Chapter 4

Hospital Mortality Rates

CONTENTS

Box

Figure

Tables

Chapter 4
Hospital Mortality Rates

INTRODUCTION

Differences in patient death rates seem on their face a valid way to distinguish good quality health care providers from poor quality providers; death is an outcome that is almost always bad,[1] and medical practice is devoted, at least in part, to postponing death, Differential mortality, or survival, has long been used as a measure of efficacy in health care technology assessments and as an indicator, albeit crude, of the health status of particular populations. Medical encounters can be dangerous (318,555,595), adding to the possibility of death from a hospital encounter.

Almost half the deaths in the United States every year occur in hospitals, although only about 3 percent of hospital admissions end in death **(667).** Although many deaths in hospitals occur because nothing more could be done for the patients involved, a substantial portion of the deaths are believed to be avoidable. Hospital-related mortality can result from various factors that are subject to control, including poor infection control, inadequate or inappropriate use of medication, falls as a result of poor supervision, mistakes during surgery, and inappropriate discharge.

Although the use of patient death rates to compare the quality of care delivered by specific health care providers has been expanding, it has also been controversial. The major problems with the use of hospital mortality rates as a quality indicator are that mortality can result from many factors other than poor quality care and that techniques to adjust for such factors are generally inadequate. In addition, there are theoretical and practical issues regarding the appropriate period of time for an analysis. Over what period of time is a death to be defined as related to hospital care? Another issue regarding time is the period covered in the analysis. Most releases of information on hospital mortality rates have included data for a single year, but critics argue that data over a longer period of time may be needed, given the uncertainties about the indicator. Yet another significant issue is the level of aggregation of hospital mortality rates. Should rates be aggregated across the hospital as a whole? If not, at what level of diagnostic coding should the data be totaled? Finally, it is important to validate hospital mortality rates against criteria related to the process of care; this validation is only beginning.

Perhaps the most visible and controversial releases of hospital mortality data have been the 1984 and 1986 analyses of the Health Care Financing Administration (HCFA), which is part of the U.S. Department of Health and Human Services (640,647). The HCFA releases illustrate well the critical issues surrounding the use of hospital mortality rates as an indicator of the quality of care. Both analyses were conducted with data derived from hospital claims filed for the purpose of Medicare reimbursement, although the 1986 analysis added information about deaths derived from Social Security Administration files (see table 4-1 for a summary of differences between the **1984** and 1986 HCFA analyses). The **1986** analysis differed in level of analysis, in the way conditions and procedures were aggregated, in the period of time after hospital admission during which hospitals were counted, in calculation methods, and in the type of information released.

A number of other analyses of hospital mortality data have been "conducted along the same basic lines as the HCFA analyses, that is, using data from hospital discharge abstracts to adjust for patients' risk of dying (80,81,189,448,462,526); other analyses have adjusted for patients' risk of dying using clinical data (190,352,353,588,589, 590) as well as proxies such as age. Few have attempted to validate statistical results against a process criterion (190,279,353,462).

OTA reviewed in depth studies whose purpose was to develop a valid technique to adjust hospital mortality statistics for patients' risk of dying. Not included were studies whose primary purpose was to test the validity of structural measures of

[1] It has been argued that in some cases death would be preferable to life; definitions of life and death are not as simple as they once seemed *(632).*

Table 4-1.—Comparlson of HCFA'S 1984 and 1986 Hospitai Mortaiity Analyses

	HCFA'S 1984 analysis[a]	HCFA'S 1988 analysis[b]
Data base	Claims filed for Medicare reimbursement	Claims filed for Medicare reimbursement and information from the Social Security Administration about date of death
Hospital population	Short-term acute care hospitals (some hospices included inadvertently)	Short-term acute care hospitals (some hospices included inadvertently)
Patient population	All Medicare patients, both aged and disabled	All Medicare patients, both aged and disabled
Period of time during which deaths were counted	In-hospital deaths	Death within 30 days of last hospital admission
Hospital "risk group"d	All discharges	Last admission
Measures used to adjust for patients' risk of dying	Average age of Medicare patients; proportion male; proportion black; proportion neither black nor white; State average length of stay; 50 most frequent diagnosis-related groups (DRGs); all cancer DRGs; 30 DRGs associated with most frequent DRGs; weighted by number of Medicare discharges	Age group; sex; comorbidities tailored to diagnostic group; prior hospital admissions in the year preceding death; whether patient was transferred from another hospital
Level of analysis	Hospital	Patient, then hospital
Levels of aggregation	Hospital overall, and 9 DRG categories	Hospital overall, and 17 diagnostic risk groups
Calculation method	Multiple linear regression	Logistic regression
Information released	Outlier hospitals only	All hospitals, with actual and expected mortality rates for each category

%.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Hospital Mortality Information 1984," Washington, DC, Mar. 10, 19S6.
bus. Department of Health and Human se~lces, Health Care Financing Administration, *Medicare Hospital Mortality Information 1986* (Washington, DC: U.S. Government Printing Office, 19S7).
cAssembled in HCFA's MEDPAR file.
'Denominator.

SOURCE: Office of Technology Assessment, 1988.

quality against hospital mortality as a standard. In addition, the OTA review included releases of crude mortality rates (55,115,116,478) to compare their results with the rates adjusted in various ways. All studies were reviewed using the procedure and checklist described in appendix C.[2] Table 4-2 lists the studies reviewed by OTA, and indicates when they were conducted, the sources of data used, the patient and hospital types that were included, and the years in which data were collected. Table 4-3 shows the diagnoses and procedures included in the analysis, when death was measured, the adjustments for patients' risk of dying, the level of analysis, and the results of each study.

The remainder of this chapter consists of an evaluation of the reliability, validity, and feasibility of using hospital mortality rates as an indicator. Conclusions and policy implications are outlined in the final section of the chapter.

'The way studies were selected for review and descriptions of the individual studies can be found in OTA'S technical working paper, "Hospital Mortality Rates as a Quality Indicator" *(187).*

RELIABILITY OF THE INDICATOR

Whether hospital mortality rates are a valid indicator of the quality of care depends on the reliability of the data on which analyses of mortality rates are performed and the reliability of the data against which results of analyses are validated. Some aspects of the data base for hospital mortality analyses have been of longstanding concern (166,167). There is reason to believe that hospital data sources vary widely in completion and accuracy; rarely have hospital mortality analy-

Table 4-2.–Characteristics of Hospital Mortality Studies Reviewed by OTA

Study[a]	Patient rJoRulation	Source of data	Hospital types included or excluded	Years data collected	Sample size
Bunker, et al , 1969 (108)	All	Hospital medical records	**Included mIlltary, Nattonal Institutes of Health,** teaching and commumty general hospitals (all volunteers)	**(1959-62) 4 years**	**34 hospttals,** 856,000 patients, 16,840 deaths
Moses and Mosteller, 1968 (441)	All	Same as Bunker, et al., 1969	Same as Bunker, et al , 1969	Same as Bunker, et al., 1969	34 hospitals; 141,914 patients, 1,844 deaths
Roemer, et al , 1968 (526)	All nonobstetrlc	State of California hospital annual reports	Hospitals m Los Angeles County, Including Veterans Admmistration and municlpal[b]	1964	33 hospitals
Goss and Reed, 1974 (259)	All nonobstetrlc	Deaths: death certificates	102 short-term general hospitals m New York City	1971	50,000 deaths
Stanford Center for Health Care Research, 1974 (588), 1976 (589).					
Extensive Study	All	Commission on Professional and Hospital Activities' Professional Activities Study	Short-term hospitals	1972	1,244 hospitals: 558,856 patients
Intensive Study	All	Same as used in Extensive Study, plus data collected at hospital sites	Short-term hospitals randomly selected from a sample stratified by size, teaching status, cost per patient day, and a crude estimate or surgical mortality	May 1973-Feb 1974 (9 months)	17 hospitals, 8,593 patients
NAS, 1977 (448)	Males	Veterans Administration Patient Treatment File	Veterans Administration hospitals, including psychiatric hospitals	1970-75 (6 years)	More than 200,000 surgeries
Knaus, et al , 1986 (353)	Adults only; no coronary artery bypass graft	Hosplal and medical records and questionnaire data	Hospitals volunteering to be in the study	Average of 5 months[c]	13 hospitals, 236 patients
US DHHS, HCFA, 1986 (640)	Medicare patients, all ages	Medicare billing file	Short-term general hospitals	1984	Not given
Blumberg, 1987 (80), 1988 (81)	All	Maryland Health Services Cost Review Commission data base (based on discharge abstracts)	All Maryland hospitals except 10	April 1984-March 1985 (1 year)	45 hospitals, 8,745 cases
New York State Department of Health, 1987 (462)	All	New York State Department of Health Statewide Planning and Research Cooperative System	Excluded children's hospitals, one maternity hospital, a cancer hospital. several rehabilitation hospitals, and an eye-ear-throat hospital	1984	Not given
Rust, et al., 1987 (545)	Newborns	Birth and death certificates, State of California	NA[d]	1980-84	340 hospitals; 2.5 million babies
Dubois, et al., 1987 (189,190)	All	Modified version of the Uniform Hospital Discharge Data Set, aggregated to the hospital level	American Medical International, Inc. hospitals selected to be geographically representative[e]	Six-month period in 1985	93 hospitals; 205,000 hospital discharges
US DHHS, HCFA, 1987 (647)	Medicare patients, all ages	Medicare billing data base (MEDPAR); Social Security Administration records (for deaths)	Short-term general hospitals	1986	10 million admissions
DesHamais, et al., 1988 (173)	a. All except newborns, transfers to other short-stay hospitals, stays of less than 1 day	a. Commission on Professional and Hospital Activities data base		a. 1983-84	a. 300 hospitals
	b. Medicare patients, all aaes	b. Medicare billing data base (MEDPAR)	Short-term general hospitals	b 1984	b. Not given

a Studies are listed in chronolo~ical order. Numbers in parentheses refer to numbered entries in the referenCe liSt at the end of this rePort.
~he hospitals were chosen to ~epresent range of medical staff organization types (loosely to highly structured).
cData were collect~ either Orl consecutive patients or on every second or third patient until a SPeCified number of PatientS Was reached.
'NA = Not apptlcable.
eHospitals were nonteaching, nongovernmental, and Proprietary.

SOURCE: Office of Technology Assessment, 1988.

