



How We Rate Hospitals

July 2013

	<u>Page</u>
1. Overview	2
2. Patient Outcomes	
2.1 Avoiding central-line bloodstream and surgical site infections	4
Appendix A: Hospital-specific public reporting of ICU bloodstream infection and surgical-site infection data	11
Appendix B: NHSN rates used for standardization of CLABSI rates	16
Appendix C: NHSN rates used for standardization of SSI rates, all risk levels combined	17
2.2 Avoiding readmissions	18
2.3 Avoiding complications	20
2.4 Avoiding adverse events following common scheduled surgeries (Surgery Ratings)	22
Appendix D: Surgical Procedures in each Surgical Group	29
3. Patient Experience	32
Appendix E: HCAHPS survey questions that comprise each Ratings category. . .	35
4. Hospital Practices	
4.1 Use of electronic health records	36
4.2 Appropriate use of abdominal and chest scanning	38
5. Safety Score	40
6 Safety Ratings in Consumer Reports magazine, August 2012	44

1. Overview

Consumer Reports' hospital Ratings (<http://www.consumerreports.org/health/doctors-hospitals/hospital-ratings.htm>) include measures of Patient Outcomes (avoiding infections, readmissions, complications, 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, and use of electronic health records). 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), state Departments of Health, 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 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 new quality 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.

Category	Measures	Source	Dates covered by data in our July 2013 update
Safety Score composite	(denoted with * below)		Varied; see below
Patient Outcomes	*Avoiding bloodstream infections *Avoiding surgical site infections	CMS; Various states (see page 11)	Varied; see Appendix A page 11
	*Avoiding readmissions	CMS	July 2008 – June 2011
	*Avoiding complications	CMS	July 2009 – June 2011
	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	January – December 2011
Hospital practices	Use of electronic health records	AHA	January – December 2010
	*Appropriate use of abdominal scanning *Appropriate use of chest scanning	CMS	January – December 2010

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-ratings/overview/index.htm) that perform coronary artery bypass (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.

2. Patient Outcomes

2.1 Avoiding Central-Line Associated Bloodstream and Surgical-Site Infections

Some healthcare-acquired infections data are reported by the Centers for Medicare and Medicaid Services (CMS), an agency of the Federal government. Most hospitals report data on Central Line Associated Bloodstream Infections (CLABSIs) to the Centers for Disease Control and Prevention (CDC); those data are then publicly reported on CMS's Hospital Compare website (hospitalcompare.hhs.gov). Beginning in 2011, reporting of select hospital-acquired infections (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 this annual payment increase; this payment structure is used as an incentive that causes virtually all of these hospitals to report. In addition, more than half of states currently have laws requiring the public reporting of data on HAIs. Reporting in these states is mandated and the measures vary somewhat from state to state, but almost all of the reporting states require hospitals to submit information about central-line associated bloodstream infections (CLABSIs) in select types of intensive care units and selected surgical site infections (SSIs).

Central-line associated bloodstream infections (CLABSI) data

For the first time, we are using intensive-care unit (ICU) CLABSI data that are publicly reported through CMS. For hospitals that do not have data available through CMS (primarily children's, critical access, and Veterans Administration hospitals), we use data reported by their states when available. For each hospital, we use the data that are most current for our Ratings (CMS or state reports).

Hospitals that report data to CMS are required to do so quarterly for every ICU in the hospital, for all ICU patients (not just Medicare patients). Data are combined for four quarters, with a delay of approximately 10 months (so for example, the CLABSI data reported by CMS in October 2012 reflects the 12-month period ending in December 2011). For the other reporting agencies (state health departments), the particular ICU types on which they report and the reporting periods vary (see Appendix A, beginning on page 11). While some of these agencies report ICU-specific rates or other measures, most of them report both the total number of CLABSIs contracted by ICU patients at each hospital, and the total number of days the hospital's ICU patients spent with central lines in place, a measure called central line-days, or CLD. Not all hospitals in a state have, or report on, all ICUs. Several states simply report the total number of CLABSIs and CLDs totaled over all ICUs, without identifying individual ICU types.

There are few public state reports that include neonatal, burn or trauma ICUs, presumably because of the complicated nature of managing lines in these ICUs and the small numbers of hospitals in which the latter two exist; when data are reported separately for any of these three types of ICUs, we do not include them in our calculations.

Surgical-site infections data

Several states publicly report data on one or more surgical-site infections in standardized ways we can analyze including different combinations of surgeries; some additional states report SSIs but not in a format we can report (see Appendix A starting on page 11).

Each state that publicly reports surgical-site infections includes two data elements for each surgical procedure they cover: the total number of surgical procedures conducted by each hospital, and the total number of infections patients contracted at the sites of those surgeries. Each state reporting agency reports data for a set of surgical procedures (see Appendix A), and a reporting period (often one year, but some states report quarterly or for other periods), which are not consistent across states. Not all hospitals conduct, or report on, all surgical procedures.

An infection is considered to be a surgical site infection if it occurs within 30 days of the surgery, or within one year if an implant is in place and the infection appears to be related to the surgery.

The basis of the Ratings: The standardized infection ratio

For each hospital, we calculate the standardized infection ratio, a measure developed by the Centers for Disease Control and Prevention (CDC) and modeled after the standardized mortality ratio (or standardized incidence ratio), a common measure in epidemiology. This measure compares data within each of several subgroups to national (or some other benchmark) data for those subgroups, and then pools the comparisons across all subgroups. The standardized infection ratio (SIR) is calculated separately for ICU bloodstream infections, standardized across ICUs, and for surgical-site infections, standardized across type of surgical procedure.

In the case of CLABSI data, the standardized infection ratio compares the number of infections reported for each individual ICU to national infection rates for that type of ICU (for example, surgical ICU or medical ICU). 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).¹ The benchmark rates are composite data from approximately 1,500 hospitals in 2006-2008 in 48 states and the District of Columbia; NHSN data for individual hospitals is publicly available through CMS. While more recent national averages are available from NHSN, like CDC, we have continued to use the data from 2006-2008 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 of ICUs, 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 data, 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.

A Standardized Infection Ratio of 1 means that the hospital's ICUs produce CLABSIs 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.

While CMS reports the SIR directly as calculated by CDC, some of the states that publicly report infection rates include various mixes of ICUs. In some cases we calculated the pooled rate for more than one NHSN ICU. For example, NHSN national data are reported separately for hospitals with three types of medical/surgical ICUs: those in major teaching hospitals, hospitals with more than 15 ICU beds, and those in other hospitals with 15 or fewer beds. For data from individual states, we generally do not have access to reliable information about hospitals' teaching status nor the number of beds; CDC determines this information from questions hospitals answer when they enter data into NHSN. We pooled the published data for these three strata of medical/surgical ICUs, and for medical ICUs for Major teaching and all other hospitals. The national rates we use, calculated from the published data from NHSN, are shown in Appendix B (page 16).

Standardized infection ratios are also calculated for surgical-site infections (SSIs), following a similar process. Since each state that publicly reports SSIs covers a potentially different set of surgical procedures (Appendix A), the standardized infection ratio compares the number of infections reported for each reported surgical procedure to national infection rates for that procedure type. While a small number of states report SSIs separately by risk category ranging from 0 (low risk) to 3 (high risk), we do not include those risk categories in this standardization, in order to maintain consistency across states. As with bloodstream infections, national data are derived from data reported to the National Healthcare Safety Network (NHSN) a data repository supported by the CDC. The national rates we used, calculated from the published data from NHSN, are in Appendix C (page 17).

¹ 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.

Assigning Ratings scores

We use the most current data available. For CLABSI data, if data are available from both CMS and state data and the dates are the same, we use CMS data. For both CLABSI and SSI data, we only use state data that is from January 2008 or later.

We report the number of infections and the number of central-line days (or surgical procedures for SSI) for any hospital that reported at least 600 CLD (or at least 50 surgical procedures for SSI), or at least one infection regardless of central-line days (or surgical procedures). In addition, we provide 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 3 infections, regardless of CLD or number of surgical procedures. 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 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 in at least 1,000 central-line days for CLABSI, or in 100 procedures for surgical-site infections. Although the standardized infection ratio 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 patients. 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 both CLABSI and surgical-site infections, we assign fractional blobs on a scale from 0.5 to 5.5 using a piecewise linear transformation as follows:²

- a. If the $SIR = 0$, then the hospital is assigned a fractional blob value of 5.5 and a blob score of 5. Only hospitals with zero reported infections ($SIR = 0$) can receive a blob score of 5.
- b. If $0 \leq SIR \leq 0.5$, then the fractional blob 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 4.
- c. If $0.5 < SIR \leq 1.0$, then the fractional blob 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 3.

² 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.

- d. If $1.0 < SIR \leq 2.0$, then the fractional blob 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 2.
- e. If $2.0 < SIR \leq 4.0$, then the fractional blob 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 1.
- f. For $SIR > 4$, the FB is set to be equal to 0.5, with a blob score of 1.

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

	Bloodstream Infection or Surgical-Site Infection Rating	Fractional Blob Range	Standardized Infection ratio range	Interpretation
Better	5-blob	FB = 5.5	SIR = 0	0 infections
	4-blob	$4.5 > FB \geq 3.5$	$0.0 < SIR \leq 0.5$	At least 50% better than average
↑ ↓	3-blob	$3.5 > FB \geq 2.5$	$0.5 < SIR \leq 1.0$	Between average and 50% better than average
	2-blob	$2.5 > FB \geq 1.5$	$1.0 < SIR \leq 2.0$	No more than 100% worse than average
Worse	1-blob	$1.5 > FB$	$2.0 < SIR$	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.

Limitations

Each reporting agency (CMS and state health agencies) reports on a somewhat different set of ICUs and surgical procedures, covering a particular period of time.

For ICU bloodstream infections, not all hospitals have the same assortment of ICUs; a smaller hospital may have only a combined medical/surgical ICU, while a larger hospital may have 6 or more ICUs, including separate medical and surgical ICUs. For example, reports from Maine and Virginia do not provide data for individual ICUs, and their overall infection rates were compared to the pooled NHSN infection rate for medical/surgical ICUs. Reports from New Jersey, Maryland, and

CMS combine all ICU types, and so we use their published predicted values (called “expected value” in statistical terms) to calculate the SIR.

Definitions of ICUs are somewhat malleable, and patients may be treated in any of a number of ICUs based on their condition and availability of beds. Some states use different definitions for classifying which infections are to be counted as CLABSIs.

For surgical-site infections, the Ratings are based on only those surgeries for which the state requires reporting, not all of the procedures performed in the hospital, and not all hospitals perform all procedures. Since the Ratings are based on a subset of all surgeries performed at a hospital, a hospital’s overall performance at preventing surgical infections may be better or worse than the particular surgical procedures that they report. Some states report surgical procedures separately by risk category, and others report for all risk categories combined. When necessary, we combine the NHSN data across risk categories so that we compare each state’s data to the comparable national rates.

Comparisons between hospitals in different states may not be reliable for SSIs, since states report data on different schedules and report infections for different surgical procedures.

Different hospitals treat different populations of patients, some of whom may be more susceptible to infection than others. The standardization we do for national rates is currently the best adjustment available to accommodate these differences.³ However, we maintain that all hospitals should be able to eliminate hospital-acquired infections, regardless of the patients they treat.

These data are historical, and reflect infections during different time periods, depending on the state. 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.

Three types of surgical-site infections are included: superficial (involves the skin or subcutaneous tissue), deep incisional surgical-site infection (involves deep soft tissues), or organ space (includes any part of the body that is opened and manipulated during surgery). The data we gather from states includes all of these types of infections combined; we have no way of determining which types of infections are reflected in the data.

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. Infections associated with implants (for example, in knee or hip replacements), which are included in our Ratings for some hospitals, can occur up to a year after surgery. 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

³ Recently the CDC has developed a model for the SIR that adjusts for patient-level risk. However, we do not have access to the patient-level data that would be required to do that risk adjustment. See (citation).

hospital's emergency department, so there are many infections that will not be identified by the hospital's reporting system.

CLABSI data reported by CMS are self-reported by hospitals, without any independent or external validation. Most states that publicly report 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. Meanwhile, if you have a choice of hospitals in states that report bloodstream or surgical infection data, you can use the information, along with other data from our hospitals Ratings (e.g., patient experience, readmissions, overuse of imaging), other data sources and recommendations from your health care provider to help find the best hospital for you.

Appendix A
Hospital-specific public reporting of ICU central-line bloodstream infection and surgical site infection data
Effective November 2012

Agency	ICUs reported	Surgical sites reported	Dates covered
Centers for Medicare and Medicaid Services (CMS)	All adult, pediatric, and neonatal ICUs		January – December 2011
California Department of Public Health	Medical Medical/surgical Pediatric Surgical	Abdominal Aortic Aneurysm Appendix Bile, Liver, Pancreatic Cardiac Surgery Chest Chest and Donor Colon Cesarean Section Exploratory Abdominal Surgery Fracture Gallbladder Gastric Hip Kidney Surgery Kidney Transplant Knee Laminectomy Liver Transplant Ovarian Surgery Pacemaker Rectal Small Bowel Spinal Fusion Spinal Refusion Thoracic Vaginal Hysterectomy	CLABSI: January – December 2011 SSI: April – December 2011

Agency	ICUs reported	Surgical sites reported	Dates covered
Colorado Department of Public Health and Environment	Medical Cardiac Surgical Surgical Cardiothoracic Medical/Surgical Medical	CABG chest and donor Knee Hip Vaginal hysterectomy Abdominal hysterectomy Hernia	CLABSI: State data are not used currently. SSI: August 2009 – July 2011. Dates vary by hospital
Connecticut Department of Public Health	Medical Medical/Surgical Pediatric Medical/Surgical		July 2011 – June 2012
Delaware Health and Social Services, Division of Public Health	Medical Medical/Surgical Pediatric Medical/Surgical		January – December 2011
Florida Department of Health*	All adult Pediatric		State data are not used currently.
Illinois Department of Public Health	Medical/Surgical Medical Medical Cardiac Surgical Surgical Cardiothoracic Pediatric	Knee CABG	CLABSI: January – December 2011 SSI: July 2010 – June 2011
Dirigo Health Agency's Maine Quality Forum	All Adult ICUs combined		October 2010 – December 2011. Dates vary by hospital
Maryland Health Care Commission	All Adult & Pediatric ICUs combined		July 2011 – June 2012

Agency	ICUs reported	Surgical sites reported	Dates covered
Massachusetts Department of Public Health	Medical Cardiac Surgical Surgical Cardiothoracic Medical/Surgical Medical Neurosurgical Pediatric Medical	Hip Knee CABG Vaginal hysterectomy Abdominal hysterectomy	CLABSI: July 2010 – June 2011 SSI: July 2009 – June 2011. Dates vary by hospital.
Minnesota Department of Health		Vaginal hysterectomy Knee	January 2010 – June 2012. Dates vary by hospital.
Missouri Department of Health and Senior Services	Coronary Surgical Medical/Surgical Medical Pediatric	Abdominal hysterectomy Hip CABG Chest and Donor	CLABSI: July 2009 – December 2011. Dates vary by hospital. SSI: January – December 2011
New Hampshire Department of Health and Human Services	Medical Cardiac Surgical Cardiothoracic Medical/Surgical Medical	CABG Colon Knee	CLABSI: No state data are used in this update. SSI: January – December 2010
State of New Jersey Department of Health and Senior Services	All adult ICUs combined	CABG Abdominal hysterectomy	CLABSI: January – December 2010 SSI: January – December 2009
New Mexico Department of Health*	Adult ICUs combined		State data are not used currently.

Agency	ICUs reported	Surgical sites reported	Dates covered
New York State Department of Health	Coronary Surgical Cardiothoracic Medical Medical/Surgical Surgical Neurosurgical Pediatric	CABG Hip Colon	CLABSI: January – December 2010. SSI: January 2009 – December 2010. Dates vary by hospital.
Ohio Department of Health		CABG C-section Knee	July 2010– June 2011
Oklahoma State Department of Health*	Medical Medical/Surgical Cardiac		State data are not used currently.
Office for Oregon Health Policy & Research (OHPR)	Medical/Surgical	Abdominal Hysterectomy CABG Colon Hip Knee Laminectomy	CLABSI: State data are not used currently. SSI: January – December 2011
Pennsylvania Department of Health*	Reported hospital wide, not by ICU.	Knee CABG Cardiac Hip Abdominal hysterectomy	State data are not used currently.
Rhode Island Department of Health	Coronary Medical Medical/Surgical Neurosurgical Pediatric Medical/Surgical Surgical Surgical Cardiothoracic		July 2011 – June 2012

Agency	ICUs reported	Surgical sites reported	Dates covered
South Carolina Department of Health and Environmental Control	All adult ICUs combined Pediatric	CABG Colon Hip Abdominal hysterectomy Knee	CLABSI: State data are not used currently. SSI: January – December 2011
Tennessee Department of Health	Medical Cardiac Surgical Cardiothoracic Medical Medical/Surgical Surgical Neurosurgical Pediatric Medical/Surgical		January – December 2010
Vermont Department of Banking, Insurance, Securities & Health Care Administration	All adult ICUs combined	Hip Abdominal Hysterectomy Knee	CLABSI: April 2011 – March 2012 SSI: October 2010 – March 2012. Dates vary by hospital.
Virginia Department of Health	All adult ICUs combined		January – December 2011
Washington State Department of Health*	Cardiothoracic Medical/Surgical Medical Coronary Neurosurgical Pediatric Cardiothoracic Pediatric Medical/Surgical		State data are not used currently.

*Data are not reported in a way in which we can use

Appendix B
NHSN rates used for standardization of CLABSI rates

ICU type	Total CLABSI reported to NHSN	Total CLD reported to NHSN	Rate (CLABSI/1000 CLD)
Medical Cardiac	876	436,409	2.007
Medical overall	2,097	911,476	2.301
Medical, teaching	1,410	549,088	2.568
Medical, others	687	362,388	1.896
Medical/Surgical overall	4,053	2,441,719	1.660
Medical/Surgical teaching	1,474	699,300	2.108
Medical/Surgical, others	2,579	1,742,419	1.480
Neurologic	61	45,153	1.351
Neurosurgical	396	160,879	2.461
Pediatric Medical/Surgical	929	314,306	2.956
Surgical	1,683	729,989	2.306
Surgical Cardiothoracic	879	632,769	1.389

Appendix C
NHSN rates used for standardization of SSI rates, All risk levels combined

Surgery_type	Total Infections reported to NHSN	Total procedures reported to NHSN	Rate (infections/ 100 surgeries)
Abdominal Aortic Aneurysm	62	1945	3.187661
Abdominal Hysterectomy	890	54078	1.645771
Appendix	83	5874	1.413006
Bile, Liver, Pancreatic	88	888	9.90991
CABG	3808	134714	2.826729
Chest and Donor	3622	123055	2.943399
Chest	186	11659	1.595334
Cardiac Surgery	369	28685	1.286387
Colon	3453	62140	5.556807
C-Section	570	30994	1.839066
Exploratory abdominal surgery	103	5099	2.020004
Fracture All	182	10478	1.736973
Gallbladder	92	14652	0.627901
Gastric All	186	8171	2.276343
Hernia All	169	7477	2.260265
Hip	1651	130391	1.266192
Kidney Surgery	10	681	1.468429
Kidney Transplant	71	1622	4.377312
Knee	1528	171183	0.892612
Laminectomy	409	40077	1.020535
Liver Transplant	113	824	13.71359
Ovarian surgery	17	3016	0.56366
Pacemaker surgery	15	3403	0.440788
Rectal Surgery	86	1167	7.369323
Small Bowel All	257	4221	6.088605
Spinal Fusion	633	41210	1.536035
Spinal Refusion	31	989	3.134479
Thoracic surgery	22	1979	1.111673
Vaginal hysterectomy	165	18869	0.87445

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 2004, Medicare reimbursement to hospitals became tied to hospitals' reporting of quality data for patients diagnosed with heart failure, heart attack, and pneumonia.

"Readmission" refers to a recently hospitalized patient who needs to go back into a hospital again. 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 heart attack, heart failure, or pneumonia. Patients may have been readmitted back to the same hospital or to a different hospital or acute care facility. They may have been readmitted for the same condition as their recent hospital stay, or for a different reason.

Starting in October of 2012, hospitals will be fined if they have too many Medicare patients readmitted to the hospital within 30 days of being discharged. About two-thirds of the hospitals serving Medicare patients, or some 2,200 facilities, will be penalized this year, with an average fine of about \$125,000, according to government estimates. Hospitals in Arkansas, the District of Columbia, Illinois, Kentucky, Massachusetts, Mississippi, New Jersey, and New York will be among the hardest hit, according to reports.⁴

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.⁵ Third, we know that, to at least some extent, readmissions reflect errors or hospital-acquired conditions in the initial hospitalization.⁶

The data

CMS publishes readmission rates after statistical adjustment for how sick patients were when they were initially admitted to the hospital and for the amount of data available for each hospital. For each individual condition (heart attack, heart failure, or pneumonia), CMS provides each hospital's 30-day risk-standardized readmission rate (RSRR). Statistical methods are described at <http://www.hospitalcompare.hhs.gov/staticpages/for-professionals/ooc/statistical-methods.aspx> and <http://www.hospitalcompare.hhs.gov/staticpages/for-consumers/ooc/readmission-measures.aspx>

Data reported on Hospital Compare cover discharges over a three-year period for over 4000 hospitals. We provide the chance of readmission for each condition (heart attack, heart failure, and pneumonia) for any hospital with at least 25 cases for that condition. In addition, we provide a

⁴ <http://news.consumerreports.org/health/2012/10/medicare-starts-fining-hospitals-for-readmitted-patients-1.html>

⁵ www.ahrq.gov/qual/impptdis.htm

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

Rating score for each hospital based on a composite of the readmission rates for whichever of these three conditions are available.

Assigning Ratings scores

For each measure, there are a number of outliers with particularly high rates. If we were to include these outliers in determining the endpoints of the linear transformation to fractional blobs they would skew the distribution to assign higher Ratings overall, which we believe would be misleading. Instead, we identify outliers at the high end (atypically high rates).⁷ These outliers are removed from the determination of the transformation to fractional blobs.

After removing outliers, we re-scale the rates for each of the three conditions (heart attack, heart failure, pneumonia) separately to the fractional blob scale, as follows:

1. The highest *non-outlier* rate within a condition is transformed to the fractional blobs value of 0.5. Removing the high rate outliers eliminates their effect of stretching out the scale.
2. The lowest readmission rate within a measure is mapped to the fractional blob value of 4.5, based on our medical expert's judgment that none of these readmission rates is low enough to qualify for our highest Rating of a 5-blob. Low-rate (best) outliers were maintained, in order to allow hospitals with low readmission rates to be differentiated from those with higher rates. The highest (worst) non-outlier readmission rate is mapped to a fractional blob of 0.5.
3. All readmission rates are linearly rescaled using the above two points as anchors. Hospitals with readmission rates that are high outliers are assigned fractional blobs of 0.5.
4. For each hospital, we use whichever of the three conditions have sufficient data (at least 25 cases), and calculate the weighted mean of the fractional blobs for those measures. Weights are proportional to the number of discharges for patients hospitalized for heart attack, heart failure, or pneumonia at that hospital.

That weighted mean is rounded to whole numbers to produce our blob scores of 1, 2, 3, and 4. A fractional blob of 4.5 is assigned a blob of 4.

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.

⁷ There are several accepted definitions of what constitutes an outlier. We use the definition that an outlier is 1.5 times the inter-quartile range greater than the 3rd quartile. This is one common definition, and is used in the creation of boxplots in many popular statistical packages. See, for example, Renze, John. "Outlier." From *MathWorld*--A Wolfram Web Resource, created by Eric W. Weisstein. <http://mathworld.wolfram.com/Outlier.html>

These data also reflect readmissions only for Medicare patients, and only those patients who were initially treated for three specific conditions, though we believe they provide a good indication of readmission rates overall.

Ratings come from recent data but it is possible that updates will show improvements or declines in performance. 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. The statistical methods used by CMS for risk adjustment are documented here: www.hospitalcompare.hhs.gov/staticpages/for-professionals/oc/statistical-methods.aspx.

2.3 Avoiding Complications

The Patient Safety Indicators (PSIs) are a set of measures that describe the incidence of hospital complications and adverse events following surgeries, procedures, and childbirth. These measures were developed by the Agency for Health Quality Research (AHRQ) for the Centers for Medicare and Medicaid Services (CMS), an agency of the federal government. Detailed methodology for the PSIs is given here:

www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/QI%20Empirical%20Methods%2005-03-11.pdf

Data come from hospital billing claims for Medicare patients paid through the Inpatient Prospective Payment System, and thus only includes hospitals that are paid through that system. Results are only reported for hospitals with at least 25 cases for an individual measure.

The data and Ratings

Avoiding Serious Complications (PSI-90) is a composite of the eight Patient Safety Indicators shown at the top of the following page, and is the basis for our complications Rating.

Results for these composites are reported on Hospital Compare indicating whether each hospital is Better than Average, Not Different from Average, or Worse than Average; more specific values are not provided. These text descriptions appear in our hospital ratings in that form.

PSI code	Description	Weight in the composite
PSI-03	Pressure Ulcer	24%
PSI-06	Iatrogenic Pneumothorax (collapsed lung)	5%
PSI-07	Central Venous Catheter related bloodstream infection	13%
PSI-08	Postoperative Hip fracture	0.1%
PSI-12	Postoperative Pulmonary Embolism or Deep Vein Thrombosis	24%
PSI-13	Postoperative Sepsis	4%
PSI-14	Postoperative Wound Dehiscence (wound opening)	1.2%
PSI-15	Accidental Puncture or Laceration	30%

Limitations

Previously Consumer Reports also reported a composite of AHRQ’s Inpatient Quality Indicators, Avoiding Mortality (IQI-91), in our hospital ratings as well. CMS removed this composite from its reporting on its Hospital Compare website⁸, and so it was removed from Consumer Reports’ Ratings because we no longer have access to current data.

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.

The patient safety indicators (PSIs) rely on the proper use of a code that indicates if the event was present when the patient arrived at the hospital - the so-called “present on admission” or “POA” indicator. Coding accuracy varies and accordingly methods have been developed to screen hospitals for accurate POA coding.⁹ These screens cannot be employed by Consumer Reports for this measure because the data required to perform them is not available from Hospital Compare.

These data also reflect outcomes only for Medicare patients, though we believe they provide a good indication of complication 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 updates will show improvements or declines in performance. 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. The statistical methods used by CMS for risk adjustment are documented here:

www.hospitalcompare.hhs.gov/staticpages/for-professionals/ooc/statistical-methods.aspx.

⁸ www.hcahpsonline.org/Files/IPPS%20FY%202013%20Proposed%20Rule.pdf

⁹ Pine et al., 2009; <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2645132/pdf/phim0006-0002.pdf>.

Finally, this measure is limited by the accuracy of coding of complications in the billing records¹⁰ and research suggests that the patient safety indicators significantly underreport the number of errors that occur in hospitals.¹¹

2.4 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 adverse event 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 adverse event rates are risk adjusted for the types of patients treated by each hospital. A description of this approach has been previously published.¹²

Risk-adjusted prolonged lengths of stay were used as an indicator for adverse events for several reasons. First, a prolonged length of stay 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 length of stay 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.^{10,11}

Overview

Data are from the 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 D).

