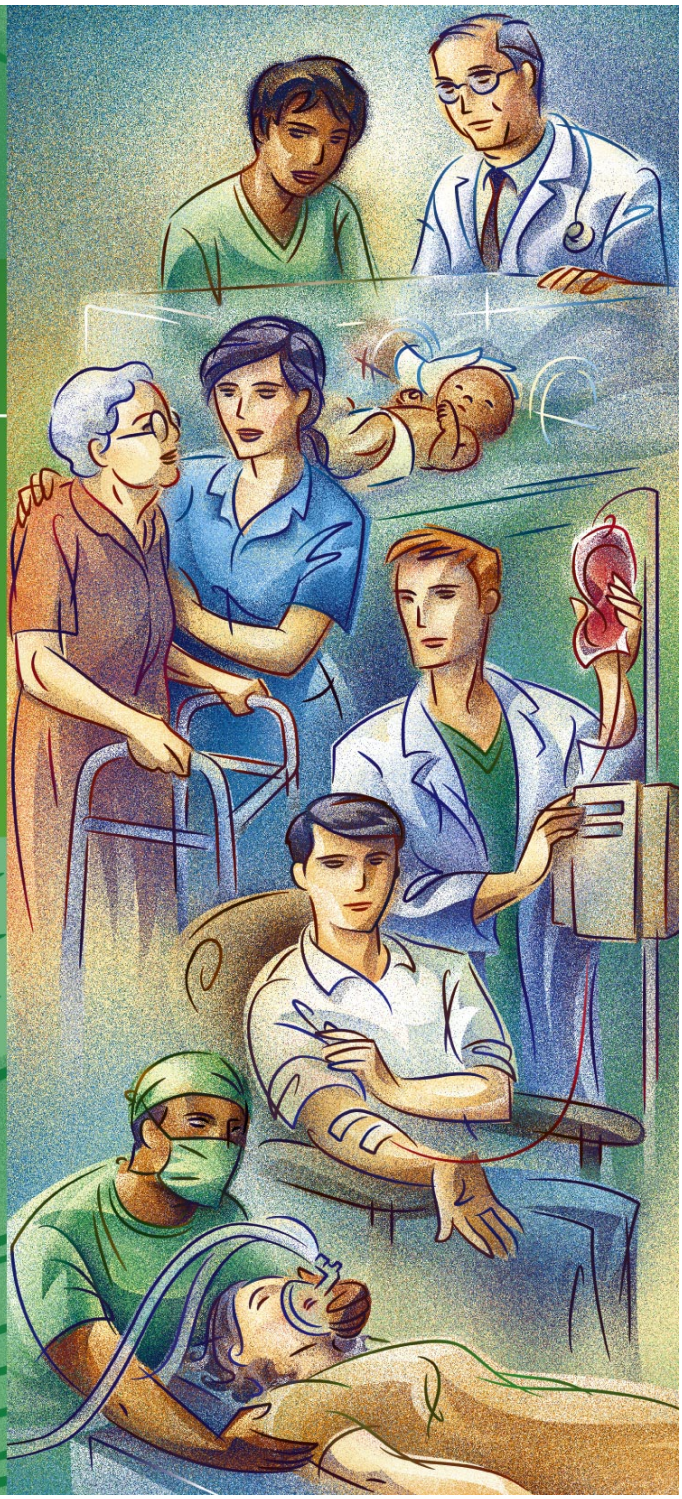
AHRQ

AHRQ QUALITY INDICATORS

Guide to Patient Safety Indicators



AHRQ Quality Indicators

Guide to Patient Safety Indicators

Department of Health and Human Services
Agency for Healthcare Research and Quality
http://www.qualityindicators.ahrq.gov

March 2003
Version 3.1 (March 12, 2007)

Preface

In health care as in other arenas, that which cannot be measured is difficult to improve. Providers, consumers, policy makers, and others seeking to improve the quality of health care need accessible, reliable indicators of quality that they can use to flag potential problems or successes; follow trends over time; and identify disparities across regions, communities, and providers. As noted in a 2001 Institute of Medicine study, *Envisioning the National Health Care Quality Report*, it is important that such measures cover not just acute care but multiple dimensions of care: staying healthy, getting better, living with illness or disability, and coping with the end of life.

The Agency for Healthcare Research and Quality (AHRQ) Quality Indicators (QIs) are one Agency response to this need for multidimensional, accessible quality indicators. They include a family of measures that providers, policy makers, and researchers can use with inpatient data to identify apparent variations in the quality of inpatient or outpatient care. AHRQ's Evidence-Based Practice Center (EPC) at the University of California San Francisco (UCSF) and Stanford University adapted, expanded, and refined these indicators based on the original Healthcare Cost and Utilization Project (HCUP) Quality Indicators developed in the early 1990s.

The new AHRQ QIs are organized into three modules: **Prevention Quality Indicators**, **Inpatient Quality Indicators**, and **Patient Safety Indicators**. AHRQ has published the three modules as a series. The first module – Prevention Quality Indicators – was released in 2001 and the second module – Inpatient Quality Indicators – was released in 2002. Both are available at AHRQ's Quality Indicators Web site at http://www.qualityindicators.ahrq.gov.

This third module focuses on potentially preventable complications and iatrogenic events for patients treated in hospitals. The Patient Safety Indicators (PSIs) are measures that screen for adverse events that patients experience as a result of exposure to the health care system; these events are likely amenable to prevention by changes at the system or provider level. The PSIs were initially released in March 2003. The PSIs now include 20 Provider-level and 7 Area-level Indicators.

Full technical information on the first two modules can be found in *Refinement of the HCUP Quality Indicators*, prepared by the UCSF-Stanford EPC. It can be accessed at AHRQ's Quality Indicators Web site (http://www.qualityindicators.ahrq.gov/downloads.htm). The technical report for the third module, entitled *Measures of Patient Safety Based on Hospital Administrative Data—The Patient Safety Indicators*, is also available on AHRQ's Quality Indicators Web site.

Improving patient safety is a critical part of efforts to provide high quality health care in the United States. This guide is intended to facilitate such efforts. As always, we would appreciate hearing from those who use our measures and tools so that we can identify how they are used, how they can be refined, and how we can measure and improve the quality of the tools themselves. You may contact us by sending an e-mail to support@qualityindicators.ahrq.gov.

Irene Fraser, Ph.D., Director
Center for Organization and Delivery Studies

The programs for the Patient Safety Indicators (PSIs) can be downloaded from http://www.qualityindicators.ahrq.gov/psi_download.htm.

Instructions on how to use the programs to calculate the PSI rates are contained in the companion text, *Patient Safety Indicators: Software Documentation (SAS* and *Windows).*

Acknowledgments

Support efforts, including refinement and enhancement of the AHRQ Quality Indicators and related products, are provided by the Support for Quality Indicators-II contract team.

The following individuals from **Battelle Memorial Institute, Stanford University,** and **University of California (UC)** constitute the Support for Quality Indicators-II core team:

Sheryl M. Davies, M.A.
Bruce Ellis, M.S.
Jeffrey Geppert, J.D.
Mark Gritz, Ph.D.
Theresa Schaaf, P.M.P.
Elaine Keller, M.Ed.
Kathryn M. McDonald, M.M.
Patrick Romano, M.D., M.P.H
Jeff Schoenborn, B.S.

The **Agency for Healthcare Research and Quality** Support for Quality Indicators team includes:

Marybeth Farquhar, Project Officer
Mamatha Pancholi, Project Officer
Mary B. Haines, Contract Officer

This product is based on the work of many individuals who contributed to its development and testing.

The following staff from the **Evidence-based Practice Center (EPC) at UCSF-Stanford** performed the evidence review, completed the empirical evaluation, and created the programming code and technical documentation for the AHRQ Quality Indicators:

Core Project Team
Kathryn M. McDonald, M.M. (Stanford), *principal investigator*

Investigators
Patrick S. Romano, M.D., M.P.H. (UC-Davis)
Jeffrey Geppert, J.D. (Stanford)
Sheryl M. Davies, M.A. (Stanford)
Bradford W. Duncan, M.D. (Stanford)
Kaveh G. Shojania, M.D. (UCSF)

Angela Hansen, B.A. (Stanford), *EPC Research Assistant*

The following staff from **Social & Scientific Systems**, **Inc.**, developed this software product, documentation, and guide:

Programmers
Leif Karell
Kathy McMillan
Fred Rohde

Technical Writer
Patricia Burgess

Graphics Designer
Laura Spofford

Contributors from the **Agency for Healthcare Research and Quality**:

Anne Elixhauser, Ph.D.
Denise Remus, Ph.D., R.N.
H. Joanna Jiang, Ph.D.
Marlene Miller, M.D., M.Sc.
Margaret Coopey, R.N., M.G.A, M.P.S.

We wish to also acknowledge the following individuals and organizations for their aid in this report: Doug Staiger, Dept. of Economics, Dartmouth College; Ros McNally, National Primary Care Research and Development Centre, University of Manchester; Rita Scichilone and the American Health Information Management Association; the various professional organizations that provided nominations for our clinical review panels; the clinical panelists; the peer reviewers of the evidence report; and the beta-testers of the software products, all of whose input was invaluable.

Table of Contents

Preface ... iii
Acknowledgments ... iv
1.0 Introduction to the AHRQ Patient Safety Indicators ... 1
1.1 What Are the Patient Safety Indicators? ... 2
1.2 How Can the PSIs Be Used to Assess Patient Safety? ... 4
1.3 What Does this Guide Contain? ... 5
1.4 Support for Potential and Current Users of the AHRQ QIs ... 5
2.0 Origins and Background of the Quality Indicators ... 6
2.1 Development of the AHRQ Quality Indicators ... 6
2.2 AHRQ Quality Indicator Modules ... 6
3.0 Methods of Identifying, Selecting, and Evaluating the Quality Indicators ... 8
3.1 Step 1: Define the Concepts and the Evaluation Framework ... 8
3.2 Step 2: Search the Literature to Identify Potential PSIs ... 10
3.3 Step 3: Develop a Candidate List of PSIs ... 11
3.4 Step 4: Review the PSIs ... 13
3.5 Step 5: Evaluate the PSIs Using Empirical Analysis ... 14
4.0 Summary Evidence on the Patient Safety Indicators ... 16
4.1 Limitations in Using the PSIs ... 19
4.2 Further Research on PSIs ... 20
4.3 Use of External Cause-of-Injury Codes ... 21
5.0 Detailed Evidence for Patient Safety Indicators ... 23
5.1 Complications of Anesthesia (PSI 1) ... 24
5.2 Death in Low-Mortality DRGs (PSI 2) ... 26
5.3 Decubitus Ulcer (PSI 3) ... 28
5.4 Failure to Rescue (PSI 4) ... 30
5.5 Foreign Body Left During Procedure, Provider Level (PSI 5) ... 32
5.6 Foreign Body Left During Procedure, Area Level (PSI 21) ... 32
5.7 Iatrogenic Pneumothorax, Provider Level (PSI 6) ... 34
5.8 Iatrogenic Pneumothorax, Area Level (PSI 22) ... 34
5.9 Selected Infections Due to Medical Care, Provider Level (PSI 7) ... 36
5.10 Selected Infections Due to Medical Care, Area Level (PSI 23) ... 36
5.11 Postoperative Hip Fracture (PSI 8) ... 38
5.12 Postoperative Hemorrhage or Hematoma (PSI 9) ... 40
5.13 Postoperative Hemorrhage or Hematoma (PSI 27) ... 40
5.14 Postoperative Physiologic and Metabolic Derangement (PSI 10) ... 43
5.15 Postoperative Respiratory Failure (PSI 11) ... 45
5.16 Postoperative Pulmonary Embolism or Deep Vein Thrombosis (PSI 12) ... 47
5.17 Postoperative Sepsis (PSI 13) ... 49
5.18 Postoperative Wound Dehiscence, Provider Level (PSI 14) ... 51
5.19 Postoperative Wound Dehiscence, Area Level (PSI 24) ... 51
5.20 Accidental Puncture or Laceration, Provider Level (PSI 15) ... 53
5.21 Accidental Puncture or Laceration, Area Level (PSI 25) ... 53
5.22 Transfusion Reaction, Provider Level (PSI 16) ... 55
5.23 Transfusion Reaction, Area Level (PSI 26) ... 55
5.24 Birth Trauma—Injury to Neonate (PSI 17) ... 57
5.25 Obstetric Trauma—Vaginal Delivery with Instrument (PSI 18) ... 59
5.26 Obstetric Trauma—Vaginal Delivery without Instrument (PSI 19) ... 61
5.27 Obstetric Trauma—Cesarean Delivery (PSI 20) ... 63
6.0 Using Different Types of QI Rates ... 65
7.0 References ... 66
Appendix A: Links ... A-1

List of Tables

Table 1: AHRQ Provider-Level Patient Safety Indicators 17
Table 2: Indicators and Use of External Cause-of-Injury Codes..... 22

1.0 Introduction to the AHRQ Patient Safety Indicators

Hospitals in the United States provide the setting for some of life's most pivotal events—the birth of a child, major surgery, treatment for otherwise fatal illnesses. These hospitals house the most sophisticated medical technology in the world and provide state-of-the-art diagnostic and therapeutic services. But access to these services comes with certain costs. About 30% of personal health care expenditures in the United States go towards hospital care,[1] and the rate of growth in spending for hospital services has only recently leveled out after several years of increases following a half a decade of declining growth.[2] Simultaneously, concerns about the quality of health care services have reached a crescendo with the Institute of Medicine's series of reports describing the problem of medical errors[3] and the need for a complete restructuring of the health care system to improve the quality of care.[4] Policymakers, employers, and consumers have made the quality of care in U.S. hospitals a top priority and have voiced the need to assess, monitor, track, and improve the quality of inpatient care.

Hospital administrative data offer a window into the medical care delivered in our nation's hospitals. These data, which are collected as a routine step in the delivery of hospital services, provide information on diagnoses, procedures, age, gender, admission source, and discharge status. From these data elements, it is possible to construct a picture of the quality of medical care. Although quality assessments based on administrative data cannot be definitive, they can be used to flag potential quality problems and success stories, which can then be further investigated and studied. Hospital associations, individual hospitals, purchasers, regulators, and policymakers at the local, State, and Federal levels can use readily available hospital administrative data to begin the assessment of quality of care. I In 2003, AHRQ first published the *National Healthcare Quality Report*[5] (NHQR) and *National Healthcare Disparities Report*[6] (NHDR) which provide a comprehensive picture of the level and variation of quality within four components of health care quality—effectiveness, safety, timeliness, and patient centeredness. These reports incorporated many Prevention Quality Indicators, Inpatient Quality Indicators, and Patient Safety Indicators. Selected mortality and utilization indicators from the IQI module will be included in the next NHQR and NHDR reports.[7]

The AHRQ Quality Indicators are now being used for applications beyond quality improvement. Some organizations have used the AHRQ Quality Indicators to produce web based, comparative reports on hospital quality, such as the Texas Department of State Health Services[8] and the Niagara Coalition[9]. These organizations also supplied users with guidance on indicator interpretation. Other organizations have incorporated selected AHRQ QIs into pay for performance demonstration projects or similar programs, such as the Centers for Medicare and Medicaid Services (CMS)[10] and Anthem Blue Cross Blue Shield of Virginia where hospitals would be financially rewarded for performance. Guidance on these alternative uses of the AHRQ QIs is summarized in an AHRQ *Summary Statement on Comparative Reporting*[11] and accompanying publication titled G*uidance for Using the AHRQ Quality Indicators for Hospital-Level Public Reporting or Payment*[12].

---

[1]. http://www.cms.hhs.gov/NationalHealthExpendData/downloads/nheprojections2004-2014.pdf: Table 2 National Health Expenditure Amounts, and Annual Percent Change by Type of Expenditure: Selected Calendar Years 1998-2014.
[2]Strunk BC, Ginsburg PB, Gabel JR. Tracking Health Care Costs. Health Affairs, 26 September 2001 (Web exclusive).
[3]Institute of Medicine. To Err is Human: Building a Safer Health System. Kohn LT, Corrigan JM, Donaldson MS (eds.) Washington DC: National Academy Press, 2000.
[4]Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Committee of Quality of Care in America. Washington DC: National Academy Press, 2001.
[5] Agency for Healthcare Research and Quality. *National Healthcare Quality Report*. Rockville, MD, U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality, December 2003.
[6] Agency for Healthcare Research and Quality. *National Healthcare Disparities Report*. Rockville, MD, U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality, July 2003.
[7] The 2005 NHQR and NHDR reports are available at http://www.qualitytools.ahrq.gov/.
[8] Texas Center for Health Statistics. *Indicators of Inpatient Care in Texas Hospitals, 2003*. http://www.dshs.state.tx.us/THCIC/Publications/Hospitals/IQIReport2003/IQIReport2003.shtm. Accessed January 2006.
[9] Niagara Health Quality Coalition. *2005 New York State Hospital Report Card,* .http://www.myhealthfinder.com/newyork05/glancechoose.htm. Accessed January 2006.
[10] Centers for Medicare & Medicaid Services. *The Premier Hospital Quality Incentive Demonstration*. http://www.cms.hhs.gov/HospitalQualityInits/downloads/HospitalPremierFactSheet.pdf. Accessed January 2006.
[11] AHRQ Summary Statement on Hospital Public Reporting.

The Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicators (PSIs) are a tool that takes advantage of hospital administrative data. The PSIs represent the current state-of-the-art in measuring the safety of hospital care through analysis of inpatient discharge data.

New micropolitan statistical areas and updated metropolitan statistical areas were established by the federal Office of Management and Budget (OMB) circular 03-04 (last revised December 4, 2005). To reflect these changes, all PSI documentation now refers to Metro Area instead of MSA. The SAS and Windows software allows users to specify stratification by county level with U.S. Census FIPS or modified FIPS, or by Metro Area with OMB 1999 or OMB 2003 definition. The AHRQ QI Windows Application allows users to generate reports stratified by all four of these, as well as by State.

See the section "What Does this Guide Contain?" for more information.

1.1 What Are the Patient Safety Indicators?

The PSIs are a set of measures that can be used with hospital inpatient discharge data to provide a perspective on patient safety. Specifically, PSIs screen for problems that patients experience as a result of exposure to the healthcare system and that are likely amenable to prevention by changes at the system or provider level. These are referred to as complications or adverse events. PSIs are defined on two levels: the provider level and the area level.

- *Provider-level Indicators* provide a measure of the potentially preventable complication for patients who received their initial care and the complication of care within the same hospitalization. Provider-level Indicators include only those cases where a secondary diagnosis code flags a potentially preventable complication.

- *Area-level Indicators* capture all cases of the potentially preventable complication that occur in a given area (e.g., metropolitan service area or county) either during hospitalization or resulting in subsequent hospitalization. Area-level Indicators are specified to include principal diagnosis, as well as secondary diagnoses, for the complications of care. This specification adds cases where a patient's risk of the complication occurred in a separate hospitalization.

Three PSIs, 27 through 29, that measured 3rd-degree obstetric trauma have been removed. A new area-level PSI, Postoperative Hemorrhage or Hematoma, has been added as PSI #27.

---

http://www.qualityindicators.ahrq.gov/news/AHRQSummaryStatement.pdf.

[12] Remus D, Fraser I. *Guidance for Using the AHRQ Quality Indicators for Hospital-level Public Reporting or Payment*. Rockville, MD: Department of Health and Human Services, Agency for Healthcare Research and Quality; 2004. AHRQ Pub. No. 04-0086-EF. The document may be downloaded from the AHRQ Quality Indicator website at http://www.qualityindicators.ahrq.gov/documentation.htm.

