

Hospital Report Cards: Questions & Answers

Prepared by the Office of Clinical Outcomes & Resource Management
The University of Iowa Hospitals and Clinics
July 2005

CORM



Dedication

This monograph is dedicated to our patients who are critical partners in our goal to improve quality and safety at UIHC.

Purpose

This monograph is designed to help our patients, staff, referring physicians, health care workers, insurers and payors to understand 1) the complexities of current quality and safety information provided to the public, and 2) our commitment to the appropriate use of such information to improve the quality and safety of care at UIHC.

Table of Contents

Dedication.....	2
Purpose.....	2
Table of Contents.....	3
Questions and Answers.....	4
Preferred Report Cards	5-6
JCAHO and <i>Hospital Compare</i>	5-6
Less Preferred Report Cards	7
AHRQ Inpatient Quality and Patient Safety Reports	7-8
Leapfrog Group.....	9
US News & World Report Best Hospitals and Best Doctors	10
Conclusion and Commitment	11

Figures and Attachments

Figure I (Hospital Compare – % of Heart Failure Patients Given ACE I for LVSD)	6
Figure II (2002 Abdominal Aortic Aneurysm Repair Volume and Mortality)	8
Figure III (Selected Annual Case Volumes).....	9
Figure IV (Best Hospitals 2005).....	10
Attachment I (Hospital & Physician Quality/Safety Report Card Web Sources)	12
Attachment II (Market Watch, June 28, 2005)	13-15

Questions and Answers

Why do we need reliable public sources of hospital quality and safety information?



The rising costs of health care and recent Institute of Medicine reports have focused national attention on an unfavorable “gap” existing between current hospital quality and safety and that which is operationally possible. Narrowing the quality and safety “gap” has become a public health priority and health policy issue in the US.

How do we close the gap? To narrow the quality and safety gap, health care providers and the public must work cooperatively and deal with the problem openly. Public confidence must not be undermined. Patient confidentiality and provider liability issues must be addressed clearly and constructively up front to prevent misunderstandings.

So why are reliable, public sources important? Simply because narrowing the quality and safety gap will require using reliable hospital quality and safety information to inform the public and support hospital improvement efforts.

Which sources of hospital quality and safety information should you rely on?

There are many different sources of hospital quality and safety information available to the public. Some of these are listed in Attachment I. These sources differ greatly in their approaches to developing information. The information presented and conclusions drawn may vary from source to source as discussed in Attachment II.

Importantly, sources differ in the degree to which: 1) they are user friendly, 2) they are readily understandable to the public and; 3) they provide important information to users on how they collect, analyze and determine the reliability of their data and conclusions.

So what is the answer to the above question about who to rely on? You should rely on public sources that are “transparent”, that discuss the reliability of their information and how they assure it.



Are some report cards better than others?

Yes. For purposes of providing public information and improving quality and safety of care, University of Iowa Health Care believes certain report cards are to be preferred to others. For purposes of discussion, we'll divide report cards into two groups: preferred and less preferred.

Preferred Report Cards

Report cards based on direct review of clinical data from health care records and services

Public report cards based upon information obtained by a process of direct review of clinical data from health care records and services are generally the most useful and reliable. The process of direct review is the only reliable means currently available to determine what particular patient care was provided and when. Direct review allows for the determination of whether the processes of evaluation and treatment were timely and met accepted guidelines for specific medical conditions. Direct review is time and labor intensive and relatively expensive. It is, therefore, usually focused on the examination of a limited number of conditions.

Direct review is employed by the Joint Commission on Accreditation of Health Care Organizations (JCAHO), as seen on the JCAHO accreditation website (www.jcaho.org/quality+check/home.htm). It is also employed by the Centers for Medicare & Medicaid Services (CMS) and can be seen on the CMS website (www.hospitalcompare.hhs.gov). Because participation is required to maintain JCAHO accreditation and full CMS payment, most acute care hospitals participate in at least one of these systems.

