



# How We Rate Hospitals

June 2014

	<u>Page</u>
1. Overview .....	2
2. Patient Outcomes	
2.1 Avoiding infections .....	7
2.2 Avoiding readmissions .....	13
2.3 Avoiding mortality - medical. . . . .	15
2.4 Avoiding mortality - surgical. . . . .	17
2.5 Avoiding adverse events following common scheduled surgeries (Surgery Ratings) ...	19
Appendix A: Surgical procedures in each surgical group .....	26
3. Patient Experience .....	29
Appendix B: HCAHPS survey questions that comprise each Ratings category. . . . .	32
4. Hospital Practices	
4.1 Appropriate use of abdominal and chest CT scanning .....	34
4.2 Avoiding C-sections. ....	36
5. Safety Score .....	39
6. Heart Surgery Ratings. ....	43

## 1. Overview

Consumer Reports' hospital Ratings (<http://www.consumerreports.org/health/doctors-hospitals/hospital-ratings.htm>) include measures of Patient Outcomes (avoiding infections, readmissions, avoiding mortality, and adverse events in surgical patients), Patient Experience (including communication about hospital discharge, communication about drug information and other measures), and Hospital Practices (appropriate use of scanning, use of electronic health records, and avoiding C-sections). Several of these measures are then combined to create our Safety Score. This document describes these Ratings in detail, starting with an overview of the Ratings on Consumer Reports online. We also periodically publish hospital Ratings in the pages of *Consumer Reports* magazine.

In constructing these Ratings, we do extensive research to bring together reliable, valid, and objective information on hospital quality. The source data come from the Centers for Medicare and Medicaid Services (CMS), the Centers for Disease Control and Prevention (CDC), state inpatient databases, and the American Hospital Association (AHA). Our research entails an in-depth evaluation of the quality and objectivity of each of these sources. If the data meet our quality standards, we then turn it into usable information that is accessible and meaningful to consumers. We routinely update our Ratings, both by updating the information that's already there and by retiring measures and adding new measures of hospital quality as they become available. Details about each measure are shown in the table on the following page.

With each set of measures, we enlist the help of external expert reviewers for feedback on measure methodology and on how we propose to turn the measures into Ratings. That feedback has been incorporated in the methods described in this document, and is a crucial part of making sure that we present information that is consistent with the current state of scientific knowledge on hospital quality.

Our Ratings use a 1-to-5 scale (corresponding to Consumer Report's well-known colored dots, called blobs), where higher numbers are better. For the components of the Safety Score and other composites, we include more significant digits in our calculations by using the "fractional blob" scale, which ranges from 0.5 to 5.5. Converting our Ratings to this scale enables us to combine and compare different quality components on a common scale. The technical details for expressing each measure on a fractional blob (FB) scale and for creating the blobs that appear in our Ratings are described in the sections of this report that follow.

## Summary of Hospital Ratings Domains

Category	Measures	Source	Dates covered by data in our March 2014 update
Safety Score composite	(denoted with * below)		Varied; see below
Patient Outcomes	*Avoiding bloodstream infections *Avoiding surgical site infections *Avoiding catheter-associated urinary tract infections (new)	CMS	April 2012 – March 2013
	*Avoiding readmissions	CMS	July 2011 – June 2012
	*Avoiding mortality – medical (new)	CMS	July 2009 – June 2012
	*Avoiding mortality – surgical (new)	CMS	July 2010 – June 2012
	Avoiding adverse events in surgical patients (surgery Ratings)	CMS	2009 – 2011
Patient Experience	Overall patient experience *Communication about hospital discharge *Communication about drug information Doctor-patient communication Nurse-patient communication Pain control Help from hospital staff Room cleanliness Room quietness	CMS (MedPAR)	April 2012 – March 2013
Hospital practices	Use of electronic health records	AHA	January – December 2010
	*Appropriate use of abdominal scanning *Appropriate use of chest scanning	CMS	January – December 2011
	Avoiding C-sections (new)	State in patient billing records (see page 36 for details).	Two years of data, spanning 2009-2012 (see page 36 for details)
Heart Surgery	Isolated heart bypass surgery *Overall Rating *Patient survival	The Society of Thoracic Surgeons	July 2012-June 2013

	<ul style="list-style-type: none"> <li>*Absence of surgical complications</li> <li>*Appropriate medications</li> <li>*Optimal surgical technique</li> </ul>		
	<p>Aortic heart valve replacement</p> <ul style="list-style-type: none"> <li>*Overall Rating</li> <li>*Patient survival</li> <li>*Absence of surgical complications</li> </ul>	<p>The Society of Thoracic Surgeons</p>	<p>July 2010 – July 2013</p>

## **Note regarding changes to measures used by Consumer Reports:**

### **March 2014 update:**

Changes to the measures reported by Consumer Reports are outlined below. More details are available in the relevant section for each measure.

March 2014 update:

- Replaced the CMS heart failure, heart attack, pneumonia readmission measures with the hospital-wide, all-cause readmission measure ("avoiding readmissions") (see page 13)
- Added the CMS heart failure, heart attack, pneumonia mortality measure ("avoiding mortality - medical") (see page 15)
- Added PSI-4 - Death among surgical patients with serious treatable complications ("avoiding mortality - surgical") (see page 17)
- Added catheter-associated urinary tract infection data to the Safety Score (see page 39)
- Removed PSI-90 "avoiding complications"

Data sources: For hospital-acquired infections we no longer use any state-based data; we used data reported to CDC's National Healthcare Safety network (NHSN) which is then reported to CMS.

### **May 2014 update:**

- Added the AHRQ IQI 33 measure (Primary c-section rate, uncomplicated ("avoiding C-sections")) for 22 states (see page 36)

### **June 2014 update**

- Added heart surgery Ratings (Isolated heart bypass surgery and aortic valve replacement surgery)
- Published infections composite and catheter-associated urinary tract infection Ratings

## **Limitations**

Unlike most other Consumer Reports' Ratings, we do not collect hospital data ourselves, and so the actual implementation of the data collection and analysis is not in our control. There may be quality control issues that do not meet the high standards that Consumer Reports generally applies to our own data. In many cases, the Consumer Reports Health Ratings Center only has access to summarized results of data analysis, preventing us from validating the data calculations or presenting data to you in alternative ways. However, we carefully review the methods of data collection, validation, and analysis used by each data provider. Based on that extensive review, we use only the highest-quality data available that provides important and useful information for consumers. Our interpretations of the data incorporate our understandings of any data limitations, which are described in greater detail in the following sections.

Our hospital Ratings are based on a range of measures that we believe reflect the quality of important dimensions of patient care. However, there are many dimensions to hospital quality, beyond those reported here. For example, there may be information available about the hospital's performance in a specific clinical area that is important to you. In fact, Consumer Reports, in collaboration with The Society of Thoracic Surgeons, publishes ratings of surgical groups ([www.consumerreports.org/health/doctors-hospitals/surgeon-ratings/heart-surgery-](http://www.consumerreports.org/health/doctors-hospitals/surgeon-ratings/heart-surgery-)

[ratings/overview/index.htm](#)) that perform coronary artery bypass surgery (CABG) who have volunteered to release their data to the public through Consumer Reports. State-based non-profit quality organizations, state departments of health and national for- and non-profit organizations publish quality data that may be helpful for you in assessing hospital quality. In addition, the Informed Patient Institute (<http://www.informedpatientinstitute.org>) publishes evaluations of medical service quality report cards.

## 2. Patient Outcomes

### 2.1 Avoiding Infections

Some healthcare-acquired infections (HAI) data are reported by the Centers for Medicare and Medicaid Services (CMS), an agency of the Federal government. Most hospitals report data on specific HAIs, such as central line associated bloodstream infections (CLABSIs), select surgical site infections, (SSIs) and catheter associated urinary tract infections (CAUTI), to the Centers for Disease Control and Prevention (CDC); those data are then publicly reported on CMS's Hospital Compare website ([hospitalcompare.hhs.gov](http://hospitalcompare.hhs.gov)). Beginning in 2011, reporting of select HAIs became linked to an annual across the board payment increase for Medicare payments to hospitals. If hospitals in the CMS Inpatient Prospective Payment System do not submit the scheduled information required on infections, they lose a portion of this annual payment increase; this payment structure is used as an incentive that causes virtually all of these hospitals to report.

#### **Central-line associated bloodstream infections (CLABSI) data**

In 2012 we started to use intensive-care unit (ICU) CLABSI data that are publicly reported by CMS, rather than to states. Hospitals are required to report infections as CLABSI if it is a laboratory-confirmed bloodstream infection where the central line (or umbilical catheter) was in place for more than two calendar days from the date of the confirmed infection and a central line (or umbilical catheter) was in place on the date of the confirmed infection or the day before.

Hospitals that report data to CDC are required to do so quarterly for every ICU and select other specialty areas in the hospital, for all these patients (not just Medicare patients). The CDC calculates a standardized infection ratio (SIR) which is reported to the public through CMS's Hospital Compare website. Data are combined for four quarters, with a delay of approximately 10 months (so for example, the CLABSI data reported by CMS in October 2013 reflects the 12-month period ending in December 2012).

#### **Surgical-site Infections (SSI) data**

For the first time, for surgical site infections (SSI), we are using only data from CMS (not state data) available from Hospital Compare. Hospitals currently report (to CMS) surgical-site infections for two types of surgeries or face a financial penalty: abdominal hysterectomy and colon. Hospitals report deep incisional primary and organ/space infections that are detected during the hospitalization during which the surgery occurred, or if the patient is readmitted to any hospital, or if the infection is identified after the patient is discharged but not rehospitalized. Hospitals are required to report infections that started 30 days or less following the surgery in patients who are 18 years or older at the time of their surgery.

The data is sent to the CDC and CDC risk adjusts the data by taking into account the patients age and the patients' overall health according to the classification system developed by the [American Society of Anesthesiologists](#). CDC this data is then reported quarterly to CMS, who publishes it on Hospital Compare.

## **Catheter-associated urinary tract infection (CAUTI) data**

For the first time our Ratings include data from CMS on catheter-associated urinary tract infections (CAUTI) that most hospitals are required to report to the government or receive a financial penalty. Hospitals must report urinary tract infections that are associated with the patient having an indwelling urinary catheter (tube inside the body inserted in the bladder) and are diagnosed based on the patients' symptoms, as well as urinary tract infections without symptoms that have caused a bloodstream infection, within 48 hours of insertion of the catheter. Hospitals are currently only required to report infections that occur in adults and children that are admitted in the intensive-care unit and select other specialty areas.

As with the other types of infections discussed above, CAUTIs are reported to the CDC's NHSN. This data is reported to the public through CMS's Hospital Compare website.

### **The basis of the Ratings: The standardized infection ratio**

For each hospital, we calculate the standardized infection ratio, a measure developed by the CDC and modeled after the standardized mortality ratio (or standardized incidence ratio), a common measure in epidemiology.