The PSIs include the following Provider-level Indicators:

Patient Safety Indicators - Provider	PSI Number
Complications of Anesthesia	1
Death in Low-Mortality DRGs	2
Decubitus Ulcer	3
Failure to Rescue	4
Foreign Body Left During Procedure	5
Iatrogenic Pneumothorax	6
Selected Infections Due to Medical Care	7
Postoperative Hip Fracture	8
Postoperative Hemorrhage or Hematoma	9
Postoperative Physiologic and Metabolic Derangements	10
Postoperative Respiratory Failure	11
Postoperative Pulmonary Embolism or Deep Vein Thrombosis	12
Postoperative Sepsis	13
Postoperative Wound Dehiscence	14
Accidental Puncture or Laceration	15
Transfusion Reaction	16
Birth Trauma – Injury to Neonate	17
Obstetric Trauma – Vaginal with Instrument	18
Obstetric Trauma – Vaginal without Instrument	19
Obstetric Trauma – Cesarean Delivery	20

In addition, the following PSIs were modified into Area-level Indicators to assess the total incidence of the adverse event within geographic areas:

Patient Safety Indicators - Area	PSI Number
Foreign Body Left During Procedure	21
Iatrogenic Pneumothorax	22
Selected Infections Due to Medical Care	23
Postoperative Wound Dehiscence	24
Accidental Puncture or Laceration	25
Transfusion Reaction	26
Postoperative Hemorrhage or Hematoma	27

1.2 How Can the PSIs Be Used to Assess Patient Safety?

Widespread consensus exists that health care organizations can reduce patient injuries by improving the environment for safety—from implementing technical changes, such as electronic medical record systems, to improving staff awareness of patient safety risks. Clinical process interventions also have strong evidence for reducing the risk of adverse events related to a patient's exposure to hospital care.[2] PSIs, which are based on computerized hospital discharge abstracts from the AHRQ's Healthcare Cost and Utilization Project (HCUP), can be used to better prioritize and evaluate local and national initiatives. Analyses of these and similar inexpensive, readily available administrative data sets may provide a screen for potential medical errors and a method for monitoring trends over time. The following scenario illustrates one potential application of the PSIs.

Evaluating and Improving Quality of Care

A hospital association recognizes its member hospitals' need for information that can help them evaluate the quality of care they provide. There is significant interest in assessing, monitoring, and improving the safety of inpatient care. After learning about the AHRQ PSIs, the association decides to apply the indicators to the discharge abstract data submitted by individual hospitals. For each hospital, the association develops a report with graphic presentation of the risk-adjusted data to show how the hospital performs on each indicator compared to its peer group, the State as a whole, and other comparable States. National and regional averages from the AHRQ Healthcare Cost and Utilization Project (HCUP) database are also provided as additional external benchmarks. Three years of trend data are included to allow the hospital to examine any changing patterns in its performance.

One member hospital, upon receiving the report, convenes an internal work group comprised of clinicians and quality improvement professionals to review the information and identify potential areas for improvement. The hospital leadership is committed to performance excellence and providing a culture supportive of systems evaluation and redesign. To begin their evaluation, they apply the AHRQ software to their internal administrative data to distinguish those patients who experienced the complication or adverse event from those who did not. This step establishes the focus for chart review.

After the initial analysis of the administrative and clinical data, the work group meets with clinical departments involved in care of these patients. They begin an in-depth analysis of the system and processes of care. Through application of process improvement concepts, they begin to identify opportunities for improvement. After selection of their priority area (for example, reduction of postoperative complications), they begin work, including:

- Review and synthesize the evidence base and best practices from scientific literature.
- Work with the multiple disciplines and departments involved in care of surgical patients to redesign care based on best practices with an emphasis on coordination and collaboration.
- Evaluate information technology solutions.
- Implement performance measurements for improvement and accountability.
- Incorporate monitoring of performance measurements in the departmental and senior leadership meetings and include in the Board quality improvement reports.

1.3 What Does this Guide Contain?

This guide provides information that hospitals, State data organizations, hospital associations, and others can use to decide how to use the PSIs. First, it describes the origin of the entire family of AHRQ Quality Indicators. Second, it provides an overview of the methods used to identify, select, and evaluate the AHRQ PSIs. Third, the guide summarizes the PSIs specifically, describes strengths and limitations of the indicators, documents the evidence that links the PSIs to the quality of health care services, and then provides in-depth descriptions of each PSI. The section, "Using Different Types of QI Rates" provides guidance in interpreting the various rates that are calculated by the QI software.

The empirical performance values that were provided in Table 1 and listed along with each indicator in previous versions have been moved to a separate document, *Patient Safety Indicators Comparative Data*. The comparative data rates as well as the means and parameter reference files used by the PSI software, reflect analyses of the 2002-2004 SID. The 2002-2004 SID is the basis for DRGs with a less than 0.05% mortality rate that are used in PSI #2, Death in Low-mortality DRGs.

The document, *Patient Safety Indicators Technical Specifications* outlines the specific definitions of each PSI, with complete ICD-9-CM and DRG coding specifications. That document also incorporates the list of operating room procedure codes that was previously a separate document. Appendix A contains links to documents and tools that may be of interest to PSI users.

The list of major operating room ICD-9-CM procedure codes now contained in the *Patient Safety Indicators Technical Specifications* document is based on the AHRQ Procedure Classes that assign all ICD-9-CM procedure codes to one of four categories:

- Minor Diagnostic - Non-operating room procedures that are diagnostic (e.g., 87.03 CT scan of head)
- Minor Therapeutic - Non-operating room procedures that are therapeutic (e.g., 02.41 Irrigate ventricular shunt)
- Major Diagnostic - All procedures considered valid operating room procedures by the Diagnosis Related Group (DRG) grouper and that are performed for diagnostic reasons (e.g., 01.14 Open brain biopsy)
- Major Therapeutic - All procedures considered valid operating room procedures by the Diagnosis Related Group (DRG) grouper and that are performed for therapeutic reasons (e.g., 39.24 Aorta-renal bypass).

For the AHRQ PSIs, major operating room procedures are ICD-9-CM procedure codes in categories #3 (major diagnostic) and #4 (major therapeutic).

1.4 Support for Potential and Current Users of the AHRQ QIs

Technical assistance is available, through an electronic user support system monitored by the QI support team, to support users in their application of the PSI software. The same e-mail address may be used to communicate to AHRQ any suggestions for PSI enhancements, general questions, and any QI related comments you may have. AHRQ welcomes your feedback. The Internet address for user support and feedback is: support@qualityindicators.ahrq.gov. AHRQ also offers a listserv to keep you informed on the Quality Indicators (QIs). The listserv is used to announce any QI changes or updates, new tools and resources, and to distribute other QI related information. This is a free service. Sign-up information is available at the QI website at http://www.qualityindicators.ahrq.gov/signup.htm.

2.0 Origins and Background of the Quality Indicators

In the early 1990s, in response to requests for assistance from State-level data organizations and hospital associations with inpatient data collection systems, AHRQ developed a set of quality measures that required only the type of information found in routine hospital administrative data—diagnoses and major procedures, along with information on patient's age, gender, source of admission, and discharge status. These States were part of the Healthcare Cost and Utilization Project, an ongoing Federal-State-private sector collaboration to build uniform databases from administrative hospital-based data.

AHRQ developed these measures, called the HCUP Quality Indicators, to take advantage of a readily available data source—administrative data based on hospital claims—and quality measures that had been reported elsewhere.[13] The 33 HCUP QIs included measures for avoidable adverse outcomes, such as in-hospital mortality and complications of procedures; use of specific inpatient procedures thought to be overused, underused, or misused; and ambulatory care sensitive conditions.

Although administrative data cannot provide definitive measures of health care quality, they can be used to provide *indicators* of health care quality that can serve as the starting point for further investigation. The HCUP QIs have been used to assess potential quality-of-care problems and to delineate approaches for dealing with those problems. Hospitals with high rates of poor outcomes on the HCUP QIs have reviewed medical records to verify the presence of those outcomes and to investigate potential quality-of-care problems.[14] For example, one hospital that detected high utilization rates for certain procedures refined patient selection criteria for these procedures to improve appropriate utilization.

2.1 Development of the AHRQ Quality Indicators

Since the original development of the HCUP QIs, the knowledge base on quality indicators has increased significantly. Risk-adjustment methods have become more readily available, new measures have been developed, and analytic capacity at the State level has expanded considerably. Based on input from current users and advances to the scientific base for specific indicators, AHRQ funded a project to refine and further develop the original QIs. The project was conducted by the UCSF-Stanford EPC.

The major constraint placed on the UCSF-Stanford EPC was that the measures could require only the type of information found in hospital discharge abstract data. Further, the data elements required by the measures had to be available from most inpatient administrative data systems. Some State data systems contain innovative data elements, often based on additional information from the medical record. Despite the value of these record-based data elements, the intent of this project was to create measures that were based on a *common denominator discharge data set*, without the need for additional data collection. This was critical for two reasons. First, this constraint would result in a tool that could be used with any inpatient administrative data, thus making it useful to most data systems. Second, this would enable national and regional benchmark rates to be provided using HCUP data, since these benchmark rates would need to be calculated using the universe of data available from the States.

2.2 AHRQ Quality Indicator Modules

The work of the UCSF-Stanford EPC resulted in the *AHRQ Quality Indicators*, which are available as separate modules:

- **Prevention Quality Indicators**. These indicators consist of "ambulatory care sensitive conditions," hospital admissions that evidence suggests could have been avoided through

[13] Ball JK, Elixhauser A, Johantgen M, et al. *HCUP Quality Indicators, Methods, Version 1.1: Outcome, Utilization, and Access Measures for Quality Improvement*. (AHCPR Publication No. 98-0035). Healthcare Cost and Utilization project (HCUP-3) Research notes: Rockville, MD: Agency for Health Care Policy and Research, 1998.
[14] *Impact: Case Studies Notebook – Documented Impact and Use of AHRQ's Research*. Compiled by Division of Public Affairs, Office of Health Care Information, Agency for Healthcare Research and Quality.

high-quality outpatient care or that reflect conditions that could be less severe, if treated early and appropriately.

- **Inpatient Quality Indicators**. These indicators reflect quality of care inside hospitals and include inpatient mortality; utilization of procedures for which there are questions of overuse, underuse, or misuse; and volume of procedures for which there is evidence that a higher volume of procedures is associated with lower mortality.

- **Patient Safety Indicators**. These indicators focus on potentially preventable instances of complications and other iatrogenic events resulting from exposure to the health care system.

- **Pediatric Quality Indicators.** This module, available in January, 2006, contains indicators that apply to the special characteristics of the pediatric population.

The core of the Pediatric Quality Indicators (PedQIs) is formed by indicators drawn from the original three modules. Some of these indicators were already geared to the pediatric population (for example, IQI 4 – Pediatric Heart Surgery Volume). These indicators are being removed from the original modules.

Others were adapted from indicators that apply to both adult and pediatric populations. These indicators remain in the original module, but will apply only to adult populations.

3.0 Methods of Identifying, Selecting, and Evaluating the Quality Indicators

Since the literature surrounding PSIs is sparse, the project team used a variety of additional techniques to identify, select, and evaluate each indicator, including clinician panels, expert coders, and empirical analyses.

3.1 Step 1: Define the Concepts and the Evaluation Framework

In approaching the task of evaluating patient safety indicators based on administrative data, the project team developed a conceptual framework and standardized definitions of commonly used terms.

3.1.1 Standardized Definitions

In the literature, the distinctions between medical error, adverse events, complications of care, and other terms pertinent to patient safety are not well established and are often used interchangeably. In this report, the terms medical error, adverse events or complications, and similar concepts are defined as follows:

Case finding indicators. Indicators for which the primary purpose is to identify specific cases in which a medical error *may* have occurred, for further investigation.

Complication or adverse event. "An injury caused by medical management rather than by the underlying disease or condition of the patient."[15] In general, adverse events prolong the hospitalization, produce a disability at the time of discharge, or both. Used in this report, complication does not refer to the sequelae of diseases, such as neuropathy as a "complication" of diabetes. Throughout the report, "sequelae" is used to refer to these conditions.

Medical error. "The failure of a planned action to be completed as intended (i.e., error of execution) or the use of a wrong plan to achieve an aim (i.e., error of planning)." The definition includes errors committed by any individual, or set of individuals, working in a health care organization.[16]

Patient safety. "Freedom from accidental injury," or "avoiding injuries or harm to patients from care that is intended to help them." Ensuring patient safety "involves the establishment of operational systems and processes that minimize the likelihood of errors and maximizes the likelihood of intercepting them when they occur."[17]

Patient safety indicators. Specific quality indicators which also reflect the quality of care inside hospitals, but focus on aspects of patient safety. Specifically, PSIs screen for problems that patients experience as a result of exposure to the healthcare system, and that are likely amenable to prevention by changes at the system or provider level.

Preventable adverse event. An adverse event attributable to error is a "preventable adverse event."[18] A condition for which reasonable steps may reduce (but not necessarily eliminate) the risk of that complication occurring.

Quality. "Quality of care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional

[15] Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991;324(6):370-6.
[16] Institute of Medicine, 2000.
[17] Envisioning the National Health Care Quality Report. Washington, DC: Institute of Medicine; 2001.
[18] Brennan et al., 1991.

knowledge." In this definition, "the term *health services* refers to a wide array of services that affect health…(and) applies to many types of health care practitioners (physicians, nurses, and various other health professionals) and to all settings of care…"[19]

Quality indicators. Screening tools for the purpose of identifying potential areas of concern regarding the quality of clinical care. For the purpose of this report, we focus on indicators that reflect the quality of care inside hospitals. Quality indicators may assess any of the four system components of health care quality, including patient safety (see below), effectiveness (i.e., "providing services based on scientific knowledge to all who could benefit, and refraining from providing services to those not likely to benefit), patient centeredness, and timeliness (i.e., "minimizing unnecessary delays").[20]

Rate based indicators. Indicators for which the primary purpose is to identify the rate of a complication rather than to identify specific cases.

While the definitions above are intended to distinguish events that are less preventable from those that are more preventable, the difference is best described as a spectrum. To conceptualize this spectrum, the project team developed the following three categories of conditions:

1. **Conditions that could be either a comorbidity or a complication.** Conditions considered comorbidities (for example, congestive heart failure) are present on admission and are not caused by medical management; rather, they are due to the patient's underlying disease. It is extremely difficult to distinguish complications from comorbidities for these conditions using administrative data. As a result, these conditions were not considered in this report.

2. **Conditions that are likely to reflect medical error.** These conditions (for example, foreign body accidentally left during a procedure) are likely to have been caused by medical error. Most of these conditions appear infrequently in administrative data, and thus rates of events lack the precision to allow for comparisons between providers. However, these conditions may be the subject of case-finding indicators.

3. **Conditions that conceivably, but not definitively reflect medical error.** These conditions (for example, postoperative DVT or PE) represent a spectrum of preventability between the previous two categories—from those that are mostly unpreventable to those that are mostly preventable. Because of the uncertainty regarding the preventability of these conditions and the likely heterogeneity of cases with the condition, indicators using these conditions are less useful as case-finding indicators. However, examining the rate of these conditions may highlight potential areas of concern.

3.1.2 Evaluation Framework

To evaluate the soundness of each indicator, the project team applied the same framework as was applied in the technical report[21] for the Prevention Quality Indicators (PQIs) and Inpatient Quality Indicators (IQIs), available at http://www.qualityindicators.ahrq.gov/downloads.htm. This included six areas of evidence:

- **Face validity.** Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control? Consensual validity expands face validity beyond one person to the opinion of a panel of experts.

[19] Measuring the Quality of Health Care: A statement of the National Roundtable on Healthcare Quality Division of Healthcare Services: National Academy Press; 1999.
[20] National Roundtable on Healthcare Quality, 1999.
[21] Davies S, Geppert J, McClellan M, McDonald KM, Romano PS, Shojania KG. Refinement of the HCUP Quality Indicators. Technical Review Number 4. Rockville, MD: (Prepared by UCSF-Stanford Evidence-based Practice Center under Contract No. 290-97-0013) Agency for Healthcare Research and Quality; 2001. Report No.: 01-0035.

- **Precision.** Is there a substantial amount of provider- or community-level variation that is not attributable to random variation?
- **Minimum bias.** Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?
- **Construct validity.** Does the indicator perform well in identifying true (or actual) quality of care problems?
- **Fosters real quality improvement.** Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?
- **Application.** Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

Face validity (consensual validity) was evaluated using a structured panel review, minimum bias was explored empirically and briefly during the panel review, and construct validity was evaluated using the limited literature available. A full discussion of this framework is available in the Stanford Technical report[22] available at http://www.qualityindicators.ahrq.gov/downloads.htm.

The relative importance of each of these evaluation areas may differ by individual PSIs. Precision and minimum bias may be less important for indicators that are primarily designed to screen only for medical error, since these events are relatively rare. In general, these indicators are better used as case-finding indicators. For these indicators, comparisons between rates are less relevant. However, for rate-based indicators, concerns of precision and minimum bias remain if indicators are used in any comparison of rates (comparison to national averages, peer group, etc.).

3.2 Step 2: Search the Literature to Identify Potential PSIs

The literature searches performed in connection with assessing potential AHRQ QIs[23] identified many references relevant to potential PSIs. In addition, the project team performed electronic searches for articles published before February 2002 followed by hand searching the bibliographies of identified references. Members of the project team were queried to supplement this list, based on their personal knowledge of recent work in the field. Because Iezzoni et al.'s Complications Screening Program (CSP)[24] included numerous candidate indicators, the team also performed an author search using her name. Forthcoming articles and Federal reports in press, but not published, were also included when identified through personal contacts.

The project team identified 326 articles from the Medline search. Articles were screened using both the titles and abstracts. To qualify for abstraction, an article must have described, evaluated, or validated a potential indicator of medical errors, patient safety, or potentially preventable complications based on International Classification for Diseases - Ninth Revision - Clinical Modifications (ICD-9-CM) coded administrative (hospital discharge or claims) data. Some indicators were also considered if they appeared to be readily translated into ICD-9-CM, even if the original authors did not use ICD-9-CM codes.

[22] McDonald KM, Romano PS, Geppert J, Davies S, Duncan BW, Shojania KG. Measures of Patient Safety Based on Hospital Administrative Data-The Patient Safety Indicators. Technical Review 5 (Prepared by the University of California San Francisco-Stanford Evidence-based Practice Center under Contract No. 290-97-0013). AHRQ Publication No. 02-0038 . Rockville, MD: Agency for Healthcare Research and Quality. August 2002.
[23] McDonald et al., 2002.
[24] Iezzoni LI, Foley SM, Heeren T, Daley J, Duncan CC, Fisher ES, et al. A method for screening the quality of hospital care using administrative data: preliminary validation results. QRB Qual Rev Bull 1992;18(11):361-71.

This search was adapted slightly and repeated using the OVID interface with EMBASE[25], limited to articles published from January 1990 through the end of first quarter 2002. The EMBASE search identified 463 references, and these articles were screened in the same manner. After elimination of articles that had already been identified using Medline[26] and the other approaches described above, only nine additional articles met the criteria for abstraction.

3.3 Step 3: Develop a Candidate List of PSIs

The project team developed a candidate list of PSIs by first reviewing the literature, then selecting a subset of indicators to undergo face validity testing by clinician panels.