Because there is overlap in information offered by the JCAHO and CMS websites, we'll focus on the CMS information available publicly on *Hospital Compare*.

Hospital Compare

Source: CMS, JCAHO, National Quality Forum for Health Care (NQF), American Hospital Association (AHA), Association of American Medical Colleges (AAMC).

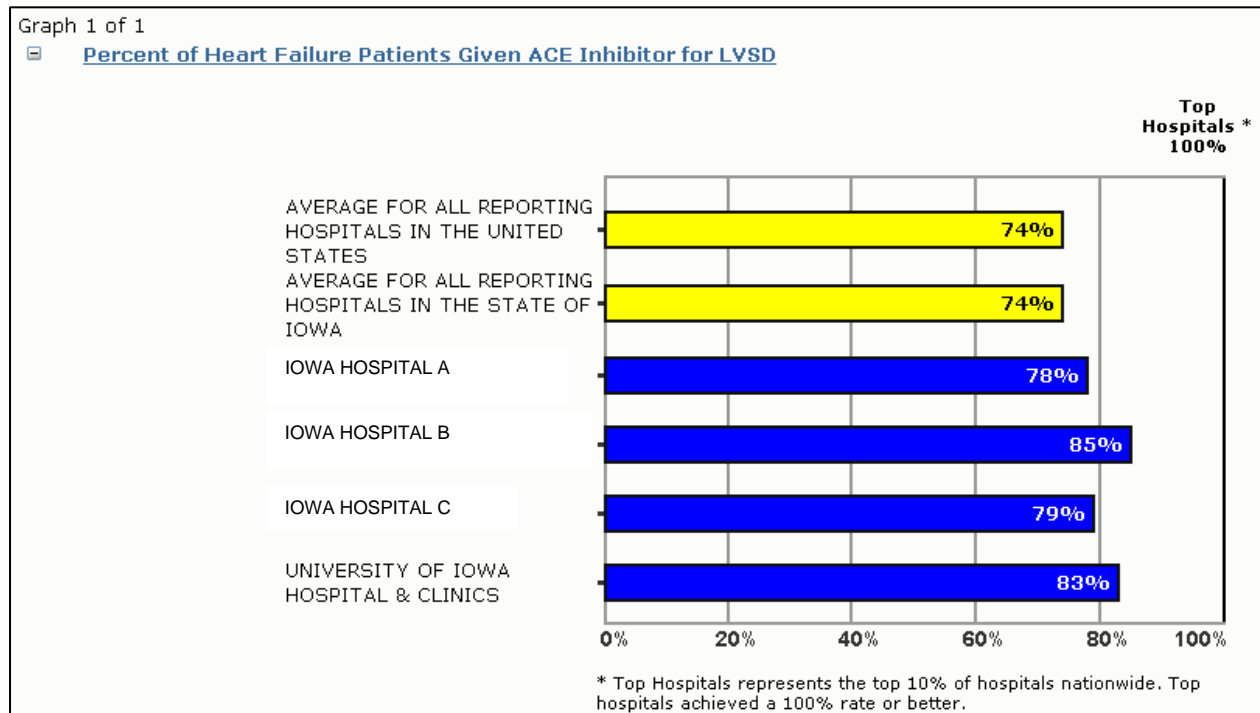
Measures: Seventeen care process measures for acute myocardial infarction (AMI), heart failure (HF), community-acquired pneumonia (CAP), are collected through medical record abstraction by hospitals and verified by CMS. More measures will be added in the future. Participation was required by 7/1/04, or hospitals would have forgone a 0.4% Medicare payment increase.

Strengths: The measures are clinically meaningful, validated and accepted by clinicians. They relate directly to processes of care which can be improved. The great majority of Iowa hospitals are participating, including all the major hospitals. It is supported by a broad range of accrediting and trade organizations and has a consumer-friendly web interface.

Weaknesses: Medical record abstraction is expensive. The measures do not assess outcomes of care (mortality, length of hospital stay, etc.).

