Failure-to-Rescue

Comparing Definitions to Measure Quality of Care

Jeffrey H. Silber, MD, PhD,†‡§ Patrick S. Romano, MD, MPH,¶‖ Amy K. Rosen, PhD,** Yanli Wang, MS,* Orit Even-Shoshan, MS,*§ and Kevin G. Volpp, MD, PhD§††‡‡*

Objectives: Use of failure-to-rescue (FTR) as an indicator of hospital quality has increased over the past decade, but recent authors have used different sets of complications and deaths to define this measure. This study examines the reliability and validity of different FTR measures currently in use.

Research Design: We studied 3 definitions: (1) "original" FTR (using all deaths); (2) FTR-N, a "nursing sensitive" definition that uses only specific complications and deaths; and (3) FTR-A [another restricted definition of FTR used by Agency for Healthcare Research and Quality (AHRQ) for analyzing Healthcare Cost and Utilization Project (HCUP) data]. Each FTR measure was applied to 403,679 general surgical patients across 1567 hospitals reported in 1999–2000 Medicare MEDPAR data.

Results: Although FTR used all deaths, FTR-N and FTR-A definitions omitted 49% and 42% of deaths, respectively. Reliability was better for FTR than FTR-A or FTR-N (ρ = 0.32 vs. 0.18 vs. 0.18, respectively).

Validity: Hospitals ranked by adjusted mortality were highly correlated with FTR (Kendall's τ = 0.83) and less correlated with FTR-A (τ = 0.43) and FTR-N (τ = 0.41). Adjusting for patient characteristics, all FTR measures showed strong associations with bed-to-nurse ratio, nursing mix, teaching status, and hospital size; however, hospital "high technology" was not as well associated with FTR-N.

Conclusions: For general surgery, more limited definitions used by FTR-N and FTR-A omit over 40% of deaths, display less reliability, and may have more questionable validity than the original FTR measure. We encourage analysts to use the original FTR definition that uses all deaths when analyzing hospital quality of care.

Key Words: failure-to-rescue, mortality rates, quality of care, patient safety indicators, outcomes research, Medicare claims data

(*Med Care* 2007;45: 918–925)

Failure-to-rescue (FTR) is a measure of hospital quality of care that has been growing in popularity. The measure has been used in numerous medical[1–13] and nursing publications,[14–23] adopted as a patient safety indicator by the Agency for Healthcare Research and Quality (AHRQ),[10] and used by various hospital groups performing hospital quality ratings, such as health grades.[24] However, the original definition of FTR[1–3,13] has been modified, and several different versions are currently in use, resulting in some confusion and problems with interpretation.[7,10,25] This article will describe the theory behind the original FTR measure, and then compare the FTR measure to recent versions, which we will define and describe. These versions will be represented as FTR-N ("N" for "nursing sensitive" set of complications,[7]) and FTR-A ("A" for AHRQ-defined patient safety indicators[10]). Because these variants of FTR represent subsets of the FTR measure, we will also define the complement set measure of FTR-N (FTR-N^C) and the complement set of FTR-A (FTR-A^C), allowing us to test whether the information not used in more limited definitions contains any important information regarding quality of care. For each measure we will compare reliability and validity. Finally, recommendations will be made concerning the appropriate measures to be used for assessing hospital quality in surgical patients.

Origins of FTR

The measure FTR was first developed by Silber et al using a special data set from the Health Care Financing Administration.[1] That data set combined administrative claims data with MedisGroups chart-based data.[26,27] Using a larger MedisGroups National Comparative Data Set of general surgical procedures, the properties of FTR were studied in a series of articles by Silber et al.[2,3,13] Subsequently, the same group has applied the measure to both state[8,9] and Medicare claims data.[5,6]

From the *Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; †Departments of Pediatrics; Anesthesiology and Critical Care, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; ‡Department of Health Care Systems, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania; §Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania; ¶Division of General Medicine and ‖Center for Health Services Research in Primary Care, University of California Davis School of Medicine, Sacramento, California; **Department of Health Services, Boston University School of Public Health, Boston, Massachusetts; ††Division of General Internal Medicine, Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; and ‡‡Veteran's Administration Hospital, Philadelphia, Pennsylvania.

Supported by a grant R01 HL082637 from the NHLBI and VA HSR&D IIR 04.202.1.

Reprints: Jeffrey H. Silber, MD, PhD, Center for Outcomes Research, The Children's Hospital of Philadelphia, 3535 Market Street, Suite 1029, Philadelphia, PA 19104. E-mail: silber@email.chop.edu.

ISSN: 0025-7079/07/4510-0918

Theory

The motivation behind the development of traditional FTR was based on 2 questions. The first was an empirical question—suppose hospitals were ranked by adjusted mortality and adjusted complication rates. Would these rankings be highly correlated? The answer is rather surprising—there is generally poor correlation or no correlation in most analyses.[3,28–32] Second, suppose 2 hospitals had identical adjusted mortality rates but different adjusted complication rates. Would one prefer care at the hospital with the higher or lower complication rate? If one believes that complications are predominantly driven by patient characteristics, then one may decide to choose the hospital with the higher complication rate, as it achieved an equivalent mortality rate with a sicker population of patients. So there is an empirical question to ask—are adjusted complication rates more related to hospital or patient factors? This has been looked at in a number of ways—and the evidence to date suggests that complication measures are less sensitive to hospital characteristics, after adjusting for severity of illness, than mortality-based measures.[1,13,30,33] This is an underlying assumption of FTR theory—complications are undesirable outcome measures because they reflect underlying patient severity and diagnosis coding more than they reflect hospital care. Instead, a hospital's quality is put to the test when a patient develops a complication, and whether a patient is salvaged after a complication will be a function of the care delivered by the hospital and its knowledge base, depth, and facilities. Thus, "good" hospitals will rescue patients by identifying complications quickly and treating them aggressively, resulting in lower FTR. Although many "failures," just like deaths, are often not preventable, we have argued that FTR may be a better measure for comparing hospital quality because of better severity adjustment properties,[1] and because of its focus on hospital actions.[13] By studying a population of patients who, by definition, have already developed a complication, the specifics of severity of illness adjustment becomes less important in failure rate analyses, because all patients have experienced complications and thus are more uniformly ill.[1,13]

METHODS

Patient Data

We obtained all US Medicare inpatient claims for general surgical admissions for the period July 1, 1999 to June 30, 2000 [Diagnosis-related Groups (DRG) procedure codes are listed in Appendix A which can be found on the *Medical Care* website, www.lww-medicalcare.com]. For each patient admitted, we selected the first admission during the study period. Data were linked to the 2000 American Hospital Association Annual Survey to better describe hospital characteristics. There were a total of 1467 hospitals and 403,679 patients. We included patients between 65 and 90 years of age.

Definitions

FTR is defined as the probability of death after a complication. To understand FTR, it helps to display the definition of the probability of death as the sum of 2 quantities shown below:

$$P(d) = P(d|c)*P(c) + P(d|\text{no } c)*P(\text{no } c) \quad (1)$$

Equation 1 states that the probability of dying is equal to the probability of dying given a complication multiplied by the probability of a complication plus the probability of dying without a complication multiplied by the probability of not developing a complication. In the case of most surgical procedures, one can envision that the second term approaches zero because one is unlikely to die without some complication, given that the surgery was performed. Hence, we can think of the probability of death as equal to the probability of dying given a complication multiplied by the probability of developing a complication. The probability of dying given a complication is defined as failure-to-rescue.

The concept of FTR focuses on the event of a complication. It is important to differentiate the ideal data set in which all complications are known and described, and the real world, in which complications are recorded in an incomplete manner. In the real world, there are deaths that may not be preceded by a complication coded in the hospital discharge record. This may be because deaths occurred outside the hospital and the complications were undocumented, or it may be that the hospital did not record some complications before death. Therefore, to conduct an FTR analysis one must define a set of complications available in the data set, with the intent of having as many deaths as possible in the data set preceded by a complication. The complications used in the traditional FTR analysis include the following categories: (1) cardiac—arrhythmias, arrest, infarction, congestive heart failure; (2) respiratory—pneumonia, pneumothorax, bronchospasm, respiratory compromise, aspiration pneumonia; (3) hypotension, shock, hypovolemia; (4) neurologic—stroke, transient ischemic attack, seizure, psychosis, coma; (5) deep vein thrombosis, pulmonary embolus, arterial clot, phlebitis; (6) internal organ damage, return to surgery; (7) infection—deep wound infection, sepsis; (8) gangrene, amputation; (9) gastrointestinal bleeding, blood loss; (10) peritonitis, intestinal obstruction; (11) renal dysfunction; (12) hepatitis; (13) pancreatitis; (14) decubitus ulcers; and (15) orthopedic complications and compartment syndromes. See Appendix B, which is also available on the *Medical Care* website, for detailed ICD-9-CM codes concerning these complications.

To construct an FTR analysis one ideally begins by constructing a 2×2 table to display the rates of in-hospital death, complication, and failure-to-rescue. Table 1 reports in-hospital deaths (yes or no) versus complications (yes or no). Table 1A displays a data set of general surgical patients using Medicare claims data from July 1, 1999 to June 30, 2000. A patient was defined to have a complication if any of the complications listed previously were identified in the index hospitalization where the surgery was performed. As can be seen in Table 1A, the failure rate is defined as the number of patients who died after a complication divided by the number of patients with complications or 17,038/189,031 = 9.0%. This FTR definition contains approximately 95% of patients who died and who had a preceding

TABLE 1. In-Hospital Deaths Versus Complications in a Random Half Sample of the 1999–2000 Medicare Claims Data Set

	Alive	Dead (%)	Total
A. FTR: before recoding undocumented complications			
No complication	213,776	872 (0.41)	214,648
Complication	171,993	17,038 (9.01)	189,031
Total	385,769	17,910 (4.44)	403,679
B. FTR: after recoding			
No complication	213,776	0 (0.00)	213,776
Complication	171,993	17,910 (9.43)	189,903
Total	385,769	17,910 (4.44)	403,679
C. FTR-N definitions			
No complication	351,456	8723 (2.42)	360,179
Complication	34,313	9187 (21.12)	43,500
Total	385,769	17,910 (4.44)	403,679
D. FTR-A definitions			
No complication	357,145	7433 (2.04)	364,578
Complication	28,624	10,477 (26.79)	39,101
Total	385,769	17,910 (4.44)	403,679

complication identified in the Medicare claims (17,038/17,910 = 95.13%), leaving only about 5% of these in-hospital deaths displaying "undocumented" complications.

To account for the deaths occurring without a documented complication, the FTR definition in Table 1B recodes all patients who died without a complication as if they died with an undocumented complication. In Table 1B, FTR can be written as $d/(c + d \mid \text{no } c)$ or the total number of deaths in the numerator divided by the number of patients with complications plus the number of deaths that did not have a documented complication. As can be seen in Table 1B versus Table 1A, the rates of complication and failure are very similar, but now the FTR analysis includes all the deaths. When constructing the failure rate, the traditional analysis aims to develop a list of complications that produces as high a precedence rate as possible (the percentage of deaths preceded by a complication) while not allowing the complication rate to become too high, because the death rate and FTR rates merge as the number of patients with complications approaches the number of total patients. If a list is comprehensive, the number of undocumented complications before death is low, and the recoding of failures needed to include these undocumented complications is minimal.

Variations on FTR

Other investigators have adapted FTR in ways that do not use all deaths. Needleman et al[7] adapted the FTR measure to "nurse sensitive complications" by selecting a limited number of complications for the FTR measure. This change in definition, which we will call FTR-N, was developed to better focus on nursing quality of care. Because only deaths after nursing sensitive complications are studied, a large number of deaths are not used in the analysis. Subsequently, AHRQ again adapted the FTR-N definition to reflect quality from a "patient safety" perspective (ie, the identification of deaths that were especially likely to be preventable). Expert panels guided both of these adaptations through consensus development panels.[34,35] The National Quality Forum, through its own process of selecting National Voluntary Consensus Standards for Nursing-Sensitive Care, endorsed Needleman et al's adaptation and assigned it to AHRQ for updating and support.[36,37]

FTR-N includes only 6 complications (pneumonia, shock, gastrointestinal bleeding, cardiac arrest, sepsis, and deep venous thrombosis) in its denominator definition, and it excludes deaths in patients without these complications (see Appendix C on the *Medical Care* website). FTR-A adds renal failure to the FTR-N list of eligible complications, and modestly alters the definition of several others (see Appendix D on the *Medical Care* website). Table 1C and 1D display the impact of restricting the denominator of FTR to more limited sets of complications, as in the FTR-N and FTR-A definitions, respectively. Note first that the number of patients defined as having a complication fell from 189,031 (46.8%) in Table 1A to 43,500 (10.8%) in Table 1C and 39,101 (9.7%) in Table 1D. However, this smaller complication rate comes at an important cost—of all deaths, the proportion coded as having a complication (the precedence rate) fell from 95% in Table 1A to only 51% in Table 1C, and 58.5% in Table 1D.

Comparing FTR Definitions

To better understand differences between FTR and FTR-N (the logic for FTR-A applies in a parallel manner) we grouped patients into 4 mutually exclusive categories (Fig. 1). Let A be the set of patients with no deaths or complications by either the traditional FTR definitions or the FTR-N definitions. Let B be the set of patients who have both FTR and FTR-N complications. Lower case b, is a subset of B and represents the number of deaths in set B. Let C equal the number of patients who have FTR complications but not FTR-N complications with lower case c representing the

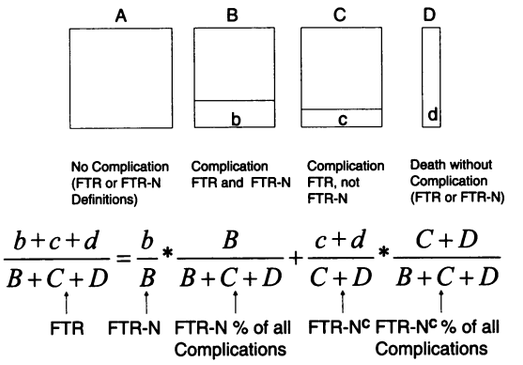

FIGURE 1. Comparing definitions of FTR, FTR-N, and the complement of FTR-N. Capital *A, B, C,* and *D* are mutually exclusive sets of patients. Lower case *b, c,* and *d* represent deaths in the sets *B, C,* and *D*.

number of patients with deaths in the set C. Finally, let D be the number of patients who died without an FTR or FTR-N complication, and let $d = D$ be the number of deaths in set D.