The SIR (which is the basis for the Rating) compares data within each of several subgroups to the national benchmark data for those subgroups, and then pools the comparisons across all subgroups. National data are derived from rates reported to the National Healthcare Safety Network (NHSN) a data repository supported by the Centers for Disease Control and Prevention (CDC).<sup>1</sup> The benchmark rates for SSI and CLABSI are composite data from approximately 1,500 hospitals in 2006-2008 in 48 states and the District of Columbia; a subset of NHSN data for individual hospitals is publicly available through CMS. CAUTI national benchmark data is from 2009. While more recent national averages are available from NHSN, like CDC, we have continued to use these benchmark data for consistency, to allow us to demonstrate changes in incidence of infections over time.

This analysis adjusts for the fact that different reporting agencies and different hospitals have data from varying mixtures requiring comparisons to different average infection rates. For instance, the average CLABSI infection rate for cardiac ICUs nationwide is two per 1,000 central-line days, so a particular cardiac ICU with a *rate* of three infections per 1,000 days has 50 percent more infections than would be predicted from the national average. For surgical ICUs, the national average CLABSI rate is 2.3 infections per 1,000 central-line days, so a surgical ICU reporting a rate of 4.6 infections per 1,000 CLD produces infections at twice the national rate, or 100 percent more infections than average. The standardized infection ratio pools these comparisons across all ICUs for which a hospital reports CLABSI rate, giving a single Rating for each hospital's reported ICUs. For more details of the calculation of the SIR, see [http://www.cdc.gov/HAI/pdfs/stateplans/SIR-2010\\_JunDec2009.pdf](http://www.cdc.gov/HAI/pdfs/stateplans/SIR-2010_JunDec2009.pdf). CAUTI and CLABSI SIRs are adjusted for patient mix by type of patient care location (ICU type), hospital affiliation with a medical school, and bed size of the patient care

---

<sup>1</sup> Edwards et. al, National Healthcare Safety Network (NHSN) Report, data summary for 2006 through 2008, issued December 2009, *Am J Infect Control*; 37: 783-805.

location. SSI uses a logistic regression model to risk adjust at the patient level before pooling data together for that procedure type.

A Standardized Infection Ratio of 1 means that the hospital's ICUs produce infections at the same rate overall as would be predicted from national rates. A SIR of more than 1 reflects more infections than predicted, and SIR less than 1 implies fewer infections than predicted.

### Assigning Individual Infection Ratings

We use the most current data available. We report the number of infections and the number of central-line days (CLDs), or surgical procedures for SSI, urinary catheter days for any hospital that reported at least 600 CLDs/catheter days (or more than 50 surgical procedures for SSI), or at least one infection regardless of CLDs (or surgical procedures). In addition, for all three types of infections (CLABSI, SSI, CAUTI) we calculate Ratings scores for all hospitals that meet *either* of the following sample size requirements:

1. At least one total predicted infection. Volumes less than this yield less reliable ratings.
2. At least three infections, regardless of CLD, number of surgical procedures, or urinary catheter days. This allows us to identify additional hospitals with high infection rates, even in small volumes.

For each hospital with sufficient data, we report the percentage different from national average rates separately for CLABSI, CAUTI, and surgical-site infections. The percentage difference from average (rounded to two decimal places) is based on the SIR, and is reported as shown in the table on the following page. SIRs are rounded to whole number percentages for display purposes; all available digits of precision are maintained for further calculations.

To receive the highest 5-blob Rating a hospital must achieve the standard of zero infections and have at least 1 predicted infection. Although the SIR on which our Ratings are based reflects comparisons with average national rates as a way for adjusting for the varying risk of infection in different ICUs or with different surgical procedures, the SIR should not be seen as a safety benchmark; average performance still infects people. All hospitals should be working toward having zero hospital-acquired infections, and there are enough hospitals reporting zero infections to expect that all hospitals can achieve this standard. For all three infection types, we assign fractional blobs (FBs) on a scale from 0.5 to 5.5 using a piecewise linear transformation as follows:<sup>2</sup>

- a. If the SIR = 0, then the hospital is assigned a FB value of 5.5 and a blob score of 5. Only hospitals with zero reported infections (SIR = 0) can receive a blob score of ●.
- b. If  $0 \leq \text{SIR} \leq 0.5$ , then the FB is calculated using a linear transformation that maps a SIR of 0 to a FB of 4.5 (note that no actual data value will exist at that point) and a SIR of 0.5 to a FB of 3.5. These hospitals will get a blob score of ●.

---

<sup>2</sup> This represents a change from our previous ratings methodology for ConsumerReports.org in July 2012 and in the August 2012 issue of *Consumer Reports* magazine.

- c. If  $0.5 < \text{SIR} \leq 1.0$ , then the FB is calculated using a linear transformation that maps a SIR of 0.5 to a FB of 3.5 and a SIR of 1.0 to a FB of 2.5. These hospitals will get a blob score of ○.
- d. If  $1.0 < \text{SIR} \leq 2.0$ , then the FB is calculated using a linear transformation that maps a SIR of 1.0 to a FB of 2.5 and a SIR of 2.0 to a FB of 1.5. These hospitals will receive a blob score of ◐.
- e. If  $2.0 < \text{SIR} \leq 4.0$ , then the FB is calculated using a linear transformation that maps a SIR of 2.0 to a FB of 1.5 and a SIR of 4.0 to a FB of 0.5. These hospitals will receive a blob score of ●.
- f. For  $\text{SIR} > 4$ , the FB is set to be equal to 0.5, with a blob score of ●.

These calculations result in ratings scores as shown in the following table:

	CLABSI, CAUTI, or SSI Rating	Fractional Blob Range	SIR range	Interpretation
Better	●	FB = 5.5	SIR = 0	0 infections
	◐	$4.5 > \text{FB} \geq 3.5$	$0.0 < \text{SIR} \leq 0.5$	At least 50% better than average
↕	○	$3.5 > \text{FB} \geq 2.5$	$0.5 < \text{SIR} \leq 1.0$	Between average and 50% better than average
	◑	$2.5 > \text{FB} \geq 1.5$	$1.0 < \text{SIR} \leq 2.0$	No more than 100% worse than average
Worse	●	$1.5 > \text{FB}$	$2.0 < \text{SIR}$	More than 100% worse than average

If you see a hospital that falls at one of the cutoff points between blob scores, you may see what looks like a discrepancy between its percentage difference from national average and its blob score. This is not an error, but results from rounding the percent difference to the nearest whole percent. For example, if a hospital has a SIR = 0.502, then it receives a 3 blob, since its SIR is greater than 0.5. This hospital is 49.8% better than national rates; since we print these percentage differences to the nearest whole number percent, it will be reported online as being 50% better than national average.

### Composite infection score

In order to calculate a hospital's Safety Score, we created a composite infection score that combines data for each hospital for CLABSI, CAUTI and SSIs. The final method to calculate the composite infection score is shown below, which weights infection types with higher SIRs more. This mitigates the composite from masking lower performance in a single infection SIR.

$$SIR_{\text{weighted}} = \frac{w_1 O_1 + w_2 O_2 + w_3 O_3}{w_1 E_1 + w_2 E_2 + w_3 E_3},$$

where

$$w_i = \begin{cases} 1 & \text{if } 0 \leq SIR_i^* < 1 \\ 1 + \left( \frac{SIR_i^* - 1}{4 - 1} \right) \times (2 - 1) & \text{if } 1 \leq SIR_i^* \leq 4 \\ 2 & \text{if } SIR_i^* > 4 \end{cases}$$

$$SIR_i^* = \frac{O_i}{\max(1, E_i)}, \quad i = \text{CLABSI, SSI, CAUTI}$$

Infection Composite Rating	Fractional Blob Range	SIR <sub>weighted</sub> range
●	5.5 ≥ FB ≥ 4.5	0.0 ≤ SIR ≤ 0.1
◐	4.5 > FB ≥ 3.5	0.1 < SIR ≤ 0.5
○	3.5 > FB ≥ 2.5	0.5 < SIR ≤ 1.0
◑	2.5 > FB ≥ 1.5	1.0 < SIR ≤ 2.0
●	1.5 > FB ≥ 0.5	2.0 < SIR

### Limitations

Although extremely serious, these infections are relatively infrequent, which makes the infection *rates* volatile, as the occurrence of one or two infections can have a large impact on reported rates. Many hospitals are working toward reducing infection rates in their ICUs, operating rooms, and throughout their facilities, so current rates may differ from those reported here. Whenever possible, we present the most current data publicly available.

Most SSIs are not identified until patients are discharged from the hospital and patients with infection do not always return to the hospital where the surgery was performed. To identify infections after discharge and accurately estimate the incidence of SSIs, hospitals use various approaches, including review of data sources for re-admission and emergency room visits, to improve the detection of SSIs. All patients who experience infections may not be re-admitted or go to the hospital's emergency department, so there are many infections that will not be identified by the hospital's reporting system.

SSI data reported to CMS includes only two surgical procedures (colon and hysterectomy), which limits the generalizability of the data. It also does not allow Consumer Reports, or consumers, to evaluate SSIs in hospitals that specialize in other areas, such as orthopedic surgery or cardiac surgery.

CLABSI, SSI and CAUTI data reported by CMS are self-reported by hospitals, without any independent or external validation. Most states that also publicly report this data are required by state law to issue valid, accurate and reliable data. But only some (in particular, New York, Tennessee, Colorado, Connecticut, and South Carolina) are doing regular evaluations or audits of the data. Consumers Union continues to advocate for laws requiring validation and auditing of hospital infection data. But we also believe that consumers have a right to the best information currently available on hospital-acquired infections, which are dangerous, costly, and even deadly.

Success at infection control requires a sustained change in the culture and approach of hospitals. As we add more safety measures it will be interesting to see if this translates to success in other aspects of hospital care. Moreover, public reporting will likely encourage hospitals to improve both their data collection and their efforts to prevent infections.

## 2.2 Avoiding Readmissions

Hospital readmissions data are collected by the Centers for Medicare and Medicaid Services (CMS), an agency of the Federal government. In 2009, CMS began reporting a 30-day readmission measure for people diagnosed with heart failure, heart attack, and pneumonia. Medicare reimbursement to hospitals paid under the Inpatient Prospective Payment System is currently tied to hospitals' reporting of this measure, as well as their performance.

To provide a broader assessment of the quality of care at hospitals, in 2013 CMS began reporting a hospital-wide, all-cause readmission rate for most hospitals in the United States. This year (as of March 2014) we replaced the three condition specific readmission measures with the new hospital-wide readmission measure. The information reported by CMS shows an estimate of the likelihood that a patient will be readmitted within 30 days of discharge from a previous hospital stay for any condition. People may have been readmitted back to the same hospital or to a different hospital. They may have had an unplanned readmission for the same condition as their recent hospital stay, or for a different reason.

Readmissions rates are important quality indicators for several reasons. First, any hospital admission has inherent risks, and hence a second admission exposes the patient to additional risk. Second, readmissions can be caused by things that go wrong in the initial discharge.<sup>3</sup> Third, we know that, to at least some extent, readmissions reflect errors or hospital-acquired conditions in the initial hospitalization.<sup>4</sup>

### The data

CMS publishes readmission rates after statistical adjustment for how sick people were when they were initially admitted to the hospital and for the amount of cases available for each hospital. CMS provides each hospital's 30-day risk-standardized readmission rate (RSRR). [Details of the measure](#) are available on the Quality Net website.

