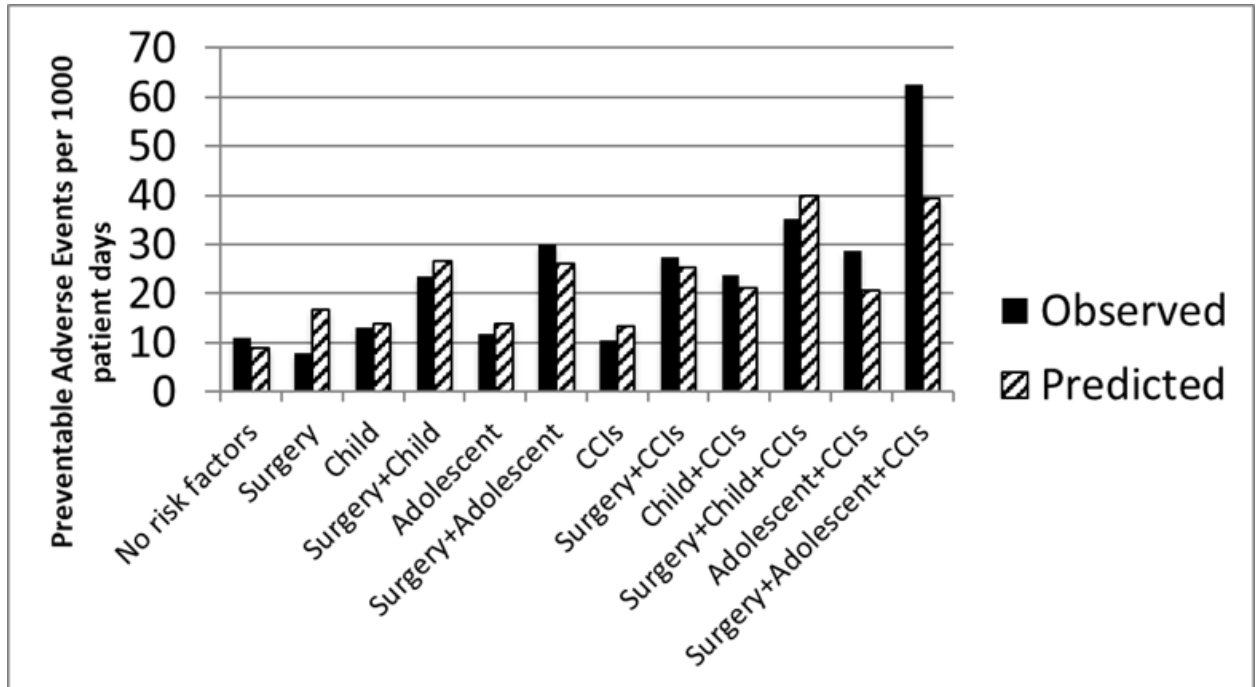
[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]

Figure 2- Calibration of Teaching Hospital Model:



2b.30) Provide the results of the risk stratification analysis.

Table 6 - Model results for teaching and community hospitals:

*	Teaching Hospitals	Teaching Hospitals	Teaching Hospitals
*	Coefficient	95% confidence interval	p-value
Number of CClIs (0 or 1=reference)	*	*	*
2+ CClIs	0.410	0.052, 0.767	0.02
Age group (< 3 years = reference)	*	*	*
≥3 years and <10 years	0.458	0.016, 0.899	0.04
≥10 years and <18 years	0.439	0.056, 0.822	0.02
Service type (medical = reference)	*	*	*
Surgery	0.638	0.225, 1.051	0.002
b ₀	-4.724	-5.000, -4.448	<0.001

This table displays the results of the risk stratification analysis for teaching hospitals.

*Cell intentionally left empty.

*	Community Hospitals	Community Hospitals	Community Hospitals
*	Coefficient	95% confidence interval	p-value
Number of CCIs (0 or 1=reference)	*	*	*
2+ CCIs	1.406	-0.055, 1.867	0.06
Age group (< 3 years = reference)	*	*	*
≥3 years and <10 years	1.450	-0.062, 2.962	0.06
≥10 years and <18 years	2.041	0.704, 3.378	0.003
Service type (medical = reference)	*	*	*
Surgery	0.676	-0.652, 2.005	0.32
b ₀	-7.060	-7.975, -6.145	<0.001

This table displays the results of the risk stratification analysis for community hospitals.

*Cell intentionally left empty.