UIHC Examples: As you will note in Figure I, the UIHC has an improvement opportunity in the category, “Percent of Heart Failure Patients Given ACE Inhibitor for LVSD.” ACE Inhibitors are important medicines used to treat patients with heart failure. Based upon these findings, UIHC is implementing performance improvement interventions for this heart failure metric. Other metrics that present opportunities for performance improvement in the care of patients with acute myocardial infarction (AMI) and Community-Acquired Pneumonia (CAP) are receiving similar attention.

Figure I



Less Preferred Report Cards



Report cards based on review of administrative data

Many report card systems use administrative information derived from electronic billing data to evaluate hospital performance. This information exists in a standard format called the UB92 form and is available for most acute care hospitals. Report cards that use this information can examine things like numbers of procedures performed, length of hospital stay, patient diagnoses and mortality reported by the hospital claims coders.

The main strength of these report cards is the ready availability and low associated costs of extraction of billing information. There are many weaknesses, however. For example, there is a paucity of information on administrative data forms about the patients' clinical condition and about non-procedural treatments performed. Thus, administrative data cannot be relied on to infer how sick or stable a patient was at hospital admission, or what drugs were given during the hospital stay, or when many treatments were performed. Further, administrative data cannot differentiate between medical problems patients had prior to being admitted and those that developed as a result of complications of care. So to re-emphasize, reliability of administrative data in judging the safety and quality of clinical care is very limited. It follows that the use of administrative data in public report cards may lead to inaccurate conclusions.

Because report cards based on billing information lack clinical details, statistical models must be used to estimate how ill patients are from hospital to hospital. These models are based on patient diagnoses, age, and demographic information. Some report card systems allow others to evaluate how well their statistical models work by making them publicly available. Other systems keep their models proprietary. Report cards that use proprietary statistical models that cannot be independently verified include Health Grades, Select Quality Care, Health Data Insights, Health Care Choices, Subimo, and Consumers' Checkbook.

AHRO Inpatient Quality and Patient Safety Reports

Source: Made available by the Agency for Healthcare Research and Quality (AHRQ), and distributed to Iowa participants by the Iowa Healthcare Collaborative. AHRQ makes all its statistical models available, so that anyone who desires can reproduce and examine its results (www.qualityindicators.ahrq.gov).

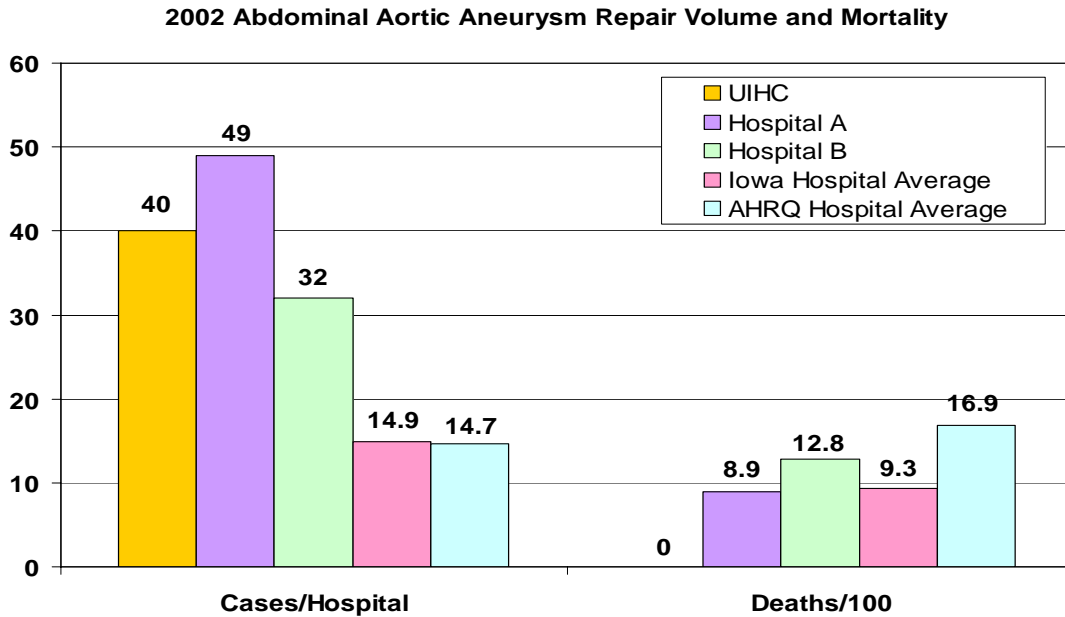
Measures: The 25 Inpatient Quality and 20 Patient Safety Measures are based upon electronic billing data.