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 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

¹⁰ Lawson et al., *Ann Surg* 2012; 256(6):973-981

¹¹ Classen et al., *Health Aff* 2011; 30 (4): 581-589

¹² Fry et al., *Am J Surg*; 2009 197(4):479–484

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 length of stay 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 adverse events predicted for the mix of patients at that hospital. Our scores are based on the degree to which observed rates of adverse events differed from those predicted by the models.

Procedure codes and diagnostic codes were grouped together into 27 groups of elective (scheduled) surgical procedures (see Appendix D). 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 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 factors¹³ for use in developing the risk-adjustment models were created based on average mortality rates, median post-operative length of stay (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.

¹³ From ICD-9-CM diagnosis and concurrent procedure codes.

- (2) *Estimation of hospital-acquired adverse events.* Cases with adverse events (PrRALOS or inpatient mortality) were counted, and a hospital-level predicted rate of adverse events was estimated based on the risk-adjustment models developed previously.
- (3) *Calculating Ratings.* Observed and predicted lengths of stay were then compared as the basis for our Ratings.

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 fifteen screens designed to assess the quality of hospital POA coding¹⁴ 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 eighty 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 records¹⁵.

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 charts¹⁶ outliers were iteratively identified and removed until all cases remaining had standardized post-operative LOS that were not outliers.¹⁷

The remaining cases comprised the analytic dataset of live discharges within good-coding hospitals and with routine LOS¹⁸. 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

¹⁴ Pine M, Fry DE, Jones BL, Meimban RJ: *Perspect Health Inf Manag* 2009; 6: 2 Epub Feb 11.

¹⁵ According to ICD-9-CM diagnosis codes

¹⁶ Fry DE, Pine M, Jones BL, Meimban RJ. *Am J Surg* 2012; 203 (3):392-6

¹⁷ Standardized risk adjusted postoperative LOS = (observed LOS) – (predicted LOS). XmR control charts were created for standardized post-operative LOS, cases above the 3 σ upper bound were removed, and a new control chart was created. This process was iterated until all cases had a standardized post-operative LOS $\leq 3\sigma$.

¹⁸ standardized post-operative LOS $\leq 3\sigma$

eliminated because they were outliers¹⁹ were classified as having had prolonged risk-adjusted post-operative length of stay.

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.

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 adverse event 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.²⁰

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

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

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

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

3. Assigning Ratings Scores

In addition to scoring the overall rate of adverse events, we assigned Ratings for 5 surgical groups (see Appendix D 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

¹⁹ above the upper control limit of 3σ

²⁰ 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.

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

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

Blobs for each of the 5 scored surgical groups, and for overall Adverse Events, are then assigned based on z*, as shown in the following table:

	Rating	Adverse Event Ratio
Better	5-blob	$z^* \leq -2.5758$
	4-blob	$-2.5758 < z^* \leq -1.64485$
	3-blob	$-1.64485 < z^* \leq 1.64485$
	2-blob	$1.64485 < z^* \leq 2.5758$
Worse	1-blob	$2.5758 < z^*$

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 (COTH).
- 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 the Centers for Medicare and Medicaid Services (CMS). Such records were intended to capture information for billing purposes rather than for identifying events involved in patient care. These data also reflect

adverse events 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 updates will show improvements or declines in performance. 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 the most current data publicly available. 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 adverse events are relative 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 adverse events. We compensate in part for this difference by assigning Ratings scores (blobs) based on the degree to which the actual rate of adverse events 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.

Research suggests that hospitals significantly underreport the number of adverse events that occur.²¹ 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 adverse events. Gaming is minimized by federal

²¹ Classen et al., *Health Aff* 2011; 30 (4): 581-589.

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 length of stay.

The overall surgery Rating we provide 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 length of stay, 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 length of stay tends to increase,²² 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.

²² Kassin et al., *J Am Coll Surg.* 2012; 215(3):322-30

Appendix D. 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.²³ 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.^{24,25} 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.²⁶ Most hospitals are currently required to report HCAHPS data to receive full payment from Medicare.²⁷ 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.²⁸

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).

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 E (page 35).

We only present Patient Experience Ratings for hospitals with at least 100 completed surveys; 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

²³ www.iom.edu/Reports/2006/Preventing-Medication-Errors-Quality-Chasm-Series.aspx

²⁴ Morrison et al, *Pain*. 2003;103(3):303-11.

²⁵ Sinatra, R. *Pain Medicine*. 2010; 11: 1859–1871.

²⁶ Moore et al., *Archives of Internal Medicine*. 2007; 167(12), 1305-1311.

²⁷ www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/HospitalRHQDAPU200808.pdf

²⁸ www.gpo.gov/fdsys/pkg/FR-2011-05-06/pdf/2011-10568.pdf

(e.g. pain management and information about new medications), and 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	5-blob	95% - 100%
	4-blob	90% - 94%
↕	3-blob	85% - 89%
	2-blob	80% - 84%
Worse	1-blob	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 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 fractional blobs 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	5-blob	85% - 100%
	4-blob	70% - 84%
↕	3-blob	55% - 69%
	2-blob	40% - 54%
Worse	1-blob	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 Centers for Medicare & Medicaid Services (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), training programs, site visits, independent data audits and analyses, and vendor certification processes (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 E: 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 Use of Electronic Health Records

The American Hospital Association (AHA) conducts an annual survey of hospitals called the AHA Annual Survey Information Technology (IT) Supplement. The survey contains dozens of questions about the use of computer-based and other electronic systems for keeping records. The federal government pays incentives to physicians and hospitals that meet specific requirements for the use of electronic health record technology. The adoption of electronic medical records could improve patient safety by improving accuracy and coordination of care, and reduce costs. However, the adoption of this technology has been shown to create new types of errors and other unintended consequences.²⁹

The data

Our Electronic Health Records score is based on 28 measures of the AHA Annual Survey Information Technology supplement that describe the extent to which the hospital has a computerized system that allows for: electronic clinical documentation (7 measures), results viewing (6 measures), computerized provider order entry (5 measures), decision support (6 measures), and bar coding (4 measures). These measures are listed on the next page. For each measure, hospitals respond about whether electronic technology is implemented on a six-point scale:

1. Fully implemented across all units
2. Fully implemented in at least one unit
3. Beginning to implement in at least one unit
4. Have resources to implement in the next year
5. Do not have resources but considering implementing, or
6. Not in place and not considering implementing.