3.3.1 Candidate List of PSIs

The literature search located relatively few patient safety indicators that could be defined using unlinked administrative data. The majority of these indicators were from the Complications Screening Program (CSP),[27] which was developed to identify potentially preventable complications of adult medical and surgical hospital care using commonly available administrative data. The algorithm uses discharge abstract data—specifically ICD-9-CM diagnosis and procedure codes, patient age, sex, diagnosis-related group (DRG), and date of procedure—to identify 28 complications that raise concern about the quality of care based on the rate of such occurrences at individual hospitals. Each of the complications is applied to some or all of the following specified "risk pools" separately: major surgery, minor surgery, invasive cardiac procedure, endoscopy, medical patients, and all patients. In addition, specified inclusion and exclusion criteria are applied to each complication to ensure that the complication developed in-hospital, as opposed to being present on admission, and that the complication was potentially preventable.

Four later studies were designed to test criterion and construct validity by validating the data used to construct CSP screens, validating the screens as a flag for actual quality problems, and validating the replicability of hospital-level results using different data sources.[28] [29] [30] [31] These studies raised concerns about the validity of the CSP, because flagged cases for most indicators were no more likely than unflagged controls to have suffered explicit process failures.

The project team also reviewed all ICD-9-CM codes implemented in or before 1999 that were identified by AHRQ as possibly describing medical errors or reflecting the consequences of such errors.[32] (This initial set of indicators is referred to as the Miller et al. indicators.) The project team added relevant codes from the 2000 and 2001 revisions of ICD-9-CM and selected codes from the CSP, such as those not clearly reflective of medical error, but representing a potentially preventable complication. This process was guided principally by conceptual considerations. For example, codes for postoperative AMI (an evaluated indicator that was not included in the final indicator set) were included in the evaluation set since recent evidence suggests that AMI is a potentially preventable complication.[33] A few codes were also deleted from the initial list based on a review of ICD-9-CM coding guidelines, described in *Coding Clinics for ICD-*

[25] EMBASE. In. The Netherlands: Elsevier Science Publishers B.V.
[26] MEDLINE [database online]. In. Bethesda (MD): National Library of Medicine.
[27] Iezzoni et al., 1992.
[28] Lawthers A, McCarthy E, Davis R, Peterson L, Palmer R, Iezzoni L. Identification of in-hospital complications from claims data: is it valid? Medical Care 2000;38(8):785-795.
[29] McCarthy EP, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamael MB, et al. Does clinical evidence support ICD-9-CM diagnosis coding of complications? Med Care 2000;38(8);868-876.
[30] Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: validation of the complications screening program Med Care 2000;38(8):796-806.
[31] Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.
[32] Miller M, Elixhauser A, Zhan C, Meyer G. Patient Safety Indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.
[33] Shojania KG, Duncan BW, McDonald KM, Wachter RM. Making health care safer: A critical analysis of patient safety practices. Evidence Report/Technology Assessment No. 43 (Prepared by the University of California at San Francisco-Stanford Evidence-based Practice Center under Contract No. 290-97-0013). Rockville, MD: Agency for Healthcare Research and Quality; 2001. Report No.: AHRQ Publication No. 01-E058.

9-CM and the *American Hospital Association's ICD-9-CM Coding Handbook*. For example, the code 2593 for hypoglycemic coma specifically excludes patients with diabetes mellitus, the population for which this complication is most preventable. This process of updating the Miller et al. PSIs resulted in a list of over 200 ICD-9-CM codes (valid in 2001) potentially related to medical error.

Codes identified in the CSP and updated from the Miller et. al. PSIs were then grouped into indicators. Where feasible, codes were compiled as they were in the CSP, or in some cases the Miller et al. PSIs, depending on which grouping yielded more clinically homogeneous groups. In most cases the resulting indicators were not identical to the CSP indicators, although they were closely related, as some of the specific codes included in the original CSP had been eliminated after the team's review of coding guidelines. The remaining codes were then incorporated into the most appropriate CSP-based indicator, or were grouped into clinically meaningful concepts to define novel indicators. Exclusion criteria were added based on CSP methods and clinical judgment. As a result, over 40 patient safety indicators were defined that, while building on prior work, reflected significantly changed measures to focus more narrowly on the most preventable complications.

Indicators were defined with both a numerator (complication of interest) and a denominator (population at risk). Different patient subpopulations have inherently different risks for developing a complication, with some patients having almost no risk. Thus, the denominator for each indicator represents the specific population at risk. The intention was to restrict the complication (and consequently the rate) to a more homogeneous population who are actually at risk for that complication. In general, the population at risk corresponded to one risk pool (e.g., major surgery) from the CSP, if applicable, or was defined more narrowly.

3.3.2 Subset Selection

After the project team developed a list of potential indicators, they selected a subset of indicators to undergo face validity testing by clinician panels, as described in Step 4. Two sources of information guided the selection process.

First, validation data from previous studies were reviewed and thresholds were set for retaining CSP-based indicators. Four studies were identified that evaluated the CSP indicators. Three of these studies, examined the predictive value of each indicator in identifying a complication that occurred in-hospital, regardless of whether this complication was due to medical error or was preventable.[34] [35] [36] In a fourth study, nurses identified specific process failures that may have contributed to complications. In order to be retained as a potential PSI, at least one of the first three studies needed to demonstrate a positive predictive value of at least 75%, meaning that 3 out of 4 patients identified by the measure did indeed have the complication of interest.[37] In addition, the positive predictive value of a "process failure" identified in the fourth study needed to reach or exceed 46%, which was the average rate for surgical cases that were not flagged by any of the CSP indicators. As a result, only CSP-derived indicators that were at least somewhat predictive of objectively defined process failures or medical errors were retained.

Second, specific changes to previous definitions or constructs of indicators fell into the following general categories:

- Changes to the denominator definitions (inclusion or exclusion criteria), intended to reduce bias due to the inclusion of atypical patients or to improve generalizability to a broader set of patients at risk.
- Elimination of selected ICD-9-CM codes from numerator definitions, intended to focus attention on more clinically significant complications or complications more likely to result from medical errors.
- Addition of selected ICD-9-CM codes to numerator definitions, intended to capture related

[34] Lawthers, et al., 2000.
[35] McCarthy, et al., 2000.
[36] Weingart et al., 2000.
[37] Iezzoni et al., 1999.

complications that could result from the same or similar medical errors.
- Division of a single indicator into two or more related indicators, intended to create more clinically meaningful and conceptually coherent indicators.
- Stratification or adjustment by relevant patient characteristics, intended to reflect fundamental clinical differences among procedures (e.g., vaginal delivery with or without instrumentation) and the complications that result from them, or fundamental differences in patient risk (e.g., decubitus ulcer in lower-risk versus high-risk patients).

A total of 34 indicators, intended to be applied to all age groups, were retained for face validity testing by clinician panels. Because the primary intent in developing these indicators was to detect potentially preventable complications related to health care exposure, the final definitions for this set of indicators represented mostly new measures that built upon previous work.

3.3.3 Coding Review

Experts in ICD-9-CM codes reviewed each code for accuracy of capturing the complication and population at risk. In some cases, additional codes or other refinements to the indicators were suggested based on current coding guidelines.

3.4 Step 4: Review the PSIs

The project team conducted a structured review of each indicator to evaluate the face validity (from a clinical perspective) of the indicators. The methodology for the structured review was adapted from the RAND/UCLA Appropriateness Method[38] and consisted of an initial independent assessment of each indicator by clinician panelists using an initial questionnaire, a conference call among all panelists, followed by a final independent assessment by clinician panelists using the same questionnaire. The review sought to establish *consensual validity,* which "extends face validity from one expert to a panel of experts who examine and rate the appropriateness of each item…."[39] The panel process served to refine definitions of some indicators, add new measures, and dismiss indicators with major concerns from further consideration.

Eight panels were formed: two panels examined complications of medical care indicators, three panels examined surgical complications indicators, one panel assessed indicators related to procedural complications, and two panels examined obstetric complications indicators.

Fifteen professional clinical organizations nominated a total of 162 clinicians to be panelists. To be eligible to participate, nominees were required to spend at least 30% of their work time on patient care, including hospitalized patients. Nominees were asked to provide information regarding their practice characteristics, including specialty, subspecialty, and setting. Fifty-seven panelists were selected to ensure that each panel had diverse membership in terms of practice characteristics and setting.

3.4.1 Initial Assessment of the Indicators

Panelists were presented with four or five indicators, including the standardized text used to describe each ICD-9-CM code, the specific numeric code, exclusion and inclusion criteria, the clinical rationale for the indicator, and the specification criteria. For each indicator, panelists completed a 10-item questionnaire that evaluated the ability of the indicator to screen out conditions present on admission, the potential preventability of the complication, and the ability of the indicator to identify medical error. In addition, the questionnaire asked panelists to consider potential bias, reporting or charting problems, potential for gaming the indicator, and adverse effects of implementing the indicator. Finally, the questionnaire provided an opportunity for panelists to suggest changes to the indicator.

---

[38] Fitch K, Bernstein J, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. the RAND/UCLA Appropriateness Method User's Manual: RAND; 2001.
[39] Green L, Lewis F. measurement and Evaluation in Health Education and Health Promotion. Mountain View, CA: Mayfield Publishing Company; 1998.

3.4.2 Conference Call Participation

After the panelists submitted the initial evaluation questionnaires, they participated in a 90-minute conference call for their panel to discuss the indicators. In general, agenda items for the conference call focused on points of disagreement among panelists. However, panelists were explicitly told that consensus was not the goal of discussion. In some cases, panelists agreed on proposed changes to the indicator definitions, and such consensus was noted and the definition was modified accordingly before the final round of rating.

Panelists were prompted throughout the process to consider the appropriate population at risk for each indicator (specifically inclusion and exclusion criteria) in addition to the complication of interest. However, if panelists wished to discuss other aspects of the indicator, this discussion was allowed within the time allotted for that indicator (approximately 15 minutes). If time remained at the end of a call, topics that were not fully addressed previously were revisited.

3.4.3 Final Evaluation and Tabulation of Results

Following each conference call, the project team made changes to each indicator suggested by panelists for changes that reached near consensus of the panelists. The indicators were then redistributed to panelists with the questionnaires used in the initial evaluation. The reason for all each indicator definition change was included, and panelists were asked to re-rate the indicator based on their current opinion. They were asked to keep in mind the discussion during the conference call.

Results from the final evaluation questionnaire were used to calculate median scores from the 9-point scale for each question and to categorize the degree of agreement among panelists. Median scores determined the level of acceptability of the indicator, and dispersion of ratings across the panel for each applicable question determined the agreement status. Therefore the median and agreement status were independent measurements for each question. Six criteria were used to identify the panel opinions (i.e., median, agreement status category) on the following aspects of the indicator:

- Overall usefulness of the indicator.
- Likelihood that the indicator measures a complication and not a comorbidity (specifically, present on admission).
- Preventability of the complication.
- Extent to which the complication is due to medical error.
- Likelihood that the complication is charted given that it occurs.
- Extent that the indicator is subject to bias (systematic differences, such as case mix that could affect the indicator, in a way not related to quality of care).

The project team used the ratings of the overall appropriateness of each indicator to assess its overall usefulness as a screen for potential patient safety problems. Indicators were triaged into three sets: Accepted Indicators (described in this guide), Experimental Indicators, and Rejected Indicators.

3.5 Step 5: Evaluate the PSIs Using Empirical Analysis

The project team conducted empirical analyses to explore the frequency and variation of the indicators, the potential bias, based on limited risk adjustment, and the relationship between indicators. The data sources used in the empirical analyses were the 1997 Florida State Inpatient Database (SID) for initial testing and development and the 1997 HCUP State Inpatient Database for 19 States (referred to in this guide as the HCUP SID) for the final empirical analyses.

All potential indicators were examined empirically by developing and conducting statistical tests for precision, bias, and relatedness of indicators. Three different estimates of hospital performance were calculated for each indicator:

1. The raw indicator rate was calculated using the number of adverse events in the numerator divided by the number of discharges in the population at risk by hospital.

2. The raw indicator was adjusted to account for differences among hospitals in age, gender, modified DRG, and comorbidities.

    - Adjacent DRG categories that were separated by the presence or absence of comorbidities or complications were collapsed to avoid adjusting for the complication being measured. Most of the super-Major Diagnostic Category (MDC) DRG categories were excluded for the same reason.
    - APR-DRG risk adjustment was not implemented because removing applicable complications from each indicator was beyond the scope of this project.
    - The ICD-9-CM codes used to define comorbidity categories were modified to exclude conditions likely to represent potentially preventable complications in certain settings.
    - "Acute on chronic" comorbidities were captured so that some patients with especially severe comorbidities would not be mislabeled as not having conditions of interest.
    - Comorbidities in obstetric patients were added.

3. Multivariate signal extraction methods were applied to adjust for reliability by estimating the amount of "noise" (i.e., variation due to random error) relative to the amount of "signal" (i.e., systematic variation in hospital performance or reliability) for each indicator.

Similar reliability adjustment has been used in the literature for similar purposes.[40] [41] The project team constructed a set of statistical tests to examine precision, bias, and relatedness of indicators for all accepted Provider-level Indicators, and precision and bias for all accepted Area-level Indicators. It should be noted that rates based on fewer than 30 cases in the numerator or the denominator are not reported. This exclusion rule serves two purposes:

- It eliminates unstable estimates based on too few cases.
- It helps protect the identities of hospitals and patients.

[40] Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease JAMA 1999;281(22):2098-105.
[41] Christiansen CL, Morris CN. Improving the statistical approach to health care provider profiling. Ann Intern Med 1997;127(8 Pt 2):764-8.

4.0 Summary Evidence on the Patient Safety Indicators

This project took a four-pronged approach to the identification, development, and evaluation of PSIs that included use of literature, clinician panels, expert coders, and empirical analyses. The literature review and the findings from the clinical panels combined with data analysis provide evidence to suggest that a number of discharge-based PSIs may be useful screens for organizations, purchasers, and policymakers to identify safety problems at the provider level, as well as to document systematic area-level differences in patient safety problems.

Most adverse events identified by the PSIs have a variety of causes in addition to potential medical error leading to the adverse event, including underlying patient health and factors that do not vary systematically. Clinician panelists rated only two of the accepted indicators as very likely to reflect medical error: (1) transfusion reaction and (2) foreign body left in during a procedure. These indicators proved to be very rare, with less than 1 per 10,000 cases at risk.

Table 1 summarizes the results of the literature review, clinician panels, and empirical analyses on the provider-level PSIs. The table lists each indicator, provides its definition, identifies any concerns about its validity based on the clinician panels, and summarizes the strength of evidence in the literature for each indicator.

The following notes about some of the terms in the table are intended to help the reader understand the context in which they are used.

Validity Concerns. The following concerns, raised during our panel review, are listed if they affect the validity of the particular indicator:

- **Rare** —This indicator is relatively rare and may not have adequate statistical power for some providers.
- **Condition definition varies**—This indicator includes conditions for which diagnosis may be subjective, depending on the threshold of the physician, and patients with the same clinical state may not have the same diagnosis.
- **Underreporting or screening**—Conditions included in this indicator may not be systematically reported (leading to an artificially low rate) or may be routinely screened for (leading to a higher rate in facilities that screen).
- **Adverse consequences**—Use of this indicator may have undesirable effects, such as increasing inappropriate antibiotic use.
- **Stratification suggested**—This indicator includes some high risk patient groups and stratification is recommended when examining rates,
- **Unclear preventability**—As compared to other PSIs, the conditions included in this indicator may be less preventable by the health system.
- **Heterogeneous severity**—This indicator includes codes that encompass several levels of severity of a condition that cannot be ascertained by the codes.
- **Case mix bias**—This indicator was felt to be particularly subject to systematic bias, and DRG and comorbidity risk adjustment may not adequately address the concern.
- **Denominator unspecific**—The denominator for this indicator is less than ideal, because the true population at risk could not be identified using ICD-9-CM codes. Some patients are likely included who are not truly at risk, or some patients who are at risk are not included.

Strength of Evidence. The following key findings represent a review of the limited literature assessing the validity of the indicators:

- **Coding**—Sensitivity is the proportion of patients who suffered an adverse event, based on detailed chart review or prospective data collection, for whom that event was coded on a discharge abstract or Medicare claim. Predictive value is the proportion of patients with a coded adverse event who were confirmed as having suffered that event, based on detailed chart review or prospective data collection.

Construct, explicit process—Adherence to specific, evidence-based or expert-endorsed processes of care, such as appropriate use of diagnostic modalities and effective therapies. The construct is that hospitals that provide better processes of care should experience fewer adverse events.

Construct, implicit process—Adherence to the "standard of care" for similar patients, based on global assessment of quality by physician chart reviewers. The construct is that hospitals that provide better overall care should experience fewer adverse events.

Construct, staffing—The construct is that hospitals that offer more nursing hours per patient day, better nursing skill mix, better physician skill mix, or more experienced physicians should have fewer adverse events.

The following distinctions were used to summarize the strength of the published evidence for each indicator:

- Published evidence suggests that the indicator lacks validity in this domain (i.e., less than 50% sensitivity or predictive value; explicit or implicit process failure rates no more frequent than among control patients).

0 No published evidence regarding this domain of validity.

± Published evidence suggests that the indicator may be valid in this domain, but different studies offer conflicting results (although study quality may account for these conflicts).

+ Published evidence suggests that the indicator is valid, or is likely to be valid, in this domain (i.e., one favorable study).

++ There is strong evidence supporting the validity of this indicator in this domain (i.e., multiple studies with consistent results, or studies showing both high sensitivity and high predictive value). When content validity is exceptionally high, as for transfusion reaction or iatrogenic pneumothorax, construct validity becomes less important.

A complete description of each PSI is included later in the guide under "Detailed Evidence for Patient Safety Indicators" and in the document *Patient Safety Indicators Technical Specifications.* See Appendix A.

