

**Agency for Healthcare Research and Quality  
AHRQ Quality Indicators (AHRQ QI)**

**Guidance on Using the AHRQ QI for Hospital-Level  
Comparative Reporting  
June 2009**

## Table of Contents

1	Introduction.....	1
2	Framework for Evaluating AHRQ QI .....	3
2.1	Evidentiary Areas.....	3
2.2	Evidence Gap Criteria by Evidentiary Areas .....	3
2.3	Indicator Tiering.....	5
3	Summary of Recommendations.....	6
3.1	Tier 1: Minor or no evidence gaps. ....	6
3.2	Tier 2: Moderate evidence gaps. ....	8
3.3	Tier 3: Significant, but addressable, evidence gaps. ....	8
3.4	Tier 4: Significant evidence gaps. ....	9
4	Recommendations by AHRQ QI Module .....	10
4.1	Inpatient Quality Indicators (IQI) .....	10
4.1.1	IQI #01 - Esophageal Resection Volume.....	10
4.1.2	IQI #02 - Pancreatic Resection Volume .....	10
4.1.3	IQI #04 - Abdominal Aortic Aneurysm (AAA) Repair Volume.....	11
4.1.4	IQI #05 - CABG Volume.....	11
4.1.5	IQI #06 - PTCA Volume.....	12
4.1.6	IQI #07 - Carotid Endarterectomy Volume .....	12
4.1.7	IQI #08 - Esophageal Resection Mortality .....	13
4.1.8	IQI #09 - Pancreatic Resection Mortality .....	13
4.1.9	IQI #11 - Abdominal Aortic Aneurysm (AAA) Repair Mortality.....	14
4.1.10	IQI #12 - CABG Mortality .....	14
4.1.11	IQI #13 - Craniotomy Mortality .....	15
4.1.12	IQI #14 - Hip Replacement Mortality.....	15
4.1.13	IQI #15 - AMI Mortality.....	16
4.1.14	IQI #16 - CHF Mortality.....	16
4.1.15	IQI #17 - Acute Stroke Mortality.....	17
4.1.16	IQI #18 - GI Hemorrhage Mortality .....	18
4.1.17	IQI #19 - Hip Fracture Mortality .....	18
4.1.18	IQI #20 - Pneumonia Mortality.....	18
4.1.19	IQI #21 and #33 - Cesarean Delivery .....	19
4.1.20	IQI #22 and #34 - VBAC Delivery.....	19
4.1.21	IQI #23 - Laparoscopic Cholecystectomy .....	20
4.1.22	IQI #24 - Incidental Appendectomy in the elderly .....	20
4.1.23	IQI #25 - Bi-lateral Catheterization .....	21
4.1.24	IQI #30 - PTCA Mortality .....	21
4.1.25	IQI #31 - Carotid Endarterectomy Mortality .....	22
4.2	Patient Safety Indicators (PSI).....	23
4.2.1	PSI #01 - Complications of Anesthesia .....	23
4.2.2	PSI #02 - Death in Low Mortality DRGs .....	23
4.2.3	PSI #03 - Pressure Ulcer .....	24
4.2.4	PSI #04 - Death among Surgical Inpatients with Treatable Serious Complications 24	
4.2.5	PSI #05 - Foreign Body left in During Procedure .....	25
4.2.6	PSI #06 - Iatrogenic Pneumothorax.....	25

4.2.7	PSI #07 - Central Line-Associated Bloodstream Infection.....	26
4.2.8	PSI #08 - Postoperative Hip Fracture .....	26
4.2.9	PSI #09 - Postoperative Hemorrhage or Hematoma.....	27
4.2.10	PSI #10 - Postoperative Physiologic and Metabolic Derangement .....	27
4.2.11	PSI #11 - Postoperative Respiratory Failure.....	28
4.2.12	PSI #12 - Postoperative DVT or PE .....	28
4.2.13	PSI #13 - Postoperative Sepsis .....	29
4.2.14	PSI #14 - Postoperative Wound Dehiscence .....	30
4.2.15	PSI #15 - Accidental Puncture or Laceration .....	30
4.2.16	PSI #16 - Transfusion Reaction .....	31
4.2.17	PSI #17 - Birth Trauma.....	31
4.2.18	PSI #18 - OB Trauma-Vaginal Delivery with Instrument.....	32
4.2.19	PSI #19 - OB Trauma-Vaginal Delivery without Instrument.....	32
4.2.20	PSI #20 - OB Trauma-Cesarean Delivery .....	32
4.3	Pediatric Quality Indicators (PDI).....	33
4.3.1	PDI #01 - Accidental Puncture or Laceration.....	33
4.3.2	PDI #02 - Pressure Ulcer .....	33
4.3.3	PDI #03 - Foreign Body left in During Procedure.....	34
4.3.4	PDI #05 - Iatrogenic Pneumothorax .....	34
4.3.5	PDI #06 - Pediatric Heart Surgery Mortality .....	35
4.3.6	PDI #07 - Pediatric Heart Surgery Volume .....	36
4.3.7	PDI #08 - Postoperative Hemorrhage or Hematoma .....	36
4.3.8	PDI #09 - Postoperative Respiratory Failure .....	37
4.3.9	PDI #10 - Postoperative Sepsis.....	37
4.3.10	PDI #11 - Postoperative Wound Dehiscence.....	38
4.3.11	PDI #12 - Central Line-Associated Bloodstream Infection.....	39
4.3.12	PDI #13 - Transfusion Reaction .....	39
4.4	Neonatal Quality Indicators (NQI).....	40
4.4.1	NQI #01 - Iatrogenic Pneumothorax in Neonates.....	40
4.4.2	NQI #02 - Neonatal Mortality.....	40
4.4.3	NQI #03 - Blood Stream Infections in Neonates .....	41

### **List of Tables**

Tier 1. AHRQ QI Suitable for Comparative Reporting .....	6
Tier 1B: Serious Reportable Events Suitable for Public Accountability .....	7
Tier 2. AHRQ QI Potentially Suitable for Comparative Reporting with <u>Some</u> Additional Development .....	8
Tier 3. AHRQ QI that are Potentially Suitable for Comparative Reporting Only with <u>Significant</u> Additional Development.....	8
Tier 4. AHRQ QI that are Not Suitable for Comparative Reporting as Currently Specified.....	9

# 1 Introduction

The Agency for Healthcare Research and Quality (AHRQ) encourages comparative reporting of hospital performance data and considers comparative reporting an important strategy to advance the quality improvement agenda in healthcare. Evidence shows that publicly reporting comparative performance data by hospitals is a key element that promotes enhanced patient care. AHRQ is an active participant in such public reporting efforts as the Hospital Quality Alliance, the AQA Alliance, the National Quality Forum (NQF), and others. In fact, AHRQ submitted the AHRQ Quality Indicators (AHRQ QI) to the NQF consensus development process for consideration for endorsement for public reporting in the summer of 2006, and has continued to respond to recent NQF call for measures on topics that overlap with AHRQ QI content areas.

There are a number of measure sets available to health care organizations, and the AHRQ QI are one set of many that may be used for comparative (public) reporting and pay-for-performance. The AHRQ QI are a unique set of measures that use readily-available hospital inpatient administrative data. Although the AHRQ QI were not initially developed for the purpose of hospital-specific comparative reporting, they have been and are currently being used for public reporting and pay-for-performance. This document provides guidance for using the AHRQ QI for public reporting and payment initiatives. Additional guidance on public reporting using a web-based approach is available from the NQF.

In order to provide guidance to users interested in applying the AHRQ QI for use in comparative reporting, and to support submission of selected AHRQ QI to the NQF endorsement process, AHRQ undertook an analysis of each indicator to determine the appropriateness for use in comparative reporting, and developed a comparative reporting template based on input from consumers, providers, experts in the field of public reporting and others. In addition, several of the indicators underwent additional validation using empirical analysis, clinical panel reviews, and medical record reviews. Finally, the indicators submitted to the NQF endorsement process underwent review by several technical advisory panels. Users of the AHRQ QI have always submitted questions and concerns through the technical support e-mail ([support@qualityindicators.ahrq.gov](mailto:support@qualityindicators.ahrq.gov)). This document summarizes the results of these analyses and reflects the current views regarding appropriate use of the AHRQ QI.

Of course, the decision on how and whether to use the AHRQ QI or any other measure set is ultimately a local matter and depends on various local issues such as data availability and data quality, legislative mandates, confidentiality issues and data use agreements, and resources to name a few. AHRQ will continue to update this document with evidence as it becomes available that will inform and further clarify hospital-specific public reporting issues and the role of the AHRQ QI in that venue.