Assigning Ratings scores

Fractional blobs and blob scores for Electronic Health Records are calculated as follows:

1. For each of the 28 measures, the hospital receives 4 points if they report that measure is fully implemented across all units, 1 point if they report it is fully implemented in at least one unit, and no points otherwise.
2. The sum of these points is calculated across all 28 measures. The maximum possible point total is 112 (28 measures x maximum 4 points/measure).
3. These sums are linearly re-scaled to the range 0.5 (for a point total of zero) through 5.5 (for hospitals with 112 total points), and then rounded to whole number blobs between 1 and 5 (5.5 is rounded down to a 5-blob).

²⁹ www.ucguide.org/index.html

Measures of hospital use of electronic records used in our Ratings from the AHA Annual Survey Information Technology Supplement

Electronic clinical documentation

- a. Patient demographics
- b. Physician notes
- c. Nursing notes
- d. Problem lists
- e. Medication lists
- f. Discharge summaries
- g. Advanced directives

Decision support

- a. Clinical guidelines
- b. Clinical reminders
- c. Drug allergy alerts
- d. Drug-drug interaction alerts
- e. Drug-lab interaction alerts
- f. Drug dosing support

Results viewing

- a. Laboratory reports
- b. Radiology reports
- c. Radiology images
- d. Diagnostic test results
- e. Diagnostic test images
- f. Consultant reports

Bar coding

- a. Laboratory specimens
- b. Tracking pharmaceuticals
- c. Pharmaceutical administration
- d. Patient ID

Computerized provider order entry

- a. Laboratory tests
- b. Radiology tests
- c. Medications
- d. Consultation requests
- e. Nursing orders

Note: We did not use the last item under “Decision Support”, “implement drug formula checks” because too few hospitals responded to this question. This question was first added in the 2010 survey; data for the 2010 survey was released in June 2012.

Limitations

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. These data are based on a voluntary survey completed by hospitals. There may be biases or errors in how these data are reported to the American Hospital Association. The survey IT supplement currently has a 60% response rate, which is considered quite high for a voluntary survey, but there are many hospitals that choose not to respond, and for which we do not have Ratings. Data are self-reported by hospitals, and are not subject to independent validation. We rely on the American Hospital

Association (AHA) to conduct the survey, affirm its accuracy and validity, and report its results, which we then convert into our unique Ratings format.

Because of the time required to conduct the survey and report its results, data are over a year old at the time of this report. Since many hospitals are continuing to update their use of electronic records, these data may not reflect current practice, although they are the most current data available.

4.2 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). 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 dataset: (1) percentage of patients 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 CT scans of the abdomen that are performed twice, once with radioactive contrast and one without; (4) the percent of all CT scans of the thorax or chest that are performed twice, once with radioactive 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.³⁰

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.

Assigning Ratings scores

We used the double-scan rates for chest and abdomen in our Ratings. To convert these rates to our fractional blob 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	5-blob	Rate \leq 5%
	4-blob	5% < rate \leq 10%
↓	3-blob	10% < rate \leq 15%
	2-blob	15% < rate \leq 20%
Worse	1-blob	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 rather than patient outcomes, but they contain significant details about a patient's stay in the hospital.

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 updates will show improvements or declines in performance.

³⁰ http://www.hospitalcompare.hhs.gov/staticpages/learn/importance_quality.aspx?measurecd=OPIMAGE

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 complications. 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 4): Central-line infections kill up to 16,250 patients a year.³¹ Even for those who survive, a central-line infection means weeks or months of debilitating treatments and side effects. Use of a simple checklist, that not all hospitals have adopted, has demonstrated effectiveness at reducing central-line infections to zero.³² 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] for more information.

Avoiding readmissions (see page 18): 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.³³ 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.³⁴ 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.³⁵ 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 to reduce hospital readmissions by 20 percent in 2013, compared with 2010. It is estimated that

³¹ *MMWR*, March 4, 2011 / 60(08);243-248;

[/www.cdc.gov/mmwr/preview/mmwrhtml/mm6008a4.htm?s_cid=mm6008a4_w](http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6008a4.htm?s_cid=mm6008a4_w)

³² Pronovost, et al., *N Engl J Med*. 2006 Dec 28;355(26):2725-32.

³³ Jencks et al., *N Engl J Med* 2009; 360:1418-1428.

³⁴ <http://oig.hhs.gov/oei/reports/oei-06-09-00090.pdf>

³⁵ www.ahrq.gov/qual/impptdis.htm

hitting this target would result in 1.6 million fewer patients being readmitted to a hospital within 30 days.³⁶

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.³⁷ 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.³⁸

Communication about medications and discharge (see page 32): 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.³⁹

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.⁴⁰ 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 38): 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.⁴¹ According to CMS, a single CT scan of 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.

Avoiding complications (see page 20): The federal Agency for Healthcare Research and Quality (AHRQ) developed a set of complication measures (patient safety indicators or PSIs) using hospital administrative data that measure in-hospital safety events. They were developed to identify potentially preventable patient safety events such as bedsores, collapsed lungs, post-operative hip fractures, infections and post-operative blood clots.⁴²

³⁶ www.healthcare.gov/compare/partnership-for-patients/

³⁷ Emerson et al., *Infect Control Hosp Epidemiol.* 2012;33(6):539-544.

³⁸ Friedman et al, *Med Care.* 2009;47(5):583-590.

³⁹ Gleason et al, *J Gen Intern Med.* 2010;25(5):441-447.

⁴⁰ Coleman et al., *Arch Intern Med.* 2005;165(16):1842-1847.

⁴¹ Berrington de González et al, *Arch Intern Med* 2009;169(22):2071-2077.

⁴² www.qualityindicators.ahrq.gov/modules/psi_overview.aspx

Infections, readmissions, communication and scanning together comprise 90 percent of the safety score; avoiding complications is worth 10 percent.⁴³ 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.

Safety score category	Components	Data Source	Weight
Avoiding infections (pages 4-10)	<ul style="list-style-type: none"> • Central-line associated bloodstream infections • Surgical-site infections 	CMS; State reported data;	22.5% of total based on combined CLABSI and SSI score, or if only one of SSI or CLABSI is available, it comprises the full infections measure.
Avoiding readmissions (pages 18-20)	30-day readmissions for: <ul style="list-style-type: none"> • Heart failure • Heart attack • Pneumonia 	CMS	22.5% of total, split among the one, two or three components which have sufficient data.
Communication (pages 32-35)	<ul style="list-style-type: none"> • Communication about discharge instructions • Communication about new medications 	CMS	22.5% of total, half for each component (discharge and medications).
Appropriate use of scanning (pages 38-39)	<ul style="list-style-type: none"> • Double chest CT scans • Double abdomen CT scans 	CMS	22.5% of total, half for each component (chest and abdomen), or if only one is available, it comprises the full scanning measure
Avoiding complications (see pages 20-22)	<ul style="list-style-type: none"> • AHRQ PSI Composite (PSI-90) 	CMS	10% of total

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.

⁴³ These weights represent a change from our earlier Safety score Ratings, in which the first four measures comprised 80% of the score, with the additional 10% being the AHRQ mortality measure that has since been removed from our Ratings.

The four measure categories that are based on interval data (infections, readmissions, communication, and scanning) 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 as (sum of reported CLABSIs and SSIs)/(sum of predicted CLABSIs and SSIs). That SIR is then transformed to our fractional blob scale using the methods described earlier (pages 5-8). 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 infections do not have a sufficient sample.
2. Readmissions is the calculated FB as described earlier.
3. Communication is the mean of the FBs for Communication about Medications and Communication about Discharge.
4. Scanning is the mean of the Chest and Abdomen fractional blobs, 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 four measure categories is then calculated using equal weights. That mean is linearly transformed to a scale from 0.5 to 90.5, so that these four measure categories combined account for 90% of the safety score, or 90 points out of a total possible of 100.

The AHRQ composite (avoiding complications) is then included to account for the final 10% of the score. This composite is reported on Hospital Compare as categorical data, with each hospital with valid data being reported as either "Better than Average", "Not Different from Average", or "Worse than Average". We gave a hospital 10 points for "Better than Average", five points for "Not Different than Average" and zero points for "Worse than Average", bringing the total possible score to 100.

Selecting weights

We examined the impact of varying the weights of the five categories on the safety score and the rank order of hospitals. For example, we considered weights of 30% for infections, 30% for readmissions, 20% for communication, 10% for scanning, and 10% for complications and mortality. With these unequal weights, the safety scores and their rank order were highly correlated ($r > 0.90$) with the scores and ranks using equal weights. 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. However, from the point of view of consumers, any data reported is of use primarily as a prediction of the care they can expect to receive if they are hospitalized in the future. The best available prediction of what a

patient can expect is the estimate of the true current value of that measure. Although none of these data are precisely current, the best available estimate of current performance in the consumer's care is based on the most recent data. While there might be better theoretical estimates of current performance, in practice the estimates described in this report and used in our Ratings are the best information we can provide consumers, regardless of the differences in time period they cover.

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 all 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.

5.1. Safety Ratings in *Consumer Reports* magazine, August 2012

The story "How Safe is Your Hospital" appeared in *Consumer Reports* in the August 2012 issue. In addition to reporting on hospital safety more broadly, the article included ratings of 1159 hospitals that had complete data for a Safety score. The Ratings were published in 4 regional editions, so that each subscriber received Ratings for hospitals in states in their region of the country. Hospitals were listed in descending rank order of the Safety score, within states.

The Safety score was presented as a bar on a 1-100 scale. In addition, four blob scores were presented for Avoiding Infections, Avoiding Readmissions, Communication, and Appropriate Use of Scanning. Each of these scores was created as a composite of several components, as described on the previous page. Each of these fractional blobs was rounded to the nearest whole number blob for presentation in *Consumer Reports*; a fractional blob of 5.5 was assigned a blob score of 5.⁴⁴ In addition to the Avoiding Complications measure from AHRQ, we also included a similar measure of Avoiding Mortality, that was available at that time (but has since been excluded from public reporting by CMS). Hospitals that were higher or lower than average in Complications or Mortality were footnoted in Ratings table in *Consumer Reports*.

⁴⁴ The fractional blobs for Infections were converted to blob scores incorrectly in this article for 133 hospitals. Corrections to those blob scores were published at <http://www.consumerreports.org/health/resources/pdf/how-we-rate-hospitals/HospitalCorrection.pdf> The safety score was not affected by this error.