We can then express the relationship between FTR and FTR-N using the equation displayed in Figure 1. As can be seen, FTR is comprised of FTR-N multiplied by a scaling parameter [the number of B patients with complications divided by the number of patients with total complications $(B + C + D)$] plus the complement of FTR-N (denoted as FTR-N^C), which are the deaths and complications not included in the FTR-N definition multiplied by the fraction of total patients with complications that are in the complement.

Measures of Reliability

We defined reliability as described by Lord and Novick[38] using a split sample methodology. Using the same population of patients and hospitals for all measures, we randomly split each hospital into 2 groups, and ranked each hospital by each measure of interest. For completeness, we evaluated the measures of death, FTR, FTR-N, and FTR-N^C, as well as FTR-A and FTR-A^C. Using identical variables for severity adjustment based on 160 patient covariates, we ranked each hospital by the observed minus expected number of events, divided by the total number of eligible patients or $(O - E)/N$.[3,13,38] We ranked hospitals with at least 50 observations per hospital. After ranking and obtaining Spearman correlations and P values,[39] we converted the split sample results using the Spearman-Brown correction.[38] The square root of these resulting reliability correlations represent the upper bound of validity for each measure.[38]

Measures of Validity

We studied validity in a number of ways. First, we examined the rank correlation between various hospital outcomes (death, FTR, FTR-N, FTR-N^C, FTR-A, and FTR-A^C). Next, we again assessed the validity of each outcome measure by studying logit models using detailed patient characteristics and 5 hospital characteristics shown to be associated with better outcomes in previous studies[8,40,41]: (1) teaching hospital status (member of the American Council of Teaching Hospitals); (2) high technology status (does the hospital perform open heart surgery or perform organ transplantation); (3) hospital size greater than 200 beds; (4) bed-to-nurse ratio (where nurses are the sum of RN plus LPN FTE positions); and (5) nursing skill mix [the ratio of RN/(RN + LPN)]. We report both "marginal" and "partial" results for each regression. The marginal results used 1 hospital characteristic at a time along with all patient characteristics. For completeness, we also reported the partial regression results, using all hospital and patient variables simultaneously. This had the disadvantage that correlation between hospital characteristics can cause difficulty in interpreting the effects of individual hospital variables.

Finally, we examined the relative contribution of patient-to-hospital characteristics that predicted each outcome of interest. The omega statistic[13,42] provided this for each logistic regression model. The omega statistic computes a ratio of the squared sum of the log odds for model patent variables divided by a similar quantity calculated for the model hospital variables. All else being equal, outcome measures that have lower omega ratios may be more desirable quality indicators, because the lower the omega, the greater the hospital's impact on outcome relative to the patient's impact. This is especially important if modeling patient severity is difficult (as with claims data) so that the lower the omega suggests the higher relative influence of hospital characteristics as compared with patient characteristics.

RESULTS

Reliability

To estimate the reliability and validity of various FTR measures, we applied the approach shown in Table 1, but used 30-day death instead of inpatient death in all measures to minimize bias due to hospital-level variation in length of stay and transfer practices. We performed a sensitivity analysis using inpatient death in all measures, and the comparative results were very similar to those reported below (results not shown). The precedence rates for FTR, FTR-N, and FTR-A defined using 30-day deaths were 88.1%, 41.4%, and 46.9% respectively.

Table 2 displays the split half sample reliability estimates for each measure. The highest reliability was found with FTR, with a correlation of 0.32, whereas AHRQ-based FTR (FTR-A) and FTR-N displayed a reliability of 0.18. Reliability was higher for the complement of both FTR-N and FTR-A, the FTR measures derived from the patients not used by FTR-N and FTR-A, than for FTR-N and FTR-A, respectively. The corresponding upper bound for validity is also displayed for each measure.

Validity

We examine the correlation between hospitals ranked by the adjusted mortality rate and those ranked by all measures of FTR as displayed in Table 3. FTR itself is highly correlated with death, with a Kendall's τ equal to 0.83, representing a probability of concordance equal to

TABLE 2. Reliability and Upper Bound of Validity Using the Spearman-Brown Half Split Sample Correction

	Death Rate (30-d)	FTR	FTR-N	FTR-N^C	FTR-A	FTR-A^C
Reliability	0.30	0.32	0.18	0.23	0.18	0.22
Upper bound on validity	0.54	0.56	0.43	0.48	0.42	0.47
*P**	<0.0001	<0.0001	<0.0001	<0.0001	0.0002	<0.0001

FTR definitions were based on deaths within 30 days of admission.
**P* values are based on the Spearman rank correlation, testing the null hypothesis of no correlation.