Given the increasingly important role of administrative data in comparative reporting and pay-for-performance efforts, it is important to closely monitor current initiatives using the AHRQ QI so that we can all continue to improve the measures and the ways they are used. Administrative data is used for national tracking, public reporting, pay-for-performance and quality improvement, so it is important to strengthen the quality and availability of such data at

the state and local levels by, for example, adding new or refining existing data elements and improving the quality and consistency of the coding of diagnosis and procedure codes.

## 2 Framework for Evaluating AHRQ QI

This document is considered a supplement to the currently published documents supporting the AHRQ Quality Indicators. These documents provide more detailed summaries of literature evidence, previous empirical analyses, and current detailed definitions. Documents may be downloaded from <http://www.qualityindicators.ahrq.gov>.

### 2.1 Evidentiary Areas

In order to evaluate an indicator for appropriateness of use for comparative reporting, five important evidentiary areas are identified.

- **Importance:** A quality indicator must reflect variation in quality or have low levels of overall performance. It should have the potential to inform consumer choice and warrant the time and effort required for providers to improve.
- **Face validity:** A quality indicator must have sound clinical and/or empirical rationale for its use. It should measure an important aspect of quality that is subject to provider or health care system control.
- **Coding or criterion validity:** The cases identified by the quality indicator reflect actual cases of the concept of interest (e.g. outcome of interest). This includes evidence of coding accuracy.
- **Construct validity:** The quality indicator should be supported by evidence of a relationship to other indicators intended to measure the same or related aspects of quality, such as process of care that are considered best practice or other outcome measures that show harm to patients.
- **Risk adjustment:** The quality indicator is not subject to substantive systematic bias due to case mix, or adequate risk adjustment can be applied using available data.

For each evidentiary area above, an indicator is identified as either having an evidentiary gap or not. These designations were made according to pre-set criteria, set forth below. It is important to remember that an evidentiary gap does not mean an indicator is not appropriate for use. Because many quality measures have not been systematically studied, most indicators have one, two or three gaps in supporting evidence. Many times, such as the case with construct validity gaps, this is simply because the studies have not been done. In other cases, coding practices may have improved since the original study or coding guidelines may have changed. Users of the AHRQ QI should take into account where the evidentiary gaps are for each measure and the degree of the evidentiary gaps when selecting measures for a given use.

### 2.2 Evidence Gap Criteria by Evidentiary Areas

Each evidentiary area has criteria for determining whether or not an evidence gap exists

- **Importance:**

- Because all indicators had to have substantial evidence for importance and face validity in the initial AHRQ QI screening process, either from prior use, published literature or clinical panel review, all indicators were assumed to pass this criterion, unless additional evidence has suggested otherwise.
- Information gathered since the development of the indicator suggests that the indicator is not viewed as important in quality improvement efforts, or is viewed as having little leverage capability.
- Face validity:
  - Because all indicators had to have substantial evidence for importance and face validity, either from prior use, published literature or clinical panel review, all indicators were assumed to pass this criterion, unless additional evidence has suggested otherwise.
  - Information gathered since the development of the indicator suggests that the indicator is not viewed as valid by a majority of clinicians or other stakeholders.
- Criterion validity:
  - Evidence from published literature or otherwise available chart review indicating poor sensitivity or specificity of coding for the complication, or other coding problems that reduce the accuracy of the indicator.
  - Evidence regarding criterion validity from literature and chart review is highly contradictory.
  - When indicator definitions differed from those used in studies, judgments regarding the impact of the differences was made, and preferences were given to studies using identical or the most similar definitions.
- Construct validity:
  - No available evidence establishes a clear link between processes of care and the outcome of interest. Volume-outcome relationships, or simple structure-outcomes relationships (such as staffing) do not qualify.
  - Evidence regarding construct validity from chart review or literature is highly contradictory.
  - When indicators definitions differed from those used in studies, judgments regarding the impact of the differences was made, and preferences were given to studies using identical or the most similar definitions.
- Risk adjustment:

- Evidence documents important risk factors which cannot be identified using administrative data, unless those risk factors have been shown not to vary systematically by provider or geographic area.

### **2.3 Indicator Tiering**

Based on the current evidence and identified gaps the indicators are assigned to one of four tiers.

- Tier 1: Minor or no evidence gaps. Indicators in this tier have the strongest evidence base, with established evidence in several or most evidentiary areas and no substantial evidence suggesting that the indicator may not be useful for comparative reporting purposes. These indicators have mostly been endorsed by the National Quality Forum. A subset of Tier 1 - labeled Tier 1B - are serious reportable events that are most suitable for public accountability (no causality) and root cause analysis, but for which comparisons between hospitals are may not be useful because the rate is not predictive of future performance.
- Tier 2: Moderate evidence gaps. Indicators in this tier have some evidence supporting their use, but all have some evidence gaps. These gaps are not judged to be serious enough to preclude use in comparative reporting, but these indicators would be stronger if highlighted issues are better understood through future validation efforts. These indicators are not currently endorsed by the National Quality Forum.
- Tier 3: Significant, but addressable, evidence gaps. Indicators in this tier have at least one serious evidentiary gap or concern, which requires further development or validation work before the indicator is considered appropriate for comparative reporting. The required improvements and validation efforts are specified and underway, and once completed these indicators will be re-tiered. These indicators are not currently endorsed by the National Quality Forum.
- Tier 4: Significant evidence gaps. The indicators in this tier have substantial gaps in the evidence that are unlikely to be addressed with further development or validation work. These indicators are not recommended for comparative reporting, but rather, can be used for internal quality improvement. These indicators are not currently endorsed by the National Quality Forum.

### 3 Summary of Recommendations

The AHRQ QI consist of three modules with indicators at the hospital level. The Inpatient Quality Indicators (IQI) include indicators of volume, mortality for procedures and conditions, and utilization. The Patient Safety Indicators (PSI) include indicators of potential preventable adverse events and serious reportable events. The Pediatric Quality Indicators (PDI) include indicators of potential preventable adverse events and serious reportable events in children, pediatric heart surgery volume and mortality and a subset of indicators specific to neonates (called the Neonatal Quality Indicators).

#### 3.1 Tier 1: Minor or no evidence gaps.

Indicators in this tier have the strongest evidence base, with established evidence in several or most evidentiary areas and no substantial evidence suggesting that the indicator may not be useful for comparative reporting purposes. These indicators have mostly been endorsed by the National Quality Forum. A subset of Tier 1 - labeled Tier 1B - are serious reportable events that are most suitable for public accountability and root cause analysis, but for which comparisons between hospitals may not be useful.

#### Tier 1. AHRQ QI Suitable for Comparative Reporting

Indicator	Label	Module	Type
IQI #01	Esophageal Resection Volume <sup>1</sup>	Inpatient Quality Indicators	Volume
IQI #02	Pancreatic Resection Volume <sup>1</sup>	Inpatient Quality Indicators	Volume
IQI #04	Abdominal Aortic Aneurysm (AAA) Repair Volume <sup>1</sup>	Inpatient Quality Indicators	Volume
IQI #08	Esophageal Resection Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #09	Pancreatic Resection Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #11	Abdominal Aortic Aneurysm (AAA) Repair Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #15	AMI Mortality	Inpatient Quality Indicators	Mortality
IQI #16	CHF Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #17	Acute Stroke Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #19	Hip Fracture Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #20	Pneumonia Mortality <sup>1</sup>	Inpatient Quality Indicators	Mortality
IQI #24	Incidental Appendectomy in the Elderly <sup>1</sup>	Inpatient Quality Indicators	Utilization
IQI #25	Bi-lateral Catheterization <sup>1</sup>	Inpatient Quality Indicators	Utilization
IQI #91	Mortality for Selected Conditions <sup>1</sup>	Inpatient Quality	Composite