Table 1: AHRQ Provider-Level Patient Safety Indicators

PSI Name	Definition	Validity Concerns	Strength of Evidence
Complications of Anesthesia (PSI 1)	Cases of anesthetic overdose, reaction, or endotrachial tube misplacement per 1,000 surgery discharges. Excludes codes for drug use and self-inflicted injury.	Condition definition varies Underreporting or screening Denominator unspecific	0 Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Death in Low Mortality DRGs (PSI 2)	In-hospital deaths per 1,000 patients in DRGs with less than 0.5% mortality.[a] Excludes trauma, immuno-compromised, and cancer patients.	Heterogeneous severity	+ Coding 0 Explicit Process + Implicit Process 0 Staffing
Decubitus Ulcer (PSI 3)	Cases of decubitus ulcer per 1,000 discharges with a length of stay of 5 or more days. Excludes patients with paralysis or in MDC 9, MDC 14, and patients admitted from a long-term care facility.	Underreporting or screening Heterogeneous severity Case mix bias	– Coding 0 Explicit Process 0 Implicit Process ± Staffing

PSI Name	Definition	Validity Concerns	Strength of Evidence
Failure to Rescue (PSI 4)	Deaths per 1,000 patients having developed specified complications of care during hospitalization. Excludes patients age 75 and older, neonates in MDC 15, patients admitted from long-term care facility and patients transferred to or from other acute care facility.	Adverse consequences Stratification suggested Unclear preventability Heterogeneous severity	+ Coding 0 Explicit Process 0 Implicit Process ++ Staffing
Foreign Body Left During Procedure (PSI 5)	Discharges with foreign body accidentally left in during procedure per 1,000 discharges	Rare Stratification suggested Denominator unspecific	0 Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Iatrogenic Pneumothorax (PSI 6)	Cases of iatrogenic pneumothorax per 1,000 discharges. Excludes trauma, thoracic surgery, lung or pleural biopsy, or cardiac surgery patients, and MDC 14.	Denominator unspecific	0 Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Selected Infections Due to Medical Care (PSI 7)	Cases of secondary ICD-9-CM codes 9993 or 00662 per 1,000 discharges. Excludes patients with immunocompromised state or cancer.	Underreporting or screening Adverse consequences	0 Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Postoperative Hip Fracture (PSI 8)	Cases of in-hospital hip fracture per 1,000 surgical discharges. Excludes patients in MDC 8, with conditions suggesting fracture present on admission and MDC 14.	Case mix bias Denominator unspecific	+ Coding + Explicit Process + Implicit Process 0 Staffing
Postoperative Hemorrhage or Hematoma (PSI 9)	Cases of hematoma or hemorrhage requiring a procedure per 1,000 surgical discharges. Excludes MDC 14.	Stratification suggested Case mix bias Denominator unspecific	± Coding ± Explicit Process + Implicit Process 0 Staffing
Postoperative Physiologic and Metabolic Derangement (PSI 10)	Cases of specified physiological or metabolic derangement per 1,000 elective surgical discharges. Excludes patients with principal diagnosis of diabetes and with diagnoses suggesting increased susceptibility to derangement. Excludes obstetric admissions.	Condition definition varies	– Coding 0 Explicit Process 0 Implicit Process – Staffing
Postoperative Respiratory Failure (PSI 11)	Cases of acute respiratory failure per 1,000 elective surgical discharges. Excludes MDC 4 and 5 and obstetric admissions.	Unclear preventability Case mix bias	+ Coding ± Explicit Process + Implicit Process ± Staffing
Postoperative PE or DVT (PSI 12)	Cases of deep vein thrombosis or pulmonary embolism per 1,000 surgical discharges. Excludes obstetric patients.	Underreporting or screening Stratification suggested	+ Coding + Explicit Process + Implicit Process ± Staffing

PSI Name	Definition	Validity Concerns	Strength of Evidence
Postoperative Sepsis (PSI 13)	Cases of sepsis per 1,000 elective surgery patients, with length of stay more than 3 days. Excludes principal diagnosis of infection, or any diagnosis of immunocompromised state or cancer, and obstetric admissions.	Condition definition varies Adverse consequences	± Coding 0 Explicit Process 0 Implicit Process – Staffing
Postoperative Wound Dehiscence (PSI 14)	Cases of reclosure of postoperative disruption of abdominal wall per 1,000 cases of abdominopelvic surgery. Excludes obstetric admissions.	Case mix bias	0 Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Accidental Puncture or Laceration (PSI 15)	Cases of technical difficulty (e.g., accidental cut or laceration during procedure) per 1,000 discharges. Excludes obstetric admissions.	Underreporting or screening Unclear preventability	± Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Transfusion Reaction (PSI 16)	Cases of transfusion reaction per 1,000 discharges.	Rare Stratification suggested	0 Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Birth Trauma— Injury to Neonate (PSI 17)	Cases of birth trauma, injury to neonate, per 1,000 liveborn births. Excludes some preterm infants and infants with osteogenic imperfecta.	Condition definition varies Unclear preventability Heterogeneous severity	– Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Obstetric Trauma— Vaginal Delivery with Instrument (PSI 18)	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 instrument-assisted vaginal deliveries.	Unclear preventability Case mix bias	+ Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Obstetric Trauma— Vaginal Delivery without Instrument (PSI 19)	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 vaginal deliveries without instrument assistance.	Unclear preventability Case mix bias	+ Coding 0 Explicit Process 0 Implicit Process 0 Staffing
Obstetric Trauma— Cesarean Delivery (PSI 20)	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 Cesarean deliveries.	Unclear preventability Case mix bias	+ Coding 0 Explicit Process 0 Implicit Process 0 Staffing

[a] DRGs that are divided into "with complications and comorbidities" and "without complications and comorbidities" are only included if both divisions have mortality rates below 0.5%.

4.1 Limitations in Using the PSIs

Many important concerns cannot currently be monitored well using administrative data, such as adverse drug events, and using these data tends to favor specific types of indicators. For example, the PSIs evaluated in this report contain a large proportion of surgical indicators, rather than medical or psychiatric, because medical complications are often difficult to distinguish from comorbidities that are present on admission. In addition, medical populations tend to be more heterogeneous than surgical, especially elective surgical populations, making it difficult to account for case-mix. Panelists often expressed that indicators were more applicable to patient safety when limited to elective surgical admissions. However, the careful use of administrative data holds promise for screening to target further data collection and

analysis. The ability to assess all patients at risk for a particular patient safety problem, along with the relative low cost, are particular strengths of these data sets.

Two broad areas of concern also hold true for these data sets.

1. Questions about the clinical accuracy of discharge-based diagnosis coding lead to concerns about the interpretation of reported diagnoses that may represent safety problems. Specifically:

    - Administrative data are unlikely to capture all cases of a complication, regardless of the preventability, without false positives and false negatives (sensitivity and specificity).

    - When the codes are accurate in defining an event, the clinical vagueness inherent in the description of the code itself (e.g., "hypotension"), may lead to a highly heterogeneous pool of clinical states represented by that code.

    - Incomplete reporting is an issue in the accuracy of any data source used for identifying patient safety problems, as medical providers might fear adverse consequences as a result of "full disclosure" in potentially public records such as discharge abstracts.

2. The information about the ability of these data to distinguish adverse events in which no error occurred from true medical errors is limited. A number of factors—such as the heterogeneity of clinical conditions included in some codes, lack of information about event timing available in these data sets, and limited clinical detail for risk adjustment—contribute to the difficulty in identifying complications that represent medical error or may be at least in some part preventable.

These factors may exist for other sources of patient safety data as well. For example, they have been raised in the context of the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) implementation of a "sentinel event" program geared at identifying serious adverse events that may be related to underlying safety problems.

4.2 Further Research on PSIs

The initial validation evaluations reviewed and performed for the PSIs leave substantial room for further research with detailed chart data and other data sources. Future validation work should focus on the following:

- The sensitivity and specificity of these indicators in detecting the occurrence of a complication.
- The extent to which failures in processes of care at the system or individual level are detected using these indicators.
- The relationship of these indicators with other measures of quality, such as mortality.
- Further explorations of bias and risk adjustment.

Enhancements to administrative data are worth exploring in the context of further validation studies that use data from other sources. For example, as with other quality indicators, the addition of timing variables may prove particularly useful in identifying whether a complication was present on admission, or whether it occurred during the hospitalization. While some of the complications that are present on admission may indeed reflect adverse events of care in a previous hospitalization or outpatient care, many may reflect comorbidities instead of complications. A second example area—linking hospital data over time and with outpatient data and other hospitalizations—would allow inclusion of complications that occur after discharge and likely would increase the sensitivity of the PSIs.

4.3 Use of External Cause-of-Injury Codes

Several of the PSIs are based on capturing external cause-of-injury (e-code) data. These codes are used to classify environmental events, circumstances, and conditions as the cause of injury, poisoning, or other adverse events. External cause-of-injury codes are critical to evaluate population-based, cause-specific data on nonfatal injuries at the state and local levels. However, not all states collect this information in their hospital discharge data programs nor do all state uniform billing committees require use of e-codes. Users of the PSIs should be knowledgeable of the e-code requirements and practices of hospitals represented in the input data file.

Table 2 provides a summary of the PSIs that are dependent on e-codes for their definition (required), the PSIs that use e-codes within their definition, and the PSIs that do not use any e-codes in their definition. If use of e-codes is not mandated or coding may be highly variable across hospitals, the PSIs that are dependent upon e-codes should not be used and the PSIs that include e-codes in their definition should be used with caution.

Table 2: Indicators and Use of External Cause-of-Injury Codes

Indicator Number (used in software)	Indicator Name	Use of External Cause-of-Injury Codes
15 & 25	Accidental Puncture or Laceration	Required. Used in both the numerator and denominator definitions.
17	Birth Trauma	Not used.
1	Complications of Anesthesia	Required. Used in the numerator definition.
2	Death in Low Mortality DRGs	Not used.
3	Decubitus Ulcer	Not used.
4	Failure to Rescue	Not used.
5 & 21	Foreign Body Left During Procedure	Required. Used in the numerator definition although the other ICD-9 CM codes may capture the same information.
6 & 22	Iatrogenic Pneumothorax	Not used.
20	Obstetric Trauma – Cesarean Section	Not used.
18	Obstetric Trauma – Vaginal with Instrument	Not used.
19	Obstetric Trauma – Vaginal without Instrument	Not used.
9	Postoperative Hemorrhage or Hematoma	Not used.
8	Postoperative Hip Fracture	Used as exclusion criteria in denominator population.
10	Postoperative Physiologic and Metabolic Derangements	Not used.
12	Postoperative Pulmonary Embolism or Deep Vein Thrombosis	Not used.
11	Postoperative Respiratory Failure	Not used.
13	Postoperative Sepsis	Not used.
14 & 24	Postoperative Wound Dehiscence	Not used.
7 & 23	Selected Infections Due to Medical Care	Not used.
16 & 26	Transfusion Reaction	Required. Used in the numerator definition although the other ICD-9 CM codes may capture the same information.
27	Postoperative Hemorrhage or Hematoma	Not used.

5.0 Detailed Evidence for Patient Safety Indicators

This section provides an abbreviated presentation of the details of the literature review and the empirical evaluation for each PSI, including:

- The definition of the indicator
- The outcome of interest (or numerator)
- The population at risk (or denominator)
- The type of indicator
- The measures of empirical performance. Rates are population rates as reported in Table 1 (PSI – Provider) and Table 2 (PSI – Area). Provider rates are per 1,000 qualifying discharges and Area rates are per 100,000 population.

The two-page descriptions for each indicator also include a more detailed discussion of the panel review, the literature review, the source of the indicator, and the results of the empirical analysis, including information related to adjustments to increase the robustness of the rates:

- Reliability. Statistics on the signal standard deviation, signal share, and signal ratio were used to examine the effect of the reliability adjustment. Multivariate methods were applied to most of the indicators, and overall the reliability adjustment reduced the provider-level variation dramatically. In general, indicators with higher rates tend to perform better on tests of reliability; as a result, obstetric indicators with high rates tend to do very well relative to other indicators.

- Bias. The effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals — compared to no risk adjustment —was assessed, if applicable. The presence of high bias suggests that risk adjustment, using administrative data elements, is necessary to interpret provider-level differences in the rates of these indicators.

A full report on the literature review and empirical evaluation can be found in *Measures of Patient Safety Based on Hospital Administrative Data — The Patient Safety Indicators* by the UCSF-Stanford EPC,

Detailed coding information for each PSI is provided in the document *Patient Safety Indicators Technical Specifications*. The software manuals *Patient Safety Indicators: SAS Software Documentation* and *QI Windows Software Documentation* provide detailed instructions on how to use the PSI software including data preparation, calculation of the PSI rates, and interpretation of output.

In the SAS version of the software, all provider level indicators are expressed as rates per 1,000 discharges. To obtain the standardized rate for each provider level PSIs, the output of the software should be multiplied by 1,000. The area level indicators are expressed as rates per 100,000 population. To obtain the standardized area rate for each area level PSIs, the output of the software should be multiplied by 100,000.

There is also a Windows version of the software that incorporates all of the QI modules into a single application. The Windows version, allows the user to select the unit to be used for rates for both provider-level and area-level PSIs.

See Appendix A for links to documents and tools.

5.1 Complications of Anesthesia (PSI 1)

Definition	Cases of anesthetic overdose, reaction, or endotrachial tube misplacement per 1,000 surgery discharges with an operating room procedure.
Numerator	Discharges with ICD-9-CM diagnosis codes for anesthesia complications in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	All surgical discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: • with ICD-9-CM diagnosis codes for anesthesia complications in the principal diagnosis field or in a secondary diagnosis field if present on admission, if known • with codes for self-inflicted injury, poisoning due to anesthetics (E8551, 9681-4, 9687) **and** any diagnosis code for active drug dependence, or active non-dependent abuse of drugs
Type of Indicator	Provider level
Empirical Performance	Bias: Not detected, but may be biased in a way undetectable by empirical tests
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to capture cases flagged by external cause-of-injury codes (e-codes) and complications codes for adverse effects from the administration of therapeutic drugs, as well as the overdose of anesthetic agents used primarily in therapeutic settings.

Panel Review

Panelists had concerns about the frequency of coding of these complications, especially since the use of e-codes is considered voluntary and appears to vary widely among providers. Plausibly, a “reaction” may be described without attributing it to anesthetic. Another concern is that some of these cases would be present on admission (e.g., due to recreational drug use).

Panelists expressed concern about the events that would be assigned to the code for incorrect placement of endotrachial tube. They noted that true misplacement does represent medical error, but they were skeptical about whether this code would be limited to those situations.

Ideally, this indicator would be used with a coding designation that distinguishes conditions present on admission from those that develop in-hospital. However, this is not available in the administrative data used to define this indicator, and so this concern was addressed by eliminating codes for drugs that are commonly used as recreational drugs. While this does not eliminate the chance that these codes represent intentional or accidental overdose on the part of the patient, it should eliminate many of these cases.

Literature Review

The literature review focused on the validity of complication indicators based on ICD-9-CM diagnosis or procedure codes. Results of the literature review indicate no published evidence for the sensitivity or predictive value of this indicator based on detailed chart review or prospective data collection. Sensitivity is the proportion of the patients who suffered an adverse event for whom that event was coded on a discharge abstract or Medicare claim. Predictive value is the proportion of patients with a coded adverse event who were confirmed as having suffered that event.

The project team found no published evidence for this indicator that supports the following

constructs: (1) that hospitals that provide better processes of care experience fewer adverse events; (2) that hospitals that provide better overall care experience fewer adverse events; and (3) that hospitals that offer more nursing hours per patient day, better nursing skill mix, better physician skill mix, or more experienced physicians have fewer adverse events.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Complications of Anesthesia generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is 75.7%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is 0.00187, indicating that the systematic differences (signal) among hospitals is lower than many indicators and less likely associated with hospital characteristics. The signal share is 0.00563, and is also lower than many indicators. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Complications of Anesthesia is low, indicating that the measures are likely not biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.)

Source

A subset of this indicator was originally proposed by Iezzoni et al.[42] as part of Complications Screening Program (CSP) (CSP 21, "Complications relating to anesthetic agents and other CNS depressants") Their definition also includes poisoning due to centrally acting muscle relaxants and accidental poisoning by nitrogen oxides, which were omitted from this PSI. Their definition excludes other codes included in the PSI, namely, poisoning by other and unspecified general anesthetics and external cause of injury codes for "endotracheal tube wrongly place during anesthetic procedure" and adverse effects of anesthetics in therapeutic use.

[42] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.

5.2 Death in Low-Mortality DRGs (PSI 2)

Definition	In-hospital deaths per 1,000 patients in DRGs with less than 0.5% mortality.
Numerator	Discharges with disposition of "deceased" among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	Discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), in DRGs with less than 0.5% mortality rate. If a DRG is divided into "without/with complications," both DRGs must have mortality rates below 0.5% to qualify for inclusion. Exclude patients with any code for trauma, immunocompromised state, or cancer.
Type of Indicator	Provider level
Empirical Performance	Bias: Substantial bias
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to identify in-hospital deaths in patients unlikely to die during hospitalization. The underlying assumption is that when patients admitted for an extremely low-mortality condition or procedure die, a health care error is more likely to be responsible. Patients experiencing trauma or having an immunocompromised state or cancer are excluded, as these patients have higher non-preventable mortality.

Panel Review

This indicator should be evaluated separately by type of DRG when used as an indicator of quality. For example, the PSI Software reports the low-mortality DRG rate for all the included DRGs and separately by DRG type: adult medical, adult surgical (with and without an operating room procedure), pediatric medical, pediatric surgical (with and without an operating room procedure), and obstetric and psychiatric. The overall usefulness of this indicator was rated as favorable by panelists. Because the denominator includes many heterogeneous patients cared for by different services, this indicator should be stratified by DRG type (i.e., medical, surgical, psychiatric, obstetric, pediatric) when used as an indicator of quality.

Panelists noted that hospital case-mix may affect the rate of death in low mortality DRGs, and patients referred from skilled nursing facilities, those with certain comorbidities, and older patients may be at higher risk of dying. They advocated risk adjustment for comorbidities and age.

Panelists advocated that this indicator not be subject to public reporting because of the potential bias and questions about the extent of preventability.

Literature Review

Based on two-stage implicit review of randomly selected deaths, Hannan et al. found that patients in low-mortality DRGs (<0.5%) were 5.2 times more likely than all other patients who died (9.8% versus 1.7%) to have received "care that departed from professionally recognized standards," after adjusting for patient demographic, geographic, and hospital characteristics.[43] In 15 of these 26 cases (58%) of substandard care, the patient's death was attributed at least partially to that care. The association with substandard care was stronger for the DRG-based definition of this indicator than for the procedure-based definition (5.7% versus 1.7%, OR=3.2). The project team was unable to find other evidence on the validity of this indicator.

[43] Hannan EL, Bernard HR, O'Donnell JF, Kilburn H, Jr. A methodology for targeting hospital cases for quality of care record reviews. Am J Public Health 1989;79(4):430-6.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Death in Low-mortality DRGs generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is high, relative to other indicators, at 94.2%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00439, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is high, relative to other indicators, at 0.04237. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Death in Low-mortality DRGs is high, indicating that the measures are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Hannan et al. as a criterion for targeting "cases that would have a higher percentage of quality of care problems than cases without the criterion, as judged by medical record review."[44] An alternative form of this indicator focused on "primary surgical procedures," rather than DRGs, with less than 0.5% inpatient mortality.

[44] Hannan et al. 1989.

5.3 Decubitus Ulcer (PSI 3)

Definition	Cases of decubitus ulcer per 1,000 discharges with a length of stay greater than 4 days.
Numerator	Discharges with ICD-9-CM code of decubitus ulcer in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	All medical and surgical discharges 18 years and older defined by specific DRGs. Exclude cases: • with length of stay of less than 5 days • with ICD-9-CM code of decubitus ulcer in the principal diagnosis field or in a secondary diagnosis field if present on admission, if known • MDC 9 (Skin, Subcutaneous Tissue, and Breast) • MDC 14 (pregnancy, childbirth, and puerperium) • with any diagnosis of hemiplegia, paraplegia, or quadriplegia • with an ICD-9-CM diagnosis code of spina bifida or anoxic brain damage • with an ICD-9-CM procedure code for debridement or pedicle graft **before or on the same day** as the major operating room procedure (surgical cases only) • admitted from a long-term care facility (SID Admission Source=3) • transferred from an acute care facility (SID Admission Source=2)
Type of Indicator	Provider level
Empirical Performance	Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to flag cases of in-hospital decubitus ulcers. Its definition is limited to decubitus ulcer as a secondary diagnosis to better screen out cases that may be present on admission. In addition, this indicator excludes patients who have a length of stay of 4 days or less, as it is unlikely that a decubitus ulcer would develop within this period of time. Finally, this indicator excludes patients who are particularly susceptible to decubitus ulcer, namely patients with major skin disorders (MDC 9) and paralysis.

Panel Review

The overall usefulness of this indicator was rated as very favorable by panelists. Concerns regarding the systematic screening for ulcers and reliability of coding, especially for early stage ulcers, brought into question that assertion. Therefore, this indicator appears to be best used as a rate-based indicator. Panelists suggested that patients admitted from a long-term care facility be excluded, as these patients may have an increased risk of having decubiti present on admission.

Panelists noted that hospitals that routinely screen for decubitus ulcers as part of a quality improvement program might have an artificially high rate of ulcers compared to other hospitals, which may cause this indicator to be somewhat biased.

This indicator includes pediatric patients. Pressure sores are very unusual in children, except among the most critically ill children (who may be paralyzed to improve ventilator management) and children with chronic neurological problems. Age stratification is recommended.

Literature Review

Coding validity. No evidence on validity is available from CSP studies. Geraci et al. confirmed only 2 of 9 episodes of pressure ulcers reported on discharge abstracts of Veterans Affairs (VA) patients hospitalized in

1987-89 for congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), or diabetes.[45] The sensitivity for a nosocomial ulcer was 40%. Among Medicare hip fracture patients, Keeler et al. confirmed 6 of 9 reported pressure ulcers, but failed to ascertain 89 additional cases (6% sensitivity) using ICD-9-CM codes.[46] In the largest study to date, Berlowitz et al. found that the sensitivity of a discharge diagnosis of pressure ulcer among all patients transferred from VA hospitals to VA nursing homes in 1996 was 31% overall, or 54% for stage IV (deep) ulcers.[47] The overall sensitivity increased modestly since 1992 (26.0%), and was slightly but statistically significantly better among medical patients than among surgical patients (33% versus 26%).

Construct validity. Needleman and Buerhaus found that nurse staffing was inconsistently associated with the occurrence of pressure ulcers among medical patients, and was independent of pressure ulcers among major surgery patients.[48] As was expected, nursing skill mix (RN hours/licensed nurse hours) was significantly associated with the pressure ulcer rate.[49] Total licensed nurse hours per acuity-adjusted patient day were inconsistently associated with the rate of pressure ulcers.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Decubitus Ulcer generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is high, relative to other indicators, at 85.6%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.0147, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.01067. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Decubitus Ulcer is high, indicating that the measure is biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al.[50] as part of the Complications Screening Program (CSP 6, "cellulitis or decubitus ulcer"). Needleman and Buerhaus identified decubitus ulcer as an "outcome potentially sensitive to nursing"[51] The American Nurses Association, its State associations, and the California Nursing Outcomes Coalition have identified the total prevalence of inpatients with Stage I, II, III, or IV pressure ulcers as a "nursing-sensitive quality indicator for acute care settings."[52]

[45] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. Med Care 1997;35(6):589-602.
[46] Keeler E, Kahn K, Bentow S. Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare Provider Analysis and Review file. Springfield, VA: NTIS; 1991.
[47] Berlowitz D, Brand H, Perkins C. Geriatric syndromes as outcome measures of hospital care: Can administrative data be used? JAGS 1999;47:692-696.
[48] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.: 230-88-0021.
[49] Lichtig LK, Knauf RA, Hilholland DK. Some impacts of nursing on acute care hospital outcomes. J Nurs Adm 1999;29(2):25-33.
[50] Iezzoni LI, Daley J, Heeren T, Foley SM, Risher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[51] Needleman et al. 2001.
[52] Nursing-Sensitive Quality Indicators for Acute Care Settings and ANA's Safety & Quality Initiative. In: American Nurses Association; 1999.

5.4 Failure to Rescue (PSI 4)

Definition	Deaths per 1,000 patients having developed specified complications of care during hospitalization.
Numerator	Discharges with a disposition of "deceased" among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	Discharges 18 years and older with potential complications of care listed in failure to rescue definition (i.e., pneumonia, DVT/PE, sepsis, acute renal failure, shock/cardiac arrest, or GI hemorrhage/acute ulcer). Exclude cases: • age 75 years and older • neonatal patients in MDC 15 • transferred to an acute care facility (SID Discharge Disposition = 2) • transferred from an acute care facility (SID Admission Source = 2) • admitted from a long-term care facility (SID Admission Source=3) Additional exclusion criteria specific to each diagnosis.
Type of Indicator	Provider level
Empirical Performance	Population Rate (2003): 127.687 per 1,000 population at risk Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to identify patients who die following the development of a complication. The underlying assumption is that good hospitals identify these complications quickly and treat them aggressively.

Failure to Rescue may be fundamentally different than other indicators reviewed in this report, as it may reflect different aspects of quality of care (effectiveness in rescuing a patient from a complication versus preventing a complication). This indicator includes pediatric patients. It is important to note that children beyond the neonatal period inherently recover better from physiological stress and thus may have a higher rescue rate.

Panel Review

Panelists expressed concern regarding patients with "do not resuscitate" (DNR) status. In cases where this DNR status is not a direct result of poor quality of care, it would be contrary to patient desire and poor quality of care to rescue a patient. In addition, very old patients—or patients with advanced cancer or HIV—may not desire or may be particularly difficult to rescue from these complications. As a result, this indicator definition was modified to exclude those patients age 75 years and older. In addition, panelists suggested the exclusion of patients admitted from long-term care facilities.

Panelists noted that several adverse incentives may be introduced by implementing this indicator. In particular, since some type of adjustment may be desirable, this indicator may encourage the upcoding of complications and comorbidities to inflate the denominator or manipulate risk adjustment. Others noted that this indicator could encourage irresponsible resource use and allocation, although this is likely to be a controversial idea. Finally, panelists emphasized that this indicator should be used internally by hospitals, as it is not validated for public reporting.

Literature Review

Construct validity. Silber and colleagues have published a series of studies establishing the construct validity of failure-to-rescue rates through their associations with hospital characteristics and other measures of hospital

performance. Among patients admitted for cholecystectomy and transurethral prostatectomy, failure to rescue was independent of severity of illness at admission, but was significantly associated with the presence of surgical house staff and a lower percentage of board-certified anesthesiologists.[53] The adverse occurrence rate was independent of this hospital characteristic. In a larger sample of patients who underwent general surgical procedures, lower failure-to-rescue rates were found at hospitals with high ratios of registered nurses to beds.[54] Failure rates were strongly associated with risk-adjusted mortality rates, as expected, but not with complication rates.[55]

More recently, Needleman and Buerhaus confirmed that higher registered nurse staffing (RN hours/adjusted patient day) and better nursing skill mix (RN hours/licensed nurse hours) were consistently associated with lower failure-to-rescue rates, even using administrative data to define complications.[56]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Failure to Rescue generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 66.6%, suggesting that observed differences in risk-adjusted rates may reflect true differences across hospitals. The signal standard deviation for this indicator is also high, relative to other indicators, at 0.04617, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.01450. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Failure to Rescue is high, indicating that the measures are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Silber et al. as a more powerful tool than the risk-adjusted mortality rate to detect true differences in patient outcomes across hospitals.[57] The underlying premise was that better hospitals are distinguished not by having fewer adverse occurrences but by more successfully averting death among (i.e., rescuing) patients who experience such complications. More recently, Needleman and Buerhaus adapted Failure to Rescue to administrative data sets, hypothesizing that this outcome might be sensitive to nurse staffing.[58]

[53] Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. Med Care 1992;30(7):615-29.
[54] Silber J, Rosenbaum P, Ross R. Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics? J Am Stat Assoc 1995;90:7-18.
[55] Silber JH, Rosenbaum PR, Williams SV, Ross RN, Schwartz JS. The relationship between choice of outcome measure and hospital rank in general surgical procedures: Implications for quality assessment. Int J Qual Health Care 1997;9(3):193-200.
[56] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston MA: Health Resources and Services Administration; 2001 February 28. Report No.:230-99-0021.
[57] Silber et al. 1992.
[58] Needleman et al. 2001.

5.5 Foreign Body Left During Procedure, Provider Level (PSI 5)

Provider Level Definition (only secondary diagnosis)

Definition	Discharges with foreign body accidentally left in during procedure per 1,000 discharges.
Numerator	Discharges with ICD-9-CM codes for foreign body left in during procedure in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator
Denominator	All medical and surgical discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), defined by specific DRGs. Exclude patients with ICD-9-CM codes for foreign body left in during procedure in the principal diagnosis field or secondary diagnosis present on admission, if known.
Type of Indicator	Provider level
Empirical Performance	Bias: Did not undergo empirical testing of bias
Risk Adjustment	No risk adjustment

5.6 Foreign Body Left During Procedure, Area Level (PSI 21)

Area Level Definition (principal or secondary diagnosis)

Definition	Discharges with foreign body accidentally left in during procedure per 100,000 population.
Numerator	Discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), with ICD-9-CM codes for foreign body left in during procedure in any diagnosis field (principal or secondary) of medical and surgical discharges defined by specific DRGs.
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.
Type of Indicator	Area level
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of a foreign body accidentally left in a patient during a procedure. This indicator is defined on both a provider level (by restricting cases to those flagged by a secondary diagnosis or procedure code) and an area level (by including all cases).

Panel Review

Panelists believed that this indicator was useful in identifying cases of a foreign body left in during a procedure. However, they suggested that each case identified be examined carefully by the hospital, because this indicator was likely to yield few cases and some automated systems report this complication when a foreign body is left in intentionally.

Panelists also noted that the population at risk included both medical and surgical patients, but not all of these patients are at risk. The panelists felt that limiting the population at risk to surgical patients would decrease the sensitivity of this indicator substantially. Since not all patients in the denominator are actually at risk, some hospitals may appear to have a lower rate if they have fewer medical patients who have undergone invasive procedures.

Literature Review

The literature review focused on the validity of complication indicators based on ICD-9-CM diagnosis or procedure codes. Results of the literature review indicate no published evidence for the sensitivity or predictive value of this indicator based on detailed chart review or prospective data collection. Sensitivity is the proportion of the patients who suffered an adverse event for whom that event was coded on a discharge abstract or Medicare claim. Predictive value is the proportion of patients with a coded adverse event who were confirmed as having suffered that event.

The project team found no published evidence for this indicator that supports the following constructs: (1) that hospitals that provide better processes of care experience fewer adverse events; (2) that hospitals that provide better overall care experience fewer adverse events; and (3) that hospitals that offer more nursing hours per patient day, better nursing skill mix, better physician skill mix, or more experienced physicians have fewer adverse events.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Foreign Body Left During Procedure generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time. Due to the rarity of this diagnosis, reliability and bias were not assessed.

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program (CSP "sentinel events").[59] It was also included as one component of a broader indicator ("adverse events and iatrogenic complications") in AHRQ's original HCUP Quality Indicators.[60] It was proposed by Miller et al. in the "Patient Safety Indicator Algorithms and Groupings."[61] Based on expert consensus panels, McKesson Health Solutions included this indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications Measures Module.

[59] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[60] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: state and national applications. Jt Comm J Qual Improv 1998;24(2):88-105.
[61] Miller M, Elixhauser A, Zhan C, Meyer G. Patient safety indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.7 Iatrogenic Pneumothorax, Provider Level (PSI 6)

Provider Level Definition (only secondary diagnosis)

Definition	Cases of iatrogenic pneumothorax per 1,000 discharges.
Numerator	Discharges with ICD-9-CM code of 512.1 in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator..
Denominator	All medical and surgical discharges age 18 years and older defined by specific DRGs. Exclude cases: • with ICD-9-CM code of 512.1 in the principal diagnosis field or secondary diagnosis present on admission, if known. • MDC 14 (pregnancy, childbirth, and puerperium) • with an ICD-9-CM diagnosis code of chest trauma or pleural effusion • with an ICD-9-CM procedure code of diaphragmatic surgery repair • with any code indicating thoracic surgery or lung or pleural biopsy or assigned to cardiac surgery DRGs
Type of Indicator	Provider level
Empirical Performance	Bias: Some bias demonstrated
Risk Adjustment	Age, sex, DRG, comorbidity categories

5.8 Iatrogenic Pneumothorax, Area Level (PSI 22)

Area Level Definition (principal or secondary diagnosis)

Definition	Cases of iatrogenic pneumothorax per 100,000 population.
Numerator	Discharges 18 years and older with ICD-9-CM code of 512.1 in any diagnosis field (principal or secondary) of medical and surgical discharges defined by specific DRGs. Exclude cases: • MDC 14 (pregnancy, childbirth, and puerperium) • with an ICD-9-CM diagnosis code of chest trauma or pleural effusion • with an ICD-9-CM procedure code of diaphragmatic surgery repair • with any code indicating thoracic surgery or lung or pleural biopsy or assigned to cardiac surgery DRGs
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.
Type of Indicator	Area level
Empirical Performance	Population Rate (2003): 7.921 per 100,000 population
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of pneumothorax caused by medical care. This indicator is defined on both a provider level (by including cases of iatrogenic pneumothorax occurring as a secondary diagnosis during hospitalization) and on an area level (by including all cases of iatrogenic pneumothorax).

Iatrogenic pneumothorax excludes all trauma patients because these patients may be more susceptible to non-preventable iatrogenic pneumothorax or may be miscoded for traumatic pneumothorax. The smaller anatomy of children, especially neonates, may increase the technical complexity of these procedures in this population (however, these procedures are less likely to be performed in unmonitored settings).

Panel Review

Panelists rated the overall usefulness of this indicator favorably. The denominator of the definition that the panelists rated was limited to patients receiving a central line, Swan-Ganz catheter, or thorocentesis. However, exploratory empirical analyses found that this definition could not be operationalized using administrative data, as these procedures appeared to be under-reported. Although the panelists noted that this complication, given the definition rated, reflected medical error, the actual final definition of this indicator includes cases that may be less reflective of medical error. Specifically, this indicator includes patients in whom a pneumothorax resulted from barotrauma, including patients with acute respiratory distress syndrome.

Panelists expressed concern that some approaches of placing a central line (e.g., subclavian) may be more likely to result in pneumothorax than other approaches (e.g., internal jugular). However, other complications—such as complications of the carotid artery—would be more common with internal jugular approaches. Thus, if providers simply change approach, they may have a decrease in pneumothorax but an increase in other unmeasured complications.

Literature Review

The literature review focused on the validity of complication indicators based on ICD-9-CM diagnosis or procedure codes. Results of the literature review indicate no published evidence for the sensitivity or predictive value of this indicator based on detailed chart review or prospective data collection. Sensitivity is the proportion of the patients who suffered an adverse event for whom that event was coded on a discharge abstract or Medicare claim. Predictive value is the proportion of patients with a coded adverse event who were confirmed as having suffered that event.

The project team found no published evidence for this indicator that supports the following constructs: (1) that hospitals that provide better processes of care experience fewer adverse events; (2) that hospitals that provide better overall care experience fewer adverse events; and (3) that hospitals that offer more nursing hours per patient day, better nursing skill mix, better physician skill mix, or more experienced physicians have fewer adverse events.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Iatrogenic Pneumothorax generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 79.9%, suggesting that observed differences in risk-adjusted rates may reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00143, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00183. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Iatrogenic Pneumothorax is moderate, indicating that the measures may or may not be substantially biased based on the characteristics observed.

Source

This diagnosis code was proposed by Miller et al. as one component of a broader indicator ("iatrogenic conditions") in the "Patient Safety Indicator Algorithms and Groupings."[62] It was also included as one component of a broader indicator ("adverse events and iatrogenic complications") in AHRQ's Version 1.3 HCUP Quality Indicators.

[62] Miller M, Elixhauser A, Zhan C, Meyer G. Patient safety indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.9 Selected Infections Due to Medical Care, Provider Level (PSI 7)

Provider Level Definition (only secondary diagnosis)

Definition	Cases of ICD-9-CM codes 9993 or 99662 per 1,000 discharges.
Numerator	Discharges with ICD-9-CM code of 9993 or 99662 in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	All medical and surgical discharges 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), defined by specific DRGs. Exclude cases: • with ICD-9-CM code of 9993 or 99662 in the principal diagnosis field or secondary diagnosis present on admission, if known • with length of stay less than 2 days • with any diagnosis code for immunocompromised state or cancer • with Cancer DRG
Type of Indicator	Provider level
Empirical Performance	Bias: Some bias demonstrated
Risk Adjustment	Age, sex, DRG, comorbidity categories

5.10 Selected Infections Due to Medical Care, Area Level (PSI 23)

Area Level Definition (principal or secondary diagnosis)

Definition	Cases of ICD-9-CM codes 9993 or 99662 per 100,000 population.
Numerator	Discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), with ICD-9-CM code of 9993 or 99662 in any diagnosis field (principal or secondary) of medical and surgical discharges defined by specific DRGs. Exclude patients with any diagnosis code for immunocompromised state or cancer.
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.
Type of Indicator	Area level
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of infection due to medical care, primarily those related to intravenous (IV) lines and catheters. This indicator is defined both on a provider level (by including cases based on secondary diagnosis associated with the same hospitalization) and on an area level (by including all cases of such infection). Patients with potential immunocompromised states (e.g., AIDS, cancer, transplant) are excluded, as they may be more susceptible to such infection. This indicator includes children and neonates. It should be noted that high-risk neonates are at particularly high risk for catheter-related infections.

Panel Review

Panelists expressed particular interest in tracking IV and catheter-related infections, despite the potential for bias due to charting or under-reporting. For the most part, they felt that these complications were important to track. As

with other indicators tracking infections, concern regarding the potential overuse of prophylactic antibiotics remains.

Literature Review

The literature review focused on the validity of complication indicators based on ICD-9-CM diagnosis or procedure codes. Results of the literature review indicate no published evidence for the sensitivity or predictive value of this indicator based on detailed chart review or prospective data collection. Sensitivity is the proportion of the patients who suffered an adverse event for whom that event was coded on a discharge abstract or Medicare claim. Predictive value is the proportion of patients with a coded adverse event who were confirmed as having suffered that event.

The project team found no published evidence for this indicator that supports the following constructs: (1) that hospitals that provide better processes of care experience fewer adverse events; (2) that hospitals that provide better overall care experience fewer adverse events; and (3) that hospitals that offer more nursing hours per patient day, better nursing skill mix, better physician skill mix, or more experienced physicians have fewer adverse events.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Selected Infections Due to Medical Care generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 70.8%, suggesting that observed differences in risk-adjusted rates may reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00134, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00095. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Selected Infections Due to Medical Care is moderate, indicating that the measures may or may not be substantially biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.)

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program (CSP 11, "miscellaneous complications").[63] The University HealthSystem Consortium adopted the CSP indicator for major (#2933) and minor (#2961) surgery patients. A much narrower definition, including only 9993 ("other infection after infusion, injection, transfusion, vaccination"), was proposed by Miller et al. in the "Patient Safety Indicator Algorithms and Groupings."[64] The American Nurses Association and its State associations have identified the number of laboratory-confirmed bacteremic episodes associated with central lines per critical care patient day as a "nursing-sensitive quality indicator for acute care settings."[65]

[63] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[64] Miller M, Elixhauser A, Zhan C, Meyer G. Patient safety indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.
[65] Nursing-Sensitive Quality Indicators for Acute Care Settings and ANA's Safety and Quality Initiative. In: American Nurses Association; 1999.

5.11 Postoperative Hip Fracture (PSI 8)

Definition	Cases of in-hospital hip fracture per 1,000 surgical discharges with an operating room procedure.
Numerator	Discharges with ICD-9-CM code for hip fracture in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	All surgical discharges 18 years and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: • with ICD-9-CM code for hip fracture in the principal diagnosis field or secondary diagnosis present on admission, if known • where the only operating room procedure is hip fracture repair • where a procedure for hip fracture repair occurs before or on the same day as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available* • with diseases and disorders of the musculoskeletal system and connective tissue (MDC 8) • with principal diagnosis (or secondary diagnosis present on admission, if known) of seizure, syncope, stroke, coma, cardiac arrest, poisoning, trauma, delirium and other psychoses, or anoxic brain injury • with any diagnosis of metastatic cancer, lymphoid malignancy or bone malignancy, or self-inflicted injury • MDC14 (Pregnancy, Childbirth and the Puerperium)
Type of Indicator	Provider level
Empirical Performance	Bias: Some bias demonstrated
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to capture cases of in-hospital fracture—specifically, hip fractures. This indicator limits diagnosis codes to secondary diagnosis codes to eliminate fractures that were present on admission. It further excludes patients in MDC 8 (musculoskeletal disorders) and patients with indications for trauma or cancer, or principal diagnoses of seizure, syncope, stroke, coma, cardiac arrest, or poisoning, as these patients may have a fracture present on admission. This indicator is limited to surgical cases since previous research suggested that these codes in medical patients often represent conditions present on admission (see Literature Review).