Data reported on Hospital Compare cover discharges over a two-year period for over 4,000 hospitals. We provide the chance of readmission for any hospital with at least 25 cases. In addition, we provide a Rating score as described below.

### Assigning Ratings scores

We re-scale the reported readmission rates on our fractional blob scale, as described in the chart below. Cut points for the blobs are based on a combination of the data distribution and on input and review by experts in quality measurement and clinical medicine.

---

<sup>3</sup> [www.ahrq.gov/qual/impptdis.htm](http://www.ahrq.gov/qual/impptdis.htm)

<sup>4</sup> Emerson et al., *Infect Control Hosp Epidemiol.* 2012;33(6):539-544.

Rating	Fractional Blob (FB)	Readmission Rate
●	$5.5 \geq \text{FB} \geq 4.5$	8.0-14.4
◐	$4.5 > \text{FB} \geq 3.5$	>14.4-15.4
○	$3.5 > \text{FB} \geq 2.5$	>15.4-16.4
◑	$2.5 > \text{FB} \geq 1.5$	>16.4-17.7
●	$1.5 > \text{FB} \geq 0.5$	>17.7-24.0

### **Limitations**

These data come from billing and other administrative data submitted by hospitals to Medicare. Such records were intended to capture information for billing purposes rather than patient outcomes, but they contain details about a patient's stay in the hospital. These data reflect readmissions only for Medicare patients. Ratings come from recent data but it is possible that performance today will show improvements or declines in performance data that is not currently available to us. The percentages reported are not exact numbers but estimates based on the statistical model used, and have some margin of error. Hospitals that have relatively low numbers of discharges have wider margins of error, and because of the statistical model CMS uses, are statistically adjusted to be closer to the average of all hospitals.

Finally, while these are the best data available for assessing readmissions, and they are adjusted for the health status of the patients discharged by each hospital, comparisons among hospitals with very different patient populations are only approximate. More about the statistical methods used by CMS can be found here:

<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/Statistical-Issues-in-Assessing-Hospital-Performance.pdf>

## 2.3 Avoiding Mortality - Medical

Mortality data are collected by the Centers for Medicare and Medicaid Services (CMS), an agency of the Federal government. CMS reports mortality rates for Medicare patients who died within 30 days of admission for patients who had been hospitalized for any of the following reasons: heart failure, heart attack and pneumonia.

### The data

CMS publishes mortality data after statistical adjustment for how sick patients were when they were initially admitted to the hospital and for the amount of cases available for each hospital. [CMS provides each hospital's 30-day risk standardized mortality rate](#), for each medical condition.

### Assigning Individual Medical Mortality Ratings

We create Ratings for each condition (heart attack, heart failure, pneumonia) and then combine them, weighted by the hospitals' mix of patients. For each hospital, we use whichever of the three conditions that have sufficient data (at least 25 cases), and calculate the weighted mean of the fractional blob for those measures.

Blob scores for the individual measures are derived as follows. (Note that the individual blob scores for each condition are not published on each hospital's report card; we report only the composite Rating.) We used different cut offs for heart failure and pneumonia vs. heart attack due to the differences in the distribution of the data for these measures. Cut points for the blobs are based on a combination of the data distribution and on input and review by experts in quality measurement and clinical medicine.

	Rating	Fractional Blob Range (FB)	Mortality Rate – heart failure and pneumonia
Better		$5.5 \geq \text{FB} \geq 4.5$	6.0-9.0
		$4.5 > \text{FB} \geq 3.5$	>9.0-10.8
↓		$3.5 > \text{FB} \geq 2.5$	>10.8-12.8
		$2.5 > \text{FB} \geq 1.5$	>12.8-15.0
Worse		$1.5 > \text{FB} \geq 0.5$	>15.0-18.0

	Rating	Fractional Blob Range (FB)	Mortality Rate – heart attack
Better		$5.5 \geq \text{FB} \geq 4.5$	9.0-13.0
		$4.5 > \text{FB} \geq 3.5$	>13.0-14.2
↓		$3.5 > \text{FB} \geq 2.5$	>14.2-16.2
		$2.5 > \text{FB} \geq 1.5$	>16.2-17.3
Worse		$1.5 > \text{FB} \geq 0.5$	>17.3-21.0

## **Composite medical mortality score**

For the medical mortality composite, weights of the individual mortality scores are proportional to the number of discharges for patients hospitalized for heart attack, heart failure, or pneumonia at that hospital.

### **Limitations**

These data come from billing and other administrative data submitted by hospitals to Medicare. Such records were intended to capture information for billing purposes rather than patient outcomes, but they contain significant details about a patient's stay in the hospital. These data reflect mortality only for Medicare patients.

Ratings come from the most recent data available, but there is a time lag in reporting these data to the public. It is possible that performance today will show improvements or declines in data that is not currently available to us. The percentages reported are not exact numbers but estimates based on the statistical model used, and have some a margin of error. Hospitals that have relatively low numbers of discharges have wider margins of error, and because of the statistical model CMS uses, are statistically adjusted to be closer to the average of all hospitals.

While these data are adjusted for the health status of the patients discharged by each hospital, comparisons among hospitals with very different patient populations are only approximate. More details about this measure can be found here:

<https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1163010421830>

## 2.4 Avoiding mortality - surgical

The Center for Medicare and Medicaid Services (CMS) publishes data that measure how often patients died who had surgery that had a serious treatable complication. With rapid identification and effective treatment a portion of these people could have been saved. Complications include pneumonia, deep vein thrombosis or pulmonary embolus, sepsis, acute renal failure, shock/cardiac arrest, or gastrointestinal hemorrhage/acute ulcer.

### The data

CMS reports the data as the number of patient deaths in the hospital for every 1,000 patients who had surgery with select complications.

CMS publishes surgical mortality rates after statistical adjustment for how sick patients were when they were initially admitted to the hospital and for the amount of data cases for each hospital.

### Assigning Ratings scores

We rescale the surgical mortality rates reported on Hospital Compare and assign them blob scores as described below. Cut points for the blobs are based on a combination of the data distribution and on input and review by experts in quality measurement and clinical medicine.

	Rating	Fractional Blob Range (FB)	Surgical Mortality (deaths per 1000 patients)
Better		$5.5 \geq \text{FB} \geq 4.5$	50.0-87.0
		$4.5 > \text{FB} \geq 3.5$	>87.0-100.0
↓		$3.5 > \text{FB} \geq 2.5$	>100.0-120.0
		$2.5 > \text{FB} \geq 1.5$	>120.0-132.0
Worse		$1.5 > \text{FB} \geq 0.5$	>132.0-180.0

### Limitations

These data come from billing and other administrative data submitted by hospitals to Medicare. Such records were intended to capture information for billing purposes rather than patient outcomes, but they contain significant details about a patient's stay in the hospital. These data reflect mortality only for Medicare patients.

Ratings come from the most recent data available, but there is a time lag in reporting these data to the public. It is possible that performance today will show improvements or declines in data that is not currently available to us. The percentages reported are not exact numbers but estimates based on the statistical model used, and have some a margin of error.

PSI data are only calculated for hospitals that are paid through the IPPS, which excludes Critical Access hospitals (CAHs), long-term care hospitals (LTCHs), Maryland waiver hospitals, cancer hospitals, children's inpatient facilities, rural health clinics, federally qualified health centers,

inpatient psychiatric hospitals, inpatient rehabilitation facilities, Veterans Administration/ Department of Defense hospitals, and religious, non-medical health care institutions.

While these data are adjusted for the health status of the patients discharged by each hospital, comparisons among hospitals with very different patient populations are only approximate.

Finally, this measure is limited by the accuracy of coding of complications in the billing records<sup>5</sup> and research suggests that the patient safety indicators significantly underreport the number of errors that occur in hospitals.<sup>6</sup> While this measure does draw on select complications to qualify cases for inclusion, the adverse event measured here is not the occurrence of these complications, but death.

---

<sup>5</sup> Lawson et al., *Ann Surg* 2012; 256(6):973-981

<sup>6</sup> Classen et al., *Health Aff* 2011; 30 (4): 581-589

## 2.5 Avoiding adverse events following common scheduled surgeries— Surgery Ratings

There is little comprehensive, validated clinical data available to enable consumers to compare hospitals with respect to patient outcomes following common scheduled (sometimes called “elective”) surgeries. To provide comparative information, therefore, we worked with MPA, a health care consulting firm with expertise in analyzing medical billing claims and clinical records, to develop our surgery Ratings. Three years of Medicare billing data (2009, 2010, 2011) were analyzed to estimate rates of adverse events (AE) for hospital inpatients undergoing common scheduled surgeries. In this analysis, patients were identified as having had an AE if they either died in the hospital or experienced a longer hospital stay than would be expected (called “prolonged” in what follows) following their surgery (for the specific surgery types analyzed). These AE rates are risk adjusted for the types of patients treated by each hospital. A description of this approach has been previously published.<sup>7</sup>

Risk-adjusted prolonged lengths of stay (LOS) were used as an indicator for AEs for several reasons. First, a prolonged LOS may be due to actual hospital-acquired complications or other circumstances, some of which are avoidable through high-quality care. Second, staying longer than expected in the hospital can itself be considered undesirable, or an “adverse event,” for the patient. Finally, a prolonged can be accurately and responsibly measured with billing data. In contrast, other attempts to use billing data to directly identify hospital-acquired complications are limited by the accuracy of coding of complications.<sup>10,11</sup>

### Overview

Data are from the Centers for Medicare and Medicaid Services (CMS) Inpatient Limited Dataset (LDS) for hospital discharges in calendar years 2009, 2010, and 2011. Only Medicare patients aged 65 and older are included. The 86 surgical procedures included in our analysis were grouped into 27 surgical groups (see Appendix A).

For each group, a statistical model was derived in the form of a predictive equation that identified which patients were in the hospital longer than would be expected for a patient with their general risk factors receiving a surgical procedure in that surgical group. Those models (one for each surgical category) were based on a national Medicare reference dataset of hospitals identified for the high quality of the coding in their billing records for conditions that were present on patient admission (“good-coding” hospitals). This subset of hospitals enabled the development of statistical models based on more accurate information about Medicare patients’ risks of adverse outcomes. Data from good-coding hospitals were used to estimate a composite rate of prolonged LOS and inpatient mortality for each hospital present in the CMS Inpatient Limited Dataset noted above. A hospital’s composite adverse event rate was compared to the rate of AEs predicted for the mix of patients at that hospital. Our scores are based on the degree to which observed rates of AEs differed from those predicted by the models.

---

<sup>7</sup> Fry et al., *Am J Surg*; 2009 197(4):479–484

Procedure codes and diagnostic codes were grouped together into 27 groups of elective (scheduled) surgical procedures (see Appendix A). If a patient had multiple surgical procedures, then that patient was assigned to a procedural category based on the earliest procedure. If the multiple procedures happened on the same date, MPA developed a hierarchy of categories to determine to which category that patient would be assigned.

