CASE FATALITY RATES: LITERATURE REVIEW – FULL REPORT

record-based

O
U
T
C
O
M
E

measures

Alastair Mason, Edel Daly and Michael Goldacre

National Centre for Health Outcomes Development

July 2000

UNIT OF HEALTH-CARE EPIDEMIOLOGY

UNIVERSITY OF OXFORD

REPORT MR5

UNIVERSITY OF OXFORD

UHCE OXFORD REPORT MR5
CASE FATALITY RATES: LITERATURE REVIEW

Alastair Mason, Edel Daly and Michael Goldacre
National Centre for Health Outcomes Development
July 2000

CONTENTS

INTRODUCTION

Over the last ten years there has been increasing interest in using case fatality rates as an outcome indicator to make comparisons over time and between hospitals. Death during or soon after an in-patient spell was considered to be a potential indicator of poor outcome in five of the ten health outcome indicator reports published by NCHOD in 1999. In view of the increasing importance that information about in-patient mortality is likely to have, the Department of Health commissioned NCHOD (Oxford) to prepare a detailed review of the use of case fatality rates as a health outcome indicator.

This report contains:

- Summary of the key issues relating to case fatality rates.
- Detailed review of the literature about the use of case fatality rates.

Key issues

Death during or soon after an in-patient spell may be a consequence of the natural course of the patient's disease or may result from sub-optimal in-patient care. Because of the possible link between mortality and sub-optimal quality of care, variation in quality of care, either over time or between hospitals, might result in variation in case fatality rates.

The *calculation* of case fatality rates requires the linking of hospital episode data with that on mortality occurring within a specified time period. The main issues relating to this are:

- methodology used to link mortality with hospital episode data
- definition of case fatality in terms of type of hospital admission and diagnostic specificity
- time period chosen within which hospital episode and death must occur and whether all deaths or only deaths in hospital are included
- risk adjustment for factors such as age, sex or case-mix
- accuracy and completeness of data required for derivation of the indicator, particularly diagnosis and procedure recording and coding, and death certification.

The *usefulness* of an outcome indicator will depend on:

- attributability of the outcome measured to the quality of health care
- reliability of the indicator
- sensitivity of the indicator to changes or variation in the quality of health care.

Key issues relating to the *interpretation* of case fatality data are:

- statistical power, relating to the adequacy of the number of events and size of the population denominators to show significant variations
- extent to which expectations of performance can be quantified by defining benchmarks
- creation of perverse incentives and games playing.

Literature search questions

The particular interest in re-admission rates currently in the NHS is related to whether they are good indicators for comparing clinical performance over time and between hospitals. Therefore, in reviewing the literature an attempt has been made to address the following questions:

- Do properly compiled case fatality rates reflect the quality of in-patient care (see Chapter 5)?
- Are case fatality rates a useful way of comparing hospital performance (see Chapter 6)?
- Does the volume of cases influence case fatality rates (see Chapter 3)?
- Does the length of stay influence case fatality rates (see Chapter 4)?
- What other factors influence case fatality rates (see Chapter 2)?
- How should case fatality rates be compiled if they are to reflect the quality of in-patient care (see Chapter 7)?

Search strategy

Various free-text searches were performed in Medline, EMBASE and PsychLIT for the years 1990 to 2000 using various combinations of the following words and phrases:
case fatality; post-surgical mortality; post-surgical death; post-operative mortality; post-operative death; patient mortality; patient death; in-patient mortality; in-patient death; in-hospital mortality; in-hospital death; or death in hospital;
combined with:
quality indicator; outcome indicator; clinical indicator; performance indicator; quality of health care; quality of care; quality assessment; outcome assessment; process assessment; outcome measurement; preventability; avoidability; quality comparisons; quality assurance; quality improvement; hospital performance; hospital standards; league table; health services research; health policy; length of stay; statistics and numerical data; surgical volume; hospital volume; mortality trends; hospital information systems

In addition, a number of other strategies were employed to identify relevant publications. These included:

- electronic searching for publications by key researchers working in the field
- electronic searching for publications which cited key papers on this subject
- hand searching of reference lists of key papers
- electronic or hand searching of recent issues of journals where relevant significant papers are most likely to appear (e.g. Medical Care).

STUDIES TO IDENTIFY PREDICTORS AND CAUSES OF CASE FATALITY

A number of studies have been done to identify:

- predictors or risk factors for mortality associated with admission
- reasons or causes for mortality associated with admission.

The studies can be classified by patient group as those involving:

- general population of patients
- patients with cardiovascular disease
- other medical patients
- elderly patients
- paediatric patients
- surgical patients.

General population of patients

The studies done on a general population of patients that have been reviewed are:

- *Beguin (1997)* undertook a systematic analysis of causes and circumstances of all in-hospital patient deaths during 1993 and 1994 in one hospital.
- *McCloskey (1998)* described the relationship between six adverse patient outcomes (including patient mortality), total hours of nursing care, and the proportion of those hours of care delivered by registered nurses (RNs) using the hospital records for 1993 for a large university hospital.
- *Czaplinski (1998)* tested the hypothesis that specialized staff nursing has an effect on length of stay (LOS) and mortality, using case-mix information and controlling for physician volume.
- *Glenn (1999)* compared in-patient mortality rates for rural hospitals with mortality rates of urban hospitals of given sizes and ranges of service. Eighty three small hospitals that had a relatively unspecialized range of services constituted the study group.

Beguin (1997) obtained the following results:

- The mean age of the patients was 60 years.
- The most frequent causes of death were malignant and cardiovascular diseases.
- At least one of the following was mentioned on 14% of the forms:
    - complications during care
    - septicaemia
    - pulmonary embolism
    - adverse effects of drugs
    - pressure ulcer.

McCloskey (1998) found that, after controlling for patient acuity:

- A direct relationship existed between total hours of care from all nursing personnel and rates of in-patient mortality.
- As the RN proportion of care rose to an 87.5% level, it related to a lower incidence of negative outcomes; however, when the RN

proportion of care went beyond that level, the adverse outcome rates also increased.

Czaplinski (1998) analysed data on 16 diagnosis related groups associated with particular specialty units in a US hospital for the period 1987-1993. Specialized nursing units were defined by the percentage of patients in a given diagnosis related group discharged from that unit. Patient age and differential intensive care unit use were used for risk adjustment. The reported findings include:

- In 13 of the 16 diagnosis related groups, patients cared for on specialized nursing units had shorter lengths of stay; the difference was statistically significant in nine.
- In the seven diagnosis related groups with any deaths, the mortality on the specialized unit(s) was lower; the difference was statistically significant in four.
- Physician volume, defined as more or fewer than 20 discharges per diagnosis related group, had little or no effect on either LOS or mortality.

The main findings from the study by *Glenn (1998)* were:

- Patients admitted to hospitals in non-metropolitan areas had a mortality rate of 0.41% compared with a mortality rate of 0.66% in peer hospitals in metropolitan areas.
- After risk adjustment of mortality rates, the urban-rural difference was not statistically significant, but rural hospitals tended to have a lower adjusted mortality rate than urban hospitals of the same size and type.
- Taken together with other studies, the data are consistent with the view that small rural hospitals generally make appropriate transfer decisions for severely ill patients and provide quality care for retained patients.

Patients with cardiovascular disease

The studies done on patients with cardiovascular disease that have been reviewed are:

- *O'Connor (1999)* analysed data on 15,331 consecutive hospital admissions to six clinical centres to identify risk factors associated with in-hospital mortality among patients undergoing percutaneous coronary interventions.
- *Ng (1999)* used Singapore population based data from 1991 to 1996 to examine recent trends in mortality, incidence and hospitalisation for acute myocardial infarction (AMI), and to explore gender differences.
- *Iezzoni (1997a)* studied all 14,083 patients admitted in 1991 for AMI to 100 hospitals nationwide, examining in-hospital death and use of invasive cardiac procedures to determine whether evaluations of gender differences in in-hospital mortality among such patients vary by data source and severity measure.
- *Pearson (1992)* assessed whether important gender differences exist in the quality of hospital care provided to a nationally representative sample of 11,242 patients aged 65 years or older who were hospitalised with one of four diseases: congestive heart failure, AMI, cerebrovascular accident, and pneumonia.
- *Lai (1999)* investigated the effect of the type of facility to which 662 patients were discharged after an acute initial stroke on survival prognosis up to four years.

- *Arboix (2000)* compared demographic characteristics, clinical features, and outcome of acute stroke events in 262 very old patients (aged 85 years or older) with those in 1,738 younger patients (< 85 years) using data from a prospective hospital based stroke registry for the period 1986 to 1995.
- *Jones (1996)* identified a set of 44 pre-operative clinical variables useful for monitoring and improving the short term mortality of patients undergoing coronary artery bypass grafting (CABG), and for risk adjustment of observed mortality outcomes.
- *Piette (1996)* examined whether patients admitted for treatment of AMI who live farther from their source of care are less likely to be followed in an out-patient clinic, and whether patients who receive follow-up care are less likely to die.
- *Wolinsky (1999)* examined the effect of gender differences among older adults (aged 70 years or over) hospitalised for AMI on outcomes in a nested case-control study of 357 cases (172 women and 185 men) compared with 3,976 hospitalised controls and 1,738 non-hospitalised controls.
- *Abrahamsson (1998)* examined risk factors for mortality within two years following AMI in an unselected population of 3,187 patients (1,039 women and 2,148 men) consecutively admitted to a coronary care unit during 1984-1991.
- *Mahon (1999)* examined the effect of patient age on the management and outcomes in a historical cohort study of consecutive unselected admissions for AMI to an acute cardiac unit.
- *Mozes (1998)* investigated predictors of 30 day post-operative mortality among 4,835 consecutive patients who underwent isolated CABG in 14 medical centres.
- *Shepardson (1999)* determined the relationship between the use of do not resuscitate (DNR) orders and in-hospital mortality, adjusting for severity of illness and other co-variates in a retrospective cohort study of 13,337 consecutive stroke admissions to 30 hospitals between 1991 and 1994.
- *Goldsmith (1999)* compared clinical characteristics and outcomes among British patients of Indo-Asian and white Caucasian descent undergoing coronary revascularization by bypass surgery between 1994 and 1997.
- *Roques (1999)* assessed risk factors for mortality in 19,030 consecutive adult patients undergoing cardiac surgery under cardiopulmonary bypass in 128 surgical centres in eight European states using logistic regression of data on 68 pre-operative and 29 operative risk factors proven or believed to influence hospital mortality.
- *Mehta (2000)* compared mortality following AMI in those with and without a diagnosis of diabetes by analysing data on all Medicare patients who were discharged with a principal diagnosis of AMI from acute care hospitals in Michigan between April 1994 and July 1995.
- *Frances (1999)* undertook a retrospective cohort study of 7,663 Medicare beneficiaries (aged 65 years and older) directly admitted to hospital in California with a confirmed AMI from April 1994 to July 1995 to evaluate whether cardiologists provide more recommended therapies to elderly patients with AMI and, if so, to determine whether variations in processes of care account for differences in patient outcome.
- *Pell (2000)* undertook a prospective cohort study, using the Heartstart (Scotland) database, on all 22,161 people suffering community based cardiopulmonary arrest in Scotland between 1988 and 1997 to determine whether men and women

suffering cardiopulmonary arrest differ in terms of survival and risk factors for survival.

- *Goldberg (1998)* examined age related differences in hospital and long term survival among 8,070 patients with validated AMI hospitalised in an acute care hospital in one metropolitan area between 1975 and 1995.

In the *O'Connor (1999)* study, logistic regression analysis identified the following variables, among others, as being associated with an increased risk of in-hospital mortality following percutaneous coronary interventions:

- older age
- congestive heart failure
- peripheral or cerebrovascular disease
- treatment of AMI
- increased creatinine levels
- lowered ejection fraction
- treatment of cardiogenic shock
- urgent priority
- emergent priority
- preprocedure insertion of an intraaortic balloon pump
- PCI of a type C lesion.

The authors conclude that accurate calculation of the risk of in-hospital mortality after a percutaneous coronary intervention is feasible and may be useful for quality improvement purposes.

Findings from the study by *Ng (1999)* were:

- Over the period 1991 to 1996 the rates of angiograms, CABG and PTCA among Singapore residents increased greatly, with the greatest increase among elderly aged 60 years and above.
- The rates of invasive cardiac procedures for AMI were all lower in females than in males.
- The population case fatality rate of AMI was higher in females than in males.
- Over the period 1991 to 1996, the population case fatality rate of AMI declined slightly only for persons below 40 years of age.

In the *Iezzoni (1997a)* study, ten severity measures were considered: four using clinical medical record data and six using diagnosis and procedure codes from discharge abstracts. The main findings were:

- After adjusting only for age, women were significantly more likely than men to die and less likely to receive CABG and coronary angiography.
- After adjusting for severity and age, women were significantly more likely than men to die in-hospital and less likely to receive coronary angiography and CABG.
- ORs reflecting gender differences in procedure use and death rates were similar across severity measures.
- Comparisons of severity adjusted in-hospital death rates and invasive procedure use between men and women yielded similar findings regardless of data source and severity measure.

The main findings from *Pearson (1992)* were:

- Sex differences in sickness at admission varied by disease.
- There was some evidence that women received worse process of care, but the difference was very small.
- There were many similarities in the process and outcomes, including mortality, of care for male and female patients.
- After controlling for sickness at admission, age, and other important co-variates, the in-hospital experiences of elderly men and women showed greater similarities than differences.

Lai (1999) investigated the survival prognosis up to four years (mean of 24 months) after an acute initial stroke depending on the facility to which the patient was discharged. Six hundred and sixty two patients were enrolled between July 1987 and August 1989, and were followed up until death, second stroke, or the end of the study. Key findings included:

- Of the 662 patients discharged alive, 128 (19%) went to a nursing home, 17 (3%) to a short term hospital, 140 (21%) to a rehabilitation facility, and 375 (57%) went home (discharge destination unknown for two patients).
- Compared to patients sent home after taking age, sex, selected baseline co-morbidities, length of hospital stay, and neurological deficits into consideration, patients sent to a nursing home had 2.6 times greater risk of dying (95% CI 1.81-4.15) while those who were discharged to a rehabilitation facility had a death hazards ratio of 1.1.
- Mortality was greatest in the early months after discharge and decreased thereafter.
- Quality of care in a nursing home setting may account for the mortality difference but other factors such as social support network and living will instructions should also be investigated.

Arboix (2000) reported that:

- Compared with younger stroke patients, very old stroke patients had:
    - greater rates of in-hospital mortality (27% vs. 13.5%, $p < 0.001$)
    - longer duration of hospital stay (22.03 ± 29.6 vs. 17.5 ± 21.5 days, $p < 0.001$)
    - lower frequency of absence of neurologic deficit at the time of hospital discharge (21.4% vs. 33.1%, $p < 0.001$).
- In multivariate analysis, independent predictors of in-hospital mortality included:
    - altered consciousness
    - limb weakness
    - sensory symptoms
    - involvement of the parietal lobe and temporal lobe
    - involvement of the internal capsule (with a protective effect)
    - intraventricular haemorrhage
    - cardiac events
    - respiratory events.

Jones (1996) identified a set of 44 pre-operative clinical variables (seven core, 13 level one and 24 level two variables depending on their importance in relation to short term mortality) useful for monitoring and improving the short term mortality of patients undergoing CABG, and for risk adjustment of observed outcomes for baseline differences in disease severity among patients. The findings include:

- Multivariable logistic regression models of the seven core variables showed all to be predictive of bypass surgery mortality in some of the seven data sets examined.
- The following variables proved to be the most important in all of the seven data sets tested:
  - acuteness
  - age
  - previous operation.
- Variables describing coronary anatomy appeared to be least significant.
- In five of the seven data sets, models including both the seven core and 13 level one variables showed the core variables to reflect 45%-83% of the predictive information. However, some level one variables were stronger than some core variables in some data sets.
- The authors conclude that a relatively small number of clinical variables provide a large amount of prognostic information in patients undergoing CABG.

Piette (1996) analysed longitudinal data from Department of Veterans Affairs (VA) databases on a national sample of 4,637 AMI patients discharged in 1992. The following results were reported:

- Patients living more than 20 miles from their admitting hospital were less likely to use ambulatory services.
- Patients receiving ambulatory care were 79% as likely to die within the year as those without any follow-up care (95% CI 0.66-0.94).
- Patients living more than 20 miles from their admitting hospital were more likely to die independent of their likelihood of receiving VA out-patient follow-up.
- Among patients who did not die in the subsequent year, those receiving ambulatory care were 33% more likely to be readmitted to a VA hospital with a cardiac diagnosis (95% CI 1.12-1.57).
- Distance may limit patients' ability to access medical care quickly in the event of a recurrent acute event.
- Ambulatory care after discharge may be an important factor determining survival for patients with cardiac disease.

Wolinsky (1999) found that the risk of all cause mortality for AMI cases was greater than that for either hospitalised controls or non-hospitalised controls (referent), and this increased risk was significantly ($p < 0.001$) stronger for women (adjusted hazards ratio (AHR) = 14.24; 95% CI 10.99-18.46) than for men (AHR = 9.91; 95% CI 7.75-12.67).

Abrahamsson (1998) reported that:

- Two year mortality following AMI decreased from 36% in 1984 to 25% in 1991.
- Factors which had a significant impact on survival after AMI included:
    - year of hospitalisation
    - age
    - diabetes mellitus
    - sex
    - prior AMI
    - indeterminable infarct location.
- Factors which had no prognostic significance included:
    - thrombolytic therapy
    - hypertension.

Mahon (1999) identified AMI cases using the Hospital In Patient Enquiry (HIPE) database and diagnoses were validated according to MONICA criteria for definite or probable AMI. The reported findings were:

- Of 1,059 patients, 309 (29.2%) were older than 75 years. Mean age in this group was 80.5 years.
- Hospital mortality was almost twice as high in those older than 75 years than in younger patients (28% vs. 15%, p <0.001), and age was an independent predictor of short and long term mortality following AMI.
- Women constituted a significantly higher proportion of older patients.
- However older patients were less likely to receive thrombolysis (13% vs. 36%, p <0.001), aspirin (76% vs. 86%, p <0.01), or beta-blockers (25% vs. 51%, p <0.001) and were less likely to undergo cardiac catheterization or revascularization.
- Only 53% of older patients were admitted to coronary care.
- Although age is an independent predictor of mortality following AMI, sub-optimal management may contribute to the high mortality in these patients.

Mozes (1998) found that the overall crude 30 day mortality rate for isolated CABG among the 4,835 patients was 3.1%; the rate varied among the 14 centres, ranging from 0.85% to 7.05%. Predictors of 30 day mortality included:

- advanced age
- female sex
- diabetes mellitus
- poor left ventricular function
- high creatinine level
- high priority of operation
- triple vessel disease (with or without left main coronary artery disease).

After adjustment for patient risk factors, two hospitals were defined as outliers.

In the study by *Shepardson (1999)*, propensity scores reflecting the likelihood of a do not resuscitate (DNR) order were developed to decrease selection bias (scores were based on nine demographic and clinical variables independently related to use of

DNR orders). The odds of death in patients with DNR orders were determined using logistic regression, with adjustment for propensity scores, severity of illness, and other factors. The reported findings include:

- DNR orders were used in 22% (n = 2,898) of patients.
- In analyses examining DNR orders written at any time during hospitalisation, unadjusted in-hospital mortality rates were higher in patients with DNR orders than in patients without orders (40% vs. 2%, p <0.001); the adjusted odds of death was 33.9 (95% CI 27.4-42.0).
- The adjusted odds of death remained higher in analyses that only considered orders written during the first 2 days (OR =3.7; 95% CI 3.2-4.4) or the first day (OR =2.4; 95% CI 2.0-2.9).
- In stratified analyses, adjusted odds of death tended to be higher in patients with lower propensity scores.
- The risk of death was substantially higher in patients with DNR orders after adjusting for propensity scores and other co-variates.
- Whereas the increased risk may reflect patient preferences for less intensive care or unmeasured prognostic factors, the current findings highlight the need for more direct evaluations of the quality and appropriateness of care of patients with DNR orders.

Goldsmith (1999) matched 194 pairs of patients for age (within three years), sex and date of admission (within three months). Compared with their white Caucasian counterparts, Indo-Asian patients:

- were more likely to undergo coronary revascularization on a non-elective basis (43% vs. 32% white Caucasian patients, p =0.018)
- had a higher prevalence of diabetes (39% vs. 12%, p =0.0001)
- had a lower prevalence of smoking (36% vs. 80%, p =0.0001)
- had a lower rate of previous AMI (47% vs. 62%, p =0.012)
- had a lower use of the arterial conduit procedure (72% vs. 81%, p =0.028), particularly for those undergoing emergency surgery (59% vs. 72%, p =0.001) and for those with a previous AMI (65% vs. 81%, p =0.01)
- had a higher in-hospital 30 day mortality (6.7% [95% CI 3.18-10.21] vs. 2.6% [95% CI 0.35-4.9], p =0.0618); however, this trend disappeared when patients in the two groups undergoing non-elective surgery only were compared (9% vs. 7%, p =0.7).

In the study by *Roques (1999)*, the mean age of the study population was 62.5 (SD 10.7) and 28% were female. The following risk factors were found to be associated with increased mortality:

- age (p =0.001)
- female gender (p =0.001)
- serum creatinine (p =0.001)
- extracardiac arteriopathy (p =0.001)
- chronic airway disease (p =0.006)
- severe neurological dysfunction (p =0.001)
- previous cardiac surgery (p =0.001)
- recent AMI (p =0.001)
- left ventricular ejection fraction (p =0.001)

- chronic congestive cardiac failure (p =0.001)
- pulmonary hypertension (p =0.001)
- active endocarditis (p =0.001)
- unstable angina (p =0.001)
- procedure urgency (p =0.001)
- critical pre-operative condition (p =0.001)
- ventricular septal rupture (p =0.002)
- non-coronary surgery (p =0.001)
- thoracic aortic surgery (p =0.001).

Findings from the study by *Mehta (2000)* included:

- Diabetic patients (who accounted for 33% of 8,455 patients with AMI) were primarily younger, female, and non-white.
- In-hospital mortality rates were higher in diabetic than in non-diabetic patients (16% vs. 13%, p <0.001), as were rates for mortality within 30 days (21% vs. 17%, p <0.001).
- Diabetic patients had a greater frequency of non-Q wave AMI and presented less often within six hours of their infarction.
- Co-morbid conditions (e.g. hypertension, prior AMI, prior stroke, prior revascularization) were more frequent and congestive heart failure occurred more frequently in diabetic than in non-diabetic patients.
- Length of stay (7.9 vs. 7.0 days, p <0.001) was higher in diabetic patients.

In the *Frances (1999)* study the main findings were:

- During hospitalisation, ‘good’ and ‘ideal’ candidates for aspirin were significantly more likely to receive aspirin if they were treated by cardiologists (87%) than by medical sub-specialists (73%), general internists (84%), or family practitioners (81%).
- Cardiologists were also significantly more likely to treat good candidates with thrombolytic therapy (51%) than were medical sub-specialists (29%), general internists (40%), or family practitioners (27%).
- Patients of cardiologists were two to fourfold more likely to undergo a revascularization procedure.
- Despite these differences in utilisation, similar 30 day mortality rates were found across physician specialties; however, one year mortality rates were greater for patients treated by medical sub-specialists (OR =1.9; 95% CI 1.6-2.3), general internists (OR =1.4; 95% CI 1.3-1.6), and family practitioners (OR =1.7; 95% CI 1.4-1.9) than for those treated by cardiologists.
- Adjusting for differences in patient and hospital characteristics markedly reduced the ORs for mortality at one year for those treated by medical sub-specialists (OR =1.2; 95% CI 0.9-1.4), general internists (OR =1.1; 95% CI 1.0-1.3), and family practitioners (OR =1.3; 95% CI 1.1-1.6), whereas further adjustment for medication use and revascularization procedures had little effect.

- In conclusion, differences in the use of recommended therapies by physician specialty are generally small and do not explain differences in patient outcome.
- In comparison, differences among patients treated by physicians of various specialties (case-mix) have a large impact on patient outcome and may account for the residual survival advantage of patients treated by cardiologists.
- With the exception of the in-hospital use of aspirin, recommended AMI therapies are markedly underused, regardless of the specialty of the physician.

Pell (2000) reported that:

- Multivariate logistic regression analysis was used to determine the association between sex and outcome after adjustment for case-mix.
- Women were older (p <0.0001) and were less likely to have shockable rhythms (p <0.0001). Despite this, they were more likely to survive to admission (p <0.0001). However, thereafter, women were more likely to die in hospital (p <0.01).
- There was no significant difference between the sexes in overall case-fatality rates to discharge.

Goldberg (1998) reported that:

- Compared with patients <55 years (n=1,326), patients 55 to 64 years (n=1,768) were 2.2 times more likely to die during hospitalisation for AMI, whereas patients 65 to 74 (n=2,325), 75 to 84 (n=1,880), and 85 years or over (n=771) were at 4.2, 7.8, and 10.2 times greater risk of dying, respectively.
- Similar age disparities in the risk of dying were seen when controlling for additional prognostic factors.
- Among discharged hospital patients, increasing age was related to a significantly poorer long term prognosis.

Other medical patients

The studies done on patients with other medical conditions that have been reviewed are:

- *Gordon (1999a)* compared in-hospital death rates between men and women after adjusting for severity of illness in a retrospective cohort study of 89,793 patients with six common non-surgical diagnoses discharged from 30 hospitals in Northeast Ohio between 1991 and 1993.

Gordon (1999a) reported that:

- Adjusted odds of death were higher (p <0.05) in men, compared with women, for four diagnoses (stroke [OR =1.60]; obstructive airway disease [OR =1.38]; gastrointestinal haemorrhage [OR =1.32]; pneumonia [OR =1.181) and similar for two diagnoses (congestive heart failure [OR =1.12]; and AMI [OR =0.97]).
- These differences were somewhat attenuated by excluding from the analyses patients discharged to skilled nursing facilities or other

    hospitals, although the odds of death in men remained higher for three diagnoses.
  - The findings indicate that in-hospital death rates are generally higher in men than in women, after adjusting for severity of illness. In addition, the risk of in-hospital death in men and women was influenced by diagnosis.
  - These differences may reflect gender related variation in the utilisation of hospital services, the effectiveness of care, over- or underestimation of severity of illness, or biological differences in men and women.

Elderly patients

The studies done on elderly patients that have been reviewed are:

- *Fox (1999)* undertook a prospective study of 923 elderly patients admitted to seven Baltimore hospitals for a hip fracture between 1984 and 1986 to determine if two types of hip fracture (inter-trochanteric and femoral neck) have different patient characteristics and sequelae.
- *Braun (1991)* undertook a retrospective cohort study of 390 veterans discharged to 11 nursing homes to investigate the relationship between nursing home quality and mortality within six months following hospital discharge.

The study by *Fox (1999)* of elderly hip fracture patients revealed that:

  - Patients with inter-trochanteric fractures were slightly older, sicker on hospital admission, had longer hospital stays, and were less likely at two months post-fracture to have recovered activities of daily living than femoral neck fracture patients.
  - Inter-trochanteric fracture patients had higher mortality rates at two and six months after fracturing than did femoral neck fracture patients. Long term recovery (one year) did not differ between fracture type.

Braun (1991) used regression analysis to determine if nursing home quality indices increased the prediction of mortality when severity of illness and case-mix differences were accounted for. Four nursing home quality indices significantly improved the prediction of mortality within six months of discharge:

  - registered nursing hours
  - nursing process
  - security
  - mean quality.

Paediatric patients

The study done on paediatric patients that has been reviewed is:

- *Jenkins (1995)* investigated independent risk factors for in-hospital mortality following paediatric congenital heart surgery in a population based retrospective cohort study of 2,833 cases at 37 acute care hospitals in California and Massachusetts.

Independent risk factors for mortality included:

  - procedure complexity category ($p < 0.0001$)

- use of cardiopulmonary bypass (p <0.0001)
- young age at surgery (p =0.001)
- transfer from another acute care hospital (p <0.0001).

Surgical patients

The studies done on surgical patients that have been reviewed are:

- *Ansari (2000a)* analysed data on 11,036 patients who underwent hemicolectomy or anterior resection between 1987 and 1996 in Victorian public hospitals to examine mortality and associated complications following large bowel resection and anastomosis.
- *Ansari (2000b)* explored associations between in-hospital mortality after transurethral resection of the prostate (TURP) and age, adverse events, type of admission (emergency/planned), location of the hospital (metropolitan/rural), teaching status of the hospital and length of stay, using Victorian hospital morbidity data from public hospitals for the period 1987 to 1995.
- *Hermansson (1999)* conducted a retrospective study to identify predictors of post-operative mortality and length of hospital stay in 246 patients with perforated peptic ulcers.
- *Pronovost (1999)* carried out a retrospective study of administrative data to evaluate the association between patient characteristics and clinical outcomes in all patients (n=2,987) who had abdominal aortic surgery in a Maryland hospital (n=46) from 1994 through 1996.
- *Khuri (1997)* examined predictors of all cause mortality within 30 days for 87,078 major non-cardiac operations performed under anaesthesia between October 1991 and December 1993 at 44 Veterans Affairs Medical Centres.
- *Longo (2000)* assessed risk factors that predict adverse events after colectomy for cancer using validated data on 5,853 patients from the National Veterans Affairs Surgical Quality Improvement Program for the years 1991 to 1995.
- *Dardik (1999)* carried out a population based study of elective abdominal aortic aneurysm (AAA) repairs performed on 2,335 patients (mean age, 70.4 years) in 46 non-federal acute care hospitals in Maryland state from 1990 to 1995 to examine factors influencing in-hospital mortality.

Findings which emerged from the study by *Ansari (2000a)* on outcomes following large bowel resection and anastomosis included:

- Mortality for surgery of the large intestine remains high, largely due to the co-morbidities of the patients.
- The patients most at risk of death were the elderly with pre-existing cardiac or respiratory disease undergoing an emergency operation.
- Certain technical complications such as leakage of an anastomosis after anterior resection are still associated with a significantly increased risk of death.

In the study of outcomes following TURP undertaken by *Ansari (2000b)*, after adjustment for age, co-morbidity, and other confounding variables, highly significant associations with in-hospital mortality were observed for:

- emergency admissions (OR = 1.99, p <0.0001)
- presence of adverse events (OR = 2.69, p <0.0001)

- length of hospital stay (p for trend <0.0001, 95% CI 1.88-2.15)
- age (p for trend <0.0001, 95% CI 1.26-1.48).

Hermansson (1999) found that the following factors had a significant influence on hospital mortality of patients with perforated peptic ulcer:

- age over 75 years (p =0.002)
- co-existing cardiac or pulmonary disease (p =0.02)
- perforation of the cardia or body of the stomach (p =0.02)
- lapse of more than 12 hours between start of symptoms and operation (p =0.006)
- type of operation (p <0.0001).

Pronovost (1999) reported that independent predictors of in-hospital mortality in patients having aortic surgery were:

- age 61 to 70 years (OR =3.1; 95% CI 1.4-6.9)
- age 71 to 84 years (OR=7.2; 95% CI 3.7-14.1)
- age 85 years or older (OR=9.3; 95% CI 3.9-21.9)
- ruptured aneurysm (OR=5.3; 95% CI 3.5-8.2)
- urgent operation (OR=2.3; 95% CI 1.1-5.2)
- emergent operation (OR=3.0; 95% CI 1.9-4.7)
- mild liver disease (OR=4.6; 95% CI 2.0-10.9)
- chronic renal disease (OR=6.9; 95% CI 3.9-12.1).

In the study by *Khuri (1997)* patient risk factors predictive of post-operative mortality included:

- serum albumin level
- American Society of Anesthesia class
- emergency operation
- 31 additional pre-operative variables.

Findings from *Longo (2000)* include:

- One or more complications were observed in 28% of cancer patients following colectomy, the most frequent of which were:
  - prolonged ileus (7.5%)
  - pneumonia (6.2%)
  - failure to wean from the ventilator (5.7%)
  - urinary tract infection (5%).
- For most complications, 30 day in-hospital mortality rates were significantly higher for patients with a complication than for those without.
- Thirty day mortality rates exceeded 50% when any of the following occurred:
  - post-operative coma
  - cardiac arrest
  - a pre-existing vascular graft prosthesis that failed after colectomy
  - renal failure
  - pulmonary embolism
  - progressive renal insufficiency.

- Pre-operative factors that predicted a high risk of 30 day mortality included:
    - ascites
    - serum sodium >145 mg/dl
    - 'do not resuscitate' status before surgery
    - American Society of Anesthesiologists classes III and IV OR V
    - low serum albumin.

Dardik (1999) reported that:

- The in-hospital mortality rate was 3.5% and increased significantly with advancing age: less than 65 years, 2.2%; 65 to 69 years, 2.5%; 70 to 79 years, 3.5%; and more than 80 years, 7.3% (p =0.002).
- Mortality rates were higher for women (4.5% vs. 3.2%, p =0.17), for blacks (6.7% vs. 3.2%, p =0.046), and for patients with renal failure (11.8% vs. 3. 4%, p =0.11), but not for patients with hypertension, diabetes, heart disease, and pulmonary disease.
- The operative mortality rate was inversely correlated with hospital volume (p =0.08), and with the experience of the individual surgeon (p =0.01).
- Multivariate analysis results identified patient age (p =0.002), low hospital volume (p =0.039), and very low surgeon volume (p =0.01) as independent predictors of operative mortality.

STUDIES TO ASSESS THE RISK EFFECT OF HOSPITAL/ SURGICAL VOLUME ON IN-PATIENT MORTALITY

Studies carried out to review the effect of hospital or surgical volume on in-patient mortality can be classified by patient group as those involving:

- general population of patients
- patients with cardiovascular disease
- patients with other medical conditions
- elderly patients
- paediatric patients
- trauma patients
- surgical patients.

General population of patients

The studies done on a general population of patients that have been reviewed are:

- *Sowden (1997)* conducted a systematic review of over 200 (mainly observational) studies to assess the evidence for a relation between hospital or doctor volume and patient outcomes (principally in-patient or 30 day mortality).
- *Sowden (1998)* published a review of the literature on the association between quality of health care and volume of activity.

The key findings from the systematic review by *Sowden (1997)* of over 200 (mainly observational) studies were:

- Routine hospital data are rarely sufficiently detailed to adjust adequately for case-mix; studies that adjust for risk of death based on detailed clinical data are the most valid.
- Much research examining the relationship between hospitals or clinician volume and health outcomes is of poor quality and does not make adequate adjustment for differences in patient case-mix.
- Most studies reported a reduction in poor health outcomes (principally in-patient or 30 day mortality) as volumes increased; however, many studies failed to control for case-mix thereby probably overestimating the effect of volume on the quality of care.
- The best research suggests that there is no general relationship between volume and quality. However, for some procedures or specialties there appear to be quality gains associated with increased hospital or clinician volume.
- The importance of adequate adjustment is illustrated in studies of CABG and also intensive care. For CABG, the size of the relation between low volume (<200 procedures per year) and increased mortality is reduced in studies which better adjust for differences in patient risk.
- In the case of adult intensive care in the UK, where well validated prognostic indicators have been developed (APACHE II), higher mortality rates found in smaller intensive care units were no longer significant after adjusting for case-mix.
- None of the research indicated that increasing the volume of activity over time resulted in changes in health outcomes.

- A positive relationship between high volume and improved outcome can be interpreted in various ways. It might support the 'practice makes perfect' hypothesis, or a selective referral hypothesis, in which hospitals or doctors with good outcomes attract more patients. It may also be the case that higher volume hospitals attract better clinicians or support staff so producing a hospital level effect.

Findings from the review by *Sowden (1998)* on the association between quality of health care and volume of activity were:

- The substantial research literature indicates a positive volume-quality relationship.
- However, these conclusions are largely based on observational studies using administrative databases which are poorly adjusted for case-mix.
- Better control for confounding shows that volume-quality effects in several cases may be an artefact.
- The research is also difficult to interpret because of the limited measurement of outcomes, poor analysis of the relative contributions of the clinician and the hospital levels, and the lack of clarity about the direction of cause and effect.
- Most research is insufficiently reliable to inform policy on the use of volume for credentialling or for the re-configuration of services.

Patients with cardiovascular disease

The studies done on patients with cardiovascular disease that have been reviewed are:

- *Sowden (1995)* undertook a systematic review of the literature to examine the evidence for a relation between volume of CABG surgery and hospital death rates in the US, and to assess the degree to which this could be due to confounding because of differences in case-mix.
- *Hartz (1997a)* compared mortality rates (ratio of the observed to the predicted rate) for patients of physicians performing CABG in three different US states with data on the physicians' reputations obtained from one national and five city surveys.
- *Hartz (1997b)* used Monte Carlo methods to assess how the number of patients per provider affected mortality rates for CABG by surgeon. Each surgeon was assigned a true or underlying probability of mortality.
- *McGrath (1998)* examined the relation between operator volume and outcomes for percutaneous coronary interventions (PCIs) using data on 12,988 PCIs performed between 1990 and 1993 by 31 cardiologists at five US hospitals supporting these procedures.
- *Thiemann (1999)* conducted a retrospective cohort study of the relation between the number of Medicare patients with AMI that each hospital in the study treated (hospital volume) and long term survival among 98,898 Medicare patients aged 65 years or older.
- *Nash (1999)* assessed the magnitude and mechanism of the influence of physician specialty on in-patient mortality in an analysis of 30,351 hospital admissions for AMI.
- *Sollano (1999)* examined the relationship between hospital volume and in-hospital deaths in three cardiovascular procedures: CABG, elective repair of abdominal

aortic aneurysms, and repair of congenital cardiac defects, using data on patients hospitalised in 230 New York State hospitals between 1990 and 1995.

- *Philbin (1999)* carried out a prospective six month follow-up study to determine whether clinical outcomes differed between patients with congestive heart failure (CHF) treated by cardiologists and by non-cardiologists in ten acute care community hospitals.
- *Ho (2000a)* examined the relation between the number of percutaneous transluminal coronary angioplasty (PTCA) procedures performed at hospitals (volume) and in-hospital bypass surgery and death for 353,488 patients treated in California between 1984 and 1996.
- *Hannan (1997a)* assessed the relationship between each of two provider volume measures (annual hospital volume and annual cardiologist volume) for percutaneous transluminal coronary angioplasty (PTCA) and two outcomes of PTCA (in-hospital mortality and same stay CABG surgery) in a cohort study of all 62,670 patients undergoing PTCA in 31 New York hospitals between 1991 and 1994.
- *Canto (2000)* analyzed data from the National Registry of Myocardial Infarction to determine the relation between the number of patients receiving reperfusion therapy (primary angioplasty or thrombolytic therapy) and subsequent in-hospital mortality.
- *Casale (1998)* determined the effect of specialty care on in-hospital mortality in patients with AMI by analysing data from 30,715 direct hospital admissions for the treatment of AMI in Pennsylvania in 1993.
- *Horner (1995)* compared the initial characteristics, in-hospital treatment, and outcomes of patients with a first stroke managed by neurologists (n=88) versus non-neurologists (n=58) in a prospective, hospital based, multi-site cohort study.

In the literature review by *Sowden (1995)* of 15 observational US studies identified, seven were included in the analysis. The reported findings were:

- The seven studies analysed reported a reduced mortality with increased volume.
- Studies with better adjustment for case-mix, however, indicated less reduction in mortality with increased volume (p =0.04).
- The apparent advantages of higher volume also decreased over time (p <0.001).

The main findings of the *Hartz (1997a)* study were:

- Mortality ratios were very similar for the 10,722 patients treated by the 31 surgeons defined as ‘best’ doctors in the surveys (MR = 0.98) and for the 74,854 patients treated by 243 other surgeons who had more than a minimal number of cases (MR = 0.96).
- The mortality ratio was 1.34 for the patients treated by surgeons with the lowest volumes and 0.87 for the surgeons who performed more than 400 CABGs in three years.
- This suggests that the quality of a coronary artery bypass surgeon may be more closely associated with patient volume than with the surgeon's reputation among peers.

Hartz (1997b) used Monte Carlo methods to assess how the number of patients per provider affected mortality rates for CABG by surgeon. Each surgeon was assigned a true or underlying probability of mortality. The findings were:

- The underlying probability of mortality explained 30% of the variation in the observed mortality rate when there were 100 patients per physician, and 63% when there were 400 patients.
- The positive predictive value of using an observed mortality rate in the bottom 10% to identify a surgeon whose underlying probability of mortality was in the bottom 10% was 31% for 100 patients and 59% for 400 patients.
- Overall, these results show that the assessment of provider outcomes may be greatly affected by random variation.

McGrath (1998) categorised surgeons into terciles based on annualized volume of procedures. Successful outcomes were defined as angiographic success (all lesions attempted dilated to <50% residual stenosis) and clinical success (at least one lesion dilated to <50% residual stenosis and no adverse outcomes); adverse outcomes were defined as undergoing CABG, AMI, and death.
It was found that, following adjustment for case-mix, as operator volume increased:

- Higher angiographic success rates were seen (low, middle and high terciles: 84.7%, 86.1% and 90.3%, p-trend 0.006).
- Clinical success rates improved (85.8%, 88.0% and 90.7%, p-trend 0.025).
- Fewer referrals to CABG were seen (4.54%, 3.75% and 2.49%, p-trend <0.001).
- In-hospital mortality rates were similar for the three tertiles (1.09%, 0.96% and 1.05%, p-trend 0.8).

In the analyses by *Thiemann (1999)*, adjustments were made for clinical, demographic, and health system related variables, including the availability of invasive procedures, the specialty of the attending physician, and the area of residence of the patient (rural, urban, or metropolitan). The key findings included:

- The patients in the quartile admitted to hospitals with the lowest volume were 17% more likely to die within 30 days after admission than patients in the quartile admitted to hospitals with the highest volume (hazard ratio =1.17; 95% CI 1.09-1.26, p <0.001), which resulted in 2.3 more deaths per 100 patients.
- The crude mortality rate at one year was 29.8% among the patients admitted to the lowest volume hospitals, as compared with 27.0% among those admitted to the highest volume hospitals.
- There was a continuous inverse dose-response relation between hospital volume and the risk of death.
- In an analysis of sub-groups defined according to age, history of cardiac disease, Killip class of infarction, presence or absence of contraindications to thrombolytic therapy, and time from the onset of symptoms, survival at high volume hospitals was consistently better than at low volume hospitals.
- The availability of technology for angioplasty and bypass surgery was not independently associated with overall mortality.

Nash (1999) reported that:

- In patients < 65 years old, the adjusted OR for mortality with cardiologist care was 0.89 (95% CI 0.64-1.24, p =0.49) relative to generalist care; in patients > or = 65 years of age, the adjusted OR was 0.86 (95% CI 0.72-1.03, p =0.10).
- Caseload was significantly higher among cardiologists and was inversely related to in-patient mortality.
- Both caseload and physician designation were significantly associated with the risk of mortality in the absence of the other variable (OR for cardiologist care 0.82, 95% CI 0.71-0.95, p =0.007; OR for patients with low volume physicians relative to high volume 1.27, 95% CI 1.05-1.51, p =0.014).

Sollano (1999) reported the following findings:

- State-wide, 31 hospitals performed 97,137 CABG procedures over the six year period (overall mortality rate, 2.75%). No correlation was found between hospital volume and in-hospital mortality following CABG.
- State-wide, 195 hospitals (out of 230) performed a total of 9,847 elective abdominal aortic aneurysm repairs with an overall mortality rate of 5.5%. A significant inverse relationship between hospital volume and in-hospital deaths was determined.
- 16 hospitals performed 7,199 repairs for congenital cardiac defects. A significant inverse relationship (which was most pronounced for neonates) was found between volume and death.
- The authors suggest that the lack of an observed relationship between volume and mortality following CABG may be largely explained by the quality improvement program in New York State for bypass operations since 1989.

Philbin (1999) found that:

- Patients (n=419) whose attending physician was a cardiologist were more likely to receive the recommended diagnostic tests and treatment strategies than patients (n=977) who were not treated by a cardiologist, although some of these differences could be explained by variations in the case-mix.
- No differences in adjusted mortality rates were observed between the two patient groups, but those treated by a cardiologist had lower CHF re-admission rates and better post-discharge quality of life measures

Ho (2000a) compared outcomes for three periods: 1984 to 1987, 1988 to 1992, and 1993 to 1996. The reported findings were:

- The in-hospital mortality rate was 2.5% for hospitals performing <200 PTCA procedures per year but only 1.3% for hospitals performing >400 procedures per year in 1984 to 1987. By 1993 to 1996, mortality rates in these two volume categories narrowed to 1.7% and 1.3%, respectively.
- Bypass surgery rates also narrowed and fell in low volume (<200 procedures) versus high volume (>400 procedures) hospitals from

12.4% versus 6.9% in 1984 to 1987 to 4.6% versus 3.3% in 1993 to 1996.
- In a logistic regression, the volume of PTCA procedures significantly predicted in-hospital mortality and bypass surgery rates in all three time periods.

Hannan (1997a) reported that:
- The overall in-hospital mortality rate for patients undergoing PTCA in New York during 1991-1994 was 0.90%, and the same stay CABG surgery rate was 3.43%.
- Patients undergoing PTCA in hospitals with annual PTCA volumes <600 experienced a significantly higher risk adjusted in-hospital mortality rate of 0.96% (95% CI 0.91-1.01) and risk adjusted same stay CABG surgery rate of 3.92% (95% CI 3.76-4.08).
- Patients undergoing PTCA by cardiologists with annual PTCA volumes <75 had mortality rates of 1.03% (95% CI 0.91-1.17) and same stay CABG surgery rates of 3.93% (95% CI 3.65-4.24); both of these rates were also significantly higher than the rates for all patients.
- In New York State, both hospital PTCA volume and cardiologist PTCA volume are significantly inversely related to in-hospital mortality rate and same stay CABG surgery rate for patients undergoing PTCA.

Canto (2000) divided a total of 450 hospitals into quartiles according to the volume of primary angioplasty and used statistical models to determine whether the volume of primary angioplasty procedures was an independent predictor of in-hospital mortality following this procedure. Similar analyses were performed for patients receiving thrombolytic therapy at 516 hospitals. It was reported that:
- In-hospital mortality was 28% lower among patients who underwent primary angioplasty at hospitals with the highest volume than among those who underwent angioplasty at hospitals with the lowest volume (adjusted relative risk =0.72; 95% CI 0.60-0.87, p <0.001). This lower rate, which represented two fewer deaths per 100 patients treated, was independent of the volume of AMI patients at each hospital, year of admission, and use or non-use of adjunctive pharmacologic therapies.
- There was no significant relation between the volume of thrombolytic interventions and in-hospital mortality among patients who received thrombolytic therapy (7.0% for patients in the highest volume hospitals vs. 6.9% for those in the lowest volume hospitals, p =0.36).

Casale (1998) developed a risk adjusted in-hospital mortality model in which 12 of 20 clinical variables were significant independent predictors of in-hospital mortality. Findings included:
- Adjusting for patient characteristics, treatment by a cardiologist (OR =0.83 CI 0.74-0.94, p <0.003), and treatment by a physician treating a high volume of AMI patients (OR =0.89 95% CI 0.80-0.99, p <0.03) were found to be independent predictors of lower in-hospital mortality.
- Treatment by a cardiologist as compared to primary care physicians was also associated with a significantly lower length of stay for both

medically treated patients (p <0.01) and those undergoing revascularization (p <0.01).

In the study by *Horner (1995)* data were collected on patient demographics, disease severity, and risk factors for stroke. The key findings were:

- The 88 patients admitted to the neurology services of the study hospitals had better prognostic profiles i.e. a lower likelihood of having either completed stroke or cardiac co-morbidity.
- Patients admitted to the neurology services of the study hospitals were less likely to die within one and six months of stroke onset.
- Substantial systematic differences in the types of patients managed by neurologists versus other specialists may preclude comparison of outcomes of care.

Patients with other medical conditions

The studies done on patients with other medical conditions that have been reviewed are:

- *Rosenthal (1997a)* compared severity adjusted mortality and length of stay (LOS) in teaching and non-teaching hospitals in a retrospective cohort study in 30 hospitals in Ohio.
- *Ward (1999)* determined whether a hospital's experience in treating patients with systemic lupus erythematosus (SLE) was associated with in-hospital mortality by analysing data on 9,989 patients with SLE admitted to acute care hospitals in California between 1991 and 1994.

Rosenthal (1997a) analysed data on a total of 89 851 patients discharged between 1991 and 1993 with AMI, congestive heart failure, obstructive airway disease, gastrointestinal haemorrhage, pneumonia, or stroke. In-hospital mortality and LOS of patients in major teaching (n=5), minor teaching (n=6), and non-teaching (n=19) hospitals were adjusted for admission severity of illness using data abstracted from patients' medical records. The reported findings were:

- The adjusted odds of death was 19% lower (95% CI 12%-34%, p =0.03) for patients in major teaching hospitals compared with non-teaching hospitals but was similar (95% CI, 7% lower to 28% higher, p =0.28) for patients in minor teaching hospitals.
- The findings were generally consistent in analyses stratified according to diagnosis, age, race, predicted risk of death, and other co-variates.
- In addition, risk adjusted LOS was 9% lower (95% CI, 8%-10%, p <0.001) among patients in major teaching hospitals relative to non-teaching hospitals but was similar (95% CI, 2% lower to 11% higher, p =0.17) in minor teaching hospitals. Major teaching hospitals also cared for higher proportions of non-white and poorly insured patients.

Ward (1999) compared the risks of in-hospital mortality between patients admitted to hospitals in which there was a high degree of experience (an average of >50 urgent or emergency SLE admissions per year) and those in which there was less experience. It was found that:

- Patients admitted on an urgent or emergency basis to hospitals in which there was a high degree of experience had a slightly lower risk of in-hospital mortality than patients admitted to hospitals in which there was less experience (mortality 3.8% vs. 5.3%; adjusted OR =0.72, 95% CI 0.50-1.04).

- In the subset of 2,372 patients hospitalised on an emergency basis, those admitted to hospitals in which there was a high degree of experience had a risk of in-hospital mortality that was 66% lower than that of patients admitted to hospitals in which there was less experience (mortality 4.2% vs. 11.3%; adjusted OR =0.34, 95% CI 0.19-0.58).
- In the subset of 405 patients hospitalised on an emergency basis due to SLE, those admitted to hospitals in which there was a high degree of experience had a risk of in-hospital mortality that was 95% lower than that of patients admitted to hospitals in which there was less experience (mortality 1.7% vs. 10.0%; adjusted OR =0.05, 95% CI 0.01-0.34).
- Among those with emergency hospitalisations for any reason or emergency hospitalisations due to SLE, lengths of stay did not differ between patients hospitalised at hospitals in which there was a high degree of experience and hospitals in which there was less experience.

Elderly patients

The studies done on elderly patients that have been reviewed are:

- *Keeler (1992)* compared quality of care at different types of hospitals measured by explicit criteria, implicit review, and mortality within 30 days of admission (adjusted for sickness at admission) using data retrospectively abstracted from the medical records of a random sample of 14,008 elderly patients with one of the following five diseases: congestive heart failure, AMI, pneumonia, stroke, or hip fracture.
- *Taylor (1999)* used data from the National Long Term Care Survey to study the effects of admission to a teaching hospital on the cost and quality of care for Medicare patients with a first hospitalisation for hip fracture, stroke, coronary heart disease, or congestive heart failure.
- *Whittle (1998)* described the relation of provider characteristics to processes and outcomes of medical care for 22,294 elderly patients hospitalised for community acquired pneumonia in 1990.

The findings reported by *Keeler (1992)* include:

- Quality of care ratings for hospital types are similar using explicit criteria, implicit review, and outcomes adjusted for sickness at admission.
- Quality differences between types of hospitals were large, with the lowest group estimated to have four percentage points higher mortality than major teaching hospitals in a cohort of patients with average mortality of 16%.
- Quality varies from state to state, but teaching, larger, and more urban hospitals have better quality in general than non-teaching, small, and rural hospitals.
- Hospital quality persists over time, but small non-teaching hospitals narrowed the gap with better quality hospitals between 1981 and 1986.

- Overall, the different measures led to consistent and plausible relationships between quality and hospital characteristics.

Taylor (1999) analysed data for Medicare patients with a first hospitalisation for hip fracture (802 patients), stroke (793), coronary heart disease (1007), or congestive heart failure (604) between January 1984 and December 1994. Survival was assessed through 1995. The findings that emerged were:

- After adjustment for patients' characteristics and social subsidies, major teaching hospitals had the lowest mortality rates (hazard ratio for death, 0.75, as compared with for profit hospitals; 95% CI 0.62-0.91).
- For individual conditions, the only significant survival advantage associated with admission to major teaching hospitals was for hip fractures (hazard ratio, 0.54, as compared with for profit hospitals; 95% CI 0.37-0.79).
- In conclusion, although admission to a major teaching hospital may be associated with increased costs, overall survival for patients with the common conditions studied was better at these hospitals, especially for patients with hip fractures.

Whittle (1998) found that overall, 30 day mortality was 17.0%. After adjusting for patient characteristics, 30 day mortality rates were unrelated to hospital teaching status or urban location or to physician specialty.

Trauma patients

The study done on trauma patients that has been reviewed is:

Cooper (2000) analysed New York State Trauma Registry data to determine whether there is a significant relationship between the volume of trauma patients treated by a trauma centre and its risk adjusted in-patient mortality rate. The main finding was that the 35 New York State trauma centres with lower annual volumes had lower, but not significantly lower, observed and risk adjusted mortality rates (7.62% and 8.25%, respectively) than the corresponding rates for the eight New York State higher volume trauma centres (9.36% and 8.83%, respectively).

Paediatric patients

The studies done on paediatric patients that have been reviewed are:

- *Jenkins (1995)* examined the impact of hospital caseload on in-hospital mortality for paediatric congenital heart surgery in a population based retrospective cohort study of 2,833 cases at 37 acute care hospitals in California and Massachusetts.
- *Pollack (1994)* analysed data on 5,415 paediatric intensive care unit admissions at 16 sites to examine the relationship between survival following paediatric intensive care and care factors including size of the intensive care unit (ICU), medical school teaching status of the hospital housing the ICU, specialist status (paediatric intensivist), and unit co-ordination.
- *Hannan (1998a)* examined the relationship between annual provider (hospital and surgeon) volume of paediatric cardiac surgery and in-hospital mortality in a retrospective cohort study of all children undergoing congenital heart surgery in 16 acute care hospitals in New York between 1992 and 1995.

Jenkins (1995) grouped cases into four categories based on the complexity of the procedure. The following findings emerged:

- Compared with centres performing > 300 cases per year, after controlling for patient characteristics, centres performing:
    - <10 cases per year had an OR for in-hospital death of 7.7 (95% CI 1.6-37.8)
    - 10 to 100 cases had an OR of 2.9 (95% CI 1.6-5.3)
    - 101 to 300 cases had an OR of 3.0 (95% CI 1.8-4.9).
- Few differences were found by hospital caseload in length of stay.
- The authors acknowledge that the study was limited by the absence of clinical detail in discharge abstract databases.

Pollack (1994) reported that:

- Adjusting for physiologic status, diagnosis, and other risk factors, mortality ranged from 2.2% to 16.4 % across the ICUs.
- The probability of patient survival after hospitalisation in an ICU located in a teaching hospital was decreased (relative odds of dying =1.79; 95% CI 1.23-2.61, p =0.002).
- In contrast, the probability of patient survival after hospitalisation in an ICU with a paediatric intensivist was improved (relative odds of dying =0.65; 95% CI 0.44-0.95, p =0.027).
- Post-hoc analysis indicated that the higher severity adjusted mortality in teaching hospitals may be explained by the presence of residents caring for ICU patients.

Hannan (1998a) reported the following results from an analysis of 7,169 cases of paediatric cardiac surgery:

- After controlling for severity of pre-procedural illness using clinical risk factors, hospitals with annual paediatric cardiac surgery volumes of fewer than 100 had significantly higher mortality rates (8.26%) than hospitals with volumes of 100 or more (5.95%).
- Surgeons with annual volumes of fewer than 75 had significantly higher mortality rates (8.77%) than surgeons with annual volumes of 75 or more (5.90%).
- These differences persisted for both high complexity and low complexity paediatric cardiac procedures.

Surgical patients

The studies done on surgical patients that have been reviewed are:

- *Kreder (1998)* undertook a retrospective population cohort study of elective total hip replacement operations during 1992 in Ontario to test the hypothesis that complication rates for this procedure are related to surgeon and hospital volumes.
- *Urschel (2000)* undertook a retrospective review of operative mortality following two high complexity thoracic surgical operations (pneumonectomy and oesophagectomy) done over a four year period at a small community hospital, by a surgeon from a tertiary care hospital that provided a community outreach satellite program in thoracic surgery. This provided a unique opportunity to study the hospital volume-outcome relationship without the confounding variable of surgeon experience.

- *Hannan (1998b)* assessed the relationship between each of two provider volume measures for carotid endarterectomies (annual hospital volume and annual surgeon volume) and in-hospital mortality among all 28,207 patients for whom carotid endarterectomy was the principal procedure performed in New York State hospitals between January 1990 and December 1995.
- *Gordon (1999b)* evaluated whether provider volume and experience were important factors influencing in-hospital mortality and other outcomes for six complex high risk gastrointestinal surgical procedures performed on 4,561 patients in Maryland between July 1989 and June 1997.
- *Johnston (2000)* assessed whether patients treated for cerebral aneurysms at institutions offering endovascular services (coil embolization and angioplasty for cerebral vasospasm) had lower rates of in-hospital mortality, and whether hospital volume was associated with in-hospital death.
- *Hillner (2000)* conducted a comprehensive review of articles published in Medline between 1988 and 1999 that considered whether hospital or physician volume or specialty affects the outcome of cancer care (including mortality).
- *Ellison (2000)* assessed whether patient outcome among 66,693 men undergoing radical prostatectomy for prostate cancer is associated with the volume of such procedures performed at hospitals annually. Patients were categorised into the following hospital volume groups: low (<25), medium (25-54), and high (>54).
- *Harmon (1999)* examined the association of surgeon and hospital case volumes with in-hospital death and length of stay among adult patients who underwent resection for colorectal carcinoma in Maryland state between 1992 and 1998. Cases were categorised by annual surgeon case volume: low (<=5), medium (5-10), and high (>10), and also by hospital volume: low (<40), medium (40-70), and high (>=70). Analyses were adjusted for variations in type of resections performed, cancer stage, patient co-morbidities, urgency of admission, and patient demographic variables.
- *Begg (1998)* undertook a retrospective cohort study of 5013 cancer patients (aged 65 years or older at diagnosis) who underwent pancreatectomy, oesophagectomy, pneumonectomy, liver resection, or pelvic exenteration, to examine whether hospital volume for these relatively complex surgical oncologic procedures was inversely associated with 30-day operative mortality, after adjusting for co-morbidity, patient age, and cancer stage.
- *Sosa (1998)* examined whether hospital volume was associated with clinical and economic outcomes among all patients with pancreatic cancer who underwent a pancreatic resection, palliative bypass, or stent procedure between 1990 and 1995 in 48 non-federal acute care hospitals in Maryland.
- *Birkmeyer (1999)* examined the relationship between hospital volume and mortality for pancreaticoduodenectomy (the Whipple procedure) in a national cohort study of 7,229 Medicare patients (aged over 65 years) undergoing pancreaticoduodenectomy between 1992 and 1995 in US hospitals.
- *Dudley (2000)* undertook a literature review to determine the difference in hospital mortality between high volume hospitals (HVHs) and low volume hospitals (LVHs) for certain conditions and to estimate how many deaths potentially would be avoided by referral to HVHs.
- *Khuri (1999)* used data from the Veterans Health Administration (VHA) National Surgical Quality Improvement Program to assess the relation of surgical volume to 30-day mortality rate for eight common surgical procedures (68,631 operations)

of intermediate complexity: non-ruptured abdominal aortic aneurysmectomy, vascular infra-inguinal reconstruction, carotid endarterectomy (CEA), lung lobectomy/pneumonectomy, open and laparoscopic cholecystectomy, partial colectomy, and total hip arthroplasty.

- *Dardik (1999)* carried out a population based study of elective abdominal aortic aneurysm (AAA) repairs performed on 2,335 patients (mean age, 70.4 years) in 46 non-federal acute care hospitals in Maryland state from 1990 to 1995 to examine the effects of hospital and surgeon volume on in-hospital mortality.
- *Glasgow (1999)* examined the volume-outcome relation for hepatectomy for hepatocellular carcinoma (HCC), a high risk procedure in general surgery, in a retrospective cohort study of 507 patients undergoing hepatectomy for HCC in 138 acute care hospitals in California between 1990 and 1994.
- *Gordon (1998)* analysed data on 795 pancreaticoduodenectomies performed in 43 Maryland hospitals between 1984 and 1995 to examine the state-wide trend toward regionalization of pancreaticoduodenectomy and its effect on state-wide in-hospital mortality rates for this procedure.
- *Patti (1998)* evaluated the relationship between a hospital's annual rate of oesophagectomy for oesophageal cancer and mortality by analysing data on 1,561 patients who underwent oesophagectomy for malignancy in 273 acute care hospitals in California between 1990 and 1994.
- *Silvestri (1998)* determined whether the outcome of lung cancer resection differs between general surgeons (GSs) and thoracic surgeons (TSs) by analysing data from a state-wide severity adjusted administrative hospital discharge database on 1,583 lung resections performed between 1991 and 1995.
- *Pearce (1999)* analysed data on 45,744 carotid endarterectomies (CEA), 31,172 lower extremity bypass graftings (LEAB), and 13,415 abdominal aortic aneurysm repairs (AAA) performed in non-federal hospitals between 1992 and 1996 to determine the impact of surgeon volume on mortality and morbidity outcomes.

Kreder (1998) reported that:

- Surgeons with patient volumes above the 80th percentile (more than 27 hip replacements annually) discharged patients approximately 2.4 days earlier ($p < 0.05$) than surgeons with volumes below the 40th percentile (less than nine hip replacements annually) even after adjusting for discharge disposition, hospital volume, patient age, sex, co-morbidity and diagnosis.
- Death rates did not differ by surgeon or hospital volume ($p > 0.05$).

Urschel (2000) reported the following findings:

- There was one in-hospital death and one 30 day mortality out of 486 thoracic surgical procedures (317 minor and 169 major cases).
- Despite having a very low volume of thoracic surgical cases the community hospital had crude outcomes comparable to those reported from high volume tertiary hospitals suggesting that the surgeon may be a more important factor in the hospital volume-outcome relