The Relationship Between Choice of Outcome Measure and Hospital Rank in General Surgical Procedures: Implications for Quality Assessment

JEFFREY H. SILBER,*‡ PAUL R. ROSENBAUM,|| SANKEY V. WILLIAMS,*‡ RICHARD N. ROSS† and J. SANFORD SCHWARTZ*‡

* The Leonard Davis Institute of Health Economics; † Division of Oncology; ‡ Division of General Internal Medicine; § Department of Pediatrics in the School of Medicine; ¶ Department of Medicine in the School of Medicine; || Department of Statistics in the Wharton School, The University of Pennsylvania, Pennsylvania, USA

INTRODUCTION

As the American health care system moves toward greater use of managed care, tools to measure quality of care are gaining importance. Since patient outcomes reflect (at least in part) quality of care, the tools of outcomes research are increasingly being utilized by both the public and private sectors to help evaluate "quality". Hospital quality rankings based on the death rate have commonly been used [1-7]. Since death is rare for many types of conditions, the death rate may have inadequate power to distinguish among hospitals for many procedures and conditions; consequently, alternative outcome measures have been suggested [8,9]. Some analysts and policy makers have proposed using complications as a measure of quality [10-17].

If adjusted complication rates are valid measures of hospital quality of care and if the aspects of hospital quality which prevent death also prevent complications, then we should observe a high correlation between hospital quality rankings based on the death rate and those based on the complication rate. Should we observe a low correlation, then we must hypothesize either that death or complication rates do not measure quality of care, or that these rates measure very different aspects of hospital quality.

For this study, we hypothesized that hospital rankings based on severity-adjusted death and failure-to-rescue rates (where "failure" is defined as a death following a complication [18-21]) are poorly correlated with rankings based on the complication rate, since factors which contribute to complication rates are different from those that contribute to death and failure rates [18,19]. To test this hypothesis, we ranked 142 hospitals by adjusted death rates, complication rates and failure rates and then examined the correlations among these rankings.

METHODS

Hospital and patient selection

We obtained the 1991 MedisGroups National Comparative Data Base, comprised of 855 548 patient admis-
sions throughout the United States. This data base serves as the reference population for MedisGroups reports [22]. The method of data collection has been described in detail elsewhere [22-31]. The reliability and validity of MedisGroups data and the MedisGroups severity score compare favorably to other data collection systems commonly used for severity adjustments [24,27-30]. The MedisGroups data base includes admissions to each of 114 hospitals during consecutive time periods in 1990. For validation purposes, we obtained the 1992 MedisGroups National Comparative Data Base, which included 743,964 patients in 108 hospitals. There are at least four months of data per hospital. The hospitals are chosen as a representative sample of MedisGroups clients, but are not a random sample of hospitals in the United States.

We wished to focus our analysis on common adult surgical procedures often performed by a general surgeon at most hospitals throughout the United States in order to evaluate the quality of each hospital's general surgery program. Therefore, we limited our analysis to patients in Medical Diagnostic Category (MDC) 6 (diseases of the gastrointestinal tract) and MDC 7 (diseases of the hepatobililiary system excluding liver transplantation) as well as those in MDC 9 who underwent breast biopsy and mastectomy procedures. These procedures are commonly performed by general surgeons throughout the United States and require non-complex, standard technologies (unlike more complex procedures such as coronary artery bypass surgery). Consequently, most hospitals that support adult operating room facilities are able to perform these procedures.

There were 41,355 patient admissions in the 1991 National Comparative data set and 33,292 patients in the 1992 National Comparative data set. We excluded hospitals that did not record any MedisGroups events (defined below), and hospitals that served primarily psychiatric, pediatric, or rehabilitation patients as defined in the American Hospital Association Annual Survey. We also excluded hospitals with sample sizes below 100 in our data base. This left 142 hospitals available for analysis, with 45 hospitals represented in both the 1991 and 1992 data sets. The same MedisGroups scoring system was used for 1991 and 1992 data sets.

Definition of complication

MedisGroups collects up to three sets of clinical variables called "key clinical findings" (KCFs) during the hospitalization of adult surgical patients. A description of the data collection system has been reported by Iezzoni et al. [27] and Steen et al. [31].

We defined two types of complications for this study:

Definition 1: A complication is a clinical finding or event that was (1) present during or after surgery, or (2) present on day 3 or later (except cardiac emergencies which could occur at any time during the hospitalization), or (3) present in association with a diagnostic or therapeutic procedure at any time during the hospitalization; and in addition was (4) serious enough to have a potentially adverse effect on the patient's outcome. Complications stemming from diagnostic tests prior to surgery are also included as part of this definition. A complete list of complications is provided in Table 1. Definition 1 tends to overestimate the "true" number of patients with complications, since abnormal findings noted later in the hospitalization are counted as complications despite the fact that such post-operative findings may have initially been comorbidities on admission. Thus, we developed Definition 2. Definition 2: The same as Definition 1, except that if a finding was present on admission (through hospital day 2), that finding cannot be used to code for a complication at a later time during the hospital stay. Definition 2 tends to underestimate the "true" number of patients with complications, since it does not allow for a worsening of the same finding to be identified as a new complication. We report final results using both definitions to allow for judgement on the sensitivity of the results with respect to alternative definitions of complications. Since both definitions yielded similar results, Definition 1 is used in this report unless otherwise specified.

We constructed an initial list of complications based on clinical knowledge and our previously published list, which we developed from a Health Care Financing Administration/MedisGroups data base using 1985 admissions [18]. We then modified the list to include selected complications noted in the 1991 National Comparative Data Set as MedisGroups events. No distinction was made between patients with multiple complications and patients with a single complication when tabulating the number of patients with a complication.

A description of each complication is provided in Table 1. Percentages do not add up to 100% since patients may have more than one complication. We formed the definitions of complications using only the 1991 data and then applied it to the 1992 data for validation purposes. There were no changes in definitions based on 1992 data. The complication rates for the 1991, 1992 and combined 1991 and 1992 data sets are described in Table 2. The robustness of these definitions is reflected by observing the stability of complication rates between the 1991 and 1992 data sets.

Statistical methods

Using the 1991 National Comparative data set, we created separate multiple logistic regression models whose outcomes were in-hospital death, complication, or failure-to-rescue following a complication (i.e. death after a complication). For each of the three models we ranked hospitals by the difference between the hospital's observed and expected number of outcomes, divided by the total number of patients at risk in each hospital.
**TABLE 1. Definitions and rates of complications and failures (deaths following complications) in 74,647 patients found in the 1991 and 1992 MedisGroups (US) National Comparative Data Base**