Strengths: Widely available and relatively inexpensive to use. Statistical models are available for testing, and revised based upon feedback from measure users.

Weaknesses: Measures are dependent on coding accuracy. Cannot measure most recommended clinical care practices.

UIHC Examples: UIHC used this report to compare its 2002 volume and mortality rate for abdominal aortic aneurysm repair to other Iowa and national hospitals:

Figure II



Similar comparisons can be made for other procedures, providing valuable comparisons in UIHC performance that can be used to focus performance improvement. However, some comparisons have been problematic. For example, one measure showed that UIHC's rate of complications from anesthesia appeared high compared to other institutions. On further examination, we found that UIHC was coding pruritis (skin itching), expected with certain pain medications, quite differently than most hospitals. This gave the false appearance of a significant clinical problem. We are working with AHRQ on this situation.

There are other potential confounding factors of concern to UIHC (Attachment II). For example, without adjustment for unique factors affecting academic medical centers, such as the proportions of transfer patients, poor patients and patients with high levels of disease complexity, the performance of hospitals like UIHC may not be accurately reflected.

Report cards based on subjective reporting by hospitals and peers

A few hospital report cards are based upon information reported by hospitals and peers. This information often has subjective elements that may be difficult to independently verify. The organizations sponsoring the report cards rely on peers and health care purchasers to rate care quality and to police the accuracy of information submitted. Two examples of this kind of report are the Leapfrog Group and the US News Best Hospitals Report.

Leapfrog Group

Source: The Leapfrog Group is a non-profit consortium of Fortune 500 companies whose goal is to improve patient safety. Its members attempt to do this by favoring hospitals that meet Leapfrog standards for health care contracts and through public reporting (www.leapfroggroup.org).

Measures: Participating hospitals self-report their compliance with safety standards and responses to the NQF Safe Practices Survey to Leapfrog Group. Safety standards include computer-assisted physician order entry, evidence-based procedure referral, and ICU physician staffing by intensivists and trained support. Measures are updated yearly.

Strengths: Leapfrog focuses on health care practices that are largely supported by research and has received endorsements from JCAHO and CMS.

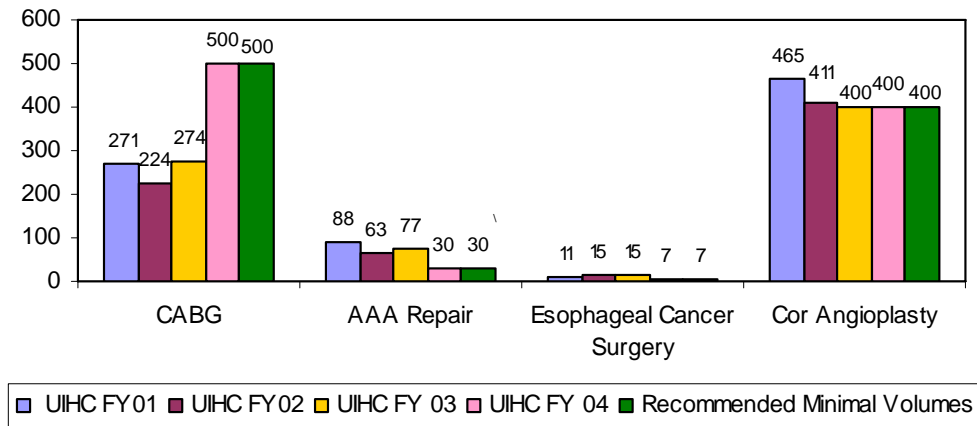
Weaknesses: Iowa is currently not a priority geographic area for the Leapfrog Group. Only Genesis in the Quad Cities is participating from Eastern Iowa. Leapfrog safety standards have a large subjective component. Most measures do not apply to rural hospitals. The importance of total institutional procedure volume, shown in Figure III, is currently under debate in the medical literature. Individual physician procedure volume is likely more important.