Based on the above models, we recommend stratification by hospital type. Additionally, we recommend use of the following three categorical variables in the GAPPs case-mix adjustment model: age, number of CCIs, and service type.

[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]

Our models showed good discrimination and were well calibrated across risk groups.

In many analyses, the goal is to explain as much of the variance as possible, in which case a high R-squared is desired. In this case, the value of the R-squared represents the extent to which case-mix adjustment affected measure scores. For example, if the case-mix adjusters had no effect (e.g., age was not predictive of measure scores), then the R-squared value would be zero. Overall, case-mix adjustments explained a small proportion of the variation in preventable adverse events.

[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?

[Response Begins]

N/A

[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.

[Response Begins]

N/A

[Response Ends]

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.

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.

[Response Begins]

Some data elements are in defined fields in electronic sources

[Response Ends]

3.03. 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 data elements not from electronic sources.

[Response Begins]

The measure currently requires a clinician to review each adverse event in order to determine whether that event may have been preventable. It may be possible in the future for this step to transition to an automated process. While capturing triggers can be automated, a completely automated approach is currently not feasible and also not likely feasible for the near future. An automated trigger identification system has no impact on the measure beyond the means by which triggers are identified. The remaining pieces of the measure process following trigger identification are exactly the same for the automated and manual trigger approaches.

[Response Ends]

3.04. Describe any efforts to develop an eQIM.

[Response Begins]

Please see response to 3.03.

[Response Ends]

3.06. 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.

[Response Begins]

Manual chart review is a labor-intensive and time-consuming process.^{1,2} Trigger tool methodology is more efficient than comprehensive chart review and still reliable in detecting harm.^{3,4,5}

REFERENCES

1. 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(4):581-589. doi:10.1377/hlthaff.2011.0190.
2. Cullen DJ, Bates DW, Small SD, Cooper JB, Nemeskal AR, Leape LL. The incident reporting system does not detect adverse drug events: a problem for quality improvement. *Jt Comm J Qual Improv*. 1995;21(10):541-548.
3. Takata GS, Mason W, Taketomo C, Logsdon T, Sharek PJ. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children's hospitals. *Pediatrics*. 2008;121(4):e927-935. doi:10.1542/peds.2007-1779.
4. Agarwal S, Classen D, Larsen G, Tofil NM, Hayes LW, Sullivan JE, Storgion SA, Coopes BJ, Craig V, Jaderlund C, Bisarya H, Parast L, Sharek P. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc*. 2010;11(5):568-578. doi:10.1097/PCC.0b013e3181d8e405.
5. Sharek PJ, Horbar JD, Mason W, Bisarya H, Thurm CW, Suresh G, Gray JE, Edwards WH, Goldmann D, Classen D. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics*. 2006;118(4):1332-1340. doi:10.1542/peds.2006-0565.

[Response Ends]

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

3.07. Detail 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),

Attach the fee schedule here, if applicable.

[Response Begins]

The codes for the GAPPs automated triggers, the GAPPs Manual of Operations, and all associated forms that reviewers complete are available to users free of charge.

[Response Ends]

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.

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

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 demonstrating performance improvement.

4a.01. Check all current uses. For each current use checked, please provide:

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

[Response Begins]

Quality Improvement (Internal to the specific organization)

[Quality Improvement (Internal to the specific organization) Please Explain]

Name of program and sponsor: West Virginia University Medicine Children's Hospital

URL: Stroupe LM, Patra KP, Dai Z, et al. Measuring Harm in Hospitalized Children via a Trigger Tool. J Pediatr Nurs. 2018;41:9-15. doi:10.1016/j.pedn.2017.09.010 (<https://doi.org/10.1016/j.pedn.2017.09.010>)

Purpose: The GAPPs trigger tool was implemented to detect adverse events in a pediatric inpatient setting. The measure's efficacy was compared to that of the hospital's internal adverse event reporting method. The team found that the GAPPs tool detected four times more adverse events than the hospital's internal patient harm reporting system.

Geographic area and number and percentage of accountable entities and patients included: We do not have access to information on the specific geographic region, though the program reported 100 patients were included in the study.

Level of measurement and setting: The GAPPs tool was implemented in a retrospective chart review of 100 patients discharged from the 531-bed children's hospital in an academic medical center.