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA

Study[a]	Diagnoses and/or [illegible]	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard for process of care
Bunker, et al., 1969 (108)	a. All operations requiring general anesthesia	a. Inhospital death within 6 weeks following surgery	a. Operation	a. Patient	a. Standardization for operation reduced variation in death rates from 25:1 (for crude rate) to 9:1	a. —
	b. Sixty-two "middle death rate" surgeries requiring general anesthesia	b. Same as "a"	b. Age, operation, physical status[c], shock-likelihood index[d]; sex; year of operation; any previous operations during same or previous month (all separately); and a multiple adjustment for age, physical status, and operation	b. Patient	b. When age, operation, and physical status were combined, the ratio for the most extreme procedure was reduced to 3:1	b. —
Moses and Mosteller, 1968 (441)	Six specific procedures combined (gallbladder surgery, gastric resection, hysterectomy, cystoscopy, open reduction of femur, thoracic or lumbar laminectomy)	Inhospital death within 6 weeks following surgery	Age, operation, physical status, frequency of all six procedures combined	Patient	10:1 ratio between hospital with highest adjusted mortality rate and hospital with lowest adjusted mortality rate for six procedures	—
Roemer, et al., 1968 (526)	Nonobstetric admissions	Inhospital death	Hospital average length of stay corrected for occupancy rate	[illegible]	[illegible]	—
Goss and Reed, 1974 (259)	Nonobstetric admissions	Inhospital death	Hospital average length of stay corrected for occupancy rate	Hospital	$R^2 = .35$	—
Stanford Center for Health Care Research, 1974 (588), 1976 (589)						
Extensive Study	a. Nine surgical categories and 2 medical categories combined	a. Inhospital death	a. Age; sex; admission findings[f]; height to weight index; severity scores (stage of disease) for principal procedures or diagnoses; comorbidities, and additional procedures, staged for severity (tailored to procedure or diagnosis)	a. Patient	a. R^2 = [illegible]	a. —
	b. 1. Intra-abdominal artery operations 2. Selected biliary tract surgery 3. Hip fracture diagnosis with no other trauma 4. Amputation of lower limb with no current trauma 5. Large bowel operations with cancer diagnosis 6. Biliary tract diagnosis-nonsurgical 7. Ulcer diagnosis-nonsurgical 8. Arthroplasty of the hip 9. Stomach operations-ulcer 10. Large bowel operations-specified diagnoses	b. Inhospital death	b. Same as "a"	b. Patient	b. 1. $R^2 = .241$ 2. $R^2 = .051$ 3. $R^2 = .089$ 4. $R^2 = .078$ 5. $R^2 = .060$ 6. $R^2 = .079$ 7. $R^2 = .110$ ISMRs were significantly different among hospitals in surgical categories 1-5 and nonsurgical categories 6 and 7 8. $R^2 = .043$ 9. $R^2 = .112$ 10. $R^2 = .087$ 11. $R^2 = .122$ 12. $R^2 = .194$ 13. $R^2 = .107$ 14. $R^2 = .125$	

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2) if available	Relation to validation standard for process of care
	11. Hip fracture diagnosis with other trauma 12. Amputation of lower limb with current trauma 13. Stomach operations with cancer diagnosis 14. Stomach operations with other diagnosis 15. Large bowel operations with other diagnosis 16. Fracture of the pelvis 17. Fracture of the shaft of the femur				15. R^2 = .201 16. R^2 = .081 17. R^2 = .156 ISMRs were not significantly different among hospitals in surgical categories 8-17	
Study	Fifteen surgical categories combined	Death within 40 days of surgery; death with 40 days of surgery or severe or moderate morbidity at 7 days after surgery	Age; sex; extent of insurance coverage; level of stress; anesthetist's rating of physical status; cardiovascular status; whether surgery was emergency; extent of surgical disease; quadratic function combining age and physical status	Patient	Logistic regression not significant for death alone[h]; Linear regression found significant main effect for hospital[i, j] Adjusted "death or morbidity"[k] ratio ranged from 0.55 to 1.41, a threefold difference	—[g]
NAS, 1977 (448)	a. Three major surgical specialties taken together (general, urologic, orthopedic)	a. Inhospital death within 40 days of surgery	a. Age, other diagnoses; emergency or not; other surgeries before or during selected operation	a. Patient	a. Indirectly standardized mortality rates varied from 0.3 to 1.6 percent. Six of 131 hospitals fell outside a range of 2 standard deviations above and below the mean	a. —[l]
	b. All surgeries performed in VA psychiatric hospitals	b. —	b. None	b. Hospital	b. Crude mortality rate was 7.1% compared to 3.3% in ...	b. —
Knaus, et al., 1986 (353)	Patients in intensive care units	Inhospital	APACHE II (acute physiologic scores; age points; chronic health points; emergency surgery status)	Patient	R^2 = .99	Number of daily therapeutic intervention points was significantly higher at hospital with lowest adjustment mortality rate[m n]
U.S. DHHS, HCFA, 1986 (640)	a. All[o]	a. Inhospital death	a. Average age of Medicare patients; proportion male, proportion black; proportion neither black nor white; State average length of stay; 50 most frequent DRGs; all cancer DRGs; 30 DRGs associated with 50 most frequent DRGs; weighted by number of Medicare discharges	a. Hospital	a. R^2 = .59[p]	a. —

Table 4-3.–Results of Hospital Mortality Studies Reviewed by OTA–Continued

Stud$	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard for process of care
	b. Nme uHti calegones: 1, Pneumonia (DRGs 089-090) 2. Coronary artery bypass surgery (DRGs 106-107) 3. Pacemaker implant (DRGs 115-116) 4. Acute myocardial infarction (DRGs 121-123) 5. Congestive heart failure (DRG 127) 6. Gastrointestinal bleeding (DRGs 174-175) 7. Major joint surgery (DRG 209) 8. Transurethral prostatectomy (DRGs 336-337)	b. Innospltal death	b Average age of Medicare patients. race, sex (all at the hospital level of aggregation)	b. Hospital	b. 1. R'= .053 2. R'= .007 3. R'= .003 4. R*= .019 5. R'= .020 6. R'= .005 7. R'= .068 8. R'= .009	b. –
Blumberg, 1987 (80), 1988 (81)	a. High-risk surgeries	a. Inhospital death	a. Age; sex; type of admission (urgent, emergency); source of admission; risk level of procedure; risk level of comorbidities	a. Patient	a One of more than 41 hospitals had death rates "deserving review' but not statistically significant. Two other hospitals had lower than expected death rates bordering on significance	a. –
	b. Trauma v. nontrauma and the following surgical categories: Nervous system Respiratory Cardiovascular Gastrointestinal Urinary Musculoskeletal	b. Inhospital death	b. Same as "a"		b Little variation in trauma cases; substantial variation in nontrauma, gastrointestinal and cardiovascular categories (Chi^2=4 or more)	b. –
New York State Department of Health, 1987 (462)	a. All medical/surgical, all ages	a. Inhospital death	a. Average age, proporhon males; proportion black; proportion neither black nor white; case-mix severity;[r] severity surrogates	a. Hospital	a. R^2 = .86[q]	a. Overall, 3Y0 of cases were found to have quality problems
	b. All Medicare discharges		b. Same as above	b. Hospital	b. R'= .781[t]	b. See "a"
	c. All medicallsurgical, under 65		c. Same as above	c. Hospital	C. R^2= .92"	c. See "a'
	d. Obstetrics-nursery		d. Same as "a" plus Medicaid as source of payment	d. –	d. R'= ,37[v]	d. –
Rust et al , 1987 (545)	Perinatal (fetal and neonatal)	Death of fetus of 20 weeks or more aestation; death within 28 dais of birth	Birthweight, sex, race, multiple births	Patient	R'= .80	—
Dubois, et al., 1987 (189,190)	a. All	a. Inhospital death	a. Age (percent older than 70); percent admitted from emergency department: percent admitted from nursing home; case-mix index based on DRG weights; average length of	a. Hospital	a. R*= .64[w]	a. See "b"

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysis	Results: percent of variance in crude mortality explained (R^2), if available	Relation to validation standard [illegible]
			stay; hospital size; percentage of Medicaid patients; occupancy rate; discharges to nursing home			
	b. 1. Myocardial infarction (DRGs 121,122,123) 2. Pneumonia (DRGs 79 and 89) 3. Cerebrovascular accident (DRG 14)		b. 1. APACHE II score; body-system score[x] 2. APACHE II score; body-system score 3. APACHE II score; body-system score; presence or absence of a mass effect or blood on head CT scan	b. Patient	b. Actual death rates were within 95% confidence intervals for predicted death rates for all 3 conditions	b. Explicit review—no differences between high and low outliers. Implicit review—after adjustment for patients' risk of dying and for oversampling. Findings were as follows between high- and low-mortality hospitals in estimated rate of preventable deaths in total patient population: 1. No significant difference 2. Significant difference 3. Significant difference
U.S. DHHS, HCFA, 1987 (647)	a. All	a. Death within 30 days of last admission	a. Age group, sex, comorbidities tailored to diagnostic group[aa]; prior hospital admissions in the year preceding death; whether patient was transferred from another hospital	a.	a. $R^2 = .59$	a. —
	b. Specific diagnostic categories High risk: 1. Cancer 2. Stroke 3. Severe acute heart disease 4. Severe chronic heart disease 5. Gastrointestinal catastrophes 6. Metabolic and electrolyte disorders 7. Pulmonary disease 8. Renal disease 9. Sepsis 10. Severe trauma Low risk:	b. Same as [illegible]			1. $R^2 = .37$[bb] 2. $R^2 = .08$ 3. $R^2 = .12$ 4. $R^2 = .07$ 5. $R^2 = .12$ 6. $R^2 = .22$ 7. $R^2 = .22$ 8. $R^2 = .26$ 9. $R^2 = .09$ 10. $R^2 = .10$ 11. $R^2 = .02$ 12. $R^2 = .04$ 13. $R^2 = .10$ 14. $R^2 = .14$ 15. $R^2 = .26$ 16. $R^2 = .10$	

Table 4-3.—Results of Hospital Mortality Studies Reviewed by OTA—Continued

Study[a]	Diagnoses and/or procedures included	Dependent variable	Adjustments	Level of analysls	Results. percent of variance m crude mortahty explained (R^2), If available	Relatlon to vahdation standard for process of care
	11. Ophthalmologic disease 12 Gynecologlc disease 13 Low nsk heart disease 14 Gastrolntestlnal disease 15 Urologic disease 16 Orthopedic conditions					
DesHarnals, et-al 1988 (73)	a All except newborns (CPHA data base)	a Inhospltal death	a (1) Age group (O-64, 65-74,75+), presence of comorbldltles modeled separately for each DRG cluster[cc]	a. Patient	a. R*= 81 (1983 data) R^2= 84 (1984 data)	a —
			(2) Age, sex: race, existence of secondary diagnoses, cancer except skm cancer as a secondary diagnosis, risk of death associated with prmc:pal diagnosis; risk of death associated with first Class i operative procedure, risk associated with comorbidity having the highest risk, number of secondary diagnoses (except complications) where the nsk of death was greater for the secondary dlagnosls than for the DRG cluster itself			
	b. All (HCFA data base)	b Inhospltal death	b Same as "a"	b. Same as "a"	b. R^2=,48	b –

=breviations' ALC =Alternative Care; CPHA = Commission on Professional and Hospital Activities; DRG = diagnosis-related group; HCFA = Health Care Financing Administration; ISMR = Indirectly Standardized Mortality Ratio

[a]Numbers in parentheses refer to numbered entries in the reference list at the end Of this report.

[b]Dash (—) indicates no attempt was made to validate results against process Of Care.

[c]Anesthetists' ratings.

[d]Combination of age and physical status.

[e]A.B Flood, Associate professor, Medical Humanities and Social Sciences program, College of Medicine, University of Illinois, Urbana, IL, personal communication, Sept. 1 i', 1987.

[f]Fo, blood pressure, temperature, hemoglobin, hematocrit, urine sugar, and albumin

[g]Another study, the Service Intensive Study (SIS), examined the variation in clinical Services received and outcomes achieved by all (N = 603,580) patients discharged from 17 IS hospitals during the study period 1970-73 (214,221). Thus, the SIS differed from the IS by: including 3 years of patient outcomes; excluding interview and other obtrusively (relative to the ES) collected IS data; and including data for all patients, not just the surgical patients whose care was emphasized i n the ES and IS. The SIS found that lower death rates were significantly related to the receipt of more intensive services, and that higher death rates were related to the duration of services (that Is, the number of days in the hospital).

[h]For th. outcome "death within 40 days of surgery or severe morbidit, on the seventh postoperative day" results were a 10:1 difference between the highest and lowest mortality hospitals (0.37 to 3.7 percent) before a Bayesian adjustment, and 31 after a Bayesian adjustment.

Table 4-3.-Results of Hospital Mortality Studies Reviewed by OTA-Continued

‘And s!gntficant interaction between hospital and difficulty of procedure.
jD@p@nd@d on surgical category, but generally, age, physical status, stage of disease, and quadratic function were Significant.
ko utcome was death within 40 days of surgery or moderate or severe morb~dity at 7 days after surgery.
{process of care e{tjons were done for a subset of hospitals and patientS (12 general hospitals, 5gFj cases), but the r@sults w@r@ not compared to the hospital mOrtatity r@ SUttS. pfOC@SS Of Car@ Crik?ria Included the fraction of surgical patients given selected initial examinations, given specific patient education, and given home-care instruction or a follow-up appointment.
mKnaus and hiS colleagues found that the major Portjon of increased therapy given at Hospital 1 (the hospital with the {OW@St mortaiity rate) came from fr@WJent laboratory testm9, dr@ssin9 chan9@s, and chest physiotherapy, which resulted from extensive reliance on a clinical protocol, and not from increased use of unique technologies such as ventilators or pulmonary artery catheters.
nKnaus and colleagues als consider interaction and coordination of staff to be process measures, but OTA considers them structural ‘easures
ONote: Medic e patients only, Note fuflher that all ;dedicar@ included Medicare patients of all ages, nOt jUSt those 65 and OV@r
PState average length of stay explained most of th@ variation in mortality. Age, sex, and race variables were flOt Significant.
qFift@en variables were significant, including proportion of transfers from long term care, average age, percent discharged to other hospitals, percent with residence (n Sam@ county aS hosPital, percent with length of stay longer than 90 days, percent with ALC days, and case-mix measure (278)
rEach of 50 DRGS with highest number of ~ths (as opposed t. admissions, as used in the 1984 HCFA analysis [640]); each DRG with the same dia9nosis as thos@ 50 DRGsi all r@mainin9 cancer DRGs; each hospitals’s predicted mortality rate based on Statewide rate for ORG.
Spropo~ions of unscheduled admissions, discharges to another acute care facility, transfers from a hospital discharges from alternate care, discharges frofrl Sarlle COUfltY &S the hospital, number Of tranSf@rS from a hospkal less number of discharges to a hospital divided by total number of discharges (’net migration”), percent of patients with length of stay greater than 90 days.
tfq}neteen “ariab\es were significant including proportion black, percent of transfers from residential health Car@ facilities, and case-mix lnd@x.
USeven teen v~riables were significant, including percent black, percent transfers from other hospitals, proportion with residence in same county as hospital, prOpOrfiOtl with length Of Stay gr@at@r than ~ days excluding ALC stay, proportion with ALC days, and case-mix.
‘Proportion males, proportion with Medicaid as primary or secondary payor, and proportion with length of stay greater than 90 days excluding ALC stay are all significant.
WFour varia&\@s were significant: Age (percent older than 70); percent @mitt@d from emergency departrnerrt; percent admitted from nursing home; &XX3-IIth4 index based On DRG weights.
‘The body-system score was a comorbidity scale for each patient that reflected the number of body systems (e.g., cardiovascular) that were affected by any of 50 comorbidities present on the day of admission.
yNot@ that analysis was done at the patient level and then aggregated to the hospital I@V@\.
zA\mOst all of the variance w s explained by 10 variables. age Over M, severe acute flfjaf'f disease (as a case mix variable), sepsis, pulmonary disease, Cancer aS a comorbidity, c@r@brOvascular accidents, r@nal disease as a comorbidity, metabolic and electrolyte disturbances, severe chronic heart disease, and age between 70 and 74.
aaAny of four additional dia~no$es (of cancer, chronic liver disease, chronic renal disease, chronic cardiovascular disease, chronic pulmOnary disease, cerebrovascular @9enerati@chronic P$Ychosis> WJer-tenswe dwease, or diabetes) beyond the principal diagnosis.
bbH Kraus, Office of Medical Review, Health Standards and Quality Bureau, Health Care Financing Admmistration, U.S. Department of Health and Human Services, personal communication Baitimor@, M“D, Mar 7, 1988.
ccComorbjdit]@S were based on ICD.9.GM cod@s Codes that w@r@ c\@arfy complications were nOt considered comorbidities.

SOURCE Office of Technology Assessment, 1988.

ses reported checking carefully the reliability of data sources. Reliability is of particular concern for hospital mortality analyses. As currently constructed, such analyses are based on small numbers and data for single years. Differences in coding, interpretation, and aggregation across time, across coders or reviewers, and across hospitals could substantially affect hospital comparisons.

Evidence indicates that errors in diagnostic labeling are fairly common (166,167,614). These findings are not surprising given the amount of subjectivity that still exists in coding (77). Errors can be made by the physicians who diagnose the patients' condition and by medical records personnel who transform the diagnoses into universal codes, such as those used in the International Classification of Diseases (ICD-9 codes) and those used for diagnosis-related groups (DRG codes). Random errors in diagnostic labeling undoubtedly exist and generally are not of concern when comparing mortality rates across hospitals, but systematic errors in diagnostic labeling could affect the comparisons. For example, a hospital would have an artificially low expected rate of death from pneumonia if it included in the diagnostic category for pneumonia patients who actually had a less serious illness, such as bronchitis (190). The relationship between tendencies to have coding errors and quality-of-care problems, however, remains unclear.

Using data reported to HCFA by hospitals seeking Medicare reimbursement, the HHS Office of the Inspector General (OIG) found a 20.3 percent error rate in coding across hospitals (304,660). The study was conducted with data from October 1984 to March 1985. A significant number of the errors favored the hospitals; that is, the hospitals were paid more for the hospital stay than they would have been if the correct codes had been submitted (so-called DRG "creep"). A common error was the transposing of principal with secondary diagnoses. A statistically nonsignificant trend was found for differences by hospital bed-size, with smaller hospitals tending to upgrade patient diagnoses. Potentially, this upgrading could lower small hospitals' adjusted mortality rates.[3] Bed-size was the only hospital characteristic used in the analysis. In another arm of the study, the OIG found a higher incidence of DRG "creep" in cases that were discharged prematurely (660a). Hospitals commenting on the 1986 HCFA analysis also reported miscoding of diagnoses so that secondary diagnoses were recorded as principal diagnoses, and vice versa.

In a study that used data from non-Medicare as well as Medicare patients, Dubois and colleagues found a rate of coding errors across hospitals similar to that found by the OIG study (20 percent); but they found that the error rate did not differ significantly between high- and low-outliefl hospitals (190). Thus, in this study, coding errors seemed not to be responsible for differences in hospital mortality rates.

Another potential source of differences among hospitals, and thus unreliability in the data, is the extent to which secondary diagnoses are recorded. Consistent recording of secondary diagnoses is essential when such diagnoses are used to indicate comorbidities, a commonly used source of information about patients' risk of dying (172,353, 588,589,590,640,647). In connection with an analysis of hospital mortality rates, the Commission on Professional and Hospital Activities found substantial variation among hospitals in the extent to which they recorded secondary diagnoses (172). When secondary diagnoses are used as proxies for comorbidities, lack of documentation could affect a hospital's expected mortality rate.

The reliability of information about the patient's clinical status on admission can be affected by incomplete entries or inconsistency across raters in recording the information that is available. Incomplete coding of clinical data is a major drawback to the use of patient classification

[3]The Inspector General's study did not include a review of mortality rates.

[4]After adjusting for patients' risk of dying, analyses estimate for each hospital an expected mortality rate. They then compare the hospital's actual mortality rate to the expected one. Typically, hospitals whose actual rates exceed the expected rates by more than 1.96 standard errors are considered high outliers, and hospitals whose actual rates fall below the expected by more than 1.96 standard errors are considered low outliers. This type of analysis assumes that hospital mortality rates follow a normal distribution, although that assumption has not been validated. See Blumberg and DesHarnais for further discussion of statistical issues surrounding hospital mortality analyses (77,172). In addition, the General Accounting Office is preparing a report on Medicare's use of patient outcome data (626).

systems based largely on clinical data (94,352, 353). When the State of Pennsylvania decided to publish outcome statistics adjusted with clinical data, for example, it simultaneously implemented a requirement that all hospitals use the same classification system, so that the needed data would be available from all hospitals (41,427). Presumably such a requirement would encourage more consistent recording of such data.

Interrater reliability for the clinically based patient classification systems that are being used in mortality analyses (94,352,353) seems good, however. Thomas, et al., found almost perfect interrater reliability for the APACHE 11[5] and MEDISGRPS[6] systems, and relatively good reliability for the Clinical Staging system of SysteMetrics (614).

Type and source of hospital admission are sometimes used as proxies for patients' risk of dying (80,81,648). Coding of such information can be another source of error. The study by the California utilization and quality control peer review organization (PRO) of premature discharge notes that guidelines for admission source are subject to interpretation by coders (117). For example, it is unclear whether the referring physician or the transferring facility takes precedence. With transfer from another hospital a surrogate for patients' risk of dying (648), errors in coding source of admission could have affected hospital results. Some hospitals responding to the 1986 HCFA analysis commented that sources of admission had been recorded incorrectly by HCFA (648). Similarly, Blumberg eliminated 10 hospitals from his analysis of Maryland hospital data because they differed from other hospitals in the way they coded whether admissions were emergent, urgent, or elective (80,81). Only nonelective surgeries were included in Blumberg's study. If elective surgeries, which presumably entail less risk of death, were included for some hospitals and not others, the results would not have been valid.

Blumberg has noted a discrepancy between in-hospital deaths reported to State agencies and those reported to HCFA in 1984, with the number reported to HCFA lower than that reported to States **(78).** Similarly, a study by the California PRO found that 23 percent of cases that had been coded as being discharged alive from California hospitals had actually been discharged dead (117). The reasons for these errors are for the most part unclear; the California PRO study did find, however, substantial miscoding in the DRG series for patients with acute myocardial infarction. In that DRG series, Medicare payment for patients who are discharged dead is lower than payment for patients discharged alive.

Differing hospital policies concerning the point at which individuals are declared dead (141) and varying do-not-resuscitate policies do not affect the coding of death, but affect the reliability of patient death information across hospitals, which in turn affects the reliability of hospital mortality rates as an indicator of quality.

Statistical analyses should be validated with reviews of medical records. A significant problem in reviews of medical records has been interrater reliability (see ch. 7). The one published study of hospital mortality rates that addressed reliability among reviewers found good interrater reliability when reviewers used explicit criteria, but poor interrater reliability for subjective judgments of care (190). Other studies comparing explicit with implicit review have found similar results (see ch. 7).

'Acute Physiology and Chronic Health Evaluation.
'Medical Illness Severity Grouping System.

VALIDITY OF THE INDICATOR

Intelligibility of Hospital Mortality Rates as an Indicator of Quality

To be useful as an indicator of the quality of care, hospital mortality should be understandable to both consumers and providers. Anecdotal evidence indicates that consumers seem well aware that a patient's inherent risk of dying is a prime contributor to whether a patient lives or dies during or soon after a hospital stay. They also seem aware, however, of the hospital errors that can result in patient death.

For providers, examination of individual patient deaths may have face validity, but aggregate hospital mortality rates may not. According to Friedman and Shorten, mortality is the outcome that always receives the most intensive scrutiny by hospital managers and clinical chiefs of staff (237). Particularly in teaching hospitals, the medical staff discusses the causes of unexpected individual patient deaths (at least those deaths among patients of interns and residents) and suggests improvements in care. There is little evidence that hospital staffs examine overall hospital mortality rates or rates within hospital departments on a systematic basis (224). To date, providers have regarded skeptically attempts such as HCFA'S to adjust statistically for patient characteristics that would explain high mortality rates so that the remaining explanation for differences among hospitals is the quality of care (97,537). It is unclear, for example, whether practicing physicians believe that a patient's likelihood of death can be predicted using systematic means. The use of mortality rates may be gaining in acceptance, however. The Commission on Professional and Hospital Activities (CPHA) reports in its mortality analysis that hospitals informally confirmed that high outliers had quality problems (172).

When To Measure Death

Researchers and policymakers do not yet (and may never) agree on when to measure an outcome of hospital care. Regional variations in lengths of stay among hospitals, differences in admitting and discharge practices, and unequal access to home care and hospice services in communities can determine whether a death occurs in the hospital or out of it (141). There seems to be considerable agreement that merely counting deaths at discharge is not a completely valid way to compare hospital mortality rates, because such a technique may reward hospitals that discharge patients in more serious condition, who may then die elsewhere. To capture a high percentage of deaths that may be attributable to poor-quality care, some analyses have used all deaths occurring within some time frame after an admission *or* a procedure, even if they did not occur in the hospital (588,589,647). This approach may, however, measure the effect of events unrelated to the quality of a hospital's care.

In empirical work relating to these issues, the Stanford Institutional Differences Study obtained essentially the same results from its Extensive Study (deaths at discharge only) as it did from its Intensive Study, which measured deaths at 40 . days (even after discharge) or severe morbidity within 7 days of surgery (215,588). DesHarnais and her colleagues analyzed HCFA'S 1986 data and found an almost perfect correlation between inhospital mortality rates and 30-day-post-admission mortality rates. DesHarnais and her colleagues concluded that "it does not matter which measure is used in terms of assessing hospitals' relative rankings" (172). It may be, however, that the conclusion would differ if all admissions rather than last admissions were included in the analysis. HCFA'S 1986 analysis used patients' last admission of the year as the denominator in its analysis. A further consideration is that the appropriate time at which to measure outcome may vary for different conditions (500); this issue has not been tested.

For practical reasons, or because no valid endpoint has been established, most analyses have measured inhospital death only (80,81,189,190, 259,353,448,526,640). Clearly, this question requires careful thought and additional study.

Adjusting for Patients' Risk of Dying

One of the most challenging questions in quality assessment is how to construct an indicator that is not confounded with the characteristics of the patients who come to the hospital. In most analyses, the patient attributes used to adjust for the risk of dying have been only rough proxies for characteristics that may be better measured by physiologic values (see table 4-3), although the physiologic values that predict death are as yet unknown (596). Studies that use claims data alone are limited to the data elements present on claims, such as Medicare's UB-82. These claims indicate patient characteristics, such as age, sex, and race; the principal diagnosis for which the patient was admitted to the hospital and up to five secondary diagnoses; the principal procedure and up to three secondary procedures; some potential sources of admission; type of admission (emergency, urgent, elective, newborn); discharge sta-

Photo credit: Foster Daily Democrat

Age is at best a crude indicator of patients' inherent risk of dying.

tus (including dead or alive and, if alive, place discharged to); and other types of information less relevant to hospital mortality analyses (657).

Age may be the most frequently used adjustment for patient mix, and there is, of course, a correlation between age and the likelihood of death. However, the relationship is not completely linear (667), and age remains at best a crude indicator of a patient's health status or physiologic reserve (76). HCFA found, for example, that "average age of Medicare patients" was not a significant predictor of mortality at the hospital level of aggregation (640). The 1986 HCFA analysis used age groupings instead of average age of patients in the hospital. At the hospital level of aggregation, several age categories were statistically significant. In other studies using data at the patient level of analysis and more refined methods of adjustment, age has been found to be significant (353,588,589). Even if measured adequately, however, a number of studies have shown that age can also be a risk factor for inadequate or poor treatment (134,318,549,700a). Similarly, adjustments for sex, race, and socioeconomic status can mask an interaction between a patient characteristic and the provision of poor-quality care (191). Average length of stay (526, 640), for example, seems particularly invalid as a hospital level adjustment for patient risk. Longer lengths of stay can themselves be indicative of poor quality. The use of easily available discharge data to adjust for case mix is a threat to the validity of the hospital mortality measure, because a patient's risk of dying cannot be adequately inferred from diagnostic categories such as DRGs or ICD-9 codes (629,630).

Measures that rely at least in part on clinical data on admission would appear to have more validity than proxy measures such as age, sex, race, source of admission, and comorbidities (352,353). A recent review of the status of severity measures concluded that "although intrinsic biological severity may one day be measurable, currently it is an abstraction" (596), but some classification systems have reported good results (93,94,190,352,353). Williams was able to explain about 80 percent of the variance in neonatal mortality using a combination of birthweight, sex, race, and whether the birth was multiple (i.e., twins); by far the best explanatory factor was birthweight (545,702). Brewster's MEDISGRPS technique relies entirely on clinical findings, while Knaus' APACHE II method includes age and some comorbidities,[7] as well as clinical findings (see figure 4-1). Perhaps in line with the conclusion that measuring intrinsic biological severity is difficult, Brewster's results are not as impressive as Knaus'. Brewster's mortality results have been published for shortness of breath (93), abdominal pain, and chest pain (94) as reasons for admission; Knaus' for patients in the intensive care unit (352,353).

Even these patient classification systems may not be able to cope with the fact that the patient's condition may change during hospitalization regardless of the medical care provided. Having some clinical information about the patient's status on admission seems clearly better than relying on comodidities and complications recorded after discharge, because existing coding schemes cannot clearly distinguish between comorbidities existing on admission and complications acquired as a result of hospital care. But a patient's status on admission to the hospital will not reflect changes in the patient's status that occur solely as a result of the trajectory of illness.

Appropriate measures of patients' risk of dying may differ considerably by disease category or patient condition. Measures that mix deaths of patients due to chronic or late-stage conditions with those of patients having more acute, less severe illnesses, and use one type of adjustment may not be nearly so valid as measures using either one or the other type of condition. Conclusions about the most appropriate aggregations and adjustments for patients' risk of dying are difficult to draw from existing studies because of the wide variation in methods used. Only one study has actually analyzed data for the hospital as a whole, with no conditions or patients excepted (189). Others have removed from consideration obstetric patients (259,526), or considered only elderly and disabled patients (640,647). The HCFA patient data base is composed primarily of patients aged 65 and over.

In general, however, analyses at the hospital-level of aggregation have been able to explain

[7] Unlike the comorbidity measure used in most adjustment methods based on claims data, the comorbidities in Knaus' APACHE II scheme must have been present within 24 hours of hospital admission (352, 353).

Figure 4-1.-Scoring of Patients Under the APACHE II System for Classifying Severity of Disease

APACHE II score = Sum of A + B + C

A — Acute physiology score

The acute physiology score is the sum of points for 12 physiologic variables:
- temperature,
- mean arterial pressure,
- heart rate,
- respiratory rate,
- oxygenation,
- arterial pH,
- serum sodium,
- serum potassium,
- serum creatinine,
- hernatocrit,
- white blood count,
- Glasgow coma score.

Each variable is scored from -4 to +4 points.

\+

B — Age points

Age points are assigned to patients according to their age as follows:

s 44-0 pts
45-54-2 pts
55-64-3 pts
6&74 -5 pts
~ 75-6 @

\+

C — Chronic health points

For patients who have a history of severe organ system insufficiency or are immunocompromised, points are assigned as fok)ws:

a. for nonoperative or emergency postoperative patients -5 pts
b. for elective postoperative patients -2 pts

SOURCE: Office of Technology Assessment, 19SS, adapted from W A Knaus, E Q Draper, D P Wagner, et al , "An Evaluation of Outcome From Intensive Care in Major Medical Centers," Annals of *Internal Medicine* 104:410-418, 1986 more of the variation in mortality than have analyses at more condition-specific levels, although there have been rather high proportions of variation explained for certain conditions. Hospital-Ievel analyses have accounted for between 35 and 93 percent of the variance. This is not surprising because random variation is less likely at the level of the institution.

Some differences in the amount of variance accounted for among diagnostic categories may be explained by the use of inappropriate variables to adjust for patients' risk of dying. Another potential explanation for differences among diagnostic categories is the extent to which medical care and its quality influence death rates. Therapy is unlikely to prevent the deaths of late-stage cancer patients, so not much variation is introduced by factors not accounted for in a regression equation (357). For early heart disease, on the other hand, good treatment does exist and its application does make a difference, so the patient's condition may account for little of the variation in patient mortality.

Validation of Hospital Mortality Rates Against the Process of Medical Care

The best way to establish hospital mortality rates as a valid indicator would be to demonstrate a link between the process of care and the outcome of death. Some studies have attempted to do this, with conflicting results. In response to HCFA'S analysis of 1984 data, which showed **29 New** York State hospitals as having higher than expected mortality rates, the New York State Department of Health conducted a regression analysis with its own set of adjustments for patients' risk of dying, modified from HCFA'S 1984 model (462). New York State found fewer outliers[8] than did HCFA. The Department then had PRO personnel examine the medical records of patients in DRGs with mortality rates above the statewide average. The reviewers concluded that only about 3 percent of these cases had quality-of-care problems (278,461).

[8]Outliers **are hospitals that have mortality rates that are significantly either higher or lower than expected.**

In 1987, New York State did not do a regression analysis, but compared the results of its targeting certain deaths for review to HCFA'S analysis of 1986 data (279,461). In this comparison, only 1 hospital of the 10 identified by HCFA as being high-mortality outliers had quality problems using New York State's standards. In general, high outliers had fewer problems than non-outliers (279).

Dubois and his colleagues used both explicit and implicit review to determine whether quality problems existed in hospitals initially identified as high or low oudiers using claims data (190). The explicit review compared the medical care provided (as reflected in the medical records) to a provisional list of criteria for quality of care in the management of patients. In the implicit review, experts read a summary of the patient's care and judged whether the death was preventable. Dubois's validation study is impressive because it was careful to test the possibility that factors other than quality, such as patients' characteristics related to their risk of dying, accounted for differences in mortality rates, for three of the most common causes of death (190).

Explicit review resulted in no apparent differences in numbers of preventable deaths, and implicit review found significant differences between high- and low-outlier hospitals in preventable deaths only for pneumonia, not for acute myocardial infarction or cerebrovascular accidents. After adjustment for differences in patients' risk of dying and for the fact that deaths were oversampled, however, the researchers found significant differences in preventable deaths between high and low outliers for cerebrovascular accidents and pneumonia. They estimated thats percent of patients with those conditions entering one of the high outlier hospitals would have a preventable death, compared to a l-percent chance of preventable death in a low-outlier hospital. The authors concluded that adjustments using claims show some promise of identifying hospitals with variations in quality, although their study should be regarded as preliminary.

Knaus, et al., found that the best ranked intensive care unit in their study of 13 hospitals used significantly more therapeutic interventions than all the other hospitals (353). Knaus realized that

84-752 0 - 88 -- 4

the amount of treatment is not a good indicator of differences in quality; he examined the components of increased treatment at the "best" hospital, and found differences in the type of treatment provided.

Somewhat similarly, the Stanford Institutional Differences Study included some crude indicators of the process of care (588,589,590). The process measures, all at the hospital level, were the rate of pathology reports, the rate of pathology reports showing the presence of disease, the rate of pathology reports showing no disease, and the autopsy ratio. The study found no significant relationships between inhospital death and any of the process measures. The study's original plan was to conduct a better validation study, but this plan was not supported because it was judged to be too lengthy and expensive (588).

The results of these studies should be regarded cautiously, however. Both New York State and Dubois and his colleagues used implicit review of records, which may be unreliable (190) (see ch. 7 of this report). New York State's targeted mortality study concentrated largely on surgical patients, while HCFA'S analysis covered all reasons for admission. New York State's 1984 model adjusted for some factors that could have been related to quality of care.

Comparisons of Hospital Mortality Rates With Other Potential Measures of Quality

Some reviewers of the literature on hospital mortality have concluded that hospital mortality may have some validity as a quality indicator because mortality showed theoretically expected relationships with other potential measures of the quality of care. In a review of 18 studies of hospital mortality, for example, Fink, Brook, and Yano found that the following hospital characteristics were associated with better outcomes: frequency of performing a procedure, size, communication among staff, commitment of staff, clinical experience, board certification of staff, and teaching status (209).

Some of the studies reviewed for this report also examined relationships between hospital mortality and potential measures of quality other than hospital mortality, primarily structural measures. One study compared mortality to scales combining mortality and morbidity (588). The results of these analyses, shown in table 4-4, indicate some significant relationships between primarily structural measures of quality, defined quite variably among studies.

Comparison of Different Hospital Mortality Analyses

Hospital-specific mortality rates have been released by a variety of sources (55,77,81,115,116, 640,647). The New York State Department of Health also conducted two analyses in response to HCFA'S releases; hospital-specific information related to these analyses were not released to the public (279,462). Some releases are of unadjusted mortality rates (55,115,116) and other analyses attempted to adjust for patient characteristics (80, 81,462,640,647). Comparisons of these analyses are instructive in several ways: they illustrate the different results obtained when mortality rates are analyzed specific to diagnoses or procedures versus aggregated by hospital; they show the potential importance of adjusting hospital mortality rates for patient characteristics; and they show the variation in results obtained when different risk-adjustment procedures are used.[9]

California

Several available data sets contained information on California hospitals: HCFA'S releases of 1984 and 1986 adjusted data (640,647) and analyses by three newspapers of unadjusted data released by California Medical Review, Inc., the California PRO for fiscal year 1985 and 1985-86 (12,359,597). Because these sources differ in several ways, some variation in results is expected. In particular, the California PRO releases were completely unadjusted for patients' risk of dying. On the other hand, all releases pertained only to Medicare patients, and the years analyzed were contiguous, so one might expect some overlap in

[9]The analysis is also limited. It recorded only the presence or absence of a hospital on a particular list. Alternative approaches would rank the hospitals or use actual mortality rates or ratios, standardized in some way. However, the large number of comparisons might also preclude tests of statistical significance.

Table 4.4.—Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA[a]

Study[b] Structural variable(s)	Variable significantly related to hospital mortality: Lower mortality	Variable significantly related to hospital mortality: Higher mortality	Variable not related to hospital mortality
Roemer, et al., 1968 (526)			
1. Technological' Adequacy Score[c]	+		
2. Hospital control:			
a. Voluntary			+
b. Proprietary			+
Roemer and Friedman, 1971 (525)[e]			
1. Medical staff organization:[d]			
a. Permissive control		+	
b. Medium control	+		
c. Strict control			+
Goss and Reed, 1974 (259)[f]			
1. Technological Adequacy Score9			
2. Hospital control:			
a. Municipal		+	
b. Voluntary			+
c. Proprietary			+
3. Teaching status:			+
a. Some commitment to teaching		+	
b. No teaching approval			+
c. Greatest commitment to teaching.	+		
d. Hospital control and teaching status combined[h]			
Stanford Center for Health Care Research, 1974 (588~1976(589~			
Flood and Scott, 1987 (215~			
Hospital Characteristics?			
l. Medical staff structure (ES)j[k]			
a. Hospital-employed physician ratio.		+	
b. Surgical-staff-to-patient ratio	+		
2. Nursing staff structure (ES)			
a. Proportion of part-time nurses	+		
b. Proportion of full-time nurses who are registered nurses	+		
c, Nurse-to-patient ratio	+		
3. Medical staff structure and nursing staff structure combined[m]			
impact ofSurgeons and Surgical Staff Organization (IS)[no]			
l. Proportion of contract physicians	+		
2. Number of surgical specialties in the department	+		
3. Average percentage of practice conducted at the study hospital			+
4. Proportion of board-certified surgeons			+
5. Strictnessof admission requirements for new members			+
Surgeon Characteristics:[n]P			
l. Percent of practice conducted at study hospital	+		
2. Number of residencies surgeon has completed	+		
3. Number of years in practice			+
4. Board certification			+
5. Surgical specialization			+
Within-Domain and Encroaching Influence (/S):[o]~			
1. "Control variables":r			
a. Percentage of surgeon's practice conducted at hospital			+
b. Hospital expenditures	+		
c. Patient's income	+		
2. "Power variables":			
a. Influence of the hospital administration within its own domain	+		
b. Encroachment by physicians on the nursing administration			
c. Influence of the nursing administration within its own domain		+	

Table 4-4.—Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA—Continued

Study[b] Structural variable(s)	Variable significantly related to hospital mortality: Lower mortality	Variable significantly related to hospital mortality: Higher mortality	Variable not related to hospital mortality
d. Encroachment by physicians on the hospital administration's domain			+
e. Influence of the surgical administration within its own domain			+
f. Encroachment by the hospital administration on the surgical administration			+
Power of Surgical Staff Over Its Own Members (IS):[o q]			
1. "Control variables":			
a. Percentage of surgeon's practice conducted at hospital			+
b. Hospital expenditures	+		
c. Patient's income	+		
2. "Power variables":			
a. Power of surgical staff over tenured surgeons	+		
b. Admission requirements for new members of the surgical staff	+		
c. Centralization of decisionmaking within the surgical staff			+
Selected Control Variables (IS and S1S):[s t]			
1. Frequency of case discussions with pathologists			+
2. Control exercised by surgical staff over tenured surgeons			+
3. Chief of surgev's administrative infiuence in own area			+
NAS, 1977 (448)"			
1. General, urologic, and orthopedic surgeries combined:			
a. Degree of affiliation with a medical school			+
b. Proportion of surgeons who are board certified			+
c. Proportion of surgeons who are residents			+
d. Average age of surgeons			+
e. Absolute number of surgical beds			+
f. Proportion of acute-care beds allocated to surgery			+
g. Complication rate"			+
2. Cardiac surgery:			
a. Volume	+		
Knaus, et al., 1986 (353)			
1. Administration of unit, scope of service[w]			+[x]
2. Teaching status			+
Blumberg, 1987, 1988 (80,81)			
1. Teaching status			+

Study	Correlations among outcome variables ([S)	Outcome A	Outcome B	Outcome C	Intermediate scaled outcome
Stanford Center for Health Care Research, 1974 (588), 1976 (589); Flood and Scott, 1987 (215)	1. Outcome A: Death within 40 days of surgery (including after discharge)	—	Moderate positive correlation	Negative correlation	Small positive correlation
	2. Outcome B: Death within 40 days of surgery or severe morbidity at 7 days after surgery	—	—	Small positive correlation	Strong positive correlation
	3. Outcome C: Death within 40 days or surgery or severe or moderate morbidity at 7 days after surgery	—	—	—	Strong positive correlation
	4. Intermediate Scaled Outcome (ISC): Dead (9 points); severe (5) or moderate morbidity (2); else (0).	—	—	—	

Table 4-4.—Comparison of Hospital Mortality With Structural and Other Outcome Indicators in Hospital Mortality Analyses Reviewed by OTA—Continued

[a]Only structural and outcome indicators are included in this table. Inclusion of process variables in studies is shown in table 4-3. Most of the analyses were a part of the primary publication reviewed by OTA. This table also includes, however, closely related studies using the hospital mortality indicator developed in the 13 analyses reviewed by OTA. For example, Roemer and Friedman (525) used the hospital mortality indicator developed in Roemer, et al (526).

[b]Numbers in parentheses refer to numbered entries in the reference list at the end of this report.

[c]The components of the Technological Adequacy Score used by Roemer and his colleagues were as follows, with points assigned to each component in parentheses

1. Accreditation by the Joint Commission on Accreditation of Hospitals (20)
2. Approved residency or internship (10)
3. Approved cancer program (8)
4. Intensive care unit (7)
5. Pathology laboratory (5)
6. Blood bank (5)
7. Therapeutic X-ray (5)
8. Postoperative recovery room (5)
9. Rehabilitation service (5)
10. Outpatient department (8)
11. Home care program (8)
12. Social service department (7)
13. Chest X-ray on admission (7)

A total of 100 points could be scored. The source of data was hospitals' reports to the American Hospital Association.

[d]Roemer and Friedman devised a typology in which they defined medical staff organizations along a continuum from loosely structured or permissive to highly structured or vigorous (see Roemer and Friedman, 1971, ch. 5) Many of the components of the medical staff organizations were subsequently disaggregate in studies using data from the Stanford Institutional Differences Study (see Flood and Scott, *Hospital Structure and Performance*, 1987).

[e]Roemer and Friedman analyzed data for only 10 general hospitals in California but included a Veterans Administration hospital, Hospitals were chosen to represent a range of medical staff organization types, from loosely to most highly structured. Hospitals were also chosen to be generally meritorious.

[f]Goss and Reed used the same severity adjustment method as Roemer and his colleagues used.

[g]Goss and Reed used the same scale for the Technology Adequacy Score as Roemer and his colleagues used, except that "chest X-ray on admission" was omitted because data were not available.

[h]Municipal hospitals with internship and/or residency approvals had the highest severity-adjusted death rates; voluntary hospitals with medical school affiliations had the lowest death rates. No statistical tests were performed for any of the structural analyses.

[i]Source. A.B Flood, W R. Scott, and W. Ewy, "Hospital Characteristics and Hospital Performance," *Hospital Structure and Performance,* A.B. Flood and W R. Scott (eds.) (Baltimore, MD: John Hopkins University Press, 1987).

[j]The initials ES, IS, and SIS in the entries that follow indicate whether the analysis was conducted with data from the Extensive Study (ES), the Intensive Study (IS), or the Service Intensity Study (SIS). The outcome in the ES and the SIS was in hospital death. The outcomes in the IS were, for the logistic regression, death within 40 days of surgery (including death after discharge), death within 40 days of surgery or severe morbidity at 7 days; death within 40 days or severe or moderate morbidity at 7 days. For the linear regression, moderate and severe morbidity at 7 days and mortality within 40 days were combined into a scaled measure. Only the Intermediate Scaled Outcome was used for most analyses (death [9 points], severe morbidity [5], moderate morbidity [2], and no or mild morbidity [0]),

[k]Adjusted for hospital size, teaching status, and expenditures, as well as patient characteristics.

[l]Note that in many of the analyses, hospital characteristics (size, teaching status, and expenditures) were controlled in addition for patient health characteristics.

[m]Results for medical staff and nursing staff combined were almost identical to those for individual variables, but when both sets were combined, the results for Proportion of full-time nurses who were registered nurses were not significant,

[n]Source. A B. Flood, W.R. Scott, W Ewy, et al , "Effectiveness in Professional Organizations," *Hospital Structure and Performance,* A B Flood and W R Scott (eds.) (Baltimore, MD John Hopkins University Press, 1987).

[o]Using the Intermediate Scaled Outcome (death [9 points], severe morbidity [5], moderate morbidity [2], and no or mild morbidity [0]),

[p]Aspects of hospital context (size, teaching status, and expenditures) were included in the analyses.

[q]Source: A.B Flood and W.R, Scott, "Professional Power and Professional Effectiveness: The Power of Surgical Staff and the Quality of Surgical Care in Hospitals, " *Hospital Structure and Performance,* A.B. Flood and W.R, Scott (eds.) (Baltimore, MD: John Hopkins University Press, 1987)

[r]Control variables entered into the analysis first

[s]The outcome in the Service Intensity Study (SIS) was inhospital death.

[t]W R Scott, A B, Flood, and W Ewy, "Organizational Determinants of services, Quality, and the Cost of care in Hospitals, " *Hospital Structure and Performance,* A B Flood and W R Scott (eds) (Baltimore, MD: John Hopkins University Press, 1987).

[u]This study used basically the same method as the Stanford Institutional Differences Study (215,588,589), apparently without the admissions data.

[v]Data were collected for only 12 hospitals. The data were not routinely available in existing reports and the researchers were required to ask various hospital personnel for parts of the record. Further contributing to the possible lack of validity of this measure, the authors note that the definition of complication was somewhat subjective

[w]Knaus, et al, based their designations of ICU levels on guidelines of the National Institutes of Health (NIH) Consensus Development conference On critical care (334) The NIH Conference included variations in technological capability in its designation of levels. The hospitals in Knaus, et al sample all had the same technological capability, however, so the assignment of levels was based on administrative structure only (353).

[x]The hospital with the lowest adjusted mortality rate was a Level I unit, and the hospital with the highest adjusted mortality rate was a Level III unit As a group, however, Level I units did not do better than Level II or III units.

SOURCE: Office of Technology Assessment, 1988.

results. The appropriate comparisons are between HCFA'S results for 1984 and 1986 and all other results, because data from the California PRO are actually for three different geographic areas.

With all sources and types of diagnoses and procedures combined, 143 (29 percent) of the approximately 490 California Medicare hospitals were either high- or low-mortality outliers in at least one analysis. Twenty-seven hospitals (I9 percent of the 143 ors percent of all California hospitals) appeared as outliers in more than one analysis.

New York

As described above, New York State undertook two types of analyses to validate HCFA'S releases. One was a regression analysis to detect outliers and the other was a targeted mortality analysis validated by PRO staff. The regression analysis was applied to the 1984 data, and the targeted mortality technique was applied to the 1986 data.

For the 1984 analysis, the New York State Department of Health used its own extensive data base to create predictor variables somewhat different from HCFA'S, although like HCFA'S, the analysis was conducted at the hospital level (462). In addition to identifying outliers for Medicare patients, New York State identified high-mortality outliers in 1984 for all patients under age 65, all discharges, and obstetrics / nursery services. The results of OTA'S comparison indicate that 52 (19 percent) of New York State's 274 Medicare hospitals were high-mortality outliers on at least one of these 1984 analyses. Twenty-nine hospitals were HCFA outliers. But only 12 (23 percent) of the 52 high-mortality outliers were high outliers in both the 1984 HCFA aggregate analysis and at least one of the 1984 New York State analyses. Only half of these 12 were both HCFA and New York State Medicare outliers, indicating that New York State was able to replicate only 20 percent of the HCFA high-mortality outliers (6 of 29 HCFA high-mortality outliers) with its model. However, none of the 18 hospitals that were low-mortality outliers in HCFA'S aggregate analysis were high-mortality outliers in the New York State analysis.

New York State used its targeted mortality method to critique HCFA'S 1986 analysis (279). The targeted mortality study approach developed a set of case characteristics that are hypothesized to have a higher than average association with quality-of-care problems (461). New York State hypothesized that reviews targeted at cases rather than outlier hospitals would be more efficient at uncovering quality problems. In the New York State study, the targeting characteristics included procedures rarely associated with death, cases within DRGs that are rarely associated with death, cases in which the patient died in the hospital within 48 hours of surgery, surgical cases with a secondary diagnosis of acute renal failure, and cases with burns or poisoning as a secondary diagnosis. Cases meeting these screening criteria were forwarded for implicit review to a registered nurse; if the nurse concluded that the care provided either might not have met professional standards or might have contributed to the death or disability of a patient, the case was reviewed by one, and possibly two, physicians.

Comparing the results of the targeted mortality study with HCFA'S *1986* analysis, New York State found only one high-outlier hospital in which there was a higher percentage of cases in which care either departed from standards or caused or contributed to patients' death than was found in the nonoutlier hospitals included in the study (279). In general, hospitals that were non-outliers in HCFA'S analysis had more quality-of-care problems than did outlier hospitals. The results of this study should be viewed somewhat cautiously, however, because the targeted mortality analysis used for comparison focused more on surgical than on medical cases, while the HCFA analysis covered all diagnoses.

It is striking that one hospital was a high-mortality outlier in almost all analyses. It did not show up as a problem in the 1984 obstetrics / nursery service data, however. [10]

Maryland

Maryland hospitals have received perhaps the most frequent examination of their mortality rates, although no attempt has been made at replication of specific analytic methods. The following analyses dealt with hospitals in Maryland: Bargmann and Grove (55), HCFA (640,647), Blumberg (80,81), and Washington Consumer Checkbook (693).

There is little convergence among the Maryland releases, which might be expected because of differences among analyses. Maryland has 58 Medicare hospitals (647). Of the 42 hospitals (72 percent) with actual mortality rates higher than those predicted by the various models, 17 hospitals (29 percent) appeared on more than one analysis. Seven of the 17 appeared as low-mortality outliers on one list and as high-mortality outliers on another list. Thus, only 10 (17 percent) appeared as high-mortality hospitals on more than one analysis, and their appearance was frequently for different procedure / condition categories. One hospital appeared as a Medicare outlier in both

10 This hospital was also a high-mortality outlier in HCFA's 1984 analyses of DRG groups 089 (pneumonia), 115 (pacemaker implants), 121 (acute myocardial infarction), 127 (congestive heart failure), and 174 (gastrointestinal hemorrhage), but not DRG groups 106 (coronary artery bypass surgery), 195 (cholecystectomy), 209 (major joint procedures), or 336 (transurethral prostatectomy). Nor was it a low-mortality outlier in the latter groups. It was also a high-mortality outlier overall in HCFA's 1987 release of 1986 data.

1984 and 1986. It was also categorized as a high-mortality hospital for two procedures in Bargmann's and Grove's analysis (55).

Coronary Artery Bypass Graft Surgery in Arizona

Patten compared coronary artery bypass graft surgery (CABG) mortality rates in Arizona for two periods: when the Arizona certificate-of-need process was still in effect (July 1, 1984 to March 15, 1985), and after it was repealed (March 15, 1985 to December 31, 1986) (478). Table 4-5, which compares Patten's data with HCFA data, shows no overlap between HCFA 1984 CABG data (640) and that published for the two periods by Patten in the Phoenix Gazette, For the latter period, there was some convergence between the HCFA 1986 results (647) and Patten's results, although only one hospital was both a HCFA and a Patten outlier in 1986. In considering the 1986 data, one should keep in mind that in 1986 HCFA did not aggregate data by procedures, such as CABG.

Coronary Artery Bypass Graft Surgery in the District of Columbia

Three Washington, DC, hospitals had the three highest crude mortality rates for CABG in a Washington *Consumer Checkbook* analysis (out of 7 hospitals studied in the Washington, DC, metropolitan area), but no Washington, DC, hospital appeared on the list of 1984 HCFA outliers for CABG, or as an outlier for severe chronic and acute heart disease in 1986 (693).

Longitudinal Analyses

For hospital mortality as a measure of quality, it is important to know if a hospital's mortality rate in the past will predict its mortality in the future. However, there are many reasons why a hospital's mortality rate (and quality of care) may change over time, including random error in measurement; changes in the types of patients served; and changes in staff, practices, or procedures. Longitudinal studies are needed to gain insight into the likely role of random error in the

Table 4-5.—Comparison of Hospital Mortality Rates: Arizona[a]

	HCFA 1984		HCFA 1986			Health Services Advisory Group — Coronary artery bypass graft surgery	
	All diagnoses	Coronary artery bypass graft surgery	All diagnoses	Severe chronic heart disease	Severe acute heart disease	7101184 to 3H5185[b]	3115185 to 12131186[c]
Hospital #l ., ...	—	—	—	–[d]	High[g]	N,A.[e f]	High[h]
Hospital #2	High[g]		—	—	—	—	—
Hospital #3			—	–[d]	—	N A.[f]	High[h]
Hospital #4	Highg	—	Highg	—	Highg	—	—
Hospital #6	—	—	—	—	—[d]	N.A.[f]	High[h]
Hospital #7	—		—	—[d]	—[d]	Low[i]	High[h]
Hospital #8	—		—	—	—	N.A.[f]	High[h]
Hospital #9	—	—	—		—	High[h]	High[h]
Hospital $10	High[g]	Highg	—	: d	—	—	—
Hospital #l 1	—	—	—	—	—	High[h]	High[h]
Hospital #12	—		—	Highg	—	—	—
Hospital #13	—		—	—	Highg	—	—

aTh, HCFA m..lit, data referred t. in this table were adjusted for patient characteristics such as age, sex, and comorbldities (see table 4-3) The Hea Hh Semlces Advisory data were not adjusted
bp..i.d when the certificate-of-need process Was In effeCt.
cPeriod when the certificate-of- need process was no longer in effect
'Near upper limit.
eN. A = Not applicable,
fDid not perform open-heart surgery in this period.
gMortality rate higher than expected
hMortality rate higher than State average.
iMortality rate lower than State average.

SOURCES: HCFA 1984 data: U.S. Department of Health and Human Services, Health Care Financing Administration, "Medicare Hospital Mortality Information, 1984, " Washington, DC: March 1986 HCFA 1988 data: U S. Department of Health and Human Services, Health Care Financing Administration, *Medicare Hospital Mortality /n formation,1986*, Washington, DC: U S Government Printing Office, Dec 1987 Health Services Advisory Group data: B. Patten, "Open Market, Open Heart Special Report," *The Phoenix [Arizona] Gazette,* p A-1, Aug 26, 1987, results, but almost no quality assessment studies have compared hospital mortality rates over time. The Service Intensity portion of the Stanford Institutional Differences Study found little difference over time in service intensity or outcome *(215,223),* but this analysis was limited to 17 hospitals that volunteered to be in the study.