Indicator	Label	Module	Type
		Indicators	
PDI #01	Accidental Puncture or Laceration <sup>1</sup>	Pediatric Quality Indicators	Adverse Event
PDI #02	Pressure Ulcer <sup>1</sup>	Pediatric Quality Indicators	Adverse Event
PDI #05	Iatrogenic Pneumothorax <sup>1</sup>	Pediatric Quality Indicators	Adverse Event
PDI #06	Pediatric Heart Surgery Mortality <sup>1</sup>	Pediatric Quality Indicators	Mortality
PDI #07	Pediatric Heart Surgery Volume <sup>1</sup>	Pediatric Quality Indicators	Volume
PDI #11	Postoperative Wound Dehiscence <sup>1</sup>	Pediatric Quality Indicators	Adverse Event
NQI #03	Blood Stream Infection in Neonates <sup>2</sup>	Neonatal Quality Indicators	Adverse Event
PDI #90	Pediatric Patient Safety for Selected Indicators <sup>1</sup>	Pediatric Quality Indicators	Composite
PSI #04	Death among Surgical Inpatients with Treatable Serious Complications <sup>1</sup>	Patient Safety Indicators	Mortality
PSI #06	Iatrogenic Pneumothorax <sup>1</sup>	Patient Safety Indicators	Adverse Event
PSI #11	Postoperative Respiratory Failure <sup>3</sup>	Patient Safety Indicators	Postoperative Complication
PSI #12	Postoperative Deep Vein Thrombosis (DVT) or Pulmonary Embolism (PE) <sup>2</sup>	Patient Safety Indicators	Postoperative Complication
PSI #14	Postoperative Wound Dehiscence <sup>1</sup>	Patient Safety Indicators	Adverse Event
PSI #15	Accidental Puncture or Laceration <sup>1</sup>	Patient Safety Indicators	Adverse Event
PSI #17	Birth Trauma <sup>2</sup>	Patient Safety Indicators	Adverse Event
PSI #18	OB Trauma-Vaginal Delivery with Instrument-assisted Delivery	Patient Safety Indicators	Adverse Event
PSI #19	OB Trauma-Vaginal Delivery without Instrument-assisted Delivery	Patient Safety Indicators	Adverse Event
PSI #90	Patient Safety for Selected Indicators <sup>1</sup>	Patient Safety Indicators	Composite

<sup>1</sup>NQF endorsed; <sup>2</sup>Time-limited NQF endorsed; <sup>3</sup>Currently under consideration for NQF endorsement

### Tier 1B: Serious Reportable Events Suitable for Public Accountability

Indicator	Label	Module	Type
PDI #03	Foreign Body left in During Procedure <sup>1</sup>	Pediatric Quality Indicators	Serious Reportable Event
PDI #13	Transfusion Reaction <sup>1</sup>	Pediatric Quality Indicators	Serious Reportable Event
PSI #02	Death in Low Mortality DRGs <sup>1</sup>	Patient Safety Indicators	Serious Reportable Event
PSI #05	Foreign Body left in During	Patient Safety Indicators	Serious Reportable Event

Indicator	Label	Module	Type
	Procedure <sup>1</sup>		
PSI #08	Postoperative Hip Fracture <sup>4</sup>	Patient Safety Indicators	Serious Reportable Event
PSI #16	Transfusion Reaction <sup>1</sup>	Patient Safety Indicators	Serious Reportable Event

Note: Serious reportable events are reported as counts, and should be used to promote public accountability and to support root-cause analysis of identified cases.

<sup>1</sup>NQF endorsed; <sup>2</sup>Time-limited NQF endorsed; <sup>3</sup>Currently under consideration for NQF endorsement; <sup>4</sup>Postoperative Hip Fracture is reported as a rate.

### 3.2 Tier 2: Moderate evidence gaps.

Indicators in this tier have some evidence supporting their use, but all have some evidence gaps. These gaps are not judged to be serious enough to preclude use in comparative reporting, but these indicators would be stronger if highlighted issues are better understood through future validation efforts. These indicators are not currently endorsed by the National Quality Forum.

#### Tier 2. AHRQ QI Potentially Suitable for Comparative Reporting with Some Additional Development

Indicator	Label	Module	Type
IQI #14	Hip Replacement Mortality	Inpatient Quality Indicators	Mortality
NQI #01	Iatrogenic Pneumothorax in Neonates	Neonatal Quality Indicators	Adverse Event
PDI #08	Postoperative Hemorrhage or Hematoma	Pediatric Quality Indicators	Postoperative Complication
PDI #09	Postoperative Respiratory Failure	Pediatric Quality Indicators	Postoperative Complication
PDI #10	Postoperative Sepsis	Pediatric Quality Indicators	Postoperative Complication
PSI #09	Postoperative Hemorrhage or Hematoma	Patient Safety Indicators	Postoperative Complication
PSI #10	Postoperative Physiologic and Metabolic Derangement	Patient Safety Indicators	Postoperative Complication
PSI #13	Postoperative Sepsis	Patient Safety Indicators	Postoperative Complication

### 3.3 Tier 3: Significant, but addressable, evidence gaps.

Indicators in this tier have at least one serious evidentiary gap or concern, which requires further development or validation work before the indicator is considered appropriate for comparative reporting. The required improvements and validation efforts are specified and underway, and once completed these indicators will be re-tiered. These indicators are not currently endorsed by the National Quality Forum.

#### Tier 3. AHRQ QI Potentially Suitable for Comparative Reporting with Significant Additional Development

Indicator	Label	Module	Type
IQI #12	CABG Mortality	Inpatient Quality Indicators	Mortality

IQI #13	Craniotomy Mortality <sup>5</sup>	Inpatient Quality Indicators	Mortality
IQI #18	GI Hemorrhage Mortality <sup>5</sup>	Inpatient Quality Indicators	Mortality
PDI #12	Central Line-Associated Bloodstream Infection (pediatrics) <sup>6</sup>	Pediatric Quality Indicators	Adverse Event
NQI #03	Neonatal Mortality <sup>7</sup>	Neonatal Quality Indicators	Mortality
PSI #03	Pressure Ulcer <sup>6</sup>	Patient Safety Indicators	Adverse Event
PSI #07	Central Line-Associated Bloodstream Infection <sup>6</sup>	Patient Safety Indicators	Adverse Event

<sup>5</sup> These indicators are undergoing re-specification as recent evidence has suggested very heterogeneous populations for which the current risk adjustment cannot account; <sup>6</sup> These indicators have recently or are planned for major re-specification resulting from the implementation of new coding guidelines and/or new available codes. The performance of these indicators cannot be assessed until these new codes are in use and data are available; <sup>7</sup> This indicator is undergoing empirical analysis of reliability.

### **3.4 Tier 4: Significant evidence gaps.**

The indicators in this tier have substantial gaps in the evidence that are unlikely to be addressed with further development or validation work. These indicators are not recommended for comparative reporting, but rather, can be used for internal quality improvement. These indicators are not currently endorsed by the National Quality Forum.

#### **Tier 4. AHRQ QI Not Suitable for Comparative Reporting as Currently Specified**

<b>Indicator</b>	<b>Label</b>	<b>Module</b>	<b>Type</b>
IQI #05	CABG Volume	Inpatient Quality Indicators	Volume
IQI #06	PTCA Volume	Inpatient Quality Indicators	Volume
IQI #07	Carotid Endarterectomy Volume	Inpatient Quality Indicators	Volume
IQI #21 and #33	Cesarean Delivery	Inpatient Quality Indicators	Utilization
IQI #22 and #34	VBAC Delivery	Inpatient Quality Indicators	Utilization
IQI #23	Laparoscopic Cholecystectomy	Inpatient Quality Indicators	Utilization
IQI #30	PTCA Mortality	Inpatient Quality Indicators	Mortality
IQI #31	Carotid Endarterectomy Mortality	Inpatient Quality Indicators	Mortality
IQI #90	Mortality for Selected Procedures	Inpatient Quality Indicators	Composite
PSI #01	Complications of Anesthesia	Patient Safety Indicators	Adverse Event
PSI #20	OB Trauma-Cesarean Delivery	Patient Safety Indicators	Adverse Event

## 4 Recommendations by AHRQ QI Module

### 4.1 Inpatient Quality Indicators (IQI)

The Inpatient Quality Indicators (IQI) consist of 28 indicators, including seven volume indicators, eight procedure-specific mortality indicators, six condition-specific mortality indicators, five utilization indicators and two composites.

#### 4.1.1 IQI #01 - Esophageal Resection Volume

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of esophageal resection in any procedure field.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence*

- Face validity: Rated as acceptable for internal quality improvement and comparative reporting by clinical panel.
- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Volume-outcome relationship supported in the literature, with a majority of studies demonstrating such a relationship.

#### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying mortality indicator. Although high volume hospitals, on average, have lower mortality, some high volume hospitals may have higher mortality (conversely some low volume hospitals may have lower mortality).

#### 4.1.2 IQI #02 - Pancreatic Resection Volume

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of pancreatic resection in any procedure field.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence*

- Face validity: Rated as acceptable for internal quality improvement and comparative reporting by clinical panel.
- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Volume-outcome relationship supported in the literature, with a majority of studies demonstrating the relationship.

#### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying mortality indicator.