Panel Review

Although this indicator was initially presented as "In-hospital hip fracture and fall," panelists unanimously suggested that falls should be eliminated from this indicator and that all in-hospital fractures should be included. The resulting indicator was termed "In-hospital fracture possibly related to falls." Children were excluded after empirical analysis revealed that they did not have a substantial number of cases in the numerator.

Panelists noted that this indicator may be slightly biased for hospitals that care for more of the elderly and frail, because they have weaker bones and are more susceptible to falls.

Panelists were interested in capturing all fractures occurring in-hospital, although it was not possible to operationalize this suggestion.

Literature Review

Coding validity. The original CSP definition had an adequate confirmation rate among major surgical cases in Medicare inpatient claims files (57% by coders' review, 71% by physicians' review), but a very poor confirmation rate among medical cases (11% by both coders' and physicians' review).[66] [67] This problem was attributable to the fact that most hip fractures among medical inpatients were actually comorbid diagnoses present at admission rather than complications of hospital care. Nurse reviews were not performed.

Construct validity. Explicit process of care failures in the CSP validation study were relatively frequent among cases with CSP 25 (76% of major surgery patients, 54% of medical patients), after excluding patients who had hip fractures at admission, but unflagged controls were not evaluated on the same criteria.[68] Physician reviewers identified potential quality problems in 24% of major surgery patients and 5% of medical patients with CSP 25 (versus 2% of unflagged controls for each risk group).[69]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative Hip Fracture generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 67.1%, suggesting that observed differences in risk-adjusted rates may reflect true differences across hospitals. The signal standard deviation for this indicator is lower than many indicators, at 0.00184, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00403. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative Hip Fracture is moderate, indicating that the measures may or may not be substantially biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.)

Source

This indicator was originally proposed by Iezzoni et al.[70] as part of the Complications Screening Program (CSP 25, "in-hospital hip fracture or fall"). Their definition also includes any documented fall, based on external cause of injury codes. Needleman and Buerhaus considered in-hospital hip fracture as an "Outcome Potentially Sensitive to Nursing," but discarded it because the "event rate was too low to be useful."[71] The American Nurses Association, its State associations, and the California Nursing Outcomes Coalition have identified the number of patient falls leading to injury per 1,000 patient days (based on clinical data collection) as a "nursing-sensitive quality indicator for acute care settings."[72]

[66] Lawthers A, McCarthy E, Davis R, Peterson L, Palmer R, Iezzoni L. Identification of in-hospital complications from claims data: Is it valid? Med Care 2000;38(8):785-795.
[67] Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: Validation of the Complications Screening Program. Med Care 2000;38(8):796-806.
[68] Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems: Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.
[69] Weingart et al. 2000.
[70] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[71] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.: 230-99-0021.
[72] Nursing-Sensitive Quality Indicators for Acute Care Settings and ANA's Safety & Quality Initiative. In: American Nurses Association; 1999.

5.12 Postoperative Hemorrhage or Hematoma (PSI 9)

Provider Level Definition

Definition	Cases of hematoma or hemorrhage requiring a procedure per 1,000 surgical discharges with an operating room procedure.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator with either of the following: • ICD-9-CM codes for postoperative hemorrhage in any secondary diagnosis field and a code for drainage of hematoma in any procedure code field. • ICD-9-CM codes for postoperative hematoma in any secondary diagnosis field and a code for postoperative control of hemorrhage in any procedure code field.
Denominator	All surgical discharges 18 years and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: • with preexisting condition (principal diagnosis or secondary diagnosis present on admission, if known) of postoperative hemorrhage or postoperative hematoma • where the only operating room procedure is postoperative control of hemorrhage or drainage of hematoma • where a procedure for postoperative control of hemorrhage or drainage of hematoma occurs before the first operating room procedure. *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* • MDC 14 (pregnancy, childbirth and the puerperium)
Type of Indicator	Provider level
Empirical Performance	Bias: Not detected in empirical tests
Risk Adjustment	Age, sex, DRG, comorbidity categories

5.13 Postoperative Hemorrhage or Hematoma (PSI 27)

Area Level Definition

Definition	Cases of hematoma or hemorrhage requiring a procedure per 100,000 population.
Numerator	All surgical discharges 18 years and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: • with preexisting condition (principal diagnosis or secondary diagnosis present on admission, if known) of postoperative hemorrhage or postoperative hematoma • where the only operating room procedure is postoperative control of hemorrhage or drainage of hematoma • where a procedure for postoperative control of hemorrhage or drainage of hematoma occurs before the first operating room procedure. *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* • MDC 14 (pregnancy, childbirth and the puerperium)
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.

Type of Indicator	Area level
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to capture cases of hemorrhage or hematoma following a surgical procedure. This indicator limits hemorrhage and hematoma codes to secondary procedure and diagnosis codes, respectively, to isolate those hemorrhages that can truly be linked to a surgical procedure.

Panel Review

Panelists noted that some patients may be at higher risk for developing a postoperative hemorrhage or hematoma. Specifically, they were concerned about patients with coagulopathies and those on anticoagulants. They suggested that where possible, this indicator be stratified for patients with underlying clotting differences. They also noted that patients admitted for trauma may be at a higher risk for developing postoperative hemorrhage or may have a hemorrhage diagnosed that occurred during the trauma. They also suggested that this indicator be stratified for trauma and non-trauma patients.

Literature Review

Coding validity. The original CSP definition had a relatively high confirmation rate among major surgical cases (83% by coders' review, 57% by physicians' review, 52% by nurse-abstracted clinical documentation, and 76% if nurses also accepted physicians' notes as adequate documentation).[73] [74] [75] Hartz and Kuhn estimated the validity of hemorrhage codes using a gold standard based on transfusion "requirement." [76] They identified only 26% of episodes of bleeding (defined as requiring return to surgery or transfusion of at least six units of blood products) by applying this indicator (9981) to Medicare patients who underwent coronary artery bypass surgery; the predictive value was 75%.

Construct Validity. Explicit process of care failures in the CSP validation study were relatively frequent among major surgical cases with CSP 24, but not among medical cases (66% and 13%, respectively), after excluding patients who had hemorrhage or hematoma at admission.[77] Cases flagged on this indicator and unflagged controls did not differ significantly on a composite of 17 generic process criteria. Similarly, cases flagged on this indicator and unflagged controls did not differ significantly on a composite of four specific process criteria for major surgical cases and two specific process criteria for medical cases in the earlier study of elderly Medicare beneficiaries.[78]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative Hemorrhage or Hematoma generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is lower than most indicators, at 8.6%, suggesting that observed differences in risk-adjusted rates may not reflect true differences across hospitals.
The signal standard deviation for this indicator is lower than most indicators, at 0.00039, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00006. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals.

[73] Lawthers A, McCarthy E, Davis R, Peterson L, Palmer R, Iezzoni L. Identification of in-hospital complications from claims data: Is it valid? Med Care 2000;38(8):785-795.
[74] McCarthy EP, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Does clinical evidence support ICD-9-CM diagnosis coding of complications? Med Care 2000;38(8):868-876.
[75] Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: Validation of the Complications Screening Program. Med Care 2000;38(8):796-806.
[76] Hartz AJ, Kuhn EM. Comparing hospitals that perform coronary artery bypass surgery: The effect of outcome measures and data sources. Am J Public Health 1994;84(10):1609-14.
[77] Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the complications Screening Program flag case with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.
[78] Iezzoni L, Lawthers A, Petersen L, McCarthy E, Palmer R, Cahalane M, et al. Project to validate the Complications Screening Program: Health Care Financing Administration; 1998 March 31. Report No: HCFA Contract 500-94-0055.

The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative Hemorrhage or Hematoma is low, indicating that the measures are likely not biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.)

Source

This indicator was originally proposed by Iezzoni et al.[79] as part of the Complications Screening Program (CSP 24, "post-procedural hemorrhage or hematoma"), although their definition allowed either procedure or diagnosis codes. By contrast, the current definition requires a hemorrhage or hematoma diagnosis with an associated procedure to either control the hemorrhage or drain the hematoma. It was also included as one component of a broader indicator ("adverse events and iatrogenic complications") in AHRQ's original HCUP Quality Indicators.[80]

[79] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[80] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-105. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

5.14 Postoperative Physiologic and Metabolic Derangement (PSI 10)

Definition	Cases of specified physiological or metabolic derangement per 1,000 elective surgical discharges with an operating room procedure.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for physiologic and metabolic derangements in any secondary diagnosis field. Discharges with acute renal failure (subgroup of physiologic and metabolic derangements) must be accompanied by a procedure code for dialysis (3995, 5498).
Denominator	All elective* surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. *Defined by admit type. Exclude cases: • with preexisting condition (principal diagnosis or secondary diagnosis present on admission, if known) of physiologic and metabolic derangements or chronic renal failure • with acute renal failure where a procedure for dialysis occurs **before or on the same day** as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available* • with both a diagnosis code of ketoacidosis, hyperosmolarity, or other coma (subgroups of physiologic and metabolic derangements coding) **and** a principal diagnosis of diabetes • with both a secondary diagnosis code for acute renal failure (subgroup of physiologic and metabolic derangements coding) **and** a principal diagnosis of acute myocardial infarction, cardiac arrhythmia, cardiac arrest, shock, hemorrhage, or gastrointestinal hemorrhage • MDC 14 (pregnancy, childbirth and the puerperium)
Type of Indicator	Provider level
Empirical Performance	Bias: Some bias demonstrated
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to flag cases of postoperative metabolic or physiologic complications. The population at risk is limited to elective surgical patients, because patients undergoing non-elective surgery may develop less preventable derangements. In addition, each diagnosis has specific exclusions, designed to reduce the number of flagged cases in which the diagnosis was present on admission or was more likely to be non-preventable.

Panel Review

Panelists expressed concern that acute renal failure suffers from the problem of varied definition: what one doctor may call acute renal failure, another may not. To ensure that the only renal failure cases that are picked up are those that are clinically severe, the panel suggested that acute renal failure be included only when it is paired with a procedure code for dialysis.

Panelists noted that coding of relatively transient metabolic and physiologic complications may be lacking, such as in cases of diabetic ketoacidosis. Conversely, some physicians may capture non-clinically significant events in this indicator.

This indicator includes pediatric patients, which was not specifically discussed by the panel. The incidence of these complications is a function of the underlying prevalence of diabetes and renal impairment, which are less common among children than among adults.

Literature Review

Coding validity. No evidence on validity is available from CSP studies. Geraci et al.[81] confirmed only 5 of 15 episodes of acute renal failure and 12 of 34 episodes of hypoglycemia reported on discharge abstracts of VA patients hospitalized for CHF, COPD, or diabetes. Romano reported no false positives in episodes of acute renal failure or hypoglycemia using discharge abstracts of diskectomy patients.[82] ICD-9-CM diagnoses (585 or 7885) had a sensitivity of 8% and a predictive value of 4% in comparison with the VA's National Surgical Quality Improvement Program database, which defines renal failure as requiring dialysis within 30 days after surgery.[83]

Construct Validity. After adjusting for patient demographic, geographic, and hospital characteristics, Hannan et al. reported that cases with a secondary diagnosis of fluid and electrolyte disorders were no more likely to have received care that departed from professionally recognized standards than cases without that code (2.2% versus 1.7%, OR=1.13).[84] However, these ICD-9-CM codes were omitted from the accepted AHRQ PSIs.

Empirical Evidence

The project team conducted extensive empirical analyses on the PSIs. Postoperative Physiologic and Metabolic Derangement generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is lower than many indicators, at 20.9%, suggesting that observed differences in risk-adjusted rates may not reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00054, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00033. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative Physiologic and Metabolic Derangement is moderate, indicating that the measures may or may not be substantially biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may or may not be related to the patient's risk of experiencing an adverse event.)

Source

This indicator was originally proposed by Iezzoni et al.[85] as part of the CSP (CSP 20, "postoperative physiologic and metabolic derangements"). The University HealthSystem Consortium adopted the CSP indicator for major surgery patients (#2945).

[81] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. Med Care 1997;35(6):589-602.

[82] Romano P. Can administrative data be used to ascertain clinically significant postoperative complications. American Journal of Medical Quality Press.

[83] Best W, Khuri S, Phelan M, Hur K, Henderson W, Demakis J, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: Results from the Department of Veterans Affairs National Surgical Quality Improvement Program. J Am Coll Surg 2002;194(3):257-266.

[84] Hannan EL, Bernard HR, O'Donnell JF, Kilburn H, Jr. A methodology for targeting hospital cases for quality of care record reviews. Am J Public Health 1989;79(4):430-6.

[85] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.

5.15 Postoperative Respiratory Failure (PSI 11)

Definition	Cases of acute respiratory failure per 1,000 elective surgical discharges with an operating room procedure.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM codes for acute respiratory failure (518.81) in any secondary diagnosis field (After 1999, include 518.84) OR Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM codes for reintubation procedure as follows: • (96.04) one or more days after the major operating room procedure code • (96.70 or 97.71) two or more days after the major operating room procedure code • (96.72) zero or more days after the major operating room procedure code
Denominator	All elective* surgical discharges age 18 and over defined by specific DRGs and an ICD-9-CM code for an operating room procedure. *Defined by admit type. Exclude cases: • with preexisting (principal diagnosis or secondary diagnosis present on admission, if known) acute respiratory failure • with an ICD-9-CM diagnosis code of neuromuscular disorder • where a procedure for tracheostomy is the only operating room procedure or tracheostomy occurs before the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available*. • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 4 (diseases/disorders of respiratory system) • MDC 5 (diseases/disorders of circulatory system)
Type of Indicator	Provider level
Empirical Performance	Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to flag cases of postoperative respiratory failure. This indicator limits the code for respiratory failure to secondary diagnosis codes to eliminate respiratory failure that was present on admission. It further excludes patients who have major respiratory or circulatory disorders and limits the population at risk to elective surgery patients.

Panel Review

Panelists rated the overall usefulness of this indicator as relatively favorable. They felt that only acute respiratory failure should be retained in this indicator and noted that this clinically significant event is at least partially preventable.

Literature Review

Coding Validity. CSP 3 had a relatively high confirmation rate among major surgical cases in the FY1994 Medicare inpatient claims files from California and Connecticut (72% by coders' review, 75% by physicians' review).[86] [87] Nurse reviews were not performed.

Geraci et al. confirmed 1 of 2 episodes of respiratory failure reported on discharge abstracts of VA patients hospitalized for CHF or diabetes; the sensitivity for respiratory

[86] Lawthers a, McCarthy E, Davis R, Peterson L, Palmer R, Iezzoni L. Identification of in-hospital complications from claims data: is it valid? Med Care 2000;38(8):785-795.
[87] Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: Validation of the Complications Screening Program. Med Care 2000;38(8):796-806.

decompensation requiring mechanical ventilation was 25%.[88]

Construct Validity. Explicit process of care failures in the CSP validation study were slightly but not significantly more frequent among major surgical cases with CSP 3 than among unflagged controls (52% versus 46%).[89] Indeed, cases flagged on this indicator were significantly less likely than unflagged controls (24% versus 64%) to have at least one of four specific process-of-care problems in the earlier study of elderly Medicare beneficiaries.[90]

Needleman and Buerhaus found that nurse staffing was independent of the occurrence of pulmonary failure among major surgery patients.[91] However, Kovner and Gergen reported that having more registered nurse hours per adjusted patient day was associated with a lower rate of "pulmonary compromise" after major surgery.[92]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative Respiratory Failure generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is lower than many indicators, at 46.6%, suggesting that observed differences in risk-adjusted rates may not reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00230, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00187. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative Respiratory Failure is high, indicating that the measures likely are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al. as part of the CSP (CSP 3, "postoperative pulmonary compromise").[93] Their definition also includes pulmonary congestion, other (or postoperative) pulmonary insufficiency, and acute pulmonary edema, which were omitted from this PSI. The University HealthSystem Consortium (#2927) and AHRQ's original HCUP Quality Indicators adopted the CSP indicator for major surgery patients.[94] Needleman and Buerhaus identified postoperative pulmonary failure as an "Outcome Potentially Sensitive to Nursing," using the original CSP definition.[95]

[88] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. In-hospital complications among survivors of admission for congestive heart failure, chronic obstructive pulmonary disease, or diabetes mellitus. J Gen Intern Med 1995;10(6):307-14.
[89] Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.
[90] Hawker GA, Coyte PC, Wright JG, Paul JE, Bombardier C. Accuracy of administrative data for assessing outcomes after knee replacement surgery. J. Clin Epidimiol 1997;50(3):265-73.
[91] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.:230-99-0021.
[92] Kovner C, Gergen PJ. Nurse staffing levels and adverse events following surgery in U.S. hospitals. Image J Nurs Sch 1998;30(4):315-21.
[93] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[94] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.
[95] Needleman et al. 2001.

5.16 Postoperative Pulmonary Embolism or Deep Vein Thrombosis (PSI 12)

Definition	Cases of deep vein thrombosis (DVT) or pulmonary embolism (PE) per 1,000 surgical discharges with an operating room procedure.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM codes for deep vein thrombosis or pulmonary embolism in any secondary diagnosis field
Denominator	All surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: • with preexisting (principal diagnosis or secondary diagnosis present on admission, if known) deep vein thrombosis or pulmonary embolism where a procedure for interruption of vena cava is the only operating room procedure • where a procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* • MDC 14 (Pregnancy, Childbirth and the Puerperium)
Type of Indicator	Provider level
Empirical Performance	Population Rate (2003): 9.830 per 1,000 population at risk Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to capture cases of postoperative venous thromboses and embolism—specifically, pulmonary embolism and deep venous thrombosis. This indicator limits vascular complications codes to secondary diagnosis codes to eliminate complications that were present on admission. It further excludes patients who have principal diagnosis of DVT, as these patients are likely to have had PE/DVT present on admission.

Panel Review

Panelists rated the overall usefulness of this indicator relatively highly as compared to other indicators. They noted that preventative techniques should decrease the rate of this indicator. This indicator includes pediatric patients. In the absence of specific thrombophilic disorders, postoperative thromboembolic complications in children are most likely to be secondary to venous catheters rather than venous stasis in the lower extremities.

Because the risk for DVT/PE varies greatly according to the type of procedure performed, panelists suggested that this indicator be adjusted or stratified according to surgical procedure types.

Literature Review

Coding validity. Geraci et al. confirmed only 1 of 6 episodes of DVT or PE reported on discharge abstracts of VA patients for CHF, COPD, or diabetes; the sensitivity was 100%.[96] Among Medicare hip fracture patients, by contrast, Keeler et al. confirmed 88% of reported PE cases, and failed to ascertain just 6 cases (65% sensitivity) using ICD-9-CM codes.[97] For DVT, they found just 1 of 6 cases using ICD-9-CM codes (but no false positive codes). Other studies have demonstrated that ICD-9-CM codes for DVT and PE have high predictive value when listed as the principal diagnosis for readmissions after major orthopedic

[96] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. In-hospital complications among survivors of admission for congestive heart failure, chronic obstructive pulmonary disease, or diabetes mellitus. J Gen Intern Med 1995;10(6):307-14.
[97] Keeler E, Kahn K, Bentow S. Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare Provider Analysis and Review File. Springfield, VA: NTIS;1991.

surgery (100%) or after inferior vena cava filter placement (98%).[98] However, these findings do not directly address the validity of DVT/PE as a secondary diagnosis among patients treated by anticoagulation.