A summary of OTA'S analyses of HCFA'S data for 1984 and 1986 for the four States and the District of Columbia is shown in table 4-6. There was little convergence between the two HCFA results for these jurisdictions, but it is difficult to say whether these differences were due to actual changes in the hospitals, the fact that HCFA used different methods, or flaws in one or both of HCFA'S methods. In California, for example, there were 18 HCFA high-outlier hospitals in 1984, and 20 in 1986, but only 2 (10 percent of the 1984 total) of the outlier hospitals were outliers in both 1984 and 1986. Another 37 hospitals, however, had actual mortality rates at or near the upper limit of the expected range of mortality rates in 1986; 7 of those had been outliers in the 1984 analysis. Thus, at most, 9 of the 18 1984 outliers in Califomia(50 percent) had potential quality problems in 1986, if HCFA'S methods are accepted as potentially valid indicators of quality problems. A number of other hospitals had actual mortality rates at or near the upper limit of the expected range of mortality rates in 1986, but *were* not high oudiers in 1984 (647).

Table 4-6.—Number of Hospitals Found To Be High= Mortality Outliers by HCFA in 1984 and 1986, Selected States[a]

State	Number of hospitals 1984[b]	1986[c]	Convergence
California.	*20*	*18*	*2* of the 18 outliers in 1986 were also outliers in 1984
New York	29	10	5 of the 10 outliers in 1986 were also outliers in 1984
Maryland	1	1	The 1 outlier in 1986 was also an outlier in 1984
Arizona	3	1	The 1 outlier in 1986 was also an outlier in 1984
District of Columbia . . .	0	1	The 1 outlier in 1986 was not an outlier in 1984

aHl*h.rn~rtalitY outliers are **hospitals with mortality rates that exceed 'he expected range of hospital mortality rates. This table shows results for overall mortality rates only, not for specific diagnostic** categories.
bus, Department of Health and Human Services, Health care Financing Administration, "Medicare Hospital Mortality Information, 1904," Washington, DC: Mar. 10, 19W
Cu s, Department of Health and Human ServiCOS, Health Care Financing Administration, *Madicare Hospital Morta/ity Information, 1%X* **(Washington, DC: U.S. Government Printing Office, Dec. 17, 1987).**

SOURCE: Office of Technology Assessment, 1988.

New York had 29 high-outlier hospitals in HCFA'S 1984 analysis of overall statistics and 10 in 1986; 5 (17 percent of 29) of the hospitals were outliers in both years. Another 21 hospitals were at or near the upper limit of the expected range in 1986. In Maryland, the same single hospital was a high-outlier hospital in 1984 and 1986. Eleven additional hospitals had mortality rates at or near the upper limit of the expected range of mortality rates in *1986.* When all diagnostic categories are considered, Arizona had two outlier hospitals in 1984 and one in 1986. One (50 percent) of these hospitals appeared on both lists. None of the District of Columbia's hospitals were high outliers in 1984; one was in 1986.

A better longitudinal analysis of the HCFA data would be based on results using the same statistical method for the 1984 and 1986 data, as well as for 1985 data. HCFA conducted such an analysis using the 1986 analytical techniques and found a 44-percent convergence of high outliers between 1984 and 1986, and a 48-percent convergence between 1985 and 1986 (647). Neither the names of the hospitals that were outliers for any 2 of the years, nor any of the 1985 data were published, however, because the analysis did not generate hospital-specific information.

Although longitudinal data on hospital mortality may seem preferable to cross-sectional (one-time) data, consumers must be careful to consider changes in other factors that may occur over time that may affect the reliability of the indicator. For example, declining admissions as a result of policy changes can make mortality rates seem to change as well, because more severely ill patients may be admitted to the hospital (192,519). Consistent use of patient adjustments could alleviate this problem in the future, but in the past, different adjustments have been used in every release (640,647).

FEASIBILITY OF USING THE INDICATOR

Construction of the Indicator

Valid hospital mortality information depends on valid and reliable adjustments for the patient characteristics that increase the likelihood of death independently of the medical care provided. Conceptually, systems to adjust for patients' risk of dying based on clinical data on admission seem to be the most nearly valid means to adjust mortality rates. However, such systems are also the most costly to use and develop, primarily because they involve the collection of patient data not currently in the discharge abstracts routinely compiled by hospitals for billing purposes.

Adjustment of hospital mortality rates may also involve the calculation of separate algorithms for individual diagnoses and procedures; these algorithms will need to be continuously updated if they are to conform to advances in statistical methods and medical practice. Such efforts will require expertise in statistical and research methods and medical practice. Similar efforts are required to devise ways of comparing hospital mortality rates to the process of care *(442).*

The Joint Commission on the Accreditation of Healthcare Organizations' clinical indicators project is assessing the feasibility of regularly collecting clinical data. Preliminary estimates indicate that the collection of such data will be relatively expensive. It will also be expensive to check mortality data against process of care information. Both the Maryland Hospital Association and CPHA have provided member hospitals with workbooks containing hospital mortality norms.

A number of providers and consumer representatives have suggested that analyses within clinically meaningful diagnosis and procedure groups are both more nearly valid and more meaningful to individual consumers than hospital-wide mortality rates *(500).* However, results aggregated by hospital may be useful for evaluating institutional performance. Individual consumers may not be sophisticated enough to distinguish among hospital services. Finally, organizational purchasers of care may contract with entire hospitals, although they do sometimes contract for specific services from different hospitals. Thus, information at the hospital level would be useful to them. There clearly seems a place for information aggregated at different levels.

Intentional Manipulation of the Indicator

The fact of death does not seem easy to manipulate intentionally, but a focus on death rates without adequate research attention to adjustments for patients' risk of dying and validation against the use of appropriate medical processes may lead hospitals to refuse to accept severely ill patients, to postpone their admission from the emergency room, to discharge them hurriedly to other facilities, or to intentionally miscode diagnoses.

The California PRO study found substantial miscoding in a DRG series that had higher paid DRGs for patients discharged alive (117). On the other hand, neither the California PRO study nor

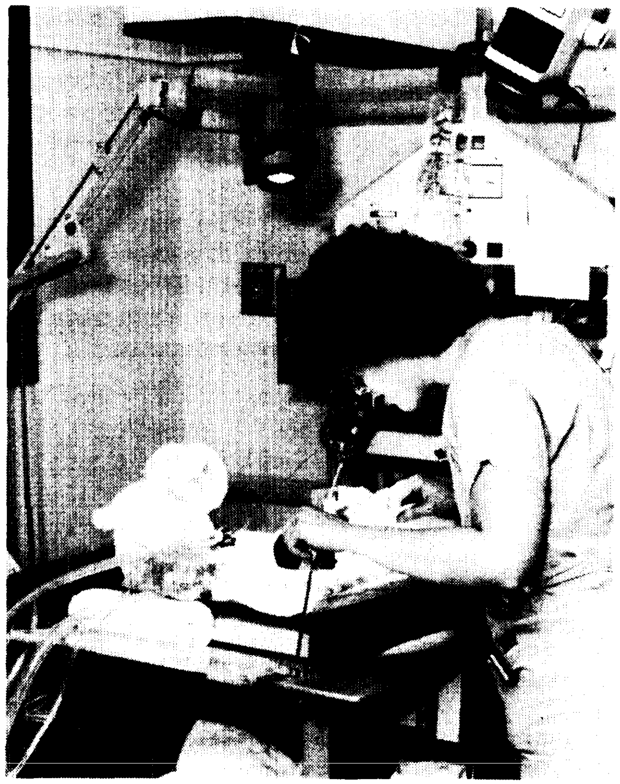

Photo credit: Strong Memorial Hospital, Rochester, New York

Consumers may wish to have hospital mortality information that is specific to particular conditions or services, such as neonatal intensive care.

an OIG study found a pervasive pattern of premature discharges within their definitions (117,660a). The OIG study covered only an early period of prospective payment implementation (October 1984 to March 1985), however, and also found that one in every five hospitals reviewed had at least one occurrence of a premature discharge; the occurrence was one in three in rural hospitals. The California PRO study, begun early in 1986, did find a higher proportion of premature discharges among patients who died within 20 days of discharge compared with premature discharges who were readmitted within the same period, and a significant pattern of premature discharges in patients readmitted within 1 day of discharge.[11] Thus, there is some evidence of premature discharge.

Some have suggested that earlier discharges of patients who are likely to die is one way to contend with the release of hospital mortality statistics, although they implied that the release be medically appropriate and to an appropriate alternative care facility (224,660a). There are also incentives and analytic procedures which may make premature discharges in the face of mortality releases unlikely, such *as* intensive PRO review and the use of *30* days post admission as the time when deaths are counted.

Another way for a hospital to reduce its mortality rate is to keep severely ill patients in the emergency room rather than admitting them to the hospital. Currently, emergency room patients are not counted as hospital admissions.

Other quality assessment/assurance mechanisms may have to be in place to prevent refusals to admit or premature discharge. For example, hospitals that participate in Medicare and transfer uninsured emergency patients before stabilizing their conditions will be fined $50,000 (Public Law 100-203). Another way to discourage hospitals from transferring patients or discharging them prematurely would be to "credit" each hospital that sees a patient during an episode of care for the patient's death.

Dissemination of Information About the Indicator

Hospital-specific mortality rates are becoming more available to the public (see box 4-A). HCFA'S releases of 1984 and 1986 Medicare data were, of course, the most prominent. The California PRO released mortality information about California Medicare patients in 1986 and 1987. The University of California Santa Barbara has available for sale hospital-specific data on infant mortality from its Maternal and Child Health Data Base *(598).* Blumberg's analysis is available to the public. Portions of all of these reports were reported in newspaper articles and in reports by consumer groups *(503,598,693).* In addition to information made available to the public, the Maryland Hospital Association and CPHA calculated hospital mortality norms and made the information available to their member organizations (141,408).[12]

Consumer advocates have applauded the availability of hospital mortality data (115), but reservations have been expressed as well (14,500). Some hospitals are reported to be making mortality data available as part of a marketing strategy (426), while others continue to criticize the release of such data in its present state (41, 426,427). Some States have mandated the collection and reporting of numerous clinical indicators and are also planning to release outcome data (41,427). The cumulation of information in its current methodological state may be helpful to consumers who are relatively sophisticated. Those who are less sophisticated will probably need a considerable amount of help interpreting the data.

HCFA used the media to disseminate to the public the 1984 and 1986 hospital mortality rates for Medicare patients (648). HCFA was reluctant to publish the 1984 data, but was pressed to do so by the possibility of a Freedom of Information Act suit (302). In the press release accompanying the 1986 analysis, HCFA characterized the release

[11] The California PRO study is flawed in that it looked only at these two groups. Presumably, there may have been patients discharged prematurely who were neither readmitted nor died soon after discharge. The OIG study recognized this problem and identified premature discharge regardless of whether the patient did well subsequently.

[12] CPHA's workbook is available for sale to the public as well, although it is not clear how useful it would be to the general consumer.

Box 4-A.—Selected Sources of Information About Hospital-Specific Mortality Rates

Type of information	Source(s)
National information about mortality rates among 1984 Medicare patients	Director, Health Standards and Quality Bureau Health Care Financing Administration U.S. Department of Health and Human Services 6325 Security Boulevard Baltimore, MD 21207
National information about mortality rates among 1986 Medicare patients	1. *Medicare Hospital Mortality Information,* Stock No. 017-060-00206-9 U.S. Government Printing Office Washington, DC 20402 Cost: $69 for a 7-volume set 2. American Association of Retired Persons regional offices; Main office: 1909 K Street, N.W. Washington, D.C. 20048 3. Utilization and quality control peer review organization (PRO) offices
Information on mortality rates among California Medicare patients for *24* diagnostic categories, April 1, 1985 through March 31, 1986; and for the *50* most common diagnosis-related groups, Federal fiscal years 1985 and 1986	California Medical Review, Inc. 1388 Sutter Street Suite 1100 San Francisco, CA 94109 Telephone: (415) 923-2000 Cost: $10 for each hospital listing
Information on California perinatal mortality rates, 1980-84	Maternal and Child Health Data Base c/o Community and Organization Research Institute University of California, Santa Barbara 2201 North Hall Santa Barbara, CA 93106 Cost: Descriptive narrative, $20; statistical appendix, including individual hospital statistics, *$301*[2]
Information on mortality rates in Maryland for nine surgical categories, 1979-80	*Surgery in Maryland Hospitals 1979 and 1980: Charges and Deaths,* by E. Bargmann and C. Grove, 1982 Public Citizen Health Research Group 2000 P St., N.W. Room 708 Washington, DC 20036 Telephone: (202) 872-0320
Information on mortality rates in Maryland for nonelective surgeries, April 1984-March 1985[3]	Maryland Health Services Cost Review Commission *201* W. Preston St., First Floor Baltimore, MD 21201

[1]Source: R. Steinbrook, "Hospital Death Rates: A Wide Variance," *Los Angeles Times,* p. 1, June 15, 1987.

[2]A summary of these data for hospitals in the Los Angeles area (including hospitals with high mortality rates in Riverside, San Bernardino, San Diego, and Ventura counties) was published in the Los *Angeles Times* in November 1987 (R. Steinbrook, "Care for Newborns Varies, Studies of Hospitals Show," Los *Angeles Times,* p. 1, Nov. 9, 1987).

[3]Mortality rates for digestive operations were reported in the *Baltimore Sun* in November 1987 (M. Knudson, "Death Rate Found To Vary for Digestive Operations," *Baltimore Sun,* p. 1A, Nov. 22, 1987).

SOURCE: Office of Technology Assessment, 1988 as an "important contribution to the existing body of knowledge about health care" but stressed that a significant shortcoming of its approach was that it did not contain an objective and direct measure of the condition of the patient at the time of hospital admission. HCFA urged health care consumers to read the explanations of the information's uses and limitations, as well as comments provided by the hospitals, and when appropriate to discuss the indicators with their physicians or with hospital administrators. HCFA also cautioned that the information should not be used to rank hospitals, and was not designed to provide a national benchmark for measuring the quality of care. An additional set of questions and answers accompanying the press release cautioned the media not to report the mortality statistics as definitive measures of the quality of care (648).

For the 1987 release, OTA examined clippings from newspapers with circulations of 50,000 or more, and transcripts from radio broadcasts to review the manner in which the HCFA data were being reported. Burrelle's clipping service found about 100 clippings (including 3 radio transcripts and several editorials, op ed pieces, and letters to the editor). Fifty-two articles describing the HCFA release were written in newspapers in 19 States and the District of Columbia.[13] Typically, stories focused on hospitals that were outliers, and stated that area hospitals were either high, average, or low on the HCFA lists. The media quoted from the HCFA press release and/or HCFA personnel about the limitations of the study, and quoted local hospital or hospital association personnel, who generally criticized the release. There was very little understanding evinced that HCFA did try to adjust for patients' risk of dying using proxy measures, such as comorbidities, transfers from other hospitals, and previous hospitalizations. Of the 10 editorials, 6 were in favor of the data release, even if only because it made some information available to consumers, and 4 opposed the release. Stories about the release of the Dubois study a week after the HCFA release gave the HCFA release more credibility as a quality indicator (190).

Only one story gave consumers information about how to get the HCFA release itself. HCFA'S press release had said that the report was available for $69 from the U.S. Government Printing Office (648). The report is sold as an entire set of seven telephone-book sized volumes with information for every State, the District of Columbia, American Samoa, and Guam. HCFA reports a steady, but not large, stream of requests for information about how to get access to the report *(98)*. By May 15, 1988(5 months after the report was released, and the last day for which data are available), the Government Printing Office had sold 236 sets (670).

Hospital mortality releases are costly. The University of California's report on perinatal mortality in California hospitals cost *$50* in 1987 (598), and each individual hospital report from the California PRO cost $10 in *1987*. HCFA sent copies of *Medicare Hospital Mortality Information, 2986* to State health offices, PROS, HHS regional offices, and **10** American Association of Retired Persons regional offices, and suggests to consumers that they also contact depositary libraries for copies of the report. HCFA acknowledges that not all of the sites they suggest for consumers to review the report are accessible to all people and not all of the sites will have copies of the report (98).

Both HCFA and the California PRO have suggested that concerns about high hospital mortality rates be the occasion for consumers to ask their physicians questions about specific hospitals (115,116,640,647). The Public Citizen Health Research Group has suggested some questions that consumers might ask, and advised that consumers be sure to get "specific and substantial answers" (503). Hospital comments on the HCFA release and other releases suggest that hospitals may respond by citing inaccuracies in the data, large numbers of admissions from the emergency room, patient characteristics, and patient and family wishes to not resuscitate, rather than errors in care, as explanatory factors (598,647).

[13]Twenty jurisdictions ran articles on the mortality release: Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Louisiana, Maryland, Massachusetts, Missouri, New Jersey, New York, Ohio, Oregon, Pennsylvania, Rhode Island, Tennessee, Texas, and Virginia.

CONCLUSIONS AND POLICY IMPLICATIONS

Given the methodological and conceptual problems associated with using hospital mortality rates as an indicator of the quality of care, the rates cannot at present be considered definitive indicators of quality. The release of hospital-specific mortality rates does have the potential, however, to initiate a dialogue between consumers and providers. Such releases can also provide national and regional information against which hospitals and others can compare results; this information may lead to identification and correction of quality problems. Physicians, hospital staff, consumers, and organizations that contract with hospitals to provide care maybe able to use hospital mortality rate information as leverage to improve care.

One of the advantages of using hospital mortality rates as an indicator of the quality of care is that it makes sense to average consumers. Consumers are also aware that patients' inherent risk of dying is usually the most important determinant of whether patients die during or soon after a hospital stay, and see the importance of adjusting for this risk.

One of the weaknesses of hospital mortality rates as a quality indicator is the current lack of methods to adjust adequately for patients' risk of dying in most analyses. In addition, methods for reviewing medical records to determine quality problems and validate statistical analyses are underdeveloped. The data on which analyses are performed may not be uniformly coded and collected, leading to spurious differences among hospitals. Hospital mortality rates are just beginning to be validated against the medical care provided. If mortality statistics are not valid but are nonetheless interpreted as indicators of the quality of care, their release could result in a breakdown of trust between patients and providers, and loss of reputation, patients, and income by hospitals.

Another problem may be that information on hospital mortality may not be available when consumers need it. An additional problem may be that individual consumers do not yet know either the appropriate questions to ask about hospital mortality statistics or the responses to accept. Information about hospital mortality rates for specific diagnoses may be as useful as information about overall institutional performance, but diagnosis-specific information is more difficult to obtain than information about the hospital as a whole. Most newspaper reports of the 1986 HCFA analysis did not include diagnosis-specific information. In any case, the diagnostic groupings included in the HCFA report may be neither understandable to consumers nor clinically meaningful in themselves.

Given the undeveloped state of hospital mortality statistics and the skepticism of some providers about the value of the information, it seems that consumers could use additional guidance about the kinds of questions to ask and the kinds of answers to accept when faced with anomalous mortality statistics. As a first step, it might be useful to develop information about the ways consumers and providers have responded so far to the mortality rate information that is already available. Are they aware of it? Are they frustrated by advice to regard the information cautiously or by hospitals' denial of its importance? The questions consumers ask about the mortality data and the responses the consumers were given could be studied. Perhaps more useful to consumers would be for HCFA to assess the validity of the mortality data so that quality concerns can be verified or dismissed. There was some consideration given to having PROS investigate the validity of the 1986 data (503), but ultimately they were not required to do so. Confidence in the results of any review would depend on the rigor of the review process.

Considerable progress seems to have been made in the development of valid hospital mortality indicators, but researchers, policymakers, and providers agree that considerable problems remain. HCFA'S 1986 mortality rate analysis was much improved over its 1984 analysis, although neither has yet been validated against an independent criterion. Dubois's study is the most careful so far in its testing of alternative hypotheses to account for mortality outliers; his finding that outlier status based on claims data for the entire hospital overall was confirmed for two conditions suggests that the use of claims data to identify hospitals with quality problems has potential (190). The results of the study by Dubois, however, can be regarded only as preliminary, particularly because of the relatively low reliability among reviewers using implicit criteria, and the fact that only three conditions were included in the research, It is also important to keep in mind that the adjustment procedure used by Dubois and his colleagues was different from those used by HCFA and by other researchers, so that Dubois's results may not be generalizable to other methods.

There are a considerable number of methodological and conceptual issues to be resolved before hospital mortality rates can be regarded as a valid indicator of the quality of care. The resolution of these issues requires a comprehensive and quite large and costly research program. The construction of valid adjustments for patients' risk of dying and establishment of a link between mortality and the process of care will depend on the establishment of an orderly, iterative process (442). Most importantly, links to process must be established, and valid techniques found to adjust crude mortality rates for patient characteristics. Needed is additional attention to research design, statistical analysis (for regression studies), and methods for confirming quality-of-care problems, Ways to link patient information across different data sets are needed. For example, ambulatory files may provide useful information about patients' status on admission to a hospital. Given the results of the Dubois and New York State studies, ways to develop explicit criteria or increase the reliability of implicit review are critical. It may not be too early to develop conventions so that the more reliable and validated sources of data and methods are used. A number of studies related to these issues are in progress (see app, E). A study of nonintrusive outcome measures, being conducted by the Rand Corporation, is examining the relationship between prior medical care and death. HCFA is comparing the results of its 1986 analysis to one using MEDISGRPS to adjust for patients' severity of illness (357).

The validation process will be expensive. Certain types of experts are needed to develop adjustment methods, and many more experts are needed to review medical records and to establish the reliability of data bases. Even so, given the limitations in ability to measure patients' risk of dying, it may be that hospital mortality will continue for quite some time to be useful only as a screen or flag for possible quality problems.

In the absence of a validated method for constructing hospital mortality as a quality indicator, is the release of the information to the public justified? Brook and Lohr suggest that it is inappropriate to identify outlier hospitals publicly before evaluating the reliability and validity of the data and giving those hospitals adequate time to review their own data (104). Brook and Lohr suggest further that outlier hospitals be given up to 6 months to correct any problems before information is released. This approach would seem to encourage a closer working relationship between releasing bodies and the hospital, and perhaps more support for the release of data by the hospitals. If, however, as HCFA seems to intend (648), hospital mortality rates for all hospitals continue to be published, HCFA might follow up with reviews of the medical care process, so that the public would know whether quality problems were in fact confirmed. To do this, HCFA would need to develop a standard review method.

Releasing hospital mortality data may provide an incentive for hospitals to look more closely at the care they provide. A recent survey of hospitals showed little use of comparative death rates (132). Hospitals that conduct appropriate investigations of the reasons for differences maybe able to improve the quality of care that they deliver. Physicians and organizations may find the data useful in referring patients or selectively contracting with hospitals. The rates of preventable deaths and the percentage of quality problems found in numerous studies (190) suggest that additional attention to patient care is warranted. Finally, once validated, adjusted hospital mortality rates and other outcome measures could be a good complement to studies based on reviews of medical care provided (process studies), and provide a good validation criterion for studies of structural properties of hospitals.