#### 4.1.3 IQI #04 - Abdominal Aortic Aneurysm (AAA) Repair Volume

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of AAA repair in any procedure field and an ICD-9-CM code of AAA in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Rated as acceptable for internal quality improvement and comparative reporting by clinical panel.
- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Volume-outcome relationship supported in the literature, with a majority of studies demonstrating this relationship. Surgeon volume accounts for some, but not all, of the hospital-volume outcome relationship.

##### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying mortality indicator.
- Data analysis suggests that stratifying by surgical approach (endoscopic vs. open) may further differentiate hospital performance. The clinical panel argues for stratification by aneurysm type (rupture vs. intact) in order to adequately account for patient risk of mortality (with volume as a proxy for mortality).

#### 4.1.4 IQI #05 - CABG Volume

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of CABG in any procedure field.

*Tier:* Not suitable for comparative reporting as currently specified (Tier 4).

##### *Supporting Evidence*

- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Volume-outcome relationship supported in the literature, although surgeon volumes may play the most important role. More recent data has shown a weakening in the relationship between hospital volume and outcome.

##### *Considerations of Use*

- This indicator should generally be used for internal quality improvement along with the accompanying mortality indicator.

##### *Future Development*

- This indicator is closely related to an existing NQF endorsed measure for CABG volume. The NQF measure has separate indicators for isolated CABG and CABG with valve surgery and is based on clinical registry data. Future development might include linking to individual surgeons and incorporating clinical registry data.

#### 4.1.5 IQI #06 - PTCA Volume

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of PTCA in any procedure field.

*Tier:* Not suitable for comparative reporting as currently specified (Tier 4).

##### *Supporting Evidence*

- Face validity: Guidelines from the American Heart Association and American College of Cardiology dictate that minimum volume standards for PTCA are essential (though the actual standards vary).
- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Volume-outcome relationship supported in the literature, with a majority of studies demonstrating this relationship. More recent data has shown a larger share of procedures performed on an outpatient basis.

##### *Considerations of Use*

- This indicator should generally be used for internal quality improvement along with the accompanying mortality indicator.

##### *Future Development*

- This indicator is closely related to an existing NQF endorsed measure for PTCA volume. The NQF measure is based on data from the National Cardiovascular Data Registry. Future development might include linking to clinical registry data and ambulatory surgery data.

#### 4.1.6 IQI #07 - Carotid Endarterectomy Volume

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of carotid endarterectomy in any procedure field.

*Tier:* Not suitable for comparative reporting as currently specified (Tier 4).

##### *Supporting Evidence*

- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Volume-outcome relationship supported in the literature, although recent literature has identified this relationship as modest and surgeon volume identified as most important factor.

##### *Considerations of Use*

- Low reliability due to infrequency makes the indicator most useful for quality improvement.
- This indicator should generally be used for internal quality improvement along with the accompanying mortality indicator.

##### *Future Development*

- Future development might include linking to individual surgeons and/or incorporating into a composite

#### 4.1.7 IQI #08 - Esophageal Resection Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of esophageal resection in any procedure field and an ICD-9-CM code of esophageal cancer in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Rated as acceptable for internal quality improvement and comparative reporting by clinical panel.
- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Volume-outcome relationship established in the literature.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

##### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying volume indicator.
- Combining several years of data may improve reliability of the indicator, since these procedures are relatively uncommon.

#### 4.1.8 IQI #09 - Pancreatic Resection Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of pancreatic resection in any procedure field and an ICD-9-CM code of pancreatic cancer in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Rated as acceptable for internal quality improvement and comparative reporting by clinical panel.
- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Volume-outcome relationship established in the literature.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

##### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying volume indicator.
- Combining several years of data may improve reliability of indicator, since these procedures are relatively uncommon.

#### 4.1.9 IQI #11 - Abdominal Aortic Aneurysm (AAA) Repair Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of AAA repair in any procedure field and an ICD-9-CM code of AAA in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Rated as acceptable for internal quality improvement but as unclear for comparative reporting by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Volume outcome relationship supported in the literature. Process factors, such as control of blood loss and perioperative beta adrenergic blockade, have been associated with better patient level outcomes. Structure factors such as volume and nursing intensity have been associated with decreased complications which in turn has been associated with decreased mortality.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass. Some users have expressed concern over the inclusion of both ruptured and un-ruptured aortic aneurysm cases in the denominator. These cases are assigned to different APR-DRG categories and the difference in underlying risk of death is address through risk-adjustment. However, users have also suggested that admission delays for ruptured cases vary systematically across hospitals, and that difference, if significant, would not be addressed through risk-adjustment. More published studies are needed in order to assess that suggestion.

##### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying mortality indicator.
- Data analysis suggests that stratifying by surgical approach (endoscopic vs. open) may further differentiate hospital performance. The clinical panel argues for stratification by aneurysm type (rupture vs. intact) in order to adequately account for patient risk of mortality (with volume as a proxy for mortality).

#### 4.1.10 IQI #12 - CABG Mortality

*Description:* Rate of in-hospital death among discharges age 40 years and older with an ICD-9-CM code of CABG in any procedure field.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

##### *Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Volume-outcome relationship supported in the literature, although surgeon volume may play the most important role. Process factors, such as cross-clamp or perfusion time,

specific surgical techniques and approaches, and postoperative beta blocker use, have been associated with reduced patient level mortality.

- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

#### *Considerations of Use*

- This indicator should generally be publically reported along with the accompanying volume indicator.

#### *Future Development*

- This indicator is closely related to an existing NQF endorsed measure for CABG mortality. The NQF measure is based on data from the New York State Cardiac Surgery Reporting System. Future development might include incorporating clinical registry data.

#### 4.1.11 IQI #13 - Craniotomy Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with a DRG code of craniotomy.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

#### *Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Volume-outcome relationship supported in the literature.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass. However, risk adjustment does not take into account the heterogeneity of procedures in the indicator.

#### *Considerations of Use*

- The current indicator combines heterogonous procedures, empirically confirmed to be at vastly different risk and as such requires extensive redefinition for risk adjustment or stratification.

#### 4.1.12 IQI #14 - Hip Replacement Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of hip replacement in any procedure field.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2)

#### *Supporting Evidence*

- Face validity: Overall usefulness rated as unclear for both quality improvement and comparative reporting purposes by clinical panel, primarily due to concerns about reliability given the rarity of this outcome.
- Criterion validity: Coding of death is highly accurate with few false positives

- Construct validity: In-hospital mortality is highly correlated with 30-day mortality
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass. In particular, risk adjustment accounts for differential risk associated with revision arthroplasty.

*Considerations of Use*

- Combining several years of data may improve reliability of the indicator, since deaths following this procedure are relatively uncommon.

4.1.13 IQI #15 - AMI Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of acute myocardial infarction (AMI) in the principal diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Process factors, such as aspirin and thromboembolic administration, early revascularization, and appropriate ACE inhibitor, have been associated with lower patient level mortality. Better processes of care have also been associated with lower hospital level 30-day mortality rates although some studies have failed to demonstrate such an association. Improving processes of care has been shown to improve 30-day mortality. Hospitals ranked as higher quality (Joint Commission and HealthGrades) have been shown to have lower 30-day mortality.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

*Considerations of Use*

- Hospitals that transfer-out a higher percentage of patients generally have lower in-hospital mortality rates, but similar 30-day mortality rates

*Future Development*

- This indicator is closely related to an existing NQF endorsed measure for AMI mortality. Future development might harmonize with the endorsed measure specifications.

4.1.14 IQI #16 - CHF Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of congestive heart failure (CHF) in the principal diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.

- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Processes of care have been shown to decrease mortality at a patient level, including appropriate use of ACE inhibitors. Mortality rates have been shown to decrease following public reporting of these data. Hospital type (e.g. teaching) has been associated, but not consistently, with lower rates.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

#### *Considerations of Use*

- Hospitals that transfer-out a higher percentage of patients generally have similar in-hospital mortality rates, and similar 30-day mortality rates
- There is a related Centers for Medicare and Medicare Services (CMS) 30-day mortality measure. The AHRQ measure is considered complementary because it is based on all-payer data and it may be used in “real-time”.

#### 4.1.15 IQI #17 - Acute Stroke Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of stroke in the principal diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence*

- Face validity: Overall usefulness rated as unclear for both quality improvement and comparative reporting purposes by clinical panel, primarily due to concerns about combining different types of stroke. NQF technical advisory panel rated as acceptable for comparative reporting.
- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: Hospital characteristics such as stroke volume, teach hospital status, staffing of vascular neurologists, and the presence of a dedicated stroke program have been associated with lower mortality rates. Randomized trials have demonstrated several effective therapies, which have been shown to reduce mortality on a patient level.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass. Some users have expressed concern that the denominator includes both ischemic cases and hemorrhagic and subarachnoid cases. These cases are assigned to different APR-DRG categories, and so the difference in mortality risk is accounted for in the risk-adjustment.