Cases were excluded from the analysis if they met any of these conditions:

- Care was not delivered at an acute care or critical access hospital
- The patient age was less than 65 years
- Gender was not coded as male or female
- Missing patient ID, Hospital ID, Principal Diagnosis, Admission Date, or Discharge Date
- Discharge status of discharged or transferred to either another short term general hospital, other institution for inpatient care, or Federal hospital; left against medical advice or discontinued care; still a patient at the hospital
- Admitted from another hospital
- The index procedure occurred more than two days after admission, indicating that there were likely other medical or surgical issues involved.

Hospitals were excluded from the analysis if they had fewer than 20 qualifying cases over the three-year interval.

Potential clinical risk factors<sup>8</sup> for use in developing the risk-adjustment models were created based on analysis involving average mortality rates, median post-operative LOS for live discharges, and clinical judgment.

## Data analysis

Data analysis occurred in three distinct stages, each done separately within each surgical procedure group, and each explained in detail in the subsequent sections.

- (1) *Model Development.* Linear models were created to estimate patient-level predicted length-of-stay (LOS). The model for predicted LOS allowed for the identification of individual cases with prolonged risk adjusted LOS (PrRALOS). Logistic models were created to estimate patient-level predicted risks of mortality and of PrRALOS.
- (2) *Estimation of hospital-acquired adverse events.* Cases with AEs (PrRALOS or inpatient mortality) were counted, and a hospital-level predicted rate of AEs was estimated based on the risk-adjustment models developed previously.
- (3) *Calculating Ratings.* Observed and predicted LOS were then compared as the basis for our Ratings.

---

<sup>8</sup> From ICD-9-CM diagnosis and concurrent procedure codes.

## 1. Model development

Hospitals were screened for the quality of their coding of co-morbidities as present on admission (POA). This coding allows the distinction between conditions that are present when a patient is admitted to the hospital, and those that are acquired in the hospital, which are presumed to be post-operative complications. Unfortunately, many hospitals are known to have poor quality of coding for POA conditions. In order to limit the development of the models to those hospitals with high-quality POA screening, a set of 15 screens designed to assess the quality of hospital POA coding<sup>9</sup> were applied to all hospitals in the analytic dataset. The results of each screen were aggregated to assign a final score to each hospital in the dataset. Only hospitals with scores greater than 80 percent (“good-coding” hospitals) were included in model development. Obvious errors in POA coding were corrected prior to performing further analyses.

Within these hospitals, the dataset was further limited to live discharges (living patients) who did not have a hospital-acquired condition coded in their medical billing records<sup>10</sup>.

Within this limited dataset and separately for each of the 27 procedure groups, preliminary risk-adjustment models were developed for predicting post-operative LOS from the potential risk factors chosen earlier, using a combination of forward stepwise regression and MPA’s clinical judgment. Using XmR control charts<sup>11</sup> outliers were iteratively identified and removed until all cases remaining had standardized post-operative LOS that were not outliers.<sup>12</sup>

The remaining cases comprised the analytic dataset of live discharges within good-coding hospitals and with routine LOS<sup>13</sup>. This dataset was used with the risk factors identified in the preliminary models to estimate new coefficients for a final predictive model for patient-level routine post-operative LOS.

This final model for routine post-operative LOS was then used to identify cases with prolonged risk-adjusted length of stay. Using this new model, MPA followed the same procedure documented above, iteratively using control charts for removing outliers. All cases that were eliminated because they were outliers<sup>14</sup> were classified as having had prolonged risk-adjusted post-operative LOS.

Finally, within each procedure group, and using cases at good-coding hospitals, MPA created predictive models for patients’ risks of inpatient mortality and of prolonged post-operative lengths of stay (two categorical dependent variables), using a combination of forward stepwise logistic regression and clinical judgment, based on the pool of all potential risk factors identified earlier. These models were used to compute estimates of each hospital’s risk-adjusted inpatient mortality rate and its risk-adjusted rate of prolonged post-operative lengths of stay.

---

<sup>9</sup> Pine M, Fry DE, Jones BL, Meimban RJ: *Perspect Health Inf Manag* 2009; 6: 2 Epub Feb 11.

<sup>10</sup> According to ICD-9-CM diagnosis codes

<sup>11</sup> Fry DE, Pine M, Jones BL, Meimban RJ. *Am J Surg* 2012; 203 (3):392-6

<sup>12</sup> Standardized risk adjusted postoperative LOS = (observed LOS) – (predicted LOS). XmR control charts were created for standardized post-operative LOS, cases above the  $3\sigma$  upper bound were removed, and a new control chart was created. This process was iterated until all cases had a standardized post-operative LOS  $\leq 3\sigma$ .

<sup>13</sup> standardized post-operative LOS  $\leq 3\sigma$

<sup>14</sup> above the upper control limit of  $3\sigma$

## 2. Computation of risk-adjusted rates of hospital-acquired adverse events

For each procedure group, MPA returned to the complete dataset of cases from both “good-coding” (see Step 1, Model development) and “poor-coding hospitals”, after excluding cases that met the exclusion criteria listed earlier. The models for predicting the two discrete dependent variables (mortality and prolonged LOS) were applied to all cases in this dataset.

A patient was classified as having had a hospital-acquired AE if the patient died in the hospital or was discharged alive but had a prolonged risk-adjusted post-operative LOS, regardless of whether they had a coded hospital-acquired complication.<sup>15</sup>

For each hospital, MPA used those models to calculate the predicted AE rate as:

$$\text{Predicted AE rate} = (\text{Predicted Mortality rate}) + \\ [(1 - \text{Predicted Mortality rate}) * (\text{Predicted Prolonged LOS rate})]$$

Predicted AEs for all hospitals were then multiplied by a constant to standardize them so that the total observed rate of AEs equals the total predicted rate of AEs over all cases in the dataset.

These methods produced predicted AEs for each hospital with sufficient data within each procedure group. For the overall rate of predicted and observed AEs, the numbers of predicted and observed AEs were summed across all procedure groups.

## 3. Assigning Ratings Scores

In addition to scoring the overall rate of AEs, we assigned Ratings for 5 surgical groups (see Appendix A for a list of the surgical procedures in each group):

- Group 7: Coronary angioplasty (Percutaneous Coronary Intervention or PCI)
- Group 12: Carotid artery surgery (Carotid Endarterectomy)
- Group 25: Back Surgery
- Group 26: Hip Replacement
- Group 27: Knee Replacement

Ratings scores were assigned separately within each procedure group and overall. Scores were only assigned if the number of predicted AEs was at least 4.5. These scores were based on an approximate z-score (which we call  $z^*$ ), calculated as:

$$z^* = ((\text{observed AEs rate}) - (\text{predicted AE rate})) / \\ \text{sqrt}((\text{predicted AE rate}) * (1 - \text{predicted AE rate}) / n)$$

---

<sup>15</sup> That is, patients were classified as having a hospital-acquired adverse event regardless of whether or not they had an ICD9-CM diagnoses code designated as hospital-acquired by POA coding.

Blobs for each of the five scored surgical groups, and for overall AEs, are then assigned based on  $z^*$ , as shown in the following table:

	Rating	AE Ratio
Better	●	$z^* \leq -2.5758$
	◐	$-2.5758 < z^* \leq -1.64485$
↑	○	$-1.64485 < z^* \leq 1.64485$
↓	◑	$1.64485 < z^* \leq 2.5758$
Worse	●	$2.5758 < z^*$

In the printed article in the September 2013 issue of Consumer Reports, hospitals are identified as “Hospitals that Make the Grade” if they have an Overall Surgery Rating of a 5-blob plus have Ratings for at least 10 procedure categories. At least 30% of the procedure groups with Ratings have to be either a 4-blob or 5-blob and no Rating can be below a 3-blob. For the analyses presented in that article, the following terms are used:

- A teaching hospital is defined to be one that is listed by the American Hospital Association as being a member of the American Medical College’s Council on Teaching Hospitals (COH).
- A specialty hospital was identified by the presence of relevant words in its name (e.g. cardiac, heart, spine, etc).
- Whether a hospital was rural or urban was determined by categories reported by the American Hospital Association based on the population size of the area in which the hospital is located.

## Limitations

These data come from billing and other administrative data submitted by hospitals to CMS. Such records were intended to capture information for billing purposes rather than for identifying events involved in patient care but they contain significant details about a patient’s stay in the hospital. These data also reflect AEs only for Medicare patients at least 65 years old, and only those patients who received elective (scheduled) surgical procedures in one of the 27 groups we investigated.

Ratings come from recent data but it is possible that performance today will show improvements or declines in data that we do not currently have access to. These data are historical, and reflect performance during a past time period. Although extremely serious, deaths and instances of prolonged lengths of stay are relatively infrequent, which makes the *rates* volatile, as the occurrence of several such instances can have a large impact on reported rates. Many hospitals are working toward reducing deaths and minimizing lengths of stays, so current rates may differ

from those reported. However, these scores are based on recent data. Using data from three years helps in providing a more stable picture of performance as it “smooths” out what may be “blips” in data if a shorter time period is used. However, three years of data include data back to a prior time period that may be less reflective of performance today.

The rates reported are not exact numbers but estimates based on the statistical model used, and have some a margin of error. Hospitals that conducted relatively low numbers of procedures have wider margins of error, which is accounted for in the way that Ratings scores are assigned.

While the data are adjusted for the health status of the patients receiving surgery at each hospital, comparisons among hospitals with very different patient populations should be done cautiously.

Hospitals vary widely in the number of elective surgeries they perform. If a hospital performs a large number of procedures, then its rates of AEs are relatively stable, in that the disposition of a single patient has a minimal impact on the Ratings. In contrast, a hospital that performs a relatively small number of procedures can have its rate impacted more strongly by the outcomes for an individual patient. Consequently, it is more likely for a small hospital than a large hospital to have an extreme rate of AEs. We compensate in part for this difference by assigning Ratings scores (blobs) based on the degree to which the actual rate of AEs can be statistically differentiated from the predicted rate, based on an estimate of its margin of error (standard deviation), making it easier to differentiate larger hospitals with above and below-predicted performance.

Ratings reflect the performance of hospitals, not of individual surgeons. The Ratings provide important information about the average performance of surgeons who practice at a given facility, but a patient may want to ask more specific questions about surgeon-specific performance.

As noted elsewhere, the data source we use is from billing claims data the hospital submits to CMS for payment. To some degree, such data are subject to what is commonly called “gaming,” in which a hospital intentionally provides an inaccurate or incomplete representation on the claim of what occurred during the hospital stay, in order to enhance their performance when the data are used to measure the occurrence of AEs. Gaming is minimized by federal oversight audits, and by the fact certain types of inaccurate claims submission are seen as fraudulent billing to Medicare and punishable by law.