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

[Response Ends]

4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

[Response Begins]

GAPPS represents a substantial advance from the means by which most hospitals track patient safety (i.e., voluntary incident reporting systems), in that it has been shown to capture far more events, with greater reliability. However, moving to this new approach remains a relatively new approach conceptually; while some hospitals routinely use trigger tool approaches to track their safety, as yet they remain a substantial minority. We believe that the continued endorsement of GAPPS by NQF will help to promote its use by Quality and Safety leaders, and that as adoption increases, it will be appropriate to begin using it as an accountability measure in the future.

[Response Ends]

4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A 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.

[Response Begins]

AHRQ and CMS intend that the GAPPS measure be available for public use with the current expectation that the full measure specifications be provided on the AHRQ website, CMS website, or both. For ease of implementation, we have prepared the GAPPS Manual of Operations and automated trigger codes for detection and analysis of preventable AE rates (see Appendix A). Our testing has shown that the measure is straightforward to implement across a variety of hospital types and on both paper and electronic medical records. In addition, we have made a series of comprehensive training videos that are easily accessible online for sites that want to learn how to utilize GAPPS.

Although GAPPS is not currently used for public reporting, continued endorsement will facilitate the measure's use by public and private payers, provider organizations, and consumer groups that require NQF endorsement of quality measures and will help support the integration of GAPPS into other patient safety measures. We anticipate that GAPPS results will be useful to everyone with a need for information on the quality of pediatric inpatient care, including patients, parents, hospitals, health plans, insurers, and policy makers. In addition, hospitals could provide GAPPS performance scores to quality organizations and purchasers. GAPPS reliably identifies preventable AEs and can be used to guide and monitor quality improvement efforts and facilitate inter-hospital comparisons.

[Response Ends]

4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

Detail 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.

[Response Begins]

During development of the original trigger list, we conducted a nine-member expert panel from top national stakeholder organizations using the RAND/UCLA Appropriateness Method. Their feedback was critical for determining which triggers to include in our final trigger list and for moving the project from development to field testing. Some of the site leads participating in the National Field Test (NFT) provided feedback on the development of the NFT materials and procedures. Following the NFT, results, data, and interpretation assistance were provided to the participating institutions.

[Response Ends]

4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

[Response Begins]

Results reports were created after development and field testing of the measure. The reports included field test summary results, brief explanations of these results, and the raw institution-specific data. These reports were sent to the NFT institutions that requested their results following completion of the field test and analysis of the data.

[Response Ends]

4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

[Response Begins]

Feedback was obtained through personal correspondence with the GAPPS team. Trained reviewers from the participating NFT sites reported understanding GAPPS and were able to identify triggers in medical records, use them to detect AEs, and assess severity and preventability. Feedback from the site leads indicated that the measure is straightforward to use and easily understandable. Our results indicated that GAPPS works for both EHRs and paper medical records.

No difficulties in implementing the GAPPS trigger tool were noted in the study evaluating the performance of the measure after implementation in an academic children's hospital.¹

REFERENCES

1. Stroupe LM, Patra KP, Dai Z, et al. Measuring Harm in Hospitalized Children via a Trigger Tool. J Pediatr Nurs. 2018;41:9-15. doi:10.1016/j.pedn.2017.09.010

[Response Ends]

4a.08. Summarize the feedback obtained from those being measured.

[Response Begins]

Please see response to 4a.07.

[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]

Throughout development and testing, the GAPPS team presented our candidate measure to our Scientific Advisory Board, consisting of representatives from Boston Children's Hospital, the larger Harvard community, and organizations such as the National Initiative for Children's Healthcare Quality, as well as to our National Stakeholder Panel, which includes representatives from diverse national organizations that represent patients and families, providers, payers, and health services researchers. Comments and feedback from these various users and stakeholders indicate that they believe such a tool is useful to pediatric medical settings. Since finalizing the GAPPS

measure, the team has presented their findings at national conferences and other public forums. We have received positive feedback and many stated that the measure would be useful to them.

[Response Ends]

4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

[Response Begins]

Based on helpful suggestions from the NFT sites and reviewers, we improved the clarity of the Manual of Operations. The GAPPs NFT also demonstrated that it is crucial to provide rigorous training and feedback to reviewers on practice cases prior to reviewers' use of the measure in order to achieve optimal standardization in AE detection.

[Response Ends]

4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, 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, provide an explanation. 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.