#### *Considerations of Use*

- A large portion of stroke deaths occur outside the acute inpatient setting.
- Transfer patterns are likely to influence the mortality rate, even after risk-adjustment; reporting the rate of transfers along with the indicator may be advisable.
- Some users may prefer to stratify by stroke type, as was advocated by a clinical panel.

#### 4.1.16 IQI #18 - GI Hemorrhage Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of gastrointestinal hemorrhage in the principal diagnosis field.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

##### *Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: New studies suggest mortality is driven by esophageal varices (veins in the lower esophagus.) which are emergent events, requiring re-specification of indicator.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

##### *Considerations of Use*

- None

#### 4.1.17 IQI #19 - Hip Fracture Mortality

*Description:* Rate of in-hospital death among discharges age 65 years and older with an ICD-9-CM code of hip fracture in the principal diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: Time to surgery, thromboembolic prophylaxis, prophylaxis against AMI, and surgical technique have all been shown to reduce mortality on a patient level.
- Risk adjustment. Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass. Some users have expressed concern about fractures that occur days prior to admission and variability across hospitals in admission delays. The APR-DRG risk-of-mortality subclass will capture co-morbidities resulting from admission delay. Risk adjustment performs well in accounting for risk due to repair type (fixation vs. replacement) and fracture location.

##### *Considerations of Use*

- The indicator was recently revised to include only discharges age 65 years and older in order to exclude hip fractures related to trauma.

#### 4.1.18 IQI #20 - Pneumonia Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of pneumonia in the principal diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Criterion validity: Coding of death is highly accurate with few false positives.

- Construct validity: In-hospital mortality is highly correlated with 30-day mortality. Processes of care have been shown to decrease mortality on a patient level, including appropriate choice and administration of antibiotics. Public reporting of hospital performance has been associated with a decrease in mortality rates.
- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

*Considerations of Use*

- There is a related CMS 30-day mortality measure. The AHRQ measure is considered complementary because it is based on all-payer data and it may be used in “real-time”.

4.1.19 IQI #21 and #33 - Cesarean Delivery

*Description:* Rate of Cesarean deliveries, defined by DRG code or an ICD-9-CM code of hysterectomy in any procedure field, among discharges with a DRG code for delivery.

*Tier:* Not Suitable for Comparative Reporting as Currently Specified (Tier 4).

*Supporting Evidence*

- The US Department of Health and Human Services and other organizations have encouraged the reduction in Cesarean section rate. Studies from the 1990s have shown that practice patterns, rather than patient characteristics, account for a significant proportion of Cesarean delivery rates.

*Considerations of Use*

- The Cesarean delivery rate provides useful information on hospital practice patterns, but high or low rates are not in themselves indicators of good or poor quality. Rather, the rates should inform decision-making by indicating to consumers how hospitals have responded to the American College of Obstetricians and Gynecologists (ACOG) recommendations by aligning their practice patterns to ensure appropriate care for obstetric patients.

*Future Development*

- This indicator is closely related to an existing NQF endorsed measure for Cesarean section delivery. Future development might harmonize with the endorsed measure specifications.

4.1.20 IQI #22 and #34 - VBAC Delivery

*Description:* Rate of vaginal deliveries, defined by DRG code, among discharges with a DRG code for delivery and an ICD-9-CM code of previous cesarean delivery in any diagnosis field.

*Tier:* Not Suitable for Comparative Reporting as Currently Specified (Tier 4).

*Supporting Evidence*

- Face validity: VBAC is a safe alternative for some women, and of interest to some women choosing to have VBAC. Some process measures have been associated with higher VBAC success. However, VBAC may also lead to serious complications and is

widely viewed as physician and patient choice, rather than higher quality of care. The correct rate of VBAC is unknown.

#### *Considerations of Use*

- The VBAC delivery rate provides useful information on hospital practice patterns, but high or low rates are not in themselves indicators of good or poor quality. Rather, the rates should inform decision-making by indicating to consumers how hospitals have responded to the American College of Obstetricians and Gynecologists (ACOG) recommendations by aligning their practice patterns to ensure appropriate care for obstetric patients.

#### *Future Development*

- This indicator was closely related to an existing NQF endorsed measure for VBAC delivery (but which is no longer endorsed). Future development might harmonize with other measure specifications, should the measure be reconsidered.

#### 4.1.21 IQI #23 - Laparoscopic Cholecystectomy

*Description:* Rate of laparoscopic cholecystectomy, defined by ICD-9-CM code, among discharges age 18 years and older with an ICD-9-CM code of cholecystectomy in any procedure field and an ICD-9-CM code of uncomplicated cholecystitis or cholelithiasis in any diagnosis field.

*Tier:* Not Suitable for Comparative Reporting as Currently Specified (Tier 4).

#### *Supporting Evidence*

- Although previous evidence supported the use of the indicator, it is superseded by evidence showing that this indicator is unreliably measured using inpatient data due to high number of procedures now done in outpatient settings.

#### *Considerations of Use*

- Many procedures are now done in an outpatient setting.

#### *Future Development*

- Future development might incorporate ambulatory surgery data.

#### 4.1.22 IQI #24 - Incidental Appendectomy in the elderly

*Description:* Rate of incidental appendectomies, defined by ICD-9-CM code, among discharges age 65 year and older with a ICD-9-CM code for intra-abdominal procedure.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence*

- Face validity: Overall usefulness rated as acceptable for both quality improvement but unclear comparative reporting purposes by clinical panel, primarily due to concerns about coding variation. Review by an NQF technical advisory panel rated acceptable for comparative public reporting. Rates vary substantially across hospitals.

- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Strong consensus that this procedure is inappropriate, particularly for the elderly.

#### *Considerations of Use*

- This is an indicator of overuse, meaning that some patients without indications are receiving the procedure. If a hospital resides in an area where overuse is common, the reference to external benchmarks may be most appropriate.

#### 4.1.23 IQI #25 - Bi-lateral Catheterization

*Description:* Rate of simultaneous right and left heart catheterizations, defined by ICD-9-CM code, among discharges age 18 years and older with an ICD-9-CM procedure code of catheterization in any procedure field and an ICD-9-CM code of coronary artery disease in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence*

- Face validity: Overall usefulness rated as unclear for both quality improvement and comparative reporting purposes by clinical panel, primarily due to concerns that the indicator is primarily one of resource use. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Coding of procedures is highly accurate with few false positives.
- Construct validity: Studies of right sided catheterization have demonstrated higher cost when the procedure is used, although the American College of Cardiology and the American Heart Association has issued guidelines stating that routine right heart catheterizations are unnecessary. One study shown that bilateral catheterization persisted among cardiologists in two large community hospitals, although a 2001 study should a decrease in the procedure.

#### *Considerations of Use*

- This is an indicator of overuse, meaning that some patients without indications are receiving the procedure. If a hospital resides in an area where overuse is common, the reference to external benchmarks may be most appropriate.

#### 4.1.24 IQI #30 - PTCA Mortality

*Description:* Rate of in-hospital death among discharges age 40 years and older with an ICD-9-CM code of PTCA in any procedure field.

*Tier:* Not suitable for comparative reporting as currently specified (Tier 4).

#### *Supporting Evidence*

- Construct validity: The literature supports a volume-outcome relationship, but the focus is on volume, rather than mortality, which is relatively rare. In addition many procedures are now done on an outpatient basis

- Risk adjustment: Hierarchical model based on age, gender, an age-gender interaction and the APR-DRG with risk-of-mortality subclass.

#### *Considerations of Use*

- This indicator is not recommended for use as a stand-alone indicator, but rather in conjunction with indicators such as PTCA volume.

#### *Future Development*

- This indicator is closely related to an existing NQF endorsed measure for PTCA mortality (HC10). The NQF measure is based on data from the National Cardiovascular Data Registry. Future development might incorporate clinical registry data and ambulatory surgery data.

#### 4.1.25 IQI #31 - Carotid Endarterectomy Mortality

*Description:* Rate of in-hospital death among discharges age 18 years and older with an ICD-9-CM code of carotid endarterectomy in any procedure field.

*Tier:* Not Suitable for Comparative Reporting as Currently Specified (Tier 4).

#### *Supporting Evidence*

- Construct validity: The literature supports a volume-outcome relationship, but the focus is on volume, rather than mortality, which is relatively rare.