Some hospitals do not appear in this report as they are not required to submit claims to the Medicare data set that is used for these Ratings. Examples of such hospitals include Veterans Administration or Pediatric hospitals, or hospitals that see mostly Medicare Advantage patients. If a hospital in which you are interested is not included in this report, you may want to inquire about its performance in areas such as mortality and prolonged LOS.

While we currently publish individual Ratings for five specific procedures, the overall surgery Rating we publish represents performance across all 27 groups of scheduled (elective) surgical procedures. This high-level general Rating may not necessarily reflect the hospital’s performance

in each of the 27 groups. Even where a hospital is highly rated, you may want to check its Ratings for the specific group in which your procedure falls, or if that rating is not shown, discuss these results with the hospital to ask about their performance for the type of procedure you will have and what systems they have in place to ensure this high level of performance across their facility.

Risk-adjusted prolonged LOS, one component that makes up the Rating, is not inherently an indication that a potentially preventable complication occurred. While there is evidence that when complications occur, the LOS tends to increase,<sup>16</sup> there has yet to be conclusive evidence published that longer than expected stays are routinely the result of a complication. Some or many prolonged lengths of stay may indeed be due to potentially preventable complications. Other instances of long stays may be due to other circumstances. Having said this, longer lengths of stay than should be anticipated are typically undesirable in several ways as long lengths of stay:

- Are generally undesirable from a quality of life perspective of the person.
- Can increase the cost and resource consumption for the payer, which may mean the patient depending on insurance.
- Increases the opportunity for bad things to happen and exposes the person to unnecessary services to be delivered.

---

<sup>16</sup> Kassir et al., *J Am Coll Surg*. 2012; 215(3):322-30

## Appendix A. Surgical procedures in each surgical group

Group 1: Brain	A: Craniectomy/Excise Lesion B: Other Excision Brain Lesion/Tissue
Group 2: Spinal Canal	A: Explore/Decompress Spinal Canal / Spondylosis B: Explore/Decompress Spinal Canal / Degen Disc C: Explore/Decompress Spinal Canal / Spinal Stenosis D: Explore/Decompress Spinal Canal / Spondylolisthesis
Group 3: Head & Neck	A: Thyroidectomy B: Parathyroidectomy C: Complete Sialoadenectomy
Group 4: Lungs & Thorax	A: Local Excision Lesion/Tissue Lung / Neoplasm B: Segmental Lung Resection / Malignancy C: Pneumonectomy/Lobectomy / Malignancy
Group 5: Cardiac Valve	A: Open Mitral Valvuloplasty B: Aortic Valve Replacement C: Mitral Valve Replacement/Annuloplasty
Group 6: CABG	A: Coronary Artery Bypass Graft
Group 7: PCI	A: Percutaneous Coronary Intervention
Group 8: Heart Lesion Excision	A: Excision Heart Lesion/Tissue / Dysrhythmia
Group 9: Auto Defib System	A: Implant Auto Cardioversion/Defib System
Group 10: Abdominal Aortic Aneurysm	A: Angioplasty/Atherectomy / Abdominal Aneurysm B: Endovasc Graft Implant Abd Aorta / Aneurysm
Group 11: Other Vascular	A: Angioplasty/Atherectomy / Renal Artery B: Angioplasty/Atherectomy / Limb Vessel C: Endovasc Repair / Head/Neck Vessel/Cerebral Aneurysm D: Percutaneous Angioplasty/Atherectomy / Precerebral
Group 12: Carotid Endarterectomy	A: Endarterectomy / Head/Neck Vessel
Group 13: Lower Limb Vascular	A: Endarterectomy Lower Limb Arteries B: Vascular Shunt/Bypass / Limb

Group 14: Other Aortic Vascular	A: Resect/Replace Thoracic Vessel / Aneurysm B: Aorta-Iliac-Femoral Bypass C: Resect/Replace Abd Aorta / Aneurysm
Group 15: Other Gastrointestinal	A: Laparoscopic Gastroenterostomy / Morbid Obesity B: Laparoscopic Gastric Restrictive Proc / Morbid Obesity C: Lap Esophagogastric Sphincter Repair / Esophagitis
Group 16: Colorectal	A: Colon Procedure / Malignancy B: Colon Procedure / Benign Neoplasm C: Colorectal Procedure / Diverticulitis/osis D: Anterior Rectal Resection / Malignancy E: Rectal Resection / Malignancy
Group 17: Lap Cholecystectomy	A: Laparoscopic Cholecystectomy
Group 18: Cholecystectomy	A: Cholecystectomy
Group 19: Hernia Graft / Prosthesis	A: Unilateral Repair Inguinal Hernia with Graft/Prosthesis B: Repair Umbilical Hernia with Prosthesis C: Repair Incisional Hernia with Prosthesis
Group 20: Renal	A: Partial Nephrectomy B: Nephroureterectomy
Group 21: Prostate	A: Transurethral Prostatectomy B: Suprapubic Prostatectomy C: Retropubic Prostatectomy D: Radical Prostatectomy
Group 22: Hysterectomy	A: Subtotal Abdominal Hysterectomy / not Ovarian/Tubal Malig B: Total Abdominal Hysterectomy / not Ovarian/Tubal Malig C: Radical Abdominal Hysterectomy / not Ovarian/Tubal Malig
Group 23: Cystocele & Rectocele	A: Repair of Cystocele and Rectocele B: Repair of Cystocele C: Repair of Rectocele
Group 24: Breast	A: Local Excise Breast Lesion B: Subtotal Mastectomy C: Bilateral Reduction Mammoplasty D: Unilateral Simple Mastectomy

- E: Bilateral Simple Mastectomy
- F: Extended Simple Mastectomy
- G: Total Breast Reconstruction
- H: Mammoplasty/not Reduction
- I: Insert Breast Tissue Expander

Group 25: Vertebra

- A: Excise Intervertebral Disc / Cerv Spondylosis/no Myelopathy
- B: Excise Intervertebral Disc / Cerv Spondylosis/Myelopathy
- C: Excise Intervertebral Disc / Degeneration Cervical Disc
- D: Excise Intervertebral Disc / Degen/Displace Thoracic Disc
- E: Excise Intervertebral Disc / Disorder Lumbar Disc/no Myelopathy
- F: Excise Intervertebral Disc / Disorder Lumbar Disc/Myelopathy
- G: Excise Intervertebral Disc / Lumbosacral Spondylosis
- H: Cervical Fusion/Anterior Technique / Spond/no Myelopathy
- I: Cervical Fusion/Anterior Technique / Spond/Myelopathy
- J: Cervical Fusion/Anterior Technique / Degen/Stenosis
- K: Cervical Fusion/Posterior Technique / Spond/no Myelopathy
- L: Cervical Fusion/Posterior Technique / Spond/Myelopathy
- M: Cervical Fusion/Posterior Technique / Degen/Stenosis
- N: Lumbar/Lumbosacral Fusion/Anterior Technique
- O: Lumbar/Lumbosacral Fusion/Lateral Transverse Technique
- P: Lumbar/Lumbosacral Fusion/Posterior Technique
- Q: Lumbar/Lumbosacral Refusion/Posterior Technique
- R: Fusion/Refusion 4-8 Vertebrae / Cervical
- S: Fusion/Refusion 4-8 Vertebrae / Lumbosacral

Group 26: Hip Replacement

- A: Total Hip Replacement

Group 27: Knee Replacement

- A: Total Knee Replacement

### 3. Patient Experience

Our Patient Experience Ratings are based on survey data collected by the Federal Government's Centers for Medicare & Medicaid Services (CMS). Hospital CAHPS, or HCAHPS, is a more recent addition to the Consumer Assessment of Healthcare Providers and Systems (CAHPS) family of surveys administered by CMS. HCAHPS measures dimensions of patient care that are important to consumers (e.g. how often the room and bathroom were kept clean; how often pain was well-controlled) and that are related to safety concerns (e.g. communication about new medications, communication about discharge). For example, the average hospital patient receives 10 different drugs, some of which might look similar or have names that sound alike, and may be prescribed by different specialists who may not communicate well with each other. In fact, the Institute of Medicine estimates that, on average, there is at least one medication error per day for every patient.<sup>17</sup> Studies have shown that pain is often not controlled well after surgery, and that uncontrolled pain increases the risk of long hospital stays and reduced quality of life.<sup>18,19</sup> The importance of proper discharge instructions is underscored by a report that found that more than a third of hospital patients fail to get needed follow-up care.<sup>20</sup> Most hospitals are currently required to report HCAHPS data to receive full payment from Medicare.<sup>21</sup> In 2013, Medicare's Hospital Value-Based Purchasing Program will make incentive payments to hospitals based on their performance on specific quality measures, including HCAHPS.<sup>22</sup>

#### The data

HCAHPS survey data are collected using a standardized survey instrument by CMS-approved and trained vendors contracted by individual hospitals (in rare occasions, the hospital serves as the approved vendor itself). Data are delivered to a centralized data bank, where they are analyzed and prepared for public reporting on CMS's Hospital Compare website ([www.hospitalcompare.hhs.gov](http://www.hospitalcompare.hhs.gov)).

The survey asks a sample of former inpatients from each hospital about various dimensions of their experiences. CMS makes HCAHPS survey results available for nine categories, some of which are composites of more than one survey question. Our Ratings of patient experience are based on these nine categories, which are shown in the table in Appendix B (page 32).

We only present Patient Experience Ratings for hospitals with at least 100 completed surveys in the most recent 12 month period; smaller samples do not produce reliable Ratings. The number of completed surveys is not the same as the number of responses to individual survey items. While most items have response rates in the range of 90-95 percent of completed surveys, a few items do not apply to all patients (e.g. pain management and information about new medications), and

---

<sup>17</sup> <http://www.iom.edu/Reports/2006/Preventing-Medication-Errors-Quality-Chasm-Series.aspx>

<sup>18</sup> Morrison et al, *Pain*. 2003;103(3):303-11.

<sup>19</sup> Sinatra, R. *Pain Medicine*. 2010; 11: 1859-1871.

<sup>20</sup> Moore et al., *Archives of Internal Medicine*. 2007; 167(12), 1305-1311.

<sup>21</sup> [www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/HospitalRHQDAPU200808.pdf](http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/HospitalRHQDAPU200808.pdf)

<sup>22</sup> [www.gpo.gov/fdsys/pkg/FR-2011-05-06/pdf/2011-10568.pdf](http://www.gpo.gov/fdsys/pkg/FR-2011-05-06/pdf/2011-10568.pdf)

have response rates as low as 65 percent of completed surveys. Individual item response rates or sample sizes are not available.

### Assigning Ratings scores

For the measures with response options of Always/Usually/Sometimes/Never, we calculated the percentage of “always” or “usually” responses (e.g. 92 percent of respondents reported that their doctors always or usually communicated well) as the sum of the “always” and “usually” percentages reported by CMS. For discharge planning, we used the percentage of patients who said they were given instructions on what to do during their recovery at home.