[Response Begins]

Key strategies for reducing preventable harms in children include early detection and treatment of potential harm¹ and identification of potentially preventable adverse events.² Use and further development of measures such as GAPPs to detect adverse events is thus a critical part of efforts to improve patient safety.³⁻⁵ By using more sensitive and reliable measures, hospitals can increase their capacity to quantify inpatient adverse events, identify priorities, and target available resources.⁶ Multiple studies have shown that hospitals with reliable means to track adverse events have experienced improvements in patient safety and associated clinical outcomes.⁷⁻¹³

REFERENCES

1. Larsen GY, Donaldson AE, Parker HB, Grant MJCP. Preventable harm occurring to critically ill children*. *Pediatr Crit Care Med* July 2007. 2007;8(4):331-6.
2. McDonald KM, Davies SM, Haberland CA, Geppert JJ, Ku A, Romano PS. Preliminary Assessment of Pediatric Health Care Quality and Patient Safety in the United States Using Readily Available Administrative Data. *Pediatrics*. 2008 Aug 1;122(2):e416-25.
3. Takata GS, Mason W, Taketomo C, Logsdon T, Sharek PJ. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children's hospitals. *Pediatrics*. 2008 Apr;121(4):e927-935.
4. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ Can Med Assoc J J Assoc Med Can*. 2004 May 25;170(11):1678-86.
5. Agarwal S, Classen D, Larsen G, Tofil NM, Hayes LW, Sullivan JE, et al. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc*. 2010 Sep;11(5):568-78.
6. Miller MR, Elixhauser A, Zhan C. Patient Safety Events During Pediatric Hospitalizations. *Pediatrics*. 2003 Jun 1;111(6):1358-66.
7. Evans RS, Pestotnik SL, Classen DC, Bass SB, Burke JP. Prevention of adverse drug events through computerized surveillance. *Proc Annu Symp Comput Appl Sic Med Care Symp Comput Appl Med Care*. 1992;437-41.

8. Kaufman J, Rannie M, Kahn MG, Vitaska M, Wathen B, Peyton C, et al. An interdisciplinary initiative to reduce unplanned extubations in pediatric critical care units. *Pediatrics*. 2012 Jun;129(6):e1594-1600.
9. Piontek F, Kohli R, Conlon P, Ellis JJ, Jablonski J, Kini N. Effects of an adverse-drug-event alert system on cost and quality outcomes in community hospitals. *Am J Health-Syst Pharm AJHP Off J Am Soc Health-Syst Pharm*. 2010 Apr 15;67(8):613–20.
10. Cohen MM, Kimmel NL, Benage MK, Cox MJ, Sanders N, Spence D, et al. Medication safety program reduces adverse drug events in a community hospital. *Qual Saf Health Care*. 2005 Jun;14(3):169–74.
11. Garrett PR Jr, Sammer C, Nelson A, Paisley KA, Jones C, Shapiro E, et al. Developing and implementing a standardized process for global trigger tool application across a large health system. *Jt Comm J Qual Patient Saf Jt Comm Resour*. 2013 Jul;39(7):292–7.
12. Miles A, Spaeder M, Stockwell D. Unplanned ICU Transfers from Inpatient Units: Examining the Prevalence and Preventability of Adverse Events Associated with ICU Transfer in Pediatrics. *J Pediatr Intensive Care*. 2015 Nov 21;5(1):021–7.
13. Smith JC, Chen Q, Denny JC, Roden DM, Johnson KB, Miller RA. Evaluation of a Novel System to Enhance Clinicians' Recognition of Preadmission Adverse Drug Reactions. *Appl Clin Inform*. 2018;9(2):313-325. doi:10.1055/s-0038-1646963

[Response Ends]

4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

[Response Begins]

No unexpected findings were identified during implementation.

[Response Ends]

4b.03. Explain any unexpected benefits realized from implementation of this measure.

[Response Begins]

No unexpected benefits were identified during implementation.

[Response Ends]

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.

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]

N/A

[Response Ends]

5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]

No

[Response Ends]

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

[Response Begins]

N/A

[Response Ends]

5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

[Response Begins]

N/A

[Response Ends]

Appendix

Supplemental materials may be provided in an appendix.:

Available in attached file

Attachment: 3136_3136_AAA_MOO_-_Appendices_FINAL_20170201_(1)-508.pdf

Contact Information

Measure Steward (Intellectual Property Owner): Center of Excellence for Pediatric Quality Measurement

Measure Steward Point of Contact: Toomey, Sara, sara.toomey@childrens.harvard.edu

Buscher, Sara, sara.buscher@childrens.harvard.edu

Baden, Asher, asher.baden@childrens.harvard.edu

Measure Developer if different from Measure Steward: Center of Excellence for Pediatric Quality Measurement

Measure Developer Point(s) of Contact: Toomey, Sara, sara.toomey@childrens.harvard.edu

Buscher, Sara, sara.buscher@childrens.harvard.edu

Baden, Asher, asher.baden@childrens.harvard.edu

Additional Information

1. Provide any supplemental materials, if needed, as an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be collated one file with a table of contents or bookmarks. If material pertains to a specific criterion, that should be indicated.