#### *Considerations of Use*

- This indicator is not recommended for use as a stand-alone indicator, but rather in conjunction with indicators such as carotid endarterectomy volume.
- Low reliability due to infrequency makes the indicator most useful for quality improvement.

## **4.2 Patient Safety Indicators (PSI)**

The Patient Safety Indicators (PSI) consist of 21 indicators, including four serious reportable event indicators, five adverse event indicators, one mortality indicator, six postoperative complication indicators, four obstetric complication indicators and one composite indicator.

### **4.2.1 PSI #01 - Complications of Anesthesia**

*Description:* Rate of discharges with an ICD-9-CM code of anesthesia complications in any secondary diagnosis field (not present on admission) among surgical discharges, defined by DRG and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 years and older.

*Tier:* Not Suitable for Comparative Reporting as Currently Specified (Tier 4).

#### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: One study showed that no cases were present on admission.
- Construct validity: None.

#### *Considerations of Use*

- The numerator is reliant on the use of E-codes, which are not collected in all states.
- National coding guidelines for E-codes do not require that a condition be an unexpected aspect of a procedure or disease in order to be coded. Therefore this indicator captures numerous minor complications and may be subject to wide variability in coding.

### **4.2.2 PSI #02 - Death in Low Mortality DRGs**

*Description:* Rate of in-hospital death among discharges with a DRG code for a low mortality DRG (defined as a mortality rate of less than 0.5%).

*Tier:* Suitable for comparative reporting (Tier 1B).

#### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting.
- Criterion validity: Coding of death is highly accurate, with relatively few false positives
- Construct validity: Validity is based on only one study from the 1980s.
- Risk adjustment: Hierarchical model based on age, gender, modified DRG and co-morbidity. The indicator is a serious reportable event, which is generally reported as a count. However, this indicator is reported as a risk-adjusted rate because the mortality rate varies by demographics and among the low mortality DRGs. In addition, reporting as a rate scales the count by the number of discharges or size of the hospital.

### Considerations of Use

- The serious reportable event indicators are not intended for use in comparing one hospital with another, but rather to promote public accountability and to support root-cause analysis of identified cases.

#### 4.2.3 PSI #03 - Pressure Ulcer

*Description:* Rate of discharges with an ICD-9-CM code of Pressure ulcer in any secondary diagnosis field (not present on admission) among medical and surgical discharges, defined by specific DRG codes, age 18 years and older.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

#### *Supporting Evidence:*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: Coding is moderately inaccurate, with many false positives (present on admission).
- Construct validity: Decubiti are largely preventable. Staffing levels have been associated, but not consistently, with decreased decubiti in studies. Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on age, gender, modified DRG and co-morbidities defined using the AHRQ co-morbidity software.

#### *Considerations of Use*

- Implementation of present on admission coding will improve this measure.
- In FY2009 there were new ICD-9-CM codes added to define pressure ulcers by stage. Excluding stage I pressure ulcers should improve the sensitivity of the indicator because stage I pressure ulcers are less likely to be coded.
- Complements an existing NQF endorsed measure that is not defined using administrative data.

#### 4.2.4 PSI #04 - Death among Surgical Inpatients with Treatable Serious Complications

*Description:* Rate of in-hospital death among surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 years and older and with a principal procedure within 2 days of admission OR admission type of elective and with an ICD-9-CM code of potential complications of care (e.g., pneumonia, DVT/PE, sepsis, shock/cardiac arrest, or GI hemorrhage/acute ulcer) in any secondary diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for

comparative reporting. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.

- Criterion validity: Several studies have noted that the complications that define the denominator may be present on admission. However, the conceptual basis of the indicator applies whether the complication occurred during or prior to admission.
- Construct validity: Staffing levels and teaching hospital status have been associated with decreased mortality rates in patients with complications.
- Risk adjustment: Hierarchical model based on age, gender, modified DRG and co-morbidities defined using the AHRQ co-morbidity software.

#### *Considerations of Use*

- Formerly known as Failure to Rescue.
- Reconciled with the existing NQF standard from the nurse-sensitive measure set.

#### 4.2.5 PSI #05 - Foreign Body left in During Procedure

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of foreign body left in during procedure in any secondary diagnosis field (not present on admission).

*Tier:* Suitable for comparative reporting (Tier 1B).

#### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Most cases flagged are not present on admission.

#### *Considerations of Use*

- The indicator is reported as a count
- The serious reportable event indicators are not intended for use in comparing one hospital with another, but rather to promote public accountability and to support root-cause analysis of identified cases.

#### 4.2.6 PSI #06 - Iatrogenic Pneumothorax

*Description:* Rate of discharges with an ICD-9-CM code of iatrogenic pneumothorax in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, age 18 years and older.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence:*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.

- Criterion validity: Coding is substantially accurate, with only some false positives. Most cases flagged are not present on admission. Almost 90% of cases are confirmed cases, and almost 70% of those were significant requiring tube thoracostomy.
- Construct validity: Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

*Considerations of Use*

- The severity of a pneumothorax varies from patient to patient.

4.2.7 PSI #07 - Central Line-Associated Bloodstream Infection

*Description:* Rate of discharges with an ICD-9-CM code of vascular catheter related infection in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, age 18 years and older.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

*Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: Most cases flagged are not present on admission. Over 60% of those cases flagged, given the pre-2008 definition, were demonstrated to be true cases, and over three-quarters are associated with catheters.
- Construct validity: Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

*Considerations of Use*

- Indicator was formerly “Selected Infection Due to Medical Care” but was recently limited to central line-associated infection and the title changed.

4.2.8 PSI #08 - Postoperative Hip Fracture

*Description:* Rate of discharges with an ICD-9-CM code of hip fracture in any secondary diagnosis field (not present on admission) among surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 and older.

*Tier:* Serious Reportable Events Suitable for Public Accountability (Tier 1B).

*Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for

comparative reporting. An independent panel (OECD) also endorsed a broader but related indicator as important and scientifically sound.

- Criterion validity: Many flagged cases are present on admission.
- Construct validity: Studies have associated specific processes of care with lower rates of in-hospital hip fracture. Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

#### *Considerations of Use*

- Use of present on admission flag will materially improve the criterion validity of this indicator.

#### 4.2.9 PSI #09 - Postoperative Hemorrhage or Hematoma

*Description:* Rate of discharges with an ICD-9-CM code of postoperative hemorrhage or postoperative hematoma in any secondary diagnosis field (not present on admission) and an ICD-9-CM code of postoperative control of hemorrhage or drainage of hematoma in any procedure field among surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 and older.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

#### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting.
- Criterion validity: Most cases flagged are not present on admission.
- Construct validity: Studies have associated specific processes of care with lower rates of hemorrhage and hematoma. Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

#### *Considerations of Use*

- None

#### 4.2.10 PSI #10 - Postoperative Physiologic and Metabolic Derangement

*Description:* Rate of discharges with an ICD-9-CM code of physiologic and metabolic derangements in any secondary diagnosis field (not present on admission) among elective surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 and older.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

#### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting.
- Criterion validity: Codes capture about half of the cases of interest with relatively few false positives. Most cases flagged are not present on admission.
- Construct validity: Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

*Considerations of Use*

- This indicator is sensitive to the number of diagnosis and procedure codes reported.

4.2.11 PSI #11 - Postoperative Respiratory Failure

*Description:* Rate of discharges with an ICD-9-CM code of acute respiratory failure in any secondary diagnosis field (not present on admission) or an ICD-9-CM code for re-intubation procedure in any procedure field that is zero, one or two or more days after the major operating room procedure (depending on the code) among elective surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 and older.

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting.
- Criterion validity: Most cases flagged are not present on admission. Codes capture over half of the cases of interest with relatively few false positives.
- Construct validity: Flagged cases have been associated with higher resource use and mortality. Poorer performance on this indicator has been associated with less favorable Joint Commission accreditation evaluations.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

*Considerations of Use*

- None.

4.2.12 PSI #12 - Postoperative DVT or PE

*Description:* Rate of discharges with an ICD-9-CM code of deep vein thrombosis or pulmonary embolism in any secondary diagnosis field (not present on admission) among surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 and older.

*Tier:* Suitable for comparative reporting (Tier 1).

### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: Many cases flagged are present on admission. Codes capture almost half cases of interest with some false positives.
- Construct validity: Process of care failures identified in over 2/3 of cases identified by early study of postoperative DVTs identified using administrative data. Nurse staffing has been associated, but not consistently with postoperative PE/DVT. Cases flagged by indicator associated with higher resource use and mortality. Poorer performance on this indicator has been associated with less favorable Joint Commission accreditation evaluations.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

### *Considerations of Use*

- Some chronic DVTs will be captured by this indicator. Most risk attributable to the procedure, which is accounted for using DRGs. Some providers screen for PVT and PE postoperatively, and may identify insignificant or old clots. Routine screening is still controversial.
- Use of the present on admission flag will improve the criterion validity of this indicator.
- In FY2010 new ICD-9-CM codes were introduced that will improve the indicator by excluding upper extremity DVTs.

#### 4.2.13 PSI #13 - Postoperative Sepsis

*Description:* Rate of discharges with an ICD-9-CM code of sepsis in any secondary diagnosis field (not present on admission) among elective surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, age 18 and older.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: Most cases flagged are not present on admission. Codes capture about 1/3 cases of interest with some false positives. Over 60% of captured cases are true cases.
- Construct validity: Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

### *Considerations of Use*

- None.

#### 4.2.14 PSI #14 - Postoperative Wound Dehiscence

*Description:* Rate of discharges with an ICD-9-CM code of reclosure of postoperative disruption of abdominal wall in any procedure field among abdominopelvic surgical discharges, defined by specific ICD-9-CM codes, age 18 and older.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Most cases flagged are not present on admission. Codes capture about 1/3 cases of interest with relatively few false positives (the other 2/3 cases might be captured with a diagnosis code, see below).
- Construct validity: Flagged cases have been associated with higher resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software.

##### *Considerations of Use*

- In FY2010 there were new ICD-9-CM codes introduced that will improve the indicator by allowing cases to be identified by diagnosis code, rather than only by procedure code.

#### 4.2.15 PSI #15 - Accidental Puncture or Laceration

*Description:* Rate of discharges with an ICD-9-CM code of accidental cut, puncture, perforation or laceration during a procedure in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, age 18 years and older.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence:*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Codes are highly accurate, with relatively few false positives, particularly in surgical cases. Most cases flagged are not present on admission. Over 90% of cases flagged are true cases, with almost 70% of those requiring a reparative procedure.
- Construct validity: Poor performance on this indicator has been correlated with poorer Joint Commission overall accreditation evaluation scores. Flagged cases have been associated with higher resource use and mortality, although less so than other PSI.

- Risk adjustment: Hierarchical model based on covariates for age, gender, an age-gender interaction and co-morbidities defined using the AHRQ co-morbidity software. .

*Considerations of Use*

- Some users have commented on the adequacy of the risk-adjustment for patients undergoing spine surgery. In anticipation of developing a more refined risk-adjustment method, these patients have been temporarily removed from the denominator.

4.2.16 PSI #16 - Transfusion Reaction

*Description:* Count of discharges age 18 years and older with an ICD-9-CM code of transfusion reaction in any secondary diagnosis field (not present on admission).

*Tier:* Suitable for comparative reporting (Tier 1B).

*Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Most cases flagged are not present on admission.

*Considerations of Use*

- This indicator is reported as a count
- Serious reportable event most useful for flagging cases for root cause analysis.

4.2.17 PSI #17 - Birth Trauma

*Description:* Rate of discharges with an ICD-9-CM code of birth trauma in any diagnosis field among newborns.

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence*

- Face validity: A clinical panel, comprised entirely of obstetric clinicians, rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. A second panel, comprised of both obstetric and neonatal clinicians, rated the indicator as “unclear” in terms of overall usefulness. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Construct validity: Some trauma associated with forceps use.

*Considerations of Use*

- The extent of trauma may be minor for some events.
- The indicator was recently harmonized with similar measure from National Perinatal Information Center (NPIC).

#### 4.2.18 PSI #18 - OB Trauma-Vaginal Delivery with Instrument

*Description:* Rate of discharges with an ICD-9-CM code of 3rd and 4th degree obstetric laceration in any diagnosis field among vaginal delivery discharges, defined by DRG, with an ICD-9-CM code of instrument-assisted delivery in any procedure field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: Codes capture almost all cases of interest with relatively few false positives.

##### *Considerations of Use*

- None.

#### 4.2.19 PSI #19 - OB Trauma-Vaginal Delivery without Instrument

*Description:* Rate of discharges with an ICD-9-CM code of 3rd and 4th degree obstetric trauma in any diagnosis field among vaginal delivery discharges, defined by DRG, without an ICD-9-CM code of instrument-assisted delivery in any procedure field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting. An independent panel (OECD) also endorsed this indicator as important and scientifically sound.
- Criterion validity: Codes capture almost all cases of interest with relatively few false positives.

##### *Considerations of Use*

- None.

#### 4.2.20 PSI #20 - OB Trauma-Cesarean Delivery

*Description:* Rate of discharges with an ICD-9-CM code of 3rd and 4th degree obstetric trauma in any diagnosis field among cesarean delivery discharges, defined by DRG.

*Tier:* Not suitable for comparative reporting as currently specified (Tier 4).

##### *Supporting Evidence*

- Face validity: Clinical panel rated the indicator as acceptable for quality improvement purposes. The panel was not asked to rate the PSI specifically on usefulness for comparative reporting.

### *Considerations of Use*

- Originally the OB trauma indicator included a broader set of complications and was stratified by delivery type. Because of concerns about coding accuracy, the set of complications was limited to 3<sup>rd</sup> and 4<sup>th</sup> degree lacerations, which are less relevant to Cesarean deliveries.

### **4.3 Pediatric Quality Indicators (PDI)**

The Pediatric Quality Indicators (PDI) consist of 14 indicators, including two serious reportable event indicators, five adverse event indicators, two mortality indicators, one volume indicator and four postoperative complication indicators.

#### 4.3.1 PDI #01 - Accidental Puncture or Laceration

*Description:* Rate of discharges with an ICD-9-CM code of accidental cut, puncture, perforation or laceration during a procedure in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, under age 18 years.

*Tier:* Suitable for comparative reporting (Tier 1).

#### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement purposes rated as acceptable by clinical panel, but panel did not recommend for comparative purposes. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Codes are highly accurate, with relatively few false positives, particularly in surgical cases. Most cases flagged are not present on admission. Over 90% of cases flagged are true cases, with almost 70% of those requiring a reparative procedure.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, co-morbidities defined using the AHRQ clinical classification software and number of and procedure type (therapeutic/diagnostic).

### *Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than community hospitals, since data show consistently higher rates amount children's hospitals even after the risk-adjustment (which may suggest the risk-adjustment is not capturing all the important risk factors).
- AHRQ QI software offers an optional stratification by clinical category of procedures (by anatomical location).

#### 4.3.2 PDI #02 - Pressure Ulcer

*Description:* Rate of discharges with an ICD-9-CM code of Pressure ulcer in any secondary diagnosis field (not present on admission) among medical and surgical discharges, defined by specific DRG codes, under age 18 years.

#### Guidance on Using the AHRQ QI

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Coding is moderately inaccurate, with many false positives (present on admission).
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender and co-morbidities defined using the AHRQ clinical classification software.

*Considerations of Use*

- The AHRQ QI software stratifies the indicator into low and high risk populations.
- Implementation of present on admission coding required for NQF endorsement.
- There is a pending coding proposal before the ICD-9-CM Coordination and Maintenance Committee to define Pressure ulcers by stage.
- Complements an existing NQF endorsed measure not based on administrative data.

#### 4.3.3 PDI #03 - Foreign Body left in During Procedure

*Description:* Count of discharges under age 18 years with an ICD-9-CM code of foreign body left in during procedure in any secondary diagnosis field (not present on admission).

*Tier:* Suitable for comparative reporting (Tier 1B).

*Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes was rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Most cases flagged are not present on admission.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use but not mortality.

*Considerations of Use*

- The serious reportable event indicators are not intended for use in comparing one hospital with another, but rather to promote public accountability and to support root-cause analysis of identified cases.

#### 4.3.4 PDI #05 - Iatrogenic Pneumothorax

*Description:* Rate of discharges with an ICD-9-CM code of iatrogenic pneumothorax in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, under age 18 (excluding neonates less than 2500g).

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Coding is substantially accurate, with only some false positives. Most cases flagged are not present on admission. Almost 90% of cases are confirmed cases, and almost 70% of those were significant requiring tube thoracostomy.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, co-morbidities defined using the AHRQ clinical classification software.

*Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.