For each of the first 8 measures, percentages are converted to fractional blobs using a piecewise linear transformation that assigns 100% a fractional blob of 5.5 and 75% a fractional blob of 0.5. Rates less than 75% are assigned a fractional blob of 0.5 and a blob score of 1. These fractional blobs are then rounded to the nearest whole number to create our blob scores; a fractional blob of 5.5 is assigned a blob score of 5. This leads to the scores shown in the following table:

	Patient Experience Rating	Adjusted percentage response
Better	●	95% - 100%
	◐	90% - 94%
↕	○	85% - 89%
	◑	80% - 84%
Worse	●	79% or below

### Overall Patient Experience

We calculate our Overall Patient Experience Rating in two stages. First, we calculate the arithmetic mean of the two overall response measures:

- The percentage of respondents who would “definitely” recommend the hospital
- The percentage of respondents who gave the hospital an overall rating of 9 or 10

These two measures are highly correlated ( $r=0.98$  for all hospitals with at least 100 completed surveys). We then transform this mean to fractional blobs (FBs) using the piecewise linear transformation that maps 100% to a FB of 5.5 and 40% to a FB of 1.5; and 40% to a FB of 1.5 and 0% to a FB of 0.5. These FBs are then rounded to blob scores, with a FB of 5.5 being assigned a blob score of 5. These transformations lead to the following ranges of scores:

	Overall Patient Experience Rating	Mean of two overall HCAHPS questions
Better	●	85% - 100%
	◐	70% - 84%
↕	○	55% - 69%
	◑	40% - 54%
Worse	●	39% or below

### **Limitations**

The survey tool and methods of data collection have been carefully researched and validated. However, unlike some other Consumer Reports Ratings, we do not collect these data ourselves, and so the actual implementation of the data collection and analysis is not in our control. We rely on the CMS, who oversees all aspects of the survey, to train hospitals and vendors in how to collect the data, to investigate how the survey is actually implemented for each hospital, and to analyze the data that we then convert into our unique Ratings format.

Data collection is decentralized—in part to accommodate the legacy of data already collected by hospitals from patients—which gives hospitals the ability to continue asking additional questions not in HCAHPS or to tailor additional questions to their specific quality improvement efforts. (If they do include additional questions on the survey, CMS requires the HCAHPS items to appear first, to reduce the chance of response bias from the other questions.) This decision is also related to cost—hospitals pay for or conduct the data collection themselves and this allows them to piggyback objectives.

To achieve standardization, CMS, the Health Services Advisory Group, and the National Committee for Quality Assurance provide detailed survey administration requirements in the HCAHPS instruction manual (Quality Assurance Guidelines, V4.0, available at [www.hcahponline.org](http://www.hcahponline.org)), training programs, site visits, independent data audits and analyses, and vendor certification processes ([www.hcahponline.org/qaguidelines.aspx](http://www.hcahponline.org/qaguidelines.aspx)).

The array of survey vendors involved in data collection introduces another set of concerns. While vendors are required to follow a strictly outlined set of procedures, there may be some inconsistencies in survey administration of which we are unaware, and over which we have no control. We do not provide Patient Experience Ratings for hospitals that are identified by CMS to have discrepancies in their data collection processes.

Finally, the Consumer Reports Health Ratings Center was only allowed access (by CMS) to the summarized results of their data analysis, preventing us from validating the data calculations or presenting data to you in alternative ways.

Despite these limitations, after our comprehensive review of the CMS survey methodology, we are confident that their stated methodologies are valid and reliable, and provide important information that allows comparison of patients' experiences in different hospitals on a common set of measures.

**Appendix B: HCAHPS survey questions that comprise each Ratings category**

Category	Response type	Survey questions
Communication about discharge	Yes/no	During this hospital stay, did doctors, nurses or other hospital staff talk with you about whether you would have the help you needed when you left the hospital? During this hospital stay, did you get information in writing about what symptoms or health problems to look out for after you left the hospital?
Communication about medications	Always Usually Sometimes Never	Before giving you any new medicine, how often did hospital staff tell you what the medicine was for? Before giving you any new medicine, how often did hospital staff describe possible side effects in a way you could understand?
Doctor-patient communication	Always Usually Sometimes Never	During this hospital stay, how often did doctors treat you with courtesy and respect? During this hospital stay, how often did doctors listen carefully to you? During this hospital stay, how often did doctors explain things in a way you could understand?
Nurse-patient communication	Always Usually Sometimes Never	During this hospital stay, how often did nurses treat you with courtesy and respect? During this hospital stay, how often did nurses listen carefully to you? During this hospital stay, how often did nurses explain things in a way you could understand?
Getting help	Always Usually Sometimes Never	During this hospital stay, after you pressed the call button, how often did you get help as soon as you wanted it? How often did you get help in getting to the bathroom or in using a bedpan as soon as you wanted?
Controlling pain	Always Usually Sometimes Never	During this hospital stay, how often was your pain well controlled? During this hospital stay, how often did the hospital staff do everything they could to help you with your pain?
Keeping room clean	Always Usually Sometimes Never	During this hospital stay, how often were your room and bathroom kept clean?
Keeping room quiet	Always Usually Sometimes Never	During this hospital stay, how often was the area around your room quiet at night?
Overall patient experience	Definitely yes Probably yes Probably no Definitely no  0-10	Would you recommend this hospital to your friends and family?  Using any number from 0 to 10, where 0 is the worst hospital possible and 10 is the best hospital possible, what number would you use to rate this hospital during your stay?

## **4. Hospital Practices**

## 4.1 Appropriate Use of Abdominal and Chest CT Scanning

Scanning data are reported by the Centers for Medicare and Medicaid Services (CMS) on their Hospital Compare website ([www.hospitalcompare.hhs.gov](http://www.hospitalcompare.hhs.gov)). Data reflect a hospital's performance for a one-year period, with generally an 18-month time lag from the end of the measurement period, and are updated annually.

There are six measures in the Hospital Compare dataset: (1) percentage of outpatients who had an MRI of the Lumbar Spine with a diagnosis of low back pain without evidence of antecedent conservative therapy; (2) percentage of outpatients with mammography screening studies that receive further screening studies (mammography or ultrasound) within 45 days; (3) the percent of all outpatient CT scans of the abdomen that are performed twice, once with contrast and one without; (4) the percent of all outpatient CT scans of the thorax or chest that are performed twice, once with contrast and one without; (5) the percent of outpatients who got cardiac imaging stress test before low-risk outpatient surgery; and (6) the percent of outpatients with brain CT scans who received a sinus CT scan at the same time. The last two of these measures (cardiac and brain) were reported for the first time in October 2012.

We currently use the two measures representing double scan rates for abdomen and chest because these represent the risk of elevated exposure to additional and unnecessary radiation.

A computerized tomography (CT) scan uses X-rays to produce detailed images inside the body. Before some CT scans, a "contrast" substance is either swallowed, or injected into a patient's vein to help make features of the body stand out more clearly. Combination or double CT scans are those scans in which a patient receives two CT scans—one scan without contrast followed by another scan with contrast.

Use of double scans exposes patients to double the radiation of a single scan. For example, radiation exposure from a single CT scan of the chest is about 350 times higher than for an ordinary chest X-ray; a double CT scan exposes a patient to 700 times more radiation than a chest X-ray. A single CT scan of the abdomen exposes the patient to 11 times more radiation than an X-ray of the abdomen; a double scan exposes the patient to 22 times more radiation than an abdominal X-ray. Additionally, the use of contrast material introduces risks of its own, such as possible harm to the kidneys or allergic reactions. Although double CT scans may be appropriate for some parts of the body and some medical conditions, they are usually not appropriate for scans of the chest or abdomen.<sup>23</sup>

### The data

These measures reflect scans on outpatients in medical imaging facilities that are part of a hospital or associated with a hospital. Data are based on Medicare claims data, and consequently represent scans only on Medicare patients. Data are not risk-adjusted, and are calculated as raw/observed rates after the exclusion and inclusion criteria are applied.

---

<sup>23</sup> [http://www.hospitalcompare.hhs.gov/staticpages/learn/importance\\_quality.aspx?measurecd=OPIMAGE](http://www.hospitalcompare.hhs.gov/staticpages/learn/importance_quality.aspx?measurecd=OPIMAGE)

## Assigning Ratings scores

We used the double-scan rates for chest and abdomen in our Ratings. To convert these rates to our fractional blob (FB) scale, we used a piecewise linear transformation that assigns a rate of zero to a fractional blob of 5.5, and a rate of 25% to a fractional blob of 0.5. Rates greater than 25% are assigned FBs of 0.5. These values are then rounded to whole number blobs for presentation (5.5 is assigned a blob score of 5).

This transformation corresponds to the Ratings scores shown in the table below.

	Rating Score	Range of double scanning rates
Better	●	Rate ≤ 5%
	◐	5% < rate ≤ 10%
↑	○	10% < rate ≤ 15%
	◑	15% < rate ≤ 20%
↓	●	20% < rate
Worse	●	20% < rate

## Limitations

These data come from billing and other administrative data submitted by hospitals to Medicare. Such records were intended to capture information for billing purposes, but they contain significant details about a patient's health status and services rendered in their outpatient encounter.

These data also reflect outcomes only for Medicare patients, though we believe they provide a good indication of scanning rates overall. Ratings come from the most recent data available, but there is a time lag in reporting these data to the public. It is possible that performance today will show improvements or declines in data that is not currently available to us.

## 4.2 Avoiding C-sections

According to the CDC, C-section rates have been rising dramatically since 1999<sup>24</sup> and have increased more than 500 percent since 1970 (total C-section rate in 1970 was 5% compared with the 2012 average of 32.8 percent).<sup>25</sup> While it is not known what the "best" C-section rate is, but there is broad agreement that current average C-section rates are too high.<sup>26</sup> In addition, The American College of Obstetricians and Gynecologists has recently released guidelines intended to reduce C-sections that are not medically needed.<sup>27</sup> Currently, there is no requirement to publicly report C-section data. The Joint Commission, however, is now collecting C-section rates from hospitals with more than 1,100 births per year and this data may eventually be reported to the public at the hospital level.

### The data

As nation-wide data on C-sections is currently not available to the public, state in-patient administrative claims data over a two-year period from 19 states was used to calculate the AHRQ IQI 33 primary C-section measure, and IQI 33 calculated values were gathered from three additional states (CA, KY, TX), for a total of over 1,500 hospitals. See the chart below for details on the data sources for each state.