<table>
<thead>
<tr>
<th>Complication</th>
<th>Definition (based on key clinical findings)</th>
<th>No. of patients (%)</th>
<th>No. of deaths (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cardiac event</td>
<td>EKG: strain pattern; ischemia; 2nd degree block; sino-atrial dysfunction; bradycardia (&lt; 50 bpm); multifocal atrial tachycardia; pacemaker malfunction</td>
<td>585 (0.8)</td>
<td>109 (18.6)</td>
</tr>
<tr>
<td>2. Cardiac emergency</td>
<td>Arrest with/without resuscitation; acute myocardial infarction; atrio-ventricular dissociation; ventricular tachycardia; ventricular standstill; or fibrillation</td>
<td>778 (1.04)</td>
<td>412 (53.0)</td>
</tr>
<tr>
<td>3. Congestive heart failure</td>
<td>As noted by presence of S3 gallop and cardiomegaly by imaging study or pulmonary edema</td>
<td>1192 (1.60)</td>
<td>346 (28.9)</td>
</tr>
<tr>
<td>4. Hypotension</td>
<td>Systolic BP &lt; 60</td>
<td>1698 (2.26)</td>
<td>896 (53.1)</td>
</tr>
<tr>
<td>5. Pulmonary embolus</td>
<td>Ventilation-perfusion scan abnormality</td>
<td>95 (0.13)</td>
<td>15 (15.8)</td>
</tr>
<tr>
<td>6. Deep vein thrombosis</td>
<td>On physical examination</td>
<td>72 (0.10)</td>
<td>18 (25.0)</td>
</tr>
<tr>
<td>7. Stroke</td>
<td>On MRI or CT scan, or acute motor/speech deficit on physical examination</td>
<td>137 (0.18)</td>
<td>55 (40.1)</td>
</tr>
<tr>
<td>8. Coma</td>
<td>As noted on examination</td>
<td>752 (1.00)</td>
<td>624 (83.0)</td>
</tr>
<tr>
<td>9. Seizure</td>
<td>As documented in chart</td>
<td>160 (0.21)</td>
<td>70 (43.8)</td>
</tr>
<tr>
<td>10. Psychosis</td>
<td>As documented in chart</td>
<td>234 (0.31)</td>
<td>22 (9.40)</td>
</tr>
<tr>
<td>11. Pneumonia</td>
<td>By chest X-ray or CT scan</td>
<td>2772 (3.71)</td>
<td>606 (21.9)</td>
</tr>
<tr>
<td>12. Pneumothorax</td>
<td>By chest X-ray or CT scan</td>
<td>191 (0.26)</td>
<td>52 (27.2)</td>
</tr>
<tr>
<td>13. Respiratory compromise</td>
<td>Requiring ventilation</td>
<td>2358 (3.15)</td>
<td>710 (30.1)</td>
</tr>
<tr>
<td>14. Internal organ damage</td>
<td>As noted in chart</td>
<td>1957 (2.6)</td>
<td>229 (11.7)</td>
</tr>
<tr>
<td>15. GI/intestinal bleed</td>
<td>As noted in chart</td>
<td>3223 (4.31)</td>
<td>484 (15.0)</td>
</tr>
<tr>
<td>16. Sepsis</td>
<td>Documented positive blood culture</td>
<td>322 (0.43)</td>
<td>110 (34.2)</td>
</tr>
<tr>
<td>17. White blood count above</td>
<td>&gt; 16,000 cells/cmm</td>
<td>3745 (5.02)</td>
<td>736 (19.7)</td>
</tr>
<tr>
<td>18. Deep wound infection</td>
<td>As noted in chart</td>
<td>1169 (1.56)</td>
<td>101 (8.64)</td>
</tr>
<tr>
<td>19. Renal dysfunction</td>
<td>Creatinine above 1.7, or oliguria</td>
<td>2158 (2.89)</td>
<td>804 (37.3)</td>
</tr>
<tr>
<td>20. Anesthesia event</td>
<td>Either bradycardia, hypotension, prolonged anesthesia, respiratory compromise</td>
<td>263 (0.35)</td>
<td>40 (15.2)</td>
</tr>
<tr>
<td>21. Transfusion</td>
<td>Units transfused after day 3 or transfusion reaction</td>
<td>109 (0.14)</td>
<td>15 (13.6)</td>
</tr>
<tr>
<td>22. Gangrene</td>
<td>As noted in chart</td>
<td>33 (0.04)</td>
<td>6 (18.1)</td>
</tr>
<tr>
<td>23. Obstruction</td>
<td>On abdominal films/X-ray studies</td>
<td>3720 (4.98)</td>
<td>332 (8.92)</td>
</tr>
<tr>
<td>24. Return to surgery</td>
<td>As noted in chart</td>
<td>700 (0.94)</td>
<td>104 (14.6)</td>
</tr>
<tr>
<td>25. Patient injury</td>
<td>As documented in chart</td>
<td>491 (0.66)</td>
<td>36 (7.30)</td>
</tr>
<tr>
<td>26. Do not resuscitate</td>
<td>If first written in chart after day 3</td>
<td>1121 (1.50)</td>
<td>731 (65.2)</td>
</tr>
</tbody>
</table>

* The denominator for this percentage is the total number of patients equal to 74,647.
† The denominator for this percentage is the number of patients with the specific complication.

Significance tests for the resulting ranking statistics were computed using the variance estimate reported by Haberman [32,33]. The correlations between hospital ranks for each pair of the three outcomes were computed using Spearman rank coefficients. A more detailed account of model construction is published elsewhere [19].

We validated the models constructed on the 1991 National Comparative data set by applying them to the 1992 National Comparative data set. We first constructed models for our three outcome variables based on the 1991 data. Next, we constructed models in an identical manner based on the 1992 data. We then applied the 1992 models to the 1991 hospital experience and re-ranked the 1991 hospitals. The correlation between 1991 hospital rankings based on the 1991 model and the same patients ranked by the 1992 model was 0.91 (P < 0.0001) for the death rate, 0.97 (P < 0.0001) for the complication rate, and 0.97 (P < 0.0001) for the failure rate. Since these rankings were well correlated, we were reassured that our model coefficients were reasonably stable across years; thus we constructed our final models using the combined 1991–1992 data sets.