TABLE 3. The Kendall's Correlation τ Between Adjusted Hospital 30-Day Death Rates and Adjusted Measures of FTR, FTR-N, FTR-N^C, FTR-A, and FTR-A^C

	FTR	FTR-N	FTR-N^C	FTR-A	FTR-A^C
Death rate (30-d)					
τ	0.83	0.41	0.57	0.43	0.55
Concordance	0.91	0.70	0.79	0.72	0.77
P	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
FTR					
τ	—	0.43	0.61	0.45	0.59
Concordance	—	0.72	0.81	0.73	0.79
P	—	<0.0001	<0.0001	<0.0001	<0.0001
FTR-N					
τ	—	—	0.08	0.67	0.12
Concordance	—	—	0.54	0.83	0.56
P	—	—	<0.0001	<0.0001	<0.0001
FTR-N^C					
τ	—	—	—	0.16	0.76
Concordance	—	—	—	0.58	0.88
P	—	—	—	<0.0001	<0.0001
FTR-A					
τ	—	—	—	—	0.09
Concordance	—	—	—	—	0.54
P	—	—	—	—	<0.0001

Results for in-hospital mortality were similar and not shown. Each correlation coefficient is based on the same number of hospitals (N = 1567). Computing Kendall's τ allows one to also display the probability of concordance between measures (concordance = $(1 + \tau)/2$) and their associated *P* values testing the null hypothesis of no concordance. All FTR measures were based on deaths within 30 days of admission.

0.91. Note that FTR-N and FTR-A both had far lower probabilities of concordance with death (0.70 and 0.72, respectively). A probability of concordance of 0.5 represents random or no association. Of interest, the complements of each measure—the FTR measures derived from the patients not used by FTR-N or FTR-A—are more highly correlated with death than FTR-N or FTR-A (0.79 and 0.77, respectively).

Table 4 describes the association between hospital characteristics and each outcome measure using a number of logistic regression models. Here, we are interested in whether specific hospital variables, widely regarded as being associated with quality,[8,40,41] are more highly associated with different ways of defining FTR. For each hospital characteristic, we report the associated odds ratio and *P* value when all patient characteristics are included in the model with only 1 hospital characteristic at a time. This describes the marginal analysis. We also report the partial analysis, which uses all patient and hospital variables simultaneously.

FTR, FTR-N, and FTR-A seem quite similar in their associations between hospital characteristics and outcome, with a few exceptions. Although FTR in the marginal analysis was highly associated, and in the correct theoretical direction (lower mortality with higher degree of teaching status, technology, nurse staffing, nurse skill mix), with all hospital characteristics, FTR-N lacked association with high technology. The complements of FTR-N and FTR-A were both highly associated with numerous hospital characteristics. That is, when utilizing the patients discarded in the FTR-N and FTR-A analyses (but included in the FTR analyses), we saw important associations between hospital characteristics and outcomes. In fact, in the marginal analysis, all hospital characteristics were significantly associated with FTR-N^C and FTR-A^C, and many were even associated using the partial analysis.

TABLE 4. Regression Model Results Using Outcomes of Death, FTR, FTR-N, FTR-N^C, FTR-A, and FTR-A^C

Variable Name	Death	FTR	FTR-N	FTR-N^C	FTR-A	FTR-A^C
Teaching status (1 = yes, 0 = no)						
Marginal	0.811[e]	0.835[e]	0.823[e]	0.846[e]	0.807[e]	0.847[e]
Partial	0.864[e]	0.886[e]	0.858[d]	0.904[c]	0.852[e]	0.898[c]
High technology (1 = yes, 0 = no)						
Marginal	0.896[e]	0.910[e]	0.950	0.904[e]	0.924[c]	0.915[e]
Partial	0.996	1.004	1.065	0.998	1.049	1.009
Beds >200 (1 = yes, 0 = no)						
Marginal	0.868[e]	0.876[e]	0.900[d]	0.867[e]	0.872[e]	0.875[e]
Partial	0.919[e]	0.918[e]	0.937	0.903[d]	0.917[b]	0.905[d]
Bed-to-nurse ratio						
Marginal	1.110[e]	1.094[e]	1.091[a]	1.117[d]	1.108[c]	1.105[c]
Partial	1.045	1.039	1.038	1.069[a]	1.044	1.062
Staff mix ratio: RN/(RN + LPN)						
Marginal	0.872[e]	0.878[e]	0.840[e]	0.896[e]	0.832[e]	0.902[e]
Partial	0.925[e]	0.923[e]	0.866[e]	0.948[a]	0.870[e]	0.951
Omega: patient-to-hospital (95% CI)	189 (141–253)	128 (93–177)	95 (60–150)	163 (103–258)	57 (7–86)	164 (100–277)
Full model C-statistics	0.86	0.79	0.77	0.81	0.72	0.80
Full model χ^2 P value	0.001	0.001	0.001	0.001	0.001	0.001
No. total patients	386,492	183,343	41,576	139,913	37,355	144,134
No. patient deaths	22,456	22,456	9298	11,279	10,535	10,042
30-day death rate (%)	5.81	12.25	22.36	8.06	28.20	6.97