State(s)	Years	Data Source
AZ,CO,FL,IA,IL,MA, MD,NC,NJ,NV, NY, OR,PA,RI,UT,VA,WA,WI	2010- 2011	State in-patient databases <sup>28,29</sup>
CA	2011- 2012	<a href="http://www.oshpd.ca.gov/HID/Products/PatDischargeData/ResearchReports/HospIPQualInd/Vol-Util_IndicatorsRpt/">http://www.oshpd.ca.gov/HID/Products/PatDischargeData/ResearchReports/HospIPQualInd/Vol-Util_IndicatorsRpt/</a>
KY	2010- 2011	<a href="http://chfs.ky.gov/ohp/healthdata/IQI.htm">http://chfs.ky.gov/ohp/healthdata/IQI.htm</a>

<sup>24</sup> [http://www.cdc.gov/nchs/data/nvsr/nvsr63/nvsr63\\_01.pdf](http://www.cdc.gov/nchs/data/nvsr/nvsr63/nvsr63_01.pdf)

<sup>25</sup> [http://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62\\_09.pdf#table21](http://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62_09.pdf#table21)

<sup>26</sup> Ye J, et al. Birth Issues in Perinatal Care, April 2014. <http://onlinelibrary.wiley.com/doi/10.1111/birt.12104/abstract>

<sup>27</sup> [http://www.acog.org/Resources\\_And\\_Publications/Obstetric\\_Care\\_Consensus\\_Series/Safe\\_Prevention\\_of\\_the\\_Primary\\_Cesarean\\_Delivery](http://www.acog.org/Resources_And_Publications/Obstetric_Care_Consensus_Series/Safe_Prevention_of_the_Primary_Cesarean_Delivery)

<sup>28</sup> Calculated IQI 33 rates are from state inpatient claims databases, courtesy of Cigna healthcare (except for CA, KY and TX). This data represents all of the state's inpatient claims that are made available to the public and does not represent only Cigna's covered population.

<sup>29</sup> The Cigna maternity hospital quality information is provided as-is and without any express or implied warranties of any kind. This information is not a guarantee or indication of the quality of care that may actually be delivered. Doctors and health care facilities are independent contractors and not agents of Cigna. "Cigna" is a registered service mark, and the "Tree of Life" logo is a service mark, of Cigna Intellectual Property, Inc., licensed for use by Cigna Corporation and its operating subsidiaries. All products and services are provided by or through such operating subsidiaries, including Connecticut General Life Insurance Company and Cigna Health and Life Insurance Company, and not by Cigna Corporation. © 2014 Cigna.

TX	2010-2011	<a href="https://www.dshs.state.tx.us/thcic/publications/hospitals/IQIReport/Indicators-of-Inpatient-Care-in-Texas-Hospitals-2011.shtm">https://www.dshs.state.tx.us/thcic/publications/hospitals/IQIReport/Indicators-of-Inpatient-Care-in-Texas-Hospitals-2011.shtm</a>
VT	2009-2010	State in-patient databases <sup>29,30</sup>

The measure includes all mothers (not just first-time mothers), and excludes all women who have previously had a C-section. It also excludes all deliveries with multiple fetuses, as well as all premature births, and deliveries with an abnormal presentation (e.g. breech and others). Please refer to the AHRQ measure specification for more details about this measure.<sup>30</sup> Online, we publish each hospital's primary C-section rate (as specified by IQI 33) as a percentage, as well as the assigned Rating, developed as described below.

### Assigning Ratings scores

Hospitals need a minimum of 100 qualified deliveries, exclusions described above, in order to receive a Rating. Hospitals that do not have two years of data available, are identified as a children's hospital in AHA, do not publically report their data, data are not reported in a usable format, or hospitals with insufficient data were not Rated. The C-section rates are rescaled using a piecewise linear transformation as described in the chart below and assign ratings on our "better to worse" scale. Hospitals with rates less than 5 receive a FB of 5.5 and hospitals with rates above 30 receive a FB of .5. Cut points for the blobs are based on published evidence, as well as input and review by experts in quality measurement and clinical medicine. The national average rate for IQI33 in 2000<sup>31</sup> (the beginning of the most recent increase in C-section rates), which was 12.6, was used to anchor the middle of our Ratings scale (○). This measure is highly correlated with total C-section rates, and with C-section rates that are calculated using an age-adjusted measure (PC-02)<sup>32</sup>.

	Rating	Fractional Blob Range	Range of IQI 33 rates
Better	●	$5.5 \geq FB \geq 4.5$	5-9.5
	◐	$4.5 > FB \geq 3.5$	>9.5-11.5
↓	○	$3.5 > FB \geq 2.5$	>11.5-15
	◑	$2.5 > FB \geq 1.5$	>15-21
Worse	●	$1.5 > FB \geq 0.5$	>21

<sup>30</sup><http://www.qualityindicators.ahrq.gov/Downloads/Modules/IQI/V45/TechSpecs/IQI%2033%20Primary%20Cesarean%20Delivery%20Rate%20Uncomplicated.pdf>

<sup>31</sup>[http://hcupnet.ahrq.gov/HCUPnet.jsp?Id=60777B3CBE9E1C79&Form=SellQIs1&JS=Y&Action=%3E%3ENext%3E%3E&\\_QITables=IQI33](http://hcupnet.ahrq.gov/HCUPnet.jsp?Id=60777B3CBE9E1C79&Form=SellQIs1&JS=Y&Action=%3E%3ENext%3E%3E&_QITables=IQI33)

<sup>32</sup><https://manual.jointcommission.org/releases/TJC2013A/MIF0167.html>

## Limitations

These data come from billing and other administrative data submitted by hospitals to each state. Such records were intended to capture information for billing purposes rather than patient outcomes, but they contain significant details about a patient's stay in the hospital.

Ratings come from recent data but it is possible that performance today will show improvements or declines in performance data that is not currently available to us.

To level the playing field, the measure controls for some things that affect C-section rates, such as not including multiple gestations and breech births. However, this measure does not account for all differences in patient characteristics (such as chronic illness) that might affect the C-section rates of an individual hospital.

This measure looks at C-section rates of all mothers that have not had a previous C-section -- it does not limit the measurement to first-time mothers. As first-time mothers have a higher risk of a C-section, rates that are calculated using data from only first-time moms (such as The Joint Commission's measure -PC-02) are higher than rates calculated with the IQI 33 method. However, as mentioned above, the rates calculated with these two different methods are highly correlated.

This measure does not assess patient outcomes following a C-section.

Looking at primary C-sections is just one dimension of how well a hospital does in maternity care. There are other measures that are emerging related to the quality of delivery and neonatal care that affect the health of the mother and newborn. Examples include neonatal infection, early elective delivery and obstetrical trauma during delivery. Consumer Reports will continue to monitor the development and availability of such measure results in the future.

## 5. Safety Score

We created a composite of measures related to hospital safety. While there are additional dimensions to hospital safety than those included here, these represent a broad range of safety factors that, combined, serve as an indicator of a hospital's commitment to the safety of their patients. We have deliberately not included dimensions about procedures a hospital can follow but that have not been shown to affect health outcomes for patients.

### The data

For the Safety Score, we use five major categories of safety-related measures, each with several components: avoiding infections, avoiding readmissions, communication about discharge and medications, appropriate use of scanning, and avoiding mortality. Details regarding the individual components of the Safety Score (including the limitations of the each) have been described earlier in this report; these sections are referenced below as appropriate.

**Avoiding infections** (see page 7): According to a recent study, hospital acquired infections affect about 650,000 patients each year; therefore on any given day, about one of every 25 hospitalized people are infected while they are in the hospital. About 12% of patients who are infected while in the hospital die in the hospital from the infections.<sup>33</sup> Hospital acquired infections are estimated to cost \$28 to \$45 billion dollars each year, in direct medical costs.<sup>34</sup> See our investigations on deadly hospital infections [[www.consumerreports.org/health/doctors-hospitals/hospital-infection/deadly-infections-hospitals-can-lower-the-danger/overview/deadly-infections-hospitals-can-lower-the-danger.htm](http://www.consumerreports.org/health/doctors-hospitals/hospital-infection/deadly-infections-hospitals-can-lower-the-danger/overview/deadly-infections-hospitals-can-lower-the-danger.htm)] for more information.

**Avoiding readmissions** (see page 13): In one study researchers found that almost one of every five Medicare patients was readmitted within 30 days of being released from the hospital and about one in three were readmitted within 90 days.<sup>35</sup> Unnecessary readmissions are tied to patient safety in several important ways.

First, any hospital admission has inherent risks. A November 2010 study by the Department of Health and Human Services' Office of the Inspector General calculated that infections, surgical mistakes, and other medical harm contribute to the deaths of 180,000 Medicare hospital patients a year, and that another 1.4 million are seriously hurt by their hospital care.<sup>36</sup> More recent estimates suggest that preventable harm contributes to the death of more than 440,000 people each year in hospitals across the United States<sup>37</sup>. Hence a second admission exposes the patient to additional safety risk.

Second, readmissions can be caused by things that go wrong in the initial discharge.<sup>38</sup> In fact, a national public-private initiative, Partnership for Patients, has set a performance target to decrease preventable complications during a transition from one care setting to another in order

---

<sup>33</sup> Magill et al., *New Engl J Med* 2014;370:1198-208

<sup>34</sup> [www.cdc.gov/HAI/pdfs/hai/Scott\\_CostPaper.pdf](http://www.cdc.gov/HAI/pdfs/hai/Scott_CostPaper.pdf)

<sup>35</sup> Jencks et al., *N Engl J Med* 2009; 360:1418-1428.

<sup>36</sup> <http://oig.hhs.gov/oei/reports/oei-06-09-00090.pdf>

<sup>37</sup> James, J, *J Patient Saf*;2013;9: 122-128.

<sup>38</sup> [www.ahrq.gov/qual/impptdis.htm](http://www.ahrq.gov/qual/impptdis.htm)

to reduce hospital readmissions by 20 percent in 2013, compared with 2010. It is estimated that hitting this target would result in 1.6 million fewer patients being readmitted to a hospital within 30 days.<sup>39</sup>

Third, we know that, to at least some extent, readmissions reflect errors in the initial hospitalization. For example, patients who develop hospital infections and other complications may end up being readmitted for further treatments.<sup>40</sup> In one study researchers found that patients who experienced specific complications were more likely to end up back in the hospital within a month than those who did not.<sup>41</sup>

**Avoiding mortality – medical (page 15) and surgical (see page 17):** Two mortality measures (30-day mortality for medical conditions and in-patient death of surgical cases who had serious complications) are included in our Safety Score. Recent estimates suggest that preventable medical harm contributes to the death of more than 440,000 people each year in hospitals across the United States.<sup>42</sup> Consumers also grossly underestimate the impact of preventable errors; in one study by the Kaiser Family Foundation, more than half of consumers who responded to a survey thought that preventable errors caused 5000 or fewer deaths each year<sup>43</sup>.

**Communication about medications and discharge (see page 29):** Two elements of the patient experience survey data—communication about new medication and communication about discharge instructions—are included in our Safety Score. Lack of communication about new medications can lead to misuse of medications or other medication errors. For example, when someone is admitted to the hospital they are likely to receive new medications. If the hospital-based physicians are not aware of the patient’s current medications there is the potential for inappropriate medications or doses to be prescribed. In fact, studies show that more than one-third of patients experience a medication error (such as omission of a required medication, an accidental duplication of a drug they were already taking, or the wrong dose of a medication) when they are admitted to the hospital.<sup>44</sup>

Lack of communication about discharge instructions can lead to errors in post-discharge care. Studies have shown that medication discrepancies (such as intentional or non-intentional non-compliance, conflicting information, duplication) occurred in 14 percent of Medicare-aged patients who were discharged from the hospital.<sup>45</sup> Patients may be discharged from the hospital without understanding the instructions for care after leaving the hospital, or may stop taking important medications that they need.