Independent variables used in the final model were as follows: age; sex; emergency admission; MedisGroups

**TABLE 2. Death, complication, and failure rates for 1991, 1992 and combined MedisGroups (US) National Comparative Data Base**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Death</td>
<td>890</td>
<td>7203</td>
<td>1562</td>
</tr>
<tr>
<td>No death</td>
<td>29</td>
<td>3323</td>
<td>41</td>
</tr>
<tr>
<td>Death rate</td>
<td>919/41,355 = 2.2%</td>
<td>684/33,292 = 2.1%</td>
<td>1603/74,647 = 2.2%</td>
</tr>
<tr>
<td>Complication rate</td>
<td>8093/41,355 = 19.6%</td>
<td>7007/33,292 = 21.1%</td>
<td>15100/74,647 = 20.2%</td>
</tr>
<tr>
<td>Failure rate</td>
<td>8090/8093 = 11.0%</td>
<td>672/7007 = 9.6%</td>
<td>1562/15,100 = 10.3%</td>
</tr>
</tbody>
</table>
severity score; history of: diabetes, cancer, stroke, angina, coronary artery bypass graft surgery, congestive heart failure, chronic obstructive pulmonary disease; DRGs 150, 157, 257, 259, 152, 166, 171, 154, 201, 148, 146, 195, 193, 197, 191. The DRGs were further subdivided into homogeneous groups by principal procedures including ICD9 codes 437, 4291, 4562, 4573, 485, and 5122. Interactions between pairs of variables were also included in the final model if they were significant at the two-sided 5% level in any model (death, failure, or complication) with the significance level based on the Bonferroni correction.

All three outcome models were highly significant and produced c-statistics of 0.92, 0.87, and 0.78 for death, complication and failure respectively. The full models are available upon request and are reported elsewhere [19,21].

**Ranking methodology**

In order to ensure the stability of correlation coefficients between hospital outcome statistics, we ranked only hospitals with at least 100 patient observations. To rank hospitals by outcome, we used the statistic (O—E)/N where O is the observed number of patients at a hospital with a specific outcome, E is the expected number of outcomes at that hospital as determined from the logistic regression model, and N is the number of patients at the hospital who underwent the procedure of interest. There is one ranking statistic (O—E)/N for each hospital for each of the three outcome measures. For example, using the death rate, if a hospital had an (O—E)/N of 3%, this would imply that the hospital has a death rate that is 3 percentage points above what would have been expected at the “typical” hospital given a case mix equivalent to the hospital in question. A hospital with an observed 5% death rate, but with an expected death rate of 2%, based on the experience of other hospitals with similar patients, would have an (O—E)/N of 3%. The (O—E)/N measures for complication and failure can be interpreted in the same manner [19,21,32,33].

**RESULTS**

Of 74,647 admissions, there were 15,100 patients with complications and 1,603 deaths. The overall death rate was 2.15%, the complication rate was 20.23% and the failure rate was 10.34%. Of patients who died in hospital, 97.44% had a preceding complication from the list of complications shown in Table 1. Individual hospital death rates ranged from 2.2 percentage points below to 1.8 percentage points above that expected. For the complication rate the range was 10.4 percentage points below to 9.5 percentage points above that expected and for the failure rate the range was −13.8 to +9.5 percentage points from that expected.

There were 142 hospitals that had greater than 100 patient observations and therefore were eligible for ranking. Figures 1–3 display the rank correlations for each pair of outcome statistics among these 142 hospitals. We found a low correlation between hospital ranks based on complication rates and hospital ranks based on both death or failure but a higher correlation between death and failure rate rankings. Only 4.3% of the variation in hospital rankings using the complication rate could be explained by rankings based on the death rate. The correlation between hospitals ranked by the death rate and those ranked by the complication rate was 0.21 (P = 0.013). A similarly low correlation was present between the complication and failure rate rankings, r = −0.09 (P = 0.174). A substantially higher correlation was observed between the death and failure rate rankings, r = 0.90 (P < 0.001).

These results were stable across definitions of complications. When Definition 2 was examined, a definition that would not count a complication if the key clinical finding used to identify that occurrence was noted on the first or second day of the hospital stay, the final results were very similar, with a correlation between death and complication rankings of 0.25 (P = 0.002); the correlation between failure and complication rates was −0.13, P = 0.12; and the correlation between death and failure was 0.85 (P = 0.0001). Furthermore, if the definition of a complication included all patients who died without a complication noted in Table 1, and counted such patients as having an “unidentified” complication, we again obtained almost exactly the same results; a correlation...
Relationship between outcome measure and hospital rank

FIGURE 2. Correlation between 142 hospital rankings using the adjusted failure rate versus the adjusted complication rate. The Spearman correlation coefficient was $-0.090$ ($P = 0.287$).

between death and complication rankings of 0.21, $P = 0.010$; a correlation between failure and complication rankings of $-0.11$, $P = 0.213$; and a correlation between the death and failure rate rankings of 0.90, $P = 0.0001$.