Each cell provides the hospital coefficient-derived odds ratio adjusted for patient comorbidities, with an associated P value. Each "marginal" (M) model was adjusted for all patient characteristics and 1 hospital characteristic at a time. Each "partial" model (P) was adjusted for all patient and hospital variables simultaneously.

All FTR measures were based on deaths within 30 days of admission. Each model included identical covariates for patient characteristics. These were 151 patient covariates, 7 regional indicators, and 1 (for the marginal analysis) or 5 (for the partial analysis) hospital characteristics. Odds ratios for continuous variables were based on their interquartile range. P values are presented without correction for clustering of patients inside hospitals, but patterns were similar when corrected, except that the high technology variable lost significance for the FTR-N marginal model, and the bed-to-nurse and staff-mix variables lost significance in the partial models using the FTR-N^C outcome.[45,46] C-statistic and χ^2 tests were based on utilizing all variables simultaneously (the partial models). The Omega statistic represents the ratio of patient-to-hospital factors associated with the contribution to predicting outcome, with geographic regions left out of the ratio.[13,42]

[a] $P < 0.05$.
[b] $P < 0.01$.
[c] $P < 0.005$.
[d] $P < 0.001$.
[e] $P < 0.0001$.

Omega Analyses

Table 4 also reports on a number of properties of the various outcome models constructed for this analysis. Omega compares the relative contribution of patient to hospital characteristics when predicting outcome, and in past work we have shown that FTR generally displays lower omegas than death. This is a desirable feature, in that patient severity assessment is difficult with claims, so we would prefer hospital factors to have higher relative influence than patient factors in predicting outcomes. In these data we found that omega for death equaled 189, whereas for FTR the value was 128. Similar to previous work,[13] omegas greater than 1 suggest that it is far better to be well and go to a hospital with poor characteristics than to be ill and go to a hospital with better characteristics, but relative to each outcome measure, FTR displayed better omega ratios than death. This was a central motivation in designing FTR and it was confirmed in this study. The better omega ratio occurred because the FTR analyses contained far more homogeneous patients than those in the death analyses. Only patients with complications are in the FTR analysis, so patient severity becomes somewhat less important. Similarly, FTR-A displays a lower omega (ω = 57) than FTR, as does FTR-N (ω = 95). Although FTR-A and FTR-N clearly have more homogeneous patients and better omega statistics than FTR, this has been achieved at the cost of excluding half the deaths, whereas FTR did not exclude any deaths.

CONCLUSIONS

FTR is a measure of hospital quality that has some advantages over simply studying mortality. The measure is less sensitive to errors in severity adjustment (because all patients in the analysis have complications) and more dependent on hospital characteristics relative to patient characteristics than the mortality rate, while having equivalent reliability. FTR has intuitive appeal as a quality marker, attempting to measure a hospital's ability to manage complications, while being less likely to confuse worse severity of illness with worse quality of care. However, the specific definition of FTR is important, and many of the desirable properties of the original FTR measure may be weakened in

the newer versions adopted by the AHRQ[34,35] and the National Quality Forum.[36,37]

The difference between FTR and the newer versions FTR-A and FTR-N rests in the exclusion of a large portion of the patients who died. FTR-N includes only 51.3% of the deaths as compared with FTR. FTR-A includes only 58.5% of the deaths used in FTR. Hence, these measures are less stable, have less reliability, and have similar or weaker associations with some hospital characteristics that are thought to be structural measures of quality. Although FTR-A and -N have a better omega profile than FTR, this occurs because they select a more homogeneous group of patients at the expense of excluding half the deaths. Is the more limited approach better related to other markers of quality of care? We saw no evidence for this. Instead, we found excellent associations between other quality markers and FTR, coupled with better reliability when a broader set of complications, and all deaths, were included in the analysis.