**Appropriate use of scanning (see page 34):** Double scans of the chest and abdomen are rarely necessary and unnecessarily expose patients to additional radiation; radiation from CT scans might contribute to an estimated 29,000 future cancers a year.<sup>46</sup> According to CMS, a single CT scan of

---

<sup>39</sup> [www.healthcare.gov/compare/partnership-for-patients/](http://www.healthcare.gov/compare/partnership-for-patients/)

<sup>40</sup> Emerson et al., *Infect Control Hosp Epidemiol.* 2012;33(6):539-544.

<sup>41</sup> Friedman et al, *Med Care.* 2009;47(5):583-590.

<sup>42</sup> James, J, *J Patient Saf* ;2013;9: 122-128.

<sup>43</sup> <http://kff.org/health-costs/poll-finding/national-survey-on-consumers-experiences-with-patient/>

<sup>44</sup> Gleason et al, *J Gen Intern Med.* 2010;25(5):441-447.

<sup>45</sup> Coleman et al., *Arch Intern Med.* 2005;165(16):1842-1847.

<sup>46</sup> Berrington de González et al, *Arch Intern Med* 2009;169(22):2071-2077.

the abdomen is 11 times higher than for an x-ray of the abdomen, and a double scan is therefore 22 times higher. A single CT scan of the chest is 350 times higher than a chest x-ray and a double scan is therefore 700 times higher.

The five components of the Safety Score (infections, readmission, mortality, communication, scanning) are equally weighted and scored on a 1-100 scale. Hospitals must have reported at least one component in each of the categories for us to calculate a Safety Score; there is no imputation of missing data. The data used in the calculation of the Safety Score are shown in the table below.

<b>Safety Score Category</b>	<b>Components</b>	<b>Data Source</b>	<b>Weight</b>
Avoiding infections (pages 7-12)	<ul style="list-style-type: none"> <li>• Central-line associated bloodstream infections</li> <li>• Surgical-site infections</li> <li>• Catheter urinary-tract infections</li> </ul>	CMS	20% of total based on combined CLABSI, SSI and CAUTI score; hospitals need sufficient data for the composite of the three infection measures. See page 10 for how the infection composite is calculated.
Avoiding readmissions (pages 13-14)	<ul style="list-style-type: none"> <li>• 30-day hospital-wide all-cause readmissions</li> </ul>	CMS	20% of total.
Avoiding mortality - (see pages 15-18 )	<ul style="list-style-type: none"> <li>• Medical: 30-day mortality for Heart attack, Heart failure, Pneumonia</li> <li>• Surgical - AHRQ PSI 4</li> </ul>	CMS	20% of total, half for each component (medical and surgical), or if only one is available it comprises the full mortality measure.
Communication (pages 29-32)	<ul style="list-style-type: none"> <li>• Communication about discharge instructions</li> <li>• Communication about new medications</li> </ul>	CMS	20% of total, half for each component (discharge and medications).
Appropriate use of scanning (pages 34-35)	<ul style="list-style-type: none"> <li>• Double chest CT scans</li> <li>• Double abdomen CT scans</li> </ul>	CMS	20% of total, half for each component (chest and abdomen), or if only one is available, it comprises the full scanning measure.

### Calculation of the Safety Score

The Safety Score is expressed on 100-point scale, where a hospital would score 100 if it earned the highest possible score in all measures (for example, 100% for patient experience measures, or zero infections), and would score 1 if it earned the lowest scores in all measures.

The measure categories that are based on interval data (infections, readmissions, communication, scanning, and mortality) and their components are expressed as fractional blobs (FBs), as described earlier in this document. Their components are combined into composites as follows:

1. Infections. A combined SIR is calculated and is then transformed to our FB scale using the methods described earlier (pages 7-12). If a hospital does not report one of the infection types then the weight for that type is zero. A hospital could have a combined SIR even if the individual infection measures alone do not have sufficient data for a Rating.
2. Readmissions is the calculated FB as described earlier.
3. Mortality is the mean of the FBs for mortality-medical and mortality-surgical (described on pages 15-18). If only one measure is available, the Mortality FB set to be equal to that measure's FB.
4. Communication is the mean of the FBs for Communication about Medications and Communication about Discharge.
5. Scanning is the mean of the Chest and Abdomen CT double scan FBs, if both measures are available. If only one measure is available, the Scanning FB set to be equal to that measure's FB.

The mean of the FBs for these five measure categories is then calculated using equal weights. That mean is linearly transformed to a scale from 0.5 to 100.5, so that these five measure categories combined account for 100% of the Safety Score.

### **Selecting weights**

We examined the impact of varying the weights of the five categories on the Safety Score and the rank order of hospitals. Several other weighting schemes we tried were also highly correlated with equal weights. Consequently, we chose to use equal weights.

### **Limitations**

Each of the categories and components are based on data and scoring methods that have limitations and weaknesses themselves. These are described in detail in the relevant sections of this report.

In addition, the component measures represent data collected in different time periods. In each case, we use the most current valid data available. The difference in time periods measured may be a limitation for hospitals looking to use these data for quality improvement. Composites are useful because they can make a complex set of data easier to understand. However, composites have their limitations. For example, hospitals that perform well on the composite do not necessarily perform well on all of the components of the composite, therefore we show consumers most of a hospital's individual Ratings on the hospital Report Card page. In addition, the composite we created for hospital safety was limited by the data that is currently available to the public.

## 6. Heart Surgery Ratings

For our heart surgery Ratings, we've teamed with the Society of Thoracic Surgeons (STS) to publish ratings of hospitals (and surgical groups) based on their performance data for heart bypass surgery and aortic heart valve replacement surgery. The Society of Thoracic Surgeons (STS) rates hospitals using standardized measures endorsed by the National Quality Forum, a nonprofit organization that has established national health-care standards for performance improvement. Using this information, consumers can see how hospitals and surgical groups compare with national benchmarks for overall performance, survival, complications, and other measures.

STS is a nonprofit organization that represents some 5,400 surgeons worldwide who operate on the thorax, or chest. Developed in 1989, the STS National Adult Cardiac Surgery Database is the largest single-specialty database in the United States, containing more than 5 million surgical records. Participating hospitals and groups add data four times a year, providing an up-to-date picture of their surgical practice. Much of the information is collected at the point of care, which has advantages over data collected for administrative or insurance reasons.

STS contracts with an independent organization, the Duke Clinical Research Institute, to analyze the data and prepare reports for participating hospitals and surgical groups, comparing their performance with national benchmarks for surgical quality. STS, hospital administrators, and surgeons from each hospital have agreed to share the reports on heart surgery with Consumer Reports as part of their ongoing commitment to improving care and helping patients make informed decisions.

Approximately 95% of the 1100 heart surgery programs in the United States are part of the STS Adult Cardiac Thoracic Database. As of June 2014, about 400 hospitals volunteered to publish their performance data for heart bypass, and aortic valve replacement surgery, publicly through Consumer Reports.

### Heart Bypass Surgery (CABG) Ratings

A hospital's rating in this measure reflects its performance in isolated CABG operations, meaning that the patient is having only that surgery, not a combination procedure. A hospital's overall score is a composite of four separate measures. Two of them—recommended medications and optimal surgical technique—reflect how well surgeons adhere to the best-established practices. The other two—patient survival and the absence of surgical complications—reflect how their patients fare.

- 1. Patient survival.** This is based on the chance that a patient will both survive at least 30 days after the surgery and be discharged from the hospital.
- 2. Absence of surgical complications.** This is based on the chance that a patient will not experience any of these five serious complications of heart-bypass surgery during their hospitalization: extended breathing support on a ventilator, an infection in the

breastbone incision, kidney failure, a stroke, or a repeat operation for postoperative bleeding or other causes.

- 3. Recommended medications.** This is based on the chance that a patient will get all of the following prescriptions: a beta-blocker before and after the procedure to prevent an abnormal heart rhythm and control blood pressure; and aspirin to prevent blood clots, and a statin or other medication to lower LDL (bad) cholesterol afterward.
- 4. Optimal surgical technique.** This is based on the chance that a patient will receive at least one graft involving an internal mammary artery, which run under the breastbone. Such grafts improve long-term survival compared with grafts taken from veins, in part because they are more resistant to cholesterol buildup and can withstand the high pressure in the heart better.

For each of the four CABG measures as well as for the overall star Rating, STS compares a hospital’s performance with the average performance of all the hospitals in their database, and assigns a “star” rating. For survival and complications, the results are statistically adjusted for the overall health of a hospital’s patients, since some hospitals treat older or sicker patients than others. (That adjustment is not necessary for medications and surgical technique, however, because the right drugs and best surgical approaches should be used with all eligible patients regardless of their health.) The star Rating calculation begins by assuming all providers are average and then determines statistically if there is at least a 99 percent probability that the performance of any specific hospital is lower than average (one star) or higher than average (three star). For the several years that the STS has been calculating these scores, about 10-15 percent of all Adult Cardiac Surgery Database participants have been one star, about 10-15 percent have been three -star, and the remainder have been two -star, or average programs.

STS Star Rating	Consumer Reports Rating	Definition
3-star		Above average
2-star		Average
1-star		Below average

## Aortic Valve Replacement Ratings

A hospital’s overall score for aortic valve replacement (AVR) is a composite of two separate measures of patient outcomes.

- **Patient survival.** This is based on the chance that a patient will both survive at least 30 days after the surgery and be discharged from the hospital.
- **Absence of surgical complications.** This is based on the chance that a patient will not experience any of these five serious complications of heart-bypass surgery during their hospitalization: extended breathing support on a ventilator, an infection in the breastbone incision, kidney failure, a stroke, or a repeat operation for postoperative bleeding or other causes.

For both AVR measures and for the overall star Rating, STS compares a hospital’s performance with the average performance of all the hospitals in their database. The results are statistically adjusted for the overall health of a hospital’s patients, since some hospitals treat older or sicker patients than others. The overall valve replacement Rating combines the scores from the two measures. For each individual performance measure, as well as the overall Rating, hospitals that score significantly above average get three stars, average performers get two stars, and those that score significantly below average receive one star. Stars are then converted to blob-scores as follows:

STS Star Rating	Consumer Reports Rating	Definition
3-star	●	Above average
2-star	○	Average
1-star	●	Below average

### Limitations

The Ratings are currently limited to hospitals that voluntarily agree to participate in the STS database, and then agree to release the data to us. Even though survival and complications are statistically adjusted for how sick a hospital’s patients are, other factors might have an impact on the differences between groups. That, together with other statistical issues, might sometimes

make it difficult to compare hospitals directly. While these Ratings are based on the most comprehensive set of data available for heart surgery, some important quality measures might not be included in the STS ratings. Some of the measures are difficult to define precisely, so differences might exist in how hospitals collect and report their data. The percentages reported are not exact numbers but estimates based on the statistical model used, and have some a margin of error. Hospitals that do a relatively small number of isolated heart operations are statistically harder to differentiate from average than those that do a larger number of them. So hospitals with fewer operations are more likely to get an average, or two-star, Rating.