Construct validity. Explicit process of care failures in the CSP validation study were relatively frequent among both major surgical and medical cases with CSP 22 (72% and 69%, respectively), after disqualifying cases in which DVT/PE was actually present at admission.[99] Needleman and Buerhaus found that nurse staffing was independent of the occurrence of DVT/PE among both major surgical or medical patients.[100] However, Kovner and Gergen reported that having more registered nurse hours and non-RN hours was associated with a lower rate of DVT/PE after major surgery.[101]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative PE or DVT generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 72.6%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00633, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00511. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative PE or DVT is high, indicating that the measures likely are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program (CSP 22, "venous thrombosis and pulmonary embolism")[102] and was one of AHRQ's original HCUP Quality Indicators for major surgery and invasive vascular procedure patients.[103] A code that maps to this indicator in the final AHRQ PSI was proposed by Miller et al. as one component of a broader indicator ("iatrogenic conditions").[104]

[98] White RH, Romano P, Zhou H, Rodrigo J, Barger W. Incidence and time course of thromboembolic outcomes following total hip or knee arthroplasty. Arch Intern Med 1998;158(14):1525-31.
[99] Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.
[100] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.:230-99-0021.
[101] Kovner C, Gergen PH. Nurse staffing levels and adverse events following surgery in U.S. hospitals. Image J Nurs Sch 1998;30(4):315-21.
[102] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[103] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.
[104] Miller M, Elixhauser A, Zhan C, Meyer G. Patient safety indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.17 Postoperative Sepsis (PSI 13)

Definition	Cases of sepsis per 1,000 elective surgery patients with an operating room procedure and a length of stay of 4 days or more.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for sepsis in any secondary diagnosis field.
Denominator	All elective* surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. *Elective - Admission type # is recorded as elective (Admission Type = 3) Exclude cases: • with preexisting (principal diagnosis or secondary diagnosis present on admission, if known) sepsis or infection • with any code for immunocompromised state or cancer • MDC 14 (pregnancy, childbirth, and puerperium) • with a length of stay of less than 4 days
Type of Indicator	Provider level
Empirical Performance	Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to flag cases of nosocomial postoperative sepsis. This indicator limits the code for sepsis to secondary diagnosis codes to eliminate sepsis that was present on admission. This indicator also excludes patients who have a principal diagnosis of infection, patients with a length of stay of less than 4 days, and patients with potential immunocompromised states (e.g., AIDS, cancer, transplant).

Panel Review

Panelists rated the overall usefulness of this indicator favorably, although they were less sure that this complication was reflective of medical error.

Literature Review

Coding validity. No evidence on validity is available from CSP studies. Barbour reported that only 38% of discharge abstracts with a diagnosis of sepsis actually had hospital-acquired sepsis.[105] However, this review was not limited to cases with a secondary diagnosis of sepsis, and sensitivity could not be evaluated. Geraci et al. confirmed (by blood culture) only 2 of 15 episodes of sepsis or "other infection" reported on discharge abstracts of VA patients hospitalized for CHF, COPD, or diabetes; the sensitivity for a positive blood culture was 50%.[106] In comparison with the VA's National Surgical Quality Improvement Program database, in which "systemic sepsis" is defined by a positive blood culture and systemic manifestations of sepsis within 30 days after surgery, the ICD-9-CM diagnosis had a sensitivity of 37% and a predictive value of 30%.[107]

Construct validity. Needleman and Buerhaus found that nurse staffing was independent of the

[105] Barbour GL. Usefulness of a discharge diagnosis of sepsis in detecting iatrogenic infection and quality of care problems. Am J Med Qual 1993;8(1):2-5.

[106] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. In-hospital complications among survivors of admission for congestive heart failure, chronic obstructive pulmonary disease, or diabetes mellitus. J Gen Intern Med 1995;10(6):307-14.

[107] Best W, Khuri S, Phelan M, Hur K, Henderson W, Demakis J, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: Results from the Department of Veterans Affairs national Surgical Quality Improvement Program. J Am Coll Surg 2002;194(3):257-266.

occurrence of sepsis among both major surgical or medical patients.[108]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative Sepsis generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is lower than many indicators, at 53.9%, suggesting that observed differences in risk-adjusted rates may not reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00869, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00790. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative Sepsis is high, indicating that the measures likely are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program (CSP 7, "septicemia").[109] Needleman and Buerhaus identified sepsis as an "Outcome Potentially Sensitive to Nursing" using the same CSP definition.[110]

[108] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.:230-99-0021.

[109] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.

[110] Needleman et al., 2001.

5.18 Postoperative Wound Dehiscence, Provider Level (PSI 14)

Provider Level Definition

Definition	Cases of reclosure of postoperative disruption of abdominal wall per 1,000 cases of abdominopelvic surgery.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for reclosure of postoperative disruption of abdominal wall (54.61) in any procedure field.
Denominator	All abdominopelvic surgical discharges age 18 and older. Exclude cases: • where a procedure for reclosure of postoperative disruption of abdominal wall occurs before or on the same day as the first abdominopelvic surgery procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available* • where length of stay is less than 2 days • with immunocompromised state • MDC 14 (pregnancy, childbirth, and puerperium).
Type of Indicator	Provider level
Empirical Performance	Bias: Some bias demonstrated
Risk Adjustment	Age, sex, DRG, comorbidity categories

5.19 Postoperative Wound Dehiscence, Area Level (PSI 24)

Area Level Definition

Definition	Cases of reclosure of postoperative disruption of abdominal wall per 100,000 population.
Numerator	Discharges with ICD-9-CM code for reclosure of postoperative disruption of abdominal wall (5461) in any procedure field. Exclude patients with immunocompromised state and MDC 14 (pregnancy, childbirth, and puerperium).
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.
Type of Indicator	Area level
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of wound dehiscence in patients who have undergone abdominal and pelvic surgery. This indicator is defined both on a provider level (by including cases based on secondary diagnosis associated with the same hospitalization) and on an area level (by including all cases of wound dehiscence).

Panel Review

Panelists suggested that postoperative wound disruption be excluded from the indicator and that trauma, cancer, and immunocompromised patients be included. They also reported that the risk of developing wound dehiscence varies with patient factors such as age and comorbidities.

Literature Review

Coding validity. No evidence on validity is available from CSP studies. Hawker et al. found that the sensitivity and predictive value of wound dehiscence were both 100%.[111] Faciszewski et al. aggregated

[111] Hawker BA, Coyte PC, Wright JG, Paul JE, Bombardier C. Accuracy of administrative data for assessing outcomes after

wound dehiscence with postoperative hemorrhage or hematoma and reported a pooled confirmation rate of 17% with 3% sensitivity of coding among patients who underwent spinal fusion.[112] In comparison with the VA's National Surgical Quality Improvement Program database, in which dehiscence is defined as fascial disruption within 30 days after surgery, the ICD-9-CM diagnosis of wound disruption had a sensitivity of 25% and a predictive value of 23%.[113] This code (9983) was ultimately removed from the accepted PSI, because the clinical panel was concerned that the diagnosis definition was too broad and failed to distinguish skin from fascial separation.

Construct validity. Based on two-stage review of randomly selected deaths, Hannan et al. reported that cases with a secondary diagnosis of wound disruption were 3.0 times more likely to have received care that departed from professionally recognized standards than cases without that code (4.3% versus 1.7%), after adjusting for patient demographic, geographic, and hospital characteristics.[114]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative Wound Dehiscence generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is related to systematic differences (signal) in hospital performance rather than random variation (noise)—is low, at 35.6%, suggesting that observed differences in risk-adjusted rates may not reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00188, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00171. Signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative Wound Dehiscence is moderate, indicating that the measures may or may not be substantially biased based on the characteristics observed.

Source

An indicator on this topic (9983) was originally proposed by Hannan et al. to target "cases that would have a higher percentage of quality of care problems than cases without the criterion, as judged by medical record review."[115] The same code was included within a broader indicator ("adverse events and iatrogenic complications") in AHRQ's original HCUP Quality Indicators.[116] Iezzoni et al. identified an associated procedure code for reclosure of an abdominal wall dehiscence (5461), and included both codes in the Complications Screening Program.[117] Miller et al. suggested the use of both codes (as "wound disruption") in the original "AHRQ PSI Algorithms and Groupings."[118]

knee replacement surgery. J Clin Epidemiol 1997;50(3):265-73.

[112] Faciszewski T, Johnson L, Noren C, Smith MD. Administrative databases' complication coding in anterior spinal fusion procedures. What does it mean? Spine 1995;20(16):1783-8.

[113] Best W, Khuri S, Phelan M, Hur K, Henderson W, Demakis J, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: Results from the Department of Veterans Affairs national Surgical Quality Improvement Program. J Am Coll Surg 2002;194(3):257-266.

[114] Hannan EL, Bernard HR, O'Donnell JF, Kilburn H, Jr. A methodology for targeting hospital cases for quality of care record reviews. Am J Public Health 1989;79(4):430-6.

[115] Hannan et al., 1989.

[116] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: state and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

[117] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.

[118] Miller M, Elixhauser A, Zhan C, Meyer G, Patient Safety Indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.20 Accidental Puncture or Laceration, Provider Level (PSI 15)

Provider Level Definition (only secondary diagnosis)

Definition	Cases of technical difficulty (e.g., accidental cut or laceration during procedure) per 1,000 discharges.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code denoting technical difficulty (e.g., accidental cut, puncture, perforation, or laceration) in any secondary diagnosis field.
Denominator	All medical and surgical discharges age 18 years and older defined by specific DRGs. Exclude cases: • with ICD-9-CM code denoting technical difficulty (e.g., accidental cut, puncture, perforation, or laceration) in the principal diagnosis field or secondary diagnosis present on admission, if known • MDC 14 (pregnancy, childbirth, and puerperium)
Type of Indicator	Provider level
Empirical Performance	Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

5.21 Accidental Puncture or Laceration, Area Level (PSI 25)

Area Level Definition (principal or secondary diagnosis)

Definition	Cases of technical difficulty (e.g., accidental cut or laceration during procedure) per 100,000 population.
Numerator	Discharges 18 years and older with ICD-9-CM code denoting technical difficulty (e.g., accidental cut, puncture, perforation, or laceration) in any diagnosis field (principal or secondary) of all medical and surgical discharges defined by specific DRGs. Exclude cases: • MDC 14 (pregnancy, childbirth, and puerperium).
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.
Type of Indicator	Area level
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of complications that arise due to technical difficulties in medical care—specifically, those involving an accidental puncture or laceration.

Panel Review

Panelists were unsure about how the culture of quality improvement in a hospital would affect the coding of this complication. Some physicians may be reluctant to record the occurrence of this complication for fear of punishment. Panelists also noted that some of these occurrences are not preventable.

Literature Review

Coding validity. No evidence on validity is available from CSP studies. A study of laparoscopic cholecystectomy found that 95% of

patients with an ICD-9 code of accidental puncture or laceration had a confirmed injury to the bile duct or gallbladder.[119] However, only 27% had a clinically significant injury that required any intervention; sensitivity of reporting was not evaluated. A similar study of cholecystectomies reported that these two ICD-9 codes had a sensitivity of 40% and a predictive value of 23% in identifying bile duct injuries.[120] Among 185 total knee replacement patients, Hawker et al. found that the sensitivity and predictive value of codes describing "miscellaneous mishaps during or as a direct result of surgery" (definition not given) were 86% and 55%, respectively.[121] Romano et al. identified 19 of 45 episodes of accidental puncture, laceration, or related procedure using discharge abstracts of diskectomy patients; there was one false positive.[122]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Accidental Puncture or Laceration generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 82.9%, suggesting that observed differences in risk-adjusted rates most likely reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00279, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00241. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Accidental Puncture or Laceration is high, indicating that the measures likely are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program, although unlike the final PSI, its codes were split between two CSP indicators (CSP 27, "technical difficulty with medical care," and "sentinel events").[123] It was also included as one component of a broader indicator ("adverse events and iatrogenic complications") in AHRQ's original HCUP Quality Indicators.[124] The University HealthSystem Consortium adopted CSP 27 as an indicator for medical (#2806) and major surgery (#2956) patients. Miller et al. also split this set of ICD-9-CM codes into two broader indicators ("miscellaneous misadventures" and "E codes") in the original "AHRQ PSI Algorithms and Groupings."[125] Based on expert consensus panels, McKesson Health Solutions included one component of this PSI (Accidental Puncture or Laceration) in its CareEnhance Resource Management Systems, Quality Profiler Complications Measures Module.

[119] Taylor B. Common bile duct injury during laparoscopic cholecystectomy in Ontario: Does ICD-9 coding indicate true incidence? CMAJ 1998;158(4):481-5.
[120] Valinsky LJ, Hockey RI, Hobbs MS, Fletcher DR, Pikora TJ, Parsons RW, et al. Finding bile duct injuries using record linkage: A validated study of complications following cholecystectomy. J Clin Epidemiol 1999;52(9):893-901.
[121] Hawker GA, Coyte PC, Wright JG, Paul JE, Bombardier C. Accuracy of administrative data for assessing outcomes after knee replacement surgery. J Clin Epidemiol 1997;50(3):265-73.
[122] Romano P. Can administrative data be used to ascertain clinically significant postoperative complications. American Journal of Medical Quality Press.
[123] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[124] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.
[125] Miller M, Elixhauser A, Zhan C, Meyer G, Patient Safety Indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.22 Transfusion Reaction, Provider Level (PSI 16)

Provider Level Definition (only secondary diagnosis)

Definition	Cases of transfusion reaction per 1,000 discharges.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for transfusion reaction in any secondary diagnosis field.
Denominator	All medical and surgical discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), defined by specific DRGs. Exclude cases: • with preexisting (principal diagnosis or secondary diagnosis present on admission, if known) transfusion reaction
Type of Indicator	Provider level
Empirical Performance	Bias: Did not undergo empirical testing of bias
Risk Adjustment	No risk adjustment

5.23 Transfusion Reaction, Area Level (PSI 26)

Area Level Definition (principal or secondary diagnosis)

Definition	Cases of transfusion reaction per 100,000 population.
Numerator	Discharges, 18 years and older **or** MDC 14 (pregnancy, childbirth, and puerperium), with ICD-9-CM code for transfusion reaction in any diagnosis field (principal or secondary) of all medical and surgical discharges defined by specific DRGs.
Denominator	Population of county or Metro Area associated with FIPS code of patient's residence or hospital location.
Type of Indicator	Area level
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of major reactions due to transfusions (ABO and Rh). This indicator is defined both on a provider level (by including cases based on secondary diagnosis associated with the same hospitalization) and on an area level (by including all cases of transfusion reactions).

Panel Review

The overall usefulness of this indicator was rated as very favorable by panelists. This indicator includes only those events that result in additional medical care. Some minor reactions may be missed, although the panel suggested that these minor reactions are less clearly due to medical error than the Rh or ABO reactions included in the indicator.

Literature Review

The project team was unable to find evidence on validity from prior studies, most likely because this complication is quite rare.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Given the low rates or occurrences for Transfusion Reaction, the project team did not measure reliability or minimum bias. The indicator could not be risk-adjusted due to the small number of numerator cases. Users of the PSI software should note the output will only contain observed rates for Transfusion Reaction.

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program (CSP "sentinel events").[126] It was also included as one component of a broader indicator ("adverse events and iatrogenic complications") in AHRQ's original HCUP Quality Indicators.[127] It was proposed by Miller et al. in the original "AHRQ PSI Algorithms and Groupings." [128]

[126] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[127] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.
[128] Miller M, Elixhauser A, Zhan C, Meyer G, Patient safety indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.24 Birth Trauma—Injury to Neonate (PSI 17)

Definition	Cases of birth trauma, injury to neonate, per 1,000 liveborn births.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for birth trauma in any diagnosis field. Exclude infants • with a subdural or cerebral hemorrhage (subgroup of birth trauma coding) **and** any diagnosis code of pre-term infant (denoting birth weight of less than 2,500 grams and less than 37 weeks gestation or 34 weeks gestation or less). • with injury to skeleton (767.3, 767.4) **and** any diagnosis code of osteogenesis imperfecta (756.51).
Denominator	All liveborn births (newborns). The definition of newborn is any neonate with either 1) an ICD-9-CM diagnosis code for an in-hospital liveborn birth or 2) an admission type of newborn (ATYPE=4), age in days at admission equal to zero, and not an ICD-9-CM diagnosis code for an out-of-hospital birth. A neonate is defined as any discharge with age in days at admission between zero and 28 days (inclusive). If age in days is missing, then a neonate is defined as any DRG in MDC 15, an admission type of newborn (ATYPE=4), or an ICD-9-CM diagnosis code for an in-hospital liveborn birth.
Type of Indicator	Provider level
Empirical Performance	Bias: Did not undergo empirical testing of bias
Risk Adjustment	Sex

Summary

This indicator is intended to flag cases of birth trauma for infants born alive in a hospital. The indicator excludes patients born pre-term, as birth trauma in these patients may be less preventable than for full-term infants.

Panel Review

The overall usefulness of this indicator was rated as favorable by panelists

Literature Review

Coding validity. A study of newborns that had a discharge diagnosis of birth trauma found that only 25% had sustained a significant injury to the head, neck, or shoulder.[129] The remaining patients either had superficial injuries or injuries inferior to the neck. The project team was unable to find other evidence on the validity of this indicator. Towner et al. linked California maternal and infant discharge abstracts from 1992 through 1994, but they used only infant discharge abstracts to describe the incidence of neonatal intracranial injury, and they did not report the extent of agreement between the two data sets.[130]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Birth Trauma generally performs well on several different dimensions, including reliability, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than

[129] Hughes C, Harley E, Milmoe G, Bala R, Martorella A. Birth trauma in the head and neck. Arch Otolaryngol Head Neck Surg 1999;125:193-199.

[130] Towner D, Castro MA, Eby-Wilkens E, Gilbert WM. Effect of mode of delivery in nulliparous women on neonatal intracranial injury. N Engl J Med 1999;341(23):1709-14.

random variation (noise)—is high, relative to other indicators, at 97.0%, suggesting that observed differences in risk-adjusted rates reflect true differences across hospitals.

The signal standard deviation for this indicator is also high, relative to other indicators, at 0.04128, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is also high, relative to other indicators, at 0.13603. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The bias for Birth Trauma was not measured, since adequate risk adjustment was not available.

Source

This indicator has been widely used in the obstetric community, although it is most commonly based on chart review rather than administrative data. It was proposed by Miller et al. in the original "AHRQ PSI Algorithms and Groupings."[131] Based on expert consensus panels, McKesson Health Solutions included a broader version of this indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications Measures Module.

[131] Miller M, Elixhauser A, Zhan C, Meyer G, Patient Safety Indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.25 Obstetric Trauma—Vaginal Delivery with Instrument (PSI 18)

Definition	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 instrument-assisted vaginal deliveries.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for 3rd and 4th degree obstetric trauma in any diagnosis or procedure field.
Denominator	All vaginal delivery discharges with any procedure code for instrument-assisted delivery.
Type of Indicator	Provider level
Empirical Performance	Bias: Did not undergo empirical testing of bias
Risk Adjustment	Age, comorbidity categories

Summary

This indicator is intended to flag cases of potentially preventable trauma during vaginal delivery with instrument.

Panel Review

The overall usefulness of an Obstetric trauma indicator was rated as favorable by panelists. After initial review, the indicator was eventually split into three separate Obstetric Trauma indicators: Vaginal Delivery with Instrument, Vaginal Delivery without Instrument, and Cesarean Delivery.

Literature Review

Coding validity. In a stratified probability sample of vaginal and Cesarean deliveries, the weighted sensitivity and predictive value of coding for third- and fourth-degree lacerations and vulvar/perineal hematomas (based on either diagnosis or procedure codes) were 89% and 90%, respectively.[158] The authors did not report coding validity for third- and fourth-degree lacerations separately. The project team was unable to find other evidence on validity from prior studies.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Obstetric Trauma—Vaginal Delivery with Instrument generally performs well on several different dimensions, including reliability, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 69.9%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is also high, relative to other indicators, at 0.09794, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is high, relative to other indicators, at 0.05539. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The bias for Obstetric Trauma—Vaginal Delivery with Instrument was not measured, since adequate risk adjustment was not available.

Source

An overlapping subset of this indicator (third- or fourth-degree perineal laceration) has been adopted by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) as a core performance measure for "pregnancy and related conditions" (PR-25). Based on expert consensus panels, McKesson Health Solutions included the JCAHO indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications

Measures Module. Fourth Degree Laceration, one of the codes mapped to this PSI, was included as one component of a broader indicator ("obstetrical complications") in AHRQ's original HCUP Quality Indicators.[132]

[132] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

5.26 Obstetric Trauma—Vaginal Delivery without Instrument (PSI 19)

Definition	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 vaginal deliveries without instrument assistance.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for 3rd and 4th degree obstetric trauma in any diagnosis or procedure field.
Denominator	All vaginal delivery discharges. Exclude cases: • with instrument-assisted delivery.
Type of Indicator	Provider level
Empirical Performance	Population Rate (2003): 46.340 per 1,000 population at risk Bias: Did not undergo empirical testing of bias
Risk Adjustment	Age, comorbidity categories

Summary

This indicator is intended to flag cases of potentially preventable trauma during a vaginal delivery without instrument.

Panel Review

The overall usefulness of an Obstetric Trauma Indicator was rated as favorable by panelists. After initial review, the indicator was split into three separate Obstetric Trauma indicators: Vaginal Delivery with Instrument, Vaginal Delivery without Instrument, and Cesarean Delivery.

Literature Review

Coding validity. In a stratified probability sample of vaginal and Cesarean deliveries, the weighted sensitivity and predictive value of coding for third- and fourth-degree lacerations and vulvar/perineal hematomas (based on either diagnosis or procedure codes) were 89% and 90%, respectively.[158] The authors did not report coding validity for third- and fourth-degree lacerations separately. The project team was unable to find other evidence on validity from prior studies.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Obstetric Trauma—Vaginal Delivery without Instrument generally performs well on several different dimensions, including reliability, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is high, relative to other indicators, at 86.4%, suggesting that observed differences in risk-adjusted rates reflect true differences across hospitals.

The signal standard deviation for this indicator is also high, relative to other indicators, at 0.04314, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is lower than many other indicators, at 0.02470. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The bias for Obstetric Trauma—Vaginal Delivery without Instrument was not measured, since adequate risk adjustment was not available.

Source

An overlapping subset of this indicator (third- or fourth-degree perineal laceration) has been adopted by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) as a core performance measure for "pregnancy and related conditions" (PR-25). Based on expert consensus panels, McKesson Health Solutions included the JCAHO indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications Measures Module. Fourth-Degree Laceration, one of the codes mapped to this PSI, was included as one component of a broader indicator ("obstetrical complications") in AHRQ's original HCUP Quality Indicators.[133]

[133] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

5.27 Obstetric Trauma—Cesarean Delivery (PSI 20)

Definition	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 Cesarean deliveries.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for obstetric trauma in any diagnosis or procedure field.
Denominator	All Cesarean delivery discharges.
Type of Indicator	Provider level
Empirical Performance	Bias: Did not undergo empirical testing of bias
Risk Adjustment	No risk adjustment

Summary

This indicator is intended to flag cases of potentially preventable trauma during Cesarean delivery.

Panel Review

The overall usefulness of an Obstetric Trauma Indicator was rated as favorable by panelists. After initial review, the indicator was eventually split into three separate Obstetric Trauma indicators: Vaginal Delivery with Instrument, Vaginal Delivery without Instrument, and Cesarean Delivery.

Literature Review

Coding validity. In a stratified probability sample of vaginal and Cesarean deliveries, the weighted sensitivity and predictive value of coding for third- and fourth-degree lacerations and vulvar/perineal hematomas (based on either diagnosis or procedure codes) were 89% and 90%, respectively.[158] The authors did not report coding validity for third- and fourth-degree lacerations separately. The project team was unable to find other evidence on validity from prior studies.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Obstetric Trauma—Cesarean Delivery generally performs well on several different dimensions, including reliability, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is lower than many indicators, at 45.9%, suggesting that observed differences in risk-adjusted rates may not reflect true differences across hospitals.

The signal standard deviation for this indicator is also lower than many indicators, at 0.00590, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00576. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The bias for Obstetric Trauma—Cesarean Delivery was not measured, since adequate risk adjustment was not available.

Source

An overlapping subset of this indicator (third- or fourth-degree perineal laceration) has been adopted by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) as a core performance measure for "pregnancy and related conditions" (PR-25). Based on expert consensus panels, McKesson Health Solutions included the JCAHO indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications Measures Module. Fourth Degree Laceration,

one of the codes mapped to this PSI, was included as one component of a broader indicator ("obstetrical complications") in AHRQ's original HCUP Quality Indicators.[134]

[134] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

6.0 Using Different Types of QI Rates

When should you use the observed, expected, risk adjusted, and/or smoothed rates generated by the AHRQ QI software? Here are some guidelines.

If the user's primary interest is to identify cases for further follow-up and quality improvement, then the observed rate would help to identify them. The observed rate is the raw rate generated by the QI software from the data the user provided. Areas for improvement can be identified by the magnitude of the observed rate compared to available benchmarks and/or by the number of patients impacted.

Additional breakdowns by the default patient characteristics used in stratified rates (e.g., age, gender, or payer) can further identify the target population. Target populations can also be identified by user-defined patient characteristics supplemented to the case/discharge level flags. Trend data can be used to measure change in the rate over time.

Another approach to identify areas to focus on is to compare the observed and expected rates. The expected rate is the rate the provider would have if it performed the same as the reference population given the provider's actual case-mix (e.g., age, gender, DRG, and comorbidity categories).

If the observed rate is higher than the expected rate (i.e., the ratio of observed/expected is greater than 1.0, or observed minus expected is positive), then the implication is that the provider performed worse than the reference population for that particular indicator. Users may want to focus on these indicators for quality improvement.

If the observed rate is lower than the expected rate (i.e., the ratio of observed/expected is less than 1.0, or observed minus expected is negative), then the implication is that the provider performed better than the reference population. Users may want to focus on these indicators for identifying best practices.

Users can also compare the expected rate to the population rate reported in the *Comparative Data* document to determine how their case-mix compares to the reference population. The population rate refers to the overall rate for the reference population. The reference population is defined in the *Comparative Data* Document. If the population rate is higher than the expected rate, then the provider's case-mix is less severe than the reference population. If the population rate is lower than the expected rate, then the provider's case-mix is more severe than the reference population.

We use this difference between the population rate and the expected rate to "adjust" the observed rate to account for the difference between the case-mix of the reference population and the provider's case-mix. This is the provider's risk-adjusted rate.

If the provider has a less severe case-mix, then the adjustment is positive (population rate > expected rate) and the risk-adjusted rate is higher than the observed rate. If the provider has a more severe case-mix, then the adjustment is negative (population rate < expected rate) and the risk-adjusted rate is lower than the observed rate. The risk-adjusted rate is the rate the provider would have if it had the same case-mix as the reference population given the provider's actual performance.

Finally, users can compare the risk-adjusted rate to the smoothed or "reliability-adjusted" rate to determine whether this difference between the risk-adjusted rate and reference population rate is likely to remain in the next measurement period. Smoothed rates are weighted averages of the population rate and the risk-adjusted rate, where the weight reflects the reliability of the provider's risk-adjusted rate.

A ratio of (smoothed rate - population rate) / (risk-adjusted rate - population rate) greater than 0.80 suggests that the difference is likely to persist (whether the difference is positive or negative). A ratio less than 0.80 suggests that the difference may be due in part to random differences in patient characteristics (patient characteristics that are not observed and controlled for in the risk-adjustment model). In general, users may want to focus on areas where the differences are more likely to persist.

7.0 References

Ball JK, Elixhauser A, Johantgen M, et al. HCUP Quality Indicators, Methods, Version 1.1: Outcome, Utilization, and Access Measures for Quality Improvement. (AHCPR Publication No. 98-0035). Healthcare Cost and Utilization project (HCUP-3) Research notes: Rockville, MD: Agency for Health Care Policy and Research, 1998.

Barbour GL. Usefulness of a discharge diagnosis of sepsis in detecting iatrogenic infection and quality of care problems. Am J Med Qual 1993;8(1):2-5.

Berlowitz D, Brand H, Perkins C. Geriatric syndromes as outcome measures of hospital care: Can administrative data be used? JAGS 1999;47:692-696.

Best W, Khuri S, Phelan M, Hur K, Henderson W, Demakis J, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: Results from the Department of Veterans Affairs National Surgical Quality Improvement Program. J Am Coll Surg 2002;194(3):257-266.

Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991;324(6):370-6.

Christiansen CL, Morris CN. Improving the statistical approach to health care provider profiling. Ann Intern Med 1997;127(8 Pt 2):764-8).

Davies S, Geppert J, McClellan M, McDonald KM, Romano PS, Shojania KG. Refinement of the HCUP Quality Indicators. Technical Review Number 4. Rockville, MD: (Prepared by UCSF-Stanford Evidence-based Practice Center under Contract No. 290-97-0013) Agency for Healthcare Research and Quality; 2001. Report No.: 01-0035.

EMBASE. In. The Netherlands: Elsevier Science Publishers B.V.

Envisioning the National Health Care Quality Report. Washington, DC: Institute of Medicine; 2001.

Faciszewski T, Johnson L, Noren C, Smith MD. Administrative databases' complication coding in anterior spinal fusion procedures. What does it mean? Spine 1995;20(16):1783-8.

Fitch K, Bernstein J, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. the RAND/UCLA Appropriateness Method User's Manual: RAND; 2001.

Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. In-hospital complications among survivors of admission for congestive heart failure, chronic obstructive pulmonary disease, or diabetes mellitus. J Gen Intern Med 1995;10(6):307-14.

Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. Med Care 1997;35(6):589-602.

Green L, Lewis F. measurement and Evaluation in Health Education and Health Promotion. Mountain View, CA: Mayfield Publishing Company; 1998.

Hannan EL, Bernard HR, O'Donnell JF, Kilburn H, Jr. A methodology for targeting hospital cases for quality of care record reviews. Am J Public Health 1989;79(4):430-6.

Hartz AJ, Kuhn EM. Comparing hospitals that perform coronary artery bypass surgery: The effect of outcome measures and data sources. Am J Public Health 1994;84(10):1609-14.

Hawker BA, Coyte PC, Wright JG, Paul JE, Bombardier C. Accuracy of administrative data for assessing

outcomes after knee replacement surgery. J Clin Epidemiol 1997;50(3):265-73.

Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician “report cards” for assessing the costs and quality of care of a chronic disease JAMA 1999;281(22):2098-105.

Hughes C, Harley E, Milmoe G, Bala R, Martorella A. Birth trauma in the head and neck. Arch Otolaryngol Head Neck Surg 1999;125:193-199.

Iezzoni L, Lawthers A, Petersen L, McCarthy E, Palmer R, Cahalane M, et al. Project to validate the Complications Screening Program: Health Care Financing Administration; 1998 March 31. Report No: HCFA Contract 500-94-0055.

Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.

Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.

Iezzoni LI, Foley SM, Heeren T, Daley J, Duncan CC, Fisher ES, et al. A method for screening the quality of hospital care using administrative data: preliminary validation results. QRB Qual Rev Bull 1992;18(11):361-71.

Impact: Case Studies Notebook – Documented Impact and Use of AHRQ's Research. Compiled by Division of Public Affairs, Office of Health Care Information, Agency for Healthcare Research and Quality.

Institute of Medicine. To Err is Human: Building a Safer Health System. Kohn LT, Corrigan JM, Donaldson MS (eds.) Washington DC: National Academy Press, 2000.

Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: state and national applications. Jt Comm J Qual Improv 1998;24(2):88-105.

Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: state and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

Keeler E, Kahn K, Bentow S. Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare Provider Analysis and Review file. Springfield, VA: NTIS; 1991.

Kovner C, Gergen PH. Nurse staffing levels and adverse events following surgery in U.S. hospitals. Image J Nurs Sch 1998;30(4):315-21.

Lawthers A, McCarthy E, Davis R, Peterson L, Palmer R, Iezzoni L. Identification of in-hospital complications from claims data: is it valid? Medical Care 2000;38(8):785-795.

Lichtig LK, Knauf RA, Hilholland DK. Some impacts of nursing on acute care hospital outcomes. J Nurs Adm 1999;29(2):25-33.

McCarthy EP, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamael MB, et al. Does clinical evidence support ICD-9-CM diagnosis coding of complications? Med Care 2000;38(8);868-876.

McDonald KM, Romano PS, Geppert J, Davies S, Duncan BW, Shojania KG. Measures of Patient Safety Based on Hospital Administrative Data-The Patient Safety Indicators. Technical Review 5 (Prepared by the University of California San Francisco-Stanford Evidence-based Practice Center under Contract No. 290-97-0013). AHRQ Publication No. 02-0038 . Rockville, MD: Agency for Healthcare Research and Quality. August 2002. (http://www.qualityindicators.ahrq.gov/downloads.htm)

Measuring the Quality of Health Care: A statement of the National Roundtable on Healthcare Quality Division of Healthcare Services: National Academy Press; 1999.

MEDLINE [database online]. In. Bethesda (MD): National Library of Medicine.

Miller M, Elixhauser A, Zhan C, Meyer G, Patient Safety Indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

National Roundtable on Healthcare Quality, 1999.

Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.: 230-88-0021.

Nursing-Sensitive Quality Indicators for Acute Care Settings and ANA's Safety & Quality Initiative. In: American Nurses Association; 1999.

Romano P. Can administrative data be used to ascertain clinically significant postoperative complications. American Journal of Medical Quality Press.

Shojania KG, Duncan BW, McDonald KM, Wachter RM. Making health care safer: A critical analysis of patient safety practices. Evidence Report/Technology Assessment No. 43 (Prepared by the University of California at San Francisco-Stanford Evidence-based Practice Center under Contract No. 290-97-0013). Rockville, MD: Agency for Healthcare Research and Quality; 2001. Report No.: AHRQ Publication No. 01-E058.

Silber J, Rosenbaum P, Ross R. Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics? J Am Stat Assoc 1995;90:7-18.

Silber JH, Rosenbaum PR, Williams SV, Ross RN, Schwartz JS. The relationship between choice of outcome measure and hospital rank in general surgical procedures: Implications for quality assessment. Int J Qual Health Care 1997;9(3):193-200.

Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. Med Care 1992;30(7):615-29.

Taylor B. Common bile duct injury during laparoscopic cholecystectomy in Ontario: Does ICD-9 coding indicate true incidence? CMAJ 1998;158(4):481-5.

Towner D, Castro MA, Eby-Wilkens E, Gilbert WM. Effect of mode of delivery in nulliparous women on neonatal intracranial injury. N Engl J Med 1999;341(23):1709-14.

Valinsky LJ, Hockey RI, Hobbs MS, Fletcher DR, Pikora TJ, Parsons RW, et al. Finding bile duct injuries using record linkage: A validated study of complications following cholecystectomy. J Clin Epidemiol 1999;52(9):893-901.

Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: Validation of the Complications Screening Program. Med Care 2000;38(8):796-806.

White RH, Romano P, Zhou H, Rodrigo J, Barger W. Incidence and time course of thromboembolic outcomes following total hip or knee arthroplasty. Arch Intern Med 1998;158(14):1525-31

.

Appendix A: Links

The following links may be helpful to users of the AHRQ Patient Safety Indicators.

Patient Safety Indicators Version 3.1 Documents and Software

Available at http://www.qualityindicators.ahrq.gov/psi_download.htm

Title	***Description***
Guide to Patient Safety Indicators	Describes how the PSIs were developed and provides detailed evidence for each indicator.
Patient Safety Indicators Technical Specifications	Provides detailed definitions of each PSI, including all ICD-9-CM and DRG codes that are included in or excluded from the numerator and denominator. Note that exclusions from the denominator are automatically applied to the numerator.
PSI Covariates used in Risk Adjustment	Tables for each PSI provide the stratification and coefficients used to calculate the risk-adjusted rate for each strata.
SAS® PSI Software Documentation	This software documentation provides detailed instructions on how to use the SAS ® version of the PSI software including data preparation, calculation of the PSI rates, and interpretation of output.
Patient Safety Indicators Comparative Data	This document provides the provider rate, provider standard deviation, area rate, and population rate, as appropriate, for each indicator.
Change Log to PSI Documents and Software	The Change Log document provides a cumulative summary of all changes to the PSI software, software documentation, and other documents made since the release of version 2.1 of the software in March 2003. Changes to indicator specifications that were not a result of new ICD-9-CM and DRG codes, are also described in the Change Log.
Fiscal year 2006 Coding Changes	This document summarizes the changes to the indicator definitions resulting from FY 2006 changes to ICD-9-CM coding and DRG changes. These changes will only affect data from FY 2006 (October 1, 2005) or later.
SAS® PSI Software	Requires the SAS® statistical program distributed by the SAS Institute, Inc. The company may be contacted directly regarding the licensing of its products: http://www.sas.com

AHRQ QI Windows Application

The AHRQ QI Windows Application calculates rates for all of the AHRQ Quality Indicators modules and does not require either SAS®. It is available at:

http://www.qualityindicators.ahrq.gov/winqi_download.htm

Additional Documents

The following documents are available within the "Documentation" section of the AHRQ QI Downloads Web page:

(http://www.qualityindicators.ahrq.gov/downloads.htm).

- *Refinement of the HCUP Quality Indicators (Technical Review), May 2001*
- *Refinement of the HCUP Quality Indicators (Summary), May 2001*
- *Measures of Patient Safety Based on Hospital Administrative Data - The Patient Safety Indicators, August 2002*
- *Measures of Patient Safety Based on Hospital Administrative Data - The Patient Safety Indicators (Summary), August 2002*

In addition, these documents may be accessed at the AHRQ QI Documentation Web page:

http://www.qualityindicators.ahrq.gov/documentation.htm

- *Guidance for Using the AHRQ Quality Indicators for Hospital-level Public Reporting or Payment, August 2004*
- *AHRQ Summary Statement on Comparative Hospital Public Reporting, December 2005*
- *Appendix A: Current Uses of AHRQ Quality Indicators and Considerations for Hospital-level*
- *Comparison of Recommended Evaluation Criteria in Five E*xisting National Frameworks

The following documents can be viewed or downloaded from the page:

http://www.qualityindicators.ahrq.gov/newsletter.htm

- *2006 Area Level Indicator Changes*
- *Considerations in Public Reporting for the AHRQ QIs*
- *June 2005 Newsletter* - Contains the article, "Using Different Types of QI Rates"

Other Tools and Information

The PSI SAS software no longer incorporates the AHRQ Comorbidity software. Before running the PSI software, the user will need to download and run program available at:

http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp

That program will create the comorbidity variables in the user's data file. These variables are only needed if the user intends to calculate risk-adjusted rates using PSSASP3.

Area indicators can be calculated using the modified Federal Information Processing Standards (FIPS) State/county code. A list of codes is available at:

http://www.census.gov/popest/geographic/codes02.pdf

AHRQ provides a free, on-line query system based on HCUP data that provides access to health statistics and information on hospital stays at the national, regional, and State level. It is available at:

http://hcup.ahrq.gov/HCUPnet.asp