UIHC Examples:

As of FY04, UIHC met recommended minimal volumes for coronary artery bypass graft surgery (CABG), abdominal aortic aneurysm repair surgery (AAA), esophageal cancer surgery and coronary angioplasty procedures (Figure III).

Figure III

Selected Annual Case Volumes



US News & World Report Best Hospitals and Best Doctors

Source: US News & World Report Magazine
 (www.usnews.com/usnews/health/hosptl/tophosp.htm and www.bestdoctors.com).

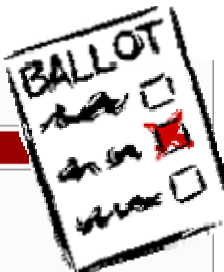
Measures: Rankings of US hospitals and physicians by specialty are based upon surveys of peers, Medicare mortality data, and other factors such as Nursing Magnet certification and are updated yearly.

Strengths: Results are very widely reported annually. Good estimate of a hospital's reputation in a particular specialty.

Weaknesses: Essentially an opinion poll and heavily influenced by past reputation.

UIHC Examples: Ophthalmology, Orthopedics and Otolaryngology Departments usually rank in the top ten nationally. 152 UI Health Care physicians are currently listed in the Best Doctors database.

Figure IV



Best Hospitals 2005

[Rankings Index](#) | [Honor Roll](#) | [Methodology](#) | [Glossary](#) | [Search](#)

EAR, NOSE, AND THROAT

Laser devices are increasingly used for diagnosis and treatment

[Printer-friendly version](#)

Rank	U.S. News Score	Reputation (%)	Hospital wide mortality ratio	Discharges (3 years)	Nurse/patient index	Nurse magnet facility	Key technologies (of 5)	Patient/community services (of 11)	Trauma center	
1.	Johns Hopkins Hospital, Baltimore	100.0	47.1	0.69	391	2.2	Yes	5.0	11	Yes
2.	Massachusetts Eye and Ear Infirmary, Boston	73.8	30.2	0.05	496	1.2	No	3.0	6	Yes
3.	University of Iowa Hospitals and Clinics, Iowa City	72.7	31.3	1.03	404	1.5	Yes	5.0	11	Yes
4.	University of Michigan Medical Center, Ann Arbor	65.0	22.9	0.73	499	2.3	No	5.0	11	Yes
5.	Mayo Clinic, Rochester, Minn.	61.6	20.2	0.80	907	2.5	Yes	5.0	11	Yes
6.	University of Pittsburgh Medical Center	59.8	21.6	0.92	546	1.9	No	5.0	10	Yes
7.	Cleveland Clinic	58.0	18.2	0.71	357	1.5	Yes	5.0	9	No
8.	Barnes - Jewish Hospital/Washington University, St. Louis	51.3	14.0	0.85	580	1.5	Yes	5.0	9	Yes
9.	UCLA Medical Center, Los Angeles	50.2	13.8	0.77	473	2.1	No	4.0	8	Yes
10.	University of Washington Medical Center, Seattle	47.9	12.1	0.78	238	2.3	Yes	4.5	9	No

Conclusion and Commitment

So what is the bottom line on report card reliability? How are the public and the quality and safety of US health care best served by report cards?

We believe the public and providers are best served by using quality and safety information:

1. that is understandable, focused and relevant to user groups, from individual members of the public to individual hospitals;
2. that is derived from report card sources committed to continuous improvement of the information they provide;
3. that is derived from clinical data (e.g., patient records), rather than from administrative data sets (e.g., UB92 forms);
4. that is derived using transparent methodology to verify and risk-adjust data; and
5. that may be directly applied by providers to improve processes of care.