Finally, when we omitted “Do Not Resuscitate” as a definition of a complication, our results remained unchanged.

DISCUSSION

The primary purpose of this study was to determine how complication- and mortality-based outcome measures relate to each other and, in so doing, to shed light on the controversy surrounding the proper choice of outcome for measuring quality of care. In this study we developed three models of outcome, all with good predictive ability. We used two definitions of complications and two methods to evaluate the outcome of mortality without complications. All methods lead to similar results: low correlations between complication and mortality measures. Such results are consistent with a number of studies, as presented in Table 3. The correlation between the hospital rank using the adjusted death rate and the hospital rank using the adjusted complication rate was less than 25% in six of eight studies listed.

We recently reported a low correlation between hospital rankings based on adjusted complications and death rates in the complex procedure of coronary artery bypass graft surgery [20]. Earlier work by Flood et al. [2] and DesHarnais et al. [10] pointed to low correlation between mortality and complication outcomes. More recently, Iezzoni et al. [12,13] have described a lack of correlation between death and complication rankings in a number of conditions, and higher complication rates seemed to be associated with a number of hospital features that would suggest better quality of care, e.g. teaching hospitals, hospitals performing open heart surgery and hospitals providing magnetic resonance imaging (MRI) [12]. Brailer et al. [16], using comorbidity adjusted complication risk (CACR) has found similar results, with correlations of 15% and 18% between mortality and complication rates. Some may conclude that these observations of uncorrelated hospital death and complication ranks are consistent with the existence of different, uncorrelated, “dimensions” of hospital quality of care.

Can uncorrelated “dimensions” of quality explain the low correlation between death and complications?

There is no a priori reason to believe that hospital characteristics that may prevent complications are the same hospital characteristics that prevent death after complications. Indeed, such components of quality may fall along different “dimensions”. Skills needed to prevent complications, if such skills exist, may be different from those skills needed to rescue patients after complications, or prevent minor complications from escalating into major ones. Prior to a complication there may be a relatively small number of caregivers involved in the care of a patient. For a routine surgical procedure this number may predominantly consist of the patient’s nurses, surgeons and anesthesiologists. After the complication, a large number of caregivers may become involved with the patient’s treatment, depending on the
type of complications and the skills required to treat them. Since there are many different types of complications, the process of treating complications will often reflect the depth and breadth of the hospital's support systems. The depth and breadth of the hospital may not be reflected in the same manner when considering the initiation of a complication as when considering the care of a patient after a complication occurs. For example, in earlier work looking at cholecystectomy and transurethral prostatectomy [18], we reported that the percentage of board certified anesthesiologists on a hospital's anesthesia staff was strongly associated with better survival rates after complications (lower failure-to-rescue rates), yet the factor was not associated with lower complication rates.

Hence, one reason for the low correlation between death and complication rates may be that hospitals differ in their ability to prevent both complications and death after complications (failure-to-rescue). However, this would not seem to explain the low correlations displayed in Table 3. To explain this factor, one must believe a rather implausible scenario — that almost no hospital characteristics that prevented deaths were associated with other characteristics that prevented complications [19]. A more plausible scenario is that hospital factors that contribute to lower mortality rates also contribute, in part, to lower complication rates. If this scenario holds, then we are forced to look elsewhere for the cause of low correlation.

Had studies consistently shown high correlation between death and complication based hospital rankings, there would be little doubt that complication based measures would become a useful tool to evaluate quality of care. However, the low correlation between death and complication based rankings suggests a deeper problem, possibly related to data quality. Believers in the complication rate may argue that a different "dimension" of quality is being measured. Critics of the complication rate argue that it is misleading, due to a lack of uniform measurement and recording of complications and difficulty with the severity adjustments of complications. If complications were indeed a measure of quality but a measure along a different dimension of quality than mortality, then we would have expected a high, but not perfect correlation.