[Response Begins]

Available in attached file

[Response Ends]

Attachment: 3136_3136_AAA_MOO_-_Appendices_FINAL_20170201_(1)-508.pdf

2. List the workgroup/panel members' names and organizations.

Describe the members' role in measure development.

[Response Begins]

The following people participated in measure development:

CORE TEAM:

Mark A. Schuster, MD, PhD (Principal Investigator, Director)
Christopher P. Landrigan, MD, MPH (Measure Co-Leader)
David C. Stockwell, MD, MBA (Measure Co-Leader)
Sara L. Toomey, MD, MPhil, MPH, MSc (Managing Director)

ADDITIONAL MEMBERS OF THE GAPPS TEAM, ESPECIALLY:

Jisun Jang, MA
Jessica A. Quinn, MS

ADDITIONAL CONTRIBUTORS:

Kimberly A. Chin, BA
Shannon M. Cottreau, BSN, RN, CPN
Colleen T. Madden, RN, BSN
Mari M. Nakamura, MD, MPH
Gareth J. Parry, PhD

RESEARCH ASSISTANTS:

Sepideh Ashrafzadeh, BS
Samuel S. Loren, BS
Sifon U. Ndon, AB
Debanjan Pain, AB
Michaela S. Tracy, BA
Michelle J. Wang, AB
Melody Wu, AB

Members of the GAPPS Advisory Committee: Hema Bisarya, MHSA, RD, David C. Classen, MD, MS, Paul J. Sharek, MD, MPH, and Rajendu Srivastava, MD, MPH

External expert primary reviewers: Kathleen M. Haig, RN, Diedre A. Rahn, RN, and Katherine R. Zigmont, RN

Expert secondary reviewers and training video instructors: Lee M. Adler, DO and Roger K. Resar, MD

Members of the GAPPs Expert Stakeholder Panel: David Bundy, MD, MPH, S. Todd Callahan, MD, MPH, Emi Datuin-Pal, RN, BSN, MSHSA, MBA, Carol Haraden, PhD, Laura Knobel, MD, FAAFP, Rita Pickler, PhD, RN, PNP-BC, FAAN, Xavier Sevilla, MD, MBA, FAAP, Jennifer Slayton, MSN, RN, and Glenn Takata, MD, MS

Reviewers and leads at all of our participating study sites: Boston Children's Hospital, Children's Hospital Colorado, Children's National Medical Center, Cincinnati Children's Hospital Medical Center, Grand View Hospital, Mary Washington Hospital, Lucile Packard Children's Hospital Stanford, Providence St. Peter Hospital, Progress West Hospital, University of Florida Health Shands Children's Hospital, Silver Cross Hospital, New York Presbyterian/Weill Cornell Medical Center, Utah Valley Regional Medical Center, Western Virginia University Hospitals, Hillcrest Hospital, and South Shore Hospital

Staff of the Center of Excellence for Pediatric Quality Measurement (CEPQM) at Boston Children's Hospital, members of CEPQM's Scientific Advisory Board, members of the National Stakeholder Panel Members of the Massachusetts Child Health Quality Coalition, and the GAPPs Expert Stakeholder Panel provided guidance and feedback on the measure.

[Response Ends]

3. Indicate the year the measure was first released.

[Response Begins]

2017

[Response Ends]

4. Indicate the month and year of the most recent revision.

[Response Begins]

07, 2017

[Response Ends]

5. Indicate the frequency of review, or an update schedule, for this measure.

[Response Begins]

N/A

[Response Ends]

6. Indicate the next scheduled update or review of this measure.

[Response Begins]

11, 2022

[Response Ends]

7. Provide a copyright statement, if applicable. Otherwise, indicate "N/A".

[Response Begins]

N/A

[Response Ends]

8. State any disclaimers, if applicable. Otherwise, indicate "N/A".

[Response Begins]

N/A

[Response Ends]

9. Provide any additional information or comments, if applicable. Otherwise, indicate "N/A".

[Response Begins]

N/A

[Response Ends]