Our concern with both FTR-N and FTR-A is that they are measures that are based on death rates that fail to count all deaths in an analysis. Although it would be impossible to ignore a death in a clinical trial simply because a preceding side effect was not thought to be associated with the study drug, the logic behind FTR-N and FTR-A presumes that only a selected subset of deaths merit attention. Although it is conceivable that these excluded deaths are less related to quality of care than the included deaths, we found that this excluded information was meaningful. We observed that the complements of FTR-A and FTR-N were at least as highly associated with some structural measures of quality as the limited measures themselves. Ignoring this information weakens the FTR-A and FTR-N measures by making them more unreliable. Restricted definitions may exclude important information on how better programs benefit patients who develop the "other" complications not on the limited set developed by AHRQ[35] and Needleman et al.[7,10,43]

There were important limitations to this study. Although we did attempt to mimic alternative definitions of FTR in a consistent manner, we did not duplicate these definitions exactly. For example, the FTR-A definition used by AHRQ restricts the age of patients to be between 18 and 75 years. We chose to compare Medicare age patients up to age 90 to be more applicable to most studies utilizing mortality information. Our intent was to be both fair (by using the same population for all definitions) and relevant (by studying an interesting population). A second limitation regards the use of in-hospital deaths for the development of the FTR definitions and 30-day mortality for the correlation and regression analyses. We could have developed FTR definitions using all 30-day deaths, rather than just in-hospital deaths, but this might have required broadening the list of complications artificially to ensure coverage of 95% of both in-hospital and postdischarge deaths. Consequently, we suggest developing the complication list on in-hospital data and then (when desired) define those who died outside the hospital within 30 days of admission but without an in-hospital complication as having died after an undocumented complication. It is reassuring, however, that when we reran all our models using only in-hospital mortality, the comparative results across measures were very similar (results not shown). Finally, FTR treats patients equally, regardless of their specific complication or number of complications. Developing a weighted complication list may be beneficial, and we have recently explored measuring the importance of the first complication.[44] However, because poor management of early complications may lead to subsequent complications, we have been hesitant to construct a complication-weighted FTR measure, although this could represent a course for future improvement of the measure.

In summary, FTR, unlike FTR-N and FTR-A, is constructed to include all the deaths. Given the superior reliability and similar construct validity of FTR, relative to FTR-N and FTR-A, we believe FTR should generally be computed with the more inclusive complication set that counts all deaths. The enhanced reliability of the original FTR measure may be particularly important for applications that involve public reporting of hospital performance, such as that advocated by the National Quality Forum.

ACKNOWLEDGMENTS

The authors thank Paul R. Rosenbaum and Richard N. Ross for their aid in conducting this research.

REFERENCES

1. Silber JH, Williams SV, Krakauer H, et al. Hospital and patient characteristics associated with death after surgery: a study of adverse occurrence and failure-to-rescue. *Med Care*. 1992;30:615–629.
2. Silber JH, Rosenbaum PR, Schwartz JS, et al. Evaluation of the complication rate as a measure of quality of care in coronary artery bypass graft surgery. *JAMA*. 1995;274:317–323.
3. Silber JH, Rosenbaum PR, Williams SV, et al. The relationship between choice of outcome measure and hospital rank in general surgical procedures: implications for quality assessment. *Int J Qual Health Care*. 1997;9:193–200.
4. McKee M, Coles J, James P. 'Failure to rescue' as a measure of quality of hospital care: the limitations of secondary diagnosis coding in English hospital data. *J Public Health Med*. 1999;21:453–458.
5. Silber JH, Kennedy SK, Even-Shoshan O, et al. Anesthesiologist direction and patient outcomes. *Anesthesiology*. 2000;93:152–163.
6. Silber JH, Kennedy SK, Even-Shoshan O, et al. Anesthesiologist board certification and patient outcomes. *Anesthesiology*. 2002;96:1044–1052.
7. Needleman J, Buerhaus P, Mattke S, et al. Nurse-staffing levels and the quality of care in hospitals. *N Engl J Med*. 2002;346:1715–1722.
8. Aiken LH, Clarke SP, Sloane DM, et al. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*. 2002; 288:1987–1993.
9. Aiken LH, Clarke SP, Cheung RB, et al. Educational levels of hospital nurses and surgical patient mortality. *JAMA*. 2003;290: 1617–1623.
10. Rosen AK, Rivard P, Zhao S, et al. Evaluating the patient safety indicators: how well do they perform on Veterans Health Administration data? *Med Care*. 2005;43:873–884.
11. Simpson KR. Failure to rescue in obstetrics. *MCN Am J Matern Child Nurs*. 2005;30:76.
12. Cote CJ, Notterman DA, Karl HW, et al. Adverse sedation events in pediatrics: a critical incident analysis of contributing factors. *Pediatrics*. 2000;105:805–814.
13. Silber JH, Rosenbaum PR, Ross RN. Comparing the contributions of groups of predictors: which outcomes vary with hospital rather than patient characteristics? *J Am Stat Assoc*. 1995;90:7–18.
14. Seago JA, Williamson A, Atwood C. Longitudinal analyses of nurse staffing and patient outcomes: more about failure to rescue. *J Nurs Adm*. 2006;36:13–21.
15. Halm M, Peterson M, Kandels M, et al. Hospital nurse staffing and