#### 4.3.5 PDI #06 - Pediatric Heart Surgery Mortality

*Description:* Rate of in-hospital death among discharges under age 18 with an ICD-9-CM code of congenital heart disease repair in any procedure field or an ICD-9-CM code of non-specific heart surgery in any procedure field and an ICD-9-CM code of congenital heart disease in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Coding of death is highly accurate with few false positives.
- Construct validity: Volume-outcome (mortality) relationship supported in the literature. Processes of care, such as bypass or cross-clamp time, have been associated with patient level mortality.
- Risk adjustment: Hierarchical model based on covariates for age, gender, and procedure complexity, based on the Risk Adjustment for Congenital Heart Surgery (RACHS) system.

*Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.

- AHRQ QI software offers an optional stratification by procedure complexity (RACHS complexity level).

#### 4.3.6 PDI #07 - Pediatric Heart Surgery Volume

*Description:* Count of discharges under age 18 with an ICD-9-CM code of congenital heart disease repair in any procedure field or an ICD-9-CM code of non-specific heart surgery in any procedure field and an ICD-9-CM code of congenital heart disease in any diagnosis field.

*Tier:* Suitable for comparative reporting (Tier 1).

##### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Coding of procedures is highly accurate.
- Construct validity: A majority of studies show that hospitals with higher volumes tend to have lower mortality.
- Risk adjustment: Volume measures do not require risk adjustment.

##### *Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.
- AHRQ QI software offers an optional stratification by procedure complexity (RACHS complexity level).

#### 4.3.7 PDI #08 - Postoperative Hemorrhage or Hematoma

*Description:* Rate of discharges with an ICD-9-CM code of postoperative hemorrhage or postoperative hematoma in any secondary diagnosis field (not present on admission) and an ICD-9-CM code of postoperative control of hemorrhage or drainage of hematoma in any procedure field among elective surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, under age 18.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

##### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Most cases flagged are not present on admission.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, co-morbidities defined using the AHRQ clinical classification software.

### *Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.
- AHRQ QI software offers stratification by high and lower risk patients.

#### 4.3.8 PDI #09 - Postoperative Respiratory Failure

*Description:* Rate of discharges with an ICD-9-CM code of acute respiratory failure in any secondary diagnosis field (not present on admission) or an ICD-9-CM code for re-intubation procedure in any procedure field that is zero, one or two or more days after the major operating room procedure (depending on the code) among elective surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, under age 18.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel.
- Criterion validity: Recent review by the National Association of Children's Hospitals and Associated Institutions (NACHRI) found a high false positive rate for this indicator, suggesting need for further specification of exclusion criteria.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, and co-morbidities defined using the AHRQ clinical classification.

### *Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.

#### 4.3.9 PDI #10 - Postoperative Sepsis

*Description:* Rate of discharges with an ICD-9-CM code of sepsis in any secondary diagnosis field (not present on admission) among surgical discharges, defined by specific DRG codes and an ICD-9-CM code of major operating room procedure in any procedure field, under age 18.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement purposes rated as acceptable by clinical panel, but panel did not recommend for comparative purposes.

- Criterion validity: Most cases flagged are not present on admission. Codes capture about 1/3 cases of interest with some false positives. Over 60% of captured cases are true cases.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, and co-morbidities defined using the AHRQ clinical classification.

*Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.

4.3.10 PDI #11 - Postoperative Wound Dehiscence

*Description:* Rate of discharges with an ICD-9-CM code of reclosure of postoperative disruption of abdominal wall in any procedure field among abdominopelvic surgical discharges, defined by specific ICD-9-CM codes, under age 18.

*Tier:* Suitable for comparative reporting (Tier 1).

*Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: Most cases flagged are not present on admission. Codes capture about 1/3 cases of interest with relatively few false positives (the other 2/3 cases might be captured with a diagnosis code, see below).
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, and co-morbidities defined using the AHRQ clinical classification software.

*Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.
- AHRQ QI software offers stratification based on estimation of surgical wound class.
- In FY2010 there were new ICD-9-CM codes introduced that will improve the indicator by allowing cases to be identified by diagnosis code, rather than only by procedure code.

#### 4.3.11 PDI #12 - Central Line-Associated Bloodstream Infection

*Description:* Rate of discharges with an ICD-9-CM code of vascular catheter related infection in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, under age 18 years.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

##### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement purposes rated as acceptable by clinical panel, but panel did not recommend for comparative purposes.
- Criterion validity: Most cases flagged are not present on admission. Over 60% of those cases flagged, given the pre-2008 definition, were demonstrated to be true cases, and over three-quarters are associated with catheters.
- Construct validity: A previous version of this indicator (PSI definition applied to children) has been associated with increased resource use and mortality.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, co-morbidities defined using the AHRQ clinical classification software and number of and procedure type (therapeutic/diagnostic).

##### *Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than to community hospitals.
- AHRQ QI software offers stratification by high, intermediate and lower risk patients.

#### 4.3.12 PDI #13 - Transfusion Reaction

*Description:* Count of discharges under age 18 years with an ICD-9-CM code of transfusion reaction in any secondary diagnosis field (not present on admission) and a count of the number of total discharges.

*Tier:* Suitable for comparative reporting (Tier 1B).

##### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel, but panel did not recommend for comparative purposes.
- Criterion validity: Most cases flagged are not present on admission.

##### *Considerations of Use*

- This indicator is reported as a count
- The serious reportable event indicators are not intended for use in comparing one hospital with another, but rather to promote public accountability and to support root-cause analysis of identified cases.

#### **4.4 Neonatal Quality Indicators (NQI)**

The Neonatal Quality Indicators (NQI) consists of 3 indicators, including two adverse event indicators and one mortality indicator.

##### **4.4.1 NQI #01 - Iatrogenic Pneumothorax in Neonates**

*Description:* Rate of discharges with an ICD-9-CM code of iatrogenic pneumothorax in any secondary diagnosis field (not present on admission) among surgical and medical discharges, defined by specific DRG codes, in neonates with birthweight less than 2500 grams.

*Tier:* Suitable for comparative reporting with some additional development (Tier 2).

##### *Supporting Evidence:*

- Face validity: Overall usefulness for quality improvement and comparative purposes rated as acceptable by clinical panel.
- Criterion validity: None.
- Construct validity: Processes of care, such as antenatal steroids, surfactant use and appropriate ventilation practices, have been shown to decrease pneumothoraces on a patient level.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, age in days, age in years, gender, co-morbidities defined using the AHRQ clinical classification software.

##### *Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children's hospitals to other children's hospitals, rather than community hospitals.
- 

##### **4.4.2 NQI #02 - Neonatal Mortality**

*Description:* Rate of in-hospital death among inborn (newborns born in this hospital) and out-born (newborns born in another hospital) neonates.

*Tier:* Suitable for comparative reporting with significant additional development (Tier 3).

##### *Supporting Evidence:*

- Face validity: Clinical panel recommended for comparative purposes, but rated the indicator as less useful for quality improvement purposes (potentially because neonatal mortality is thought of as a structural marker for the level of care provided)
- Criterion validity: None.
- Construct validity: . Interventions to reduce neonatal mortality focus on improving outcomes for low birth-weight infants, higher risk larger neonates and delaying premature delivery when possible; although the latter has fewer known processes of care associated with it.

- Risk adjustment: Hierarchical model based on covariates for birth-weight, gestational age, age in days, gender, and congenital anomalies.

*Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children’s hospitals to other children’s hospitals, rather than to community hospitals.
- AHRQ QI software offers stratification based on birth-weight

4.4.3 NQI #03 - Blood Stream Infections in Neonates

*Description:* Rate of discharges with an ICD-9-CM code of nosocomial blood stream infections in any secondary diagnosis field (not present on admission) among inborn and out-born neonates with birth-weight 500 to 1499g, with gestational age between 24 and 30 weeks, or with a birth-weight greater than or equal to 1500g, if the infant experienced death, major surgery, mechanical ventilation, or transfer in or out from/to an acute care facility...

*Tier:* Suitable for comparative reporting (Tier 1)

*Supporting Evidence:*

- Face validity: Clinical panel recommended for quality improvement and comparative purposes. Review by an NQF technical advisory panel rated acceptable for comparative public reporting.
- Criterion validity: None.
- Construct validity: Processes of care, such as hand washing, have been shown to decrease infection rates. Some hospitals have been able to reduce their infection rates following implementation of quality improvement interventions.
- Risk adjustment: Hierarchical model based on covariates for birth-weight, gestational age, age in days, gender, and congenital anomalies.

*Considerations of Use*

- With all the pediatric indicators, when comparing performance between general acute care hospitals and pediatric hospitals, the comparison should be among pediatric populations only. It is most appropriate to compare children’s hospitals to other children’s hospitals, rather than to community hospitals.
- AHRQ QI software offers stratification based on birth-weight..