It is difficult for us to imagine a dimension of hospital quality that accounts for a hospital's complication rate but is almost entirely unrelated to its mortality rate. In contrast, note the relationship of mortality to failure in the present study. We observed discrepancies between death and failure rate rankings, but these were small in comparison to those based on complications. The reported correlation of 0.90 implies that 81% of the variation in failure rates can be explained by the death rate. It would appear from this high correlation that both death and failure are measuring similar characteristics. However, inspection of Fig. 3 reveals considerable differences between death and failure rankings for some hospitals and these differences may be clinically meaningful. Had the complication rate performed in a similar manner, we would have gained confidence that complications were related to quality. As it stands, the low correlations observed would seem to imply an etiology rooted in problems of data, not uncorrelated "dimensions" of quality.

Can inadequate data explain the low correlation between death and complications?

Inadequate data, whether inconsistent ascertainment or recording of severity adjustment data, or unreliable complication event data across hospitals, is a serious problem for those preferring to use complications to determine quality of care. A low correlation between complication and death rate measures may relate to inadequate, asymmetrical, severity of illness correction. By asymmetrical, we mean that adjustments are adequate for death and failure but inadequate for complications. There is some evidence to support this asymmetry. The complication rate appears to be more influenced by patient severity and case mix than the failure rate in simple surgical procedures. Our earlier work reported differences in the contribution of hospital and patient characteristics to outcome across outcome measures, with the failure rate being more influenced by hospital characteristics than the complication rate, and the complication rate being more influenced by patient characteristics than the failure rate [18,19]. Such a result suggested that: (1) ranking of hospitals by adjusted complication rates would be poorly correlated with

![FIGURE 3. Correlation between 142 hospital rankings using the adjusted death rate versus the adjusted failure rate. The Spearman correlation coefficient was 0.90 ($P < 0.001$).](image)
rankings performed using the failure rate and (2) the complication rate may be highly influenced by patient characteristics and hence may not adequately reflect hospital quality, but instead it may reflect a hospital population's case-mix and severity (a very undesirable property for a quality measure).

If low correlation exists between death and complication rates, which measure should be used?

Increased statistical power is a strong motivating force for using complication measures rather than mortality based measures. All agree that there are more complications than deaths and hence power may be improved when using the complication rate. The debate centers on the information gained from the complication rate. A measure with increased power is not preferable if that measure does not reflect hospital quality of care, but rather reflects hospital case-mix, patient severity of illness, or other factors which may mislead decision makers. The question still remains: if a hospital has low death and failure rates but a high complication rate, which measure should be used to guide channelling decisions for patients?

Other studies may provide some insight to aid in answering this question. In earlier work on the same surgical patients reported here, we have shown that some hospital characteristics that on face value may improve quality of care are associated with increased complication rates, but are associated with lower failure-to-rescue and death rates [19]. For example, we found that the hospital characteristics of higher rates of board certification rates among anesthesia staff, and higher nurse-to-bed ratios, seemed to be well correlated with lower death and failure rates, yet did not correlate with lower complication rates. We also found similar correlations among ranking measures for coronary artery bypass surgery [20]. Hospitals performing bone marrow transplant, with MRI facilities, or with accredited residency training programs had higher than average adjusted complication rates, but either average or lower than average adjusted death and failure rates.

From our observations both on general surgical procedures presented here and on the more technologically intensive CABG procedures, we have serious concerns about using complications to compare quality of care across hospitals. That is not to say that monitoring complications inside hospitals should be abandoned. On the contrary, we believe this is an essential part of any excellent surgical or medical program. However, following an institution's complication rates over time is very different from using them to conduct comparisons across institutions, where differences across hospitals in patient monitoring protocols, data collection systems and chart documentation may lead to incorrect inferences concerning quality of care.

CONCLUSIONS

We conclude that for general surgical procedures, the complication rate produces very different hospital rankings than does the death or failure rate and the death and failure rate rankings are far from perfectly correlated. Given (1) the low correlation between the complication rate and mortality-based measures; (2) the fact that mortality measures are often associated with plausible hospital characteristics, we believe that the risk of incorrect inference derived from the complication rate may outweigh the benefit obtained in terms of gains in statistical power. We therefore advise that the complication rate should be interpreted with great caution and should not be used in isolation when assessing quality of care for general surgical procedures across hospitals.

Acknowledgements: This research was funded by the Agency for Health Care Policy and Research grant number RO1-HS06560.

REFERENCES