patient mortality, emotional exhaustion, and job dissatisfaction. *Clin Nurse Spec*. 2005;19:241–251.
16. Simpson KR. Failure to rescue: implications for evaluating quality of care during labor and birth. *J Perinat Neonatal Nurs*. 2005;19: 24–34.
17. Ashcraft AS. Differentiating between pre-arrest and failure-to-rescue. *Medsurg Nurs*. 2004;13:211–215.
18. Lang TA, Hodge M, Olson V, et al. Nurse-patient ratios: a systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. *J Nurs Adm*. 2004;34:326–337.
19. Boyle SM. Nursing unit characteristics and patient outcomes. *Nurs Econ*. 2004;22:111–119, 123, 107.
20. Long KA, Bernier S. RN education: a matter of degrees. *Nursing*. 2004;34:48–51.
21. Clarke SP. Failure to rescue: lessons from missed opportunities in care. *Nurs Inq*. 2004;11:67–71.
22. Ashcraft AS. Failure to rescue. *Am J Nurs*. 2003;103:15–16.
23. Clarke SP, Aiken LH. Failure to rescue. *Am J Nurs*. 2003;103:42–47.
24. Davies P. Fatal medical errors said to be more widespread. *Wall Street Journal*. July 27, 2004.
25. Sedman A, Harris JM, Schulz K, et al. Relevance of the Agency for Healthcare Research and Quality Patient Safety Indicators for children's hospitals. *Pediatrics*. 2005;115:135–145.
26. Brewster AC, Karlin BG, Hyde LA, et al. MedisGroups: a clinically based approach to classifying hospital patients at admission. *Inquiry*. 1985;22:377–387.
27. Steen PM, Brewster AC, Bradbury RC, et al. Predicted probabilities of hospital death as a measure of admission severity of illness. *Inquiry*. 1993;30:128–141.
28. Silber JH, Rosenbaum PR. A spurious correlation between hospital mortality and complication rates: the importance of severity adjustment. *Med Care*. 1997;35:OS77–OS92.
29. DesHarnais SI, McMahon LF, Wroblewski RT, et al. Measuring hospital performance. The development and validation of risk-adjusted indexes of mortality, readmissions, and complications. *Med Care*. 1990;28: 1127–1141.
30. DesHarnais S, McMahon LF, Wroblewski R. Measuring outcomes of hospital care using multiple risk-adjusted indexes. *Health Serv Res*. 1991;26:425–445.
31. Shroyer AL, Coombs LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. *Ann Thorac Surg*. 2003;75:1856–1864.
32. Arozullah AM, Henderson WG, Khuri SF, et al. Postoperative mortality and pulmonary complication rankings: how well do they correlate at the hospital level? *Med Care*. 2003;41:979–991.
33. Iezzoni LI, Daley J, Heeren T, et al. Using administrative data to screen hospitals for high complication rates. *Inquiry*. 1994;31:40–55.
34. Department of Health and Human Services. *Guide to Patient Safety Indicators*. Rockville, MD: Agency for Healthcare Research and Quality; 2006.
35. McDonald KM, Romano PS, Geppert J, et al. *Measures of Patient Safety Based on Hospital Administrative Data—The Patient Safety Indicators*. AHRQ Publication No. 02-0038. Rockville, MD: Agency for Healthcare Research and Quality; 2002.
36. Nursing Care Quality at NQF. [National Quality Forum web site]. Available at: http://216.122.138.39/nursing/. Accessed November 3, 2006.
37. *National Voluntary Consensus Standards for Nursing-Sensitive Care: An Initial Performance Measure Set*. Washington, DC: National Quality Forum; 2004.
38. Lord FM, Novick MR. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley; 1968.
39. Hollander M, Wolfe DA. *Nonparametric Statistical Methods*. New York, NY: Wiley & Sons; 1999:394–405.
40. Hartz AJ, Krakauer H, Kuhn EM. Hospital characteristics and mortality rates. *N Engl J Med*. 1989;321:1720–1725.
41. Rosenthal GE, Harper DL, Quinn LM, et al. Severity-adjusted mortality and length of stay in teaching and nonteaching hospitals. Results of a regional study. *JAMA*. 1997;278:485–480.
42. Polsky D, Armstrong KA, Randall TC, et al. Variation in chemotherapy utilization in ovarian cancer: the relative contribution of geography. *Health Serv Res*. 2006;41:2201–2218.
43. Romano PS, Geppert JJ, Davies S, et al. A national profile of patient safety in U.S. hospitals. *Health Aff*. 2003;22:154–166.
44. Silber JH, Rosenbaum PR, Trudeau ME, et al. Changes in prognosis after the first postoperative complication. *Med Care*. 2005;43:122–131.
45. *Stata Users Guide Release 9*. College Station, TX: Stata Corporation; 2005:275–280.
46. Freedman DA. On the so-called "Huber Sandwich Estimator" and "Robust Standard Errors". *Am Stat*. 2006;60:299–302.