What is UI Health Care committed to do to improve the quality and safety of the care we provide?

UI Health Care will continue to monitor and improve the quality and safety of its care focusing on report card sources that meet the above criteria, such as *Hospital Compare* and the AHRQ Inpatient Quality and Patient Safety Indicators.

When other report card sources suggest significant opportunities for improvement, we will examine the data and, if validated and appropriate, implement performance improvement activities to be sure our patients receive the best possible care.

Hospital & Physician Quality/Safety Report Card Web Sources					
Web Sources	Web Address	Publicly Available Reports*	Based on Direct Review of Healthcare records	Based on Electronic Billing Information	Based on Subjective Reporting by hospitals/peers
AHRQ	http://qualityindicators.ahrq.gov			✓	
Americas Top Doctors	http://www.americastopdoctors.com				✓
Best Doctors	http://www.bestdoctors.com	✓			✓
CMS	http://www.hospitalcompare.hhs.gov	✓	✓		
Health Care Choices and Consumers' Checkbook	http://www.checkbook.org/			✓	
Health Data Insights	http://www.selectqualitycare.com			✓	
HealthGrades	http://healthgrades.com	✓		✓	
JCAHO	http://www.jcaho.org/index.htm	✓	✓		
Leapfrog Group	http://www.leapfroggroup.org/	✓		✓	✓
Select Quality Care	http://www.selectqualitycare.com			✓	
Subimo	http://www.subimo.com			✓	
Total Benchmark Solution, LLC	http://www.totalbenchmarksolution.com	✓	✓ (based on CMS report)		
US News Best Hospitals Report	http://www.usnews.com/usnews/health/best-hospitals/tophosp.htm	✓			✓

*Publicly available reports: web based hospital and/or physician report available without registration or subscription

Periodic Briefing on Health Care Market Issues

Attachment II

Market Watch

Health Care Report Cards: Slow Learners?



Exhibit 1



Exhibit 2

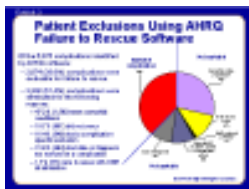


Exhibit 3

In June 2003, a UHC-published report criticized the proliferation of health care report cards that were confusing the market with conflicting ratings of hospitals and physicians. Relying on administrative data that lacked both the timeliness and the information necessary to differentiate clinical performance, report cards often excluded many better performers and included some of the worst performers in their honor rolls. During the intervening years, some of the rating methodologies have changed, but the most important criticism of 2 years ago still applies: Health care report cards are misdirecting consumers who rely on faulty ratings.

A consistent criticism of health care report cards is their inability to differentiate between providers in a head-to-head comparison. In the March 13, 2002 issue of the *Journal of the American Medical Association* (JAMA), Krumholz et al concluded that HealthGrades.com, an Internet-based report card, used ratings that provided “little meaningful discrimination between individual hospital performance in a manner sufficient for making informed hospital choices.”¹

The JAMA study found that patients treated in higher-rated hospitals were, on average, more likely to receive aspirin and beta-blockers and had lower risk-adjusted mortality rates for acute myocardial infarction than patients treated in lower-rated hospitals. On the surface, such statistics appeared to support the 5-star rating system adopted by the report card. A more careful analysis, however, showed significant variation within rating categories, with substantial overlap between the high and low ranges of each category. When hospitals from different rating categories were compared individually, risk-adjusted **mortality rates were either comparable or even better in the lower-rated hospitals in more than 90% of comparisons.**

Intra-rating variation is the Achilles’ heel of health care report cards,

and the problem is not limited to HealthGrades.com. In an investigation of the *U.S. News & World Report* rankings, Chen et al found a similar issue.² Average risk-adjusted mortality rates for top-ranked hospitals were slightly lower than for lower-ranked hospitals, but wide variation within rating categories meant that in head-to-head comparisons, lower-rated hospitals would commonly have superior-quality performance ([Exhibit 1](#)).

Attachment II

Generalizations based on averages that fail to accurately account for individual provider performance can misdirect consumers. This problem is not new to health care report cards. In 2003, UHC reported that the minimum-volume threshold adopted by the Leapfrog Group for coronary artery bypass graft (CABG) excluded a greater number of better performers than it included. The Leapfrog threshold excluded two thirds of the best performers (the top 10% of all hospitals studied), while nearly half of all CABG procedures carried out at poorer performers were at hospitals that met the Leapfrog volume threshold ([Exhibit 2](#)). It is troubling to note that health care report cards have continued to promulgate individual provider ratings that are inconsistent with more reliable clinical data, misleading unsuspecting consumers.

The latest example of the questionable use of publicly available administrative data sets comes from HealthGrades.com and was reported in the May 2, 2005, issue of *Modern Healthcare*. After analyzing more than 38 million Medicare discharge abstracts from 2001 to 2003, HealthGrades.com asserted that patient safety was getting worse in American hospitals and that more than 240,000 patients died as a result of medical mistakes during the study period. Unfortunately, the statistical tools used to estimate these errors came from an Agency for Healthcare Research and Quality (AHRQ) software program that was never validated for estimating error rates but instead was designed as a screening tool to identify potential cases for further review.

Nearly 200,000 of the 241,000 deaths in the HealthGrades.com analysis were classified as “failure to rescue,” which is the term for a fatality resulting from an avoidable complication of hospitalization. A recent UHC study of this AHRQ metric found that it has a false-positive rate of 60% ([Exhibit 3](#)). Cases from UHC’s Clinical Data Base that were initially identified by the software as failure to rescue were subsequently excluded upon medical record review for a number of reasons, most commonly because the condition in question was present at admission and therefore a comorbidity, not a complication. The remaining 40% of the failure to rescue cases were properly identified by the software and formed the basis of the UHC benchmarking study. AHRQ makes no claims about its software’s ability to estimate error rates, and it is unfortunate that a potentially useful screening device was misused in such a public way by HealthGrades.com, resulting in at least a two-fold overestimation of deaths from medical error.

Careless use of administrative data undermines the presenter’s credibility and can lead to distracting arguments over risk

adjustment. Done properly, risk adjustment assists in the interpretation of the outcome in question, such as mortality, by estimating the effect of confounding variables. However, all risk-adjustment methods are not alike. Most risk-adjustment methods correct for age and gender and some comorbidities, but many do not account for socioeconomic status, admission source (transfer versus direct admit), or admission status (emergency versus elective). A soon-to-be-published UHC study involving patients from an intensive care unit benchmarking project indicates that an adjustment for transfer status increased expected mortality by 60%; this variable is rarely included in publicly available databases. Neither socioeconomic status nor transfer status is in the HealthGrades.com models, placing at a disadvantage those organizations that see a disproportionate number of patients who either are transfers or are in lower socioeconomic groups. In addition to being inflammatory, the misuse of administrative data (especially when packaged and communicated directly to consumers rather than through peer-reviewed processes) misleads the market.

It is interesting to revisit UHC's treatise on this topic, published in June 2003. The report concluded with an observation that is equally applicable today:

At best, the market lacks a trusted voice with respect to risk-adjusted quality outcomes measurement. At worst, some health care report cards steer consumers away from the best providers and toward providers with substandard outcomes. Until consensus emerges regarding quality measurement methodologies, it is too soon to give passing grades to most health care report cards.

For more information, contact [UHC staff](#).

References

1. Krumholz HM, Rathore SS, Chen J, Wang Y, Radford MJ. Evaluation of a consumer-oriented Internet health care report card: the risk of quality ratings based on mortality data. *JAMA*. 2002;287(10):1277-1287.
2. Chen J, Radford MJ, Wang Y, Marciniak TA, Krumholz HM. Do "America's best hospitals" perform better for acute myocardial infarction? *N Engl J Med*. 1999;340(4):286-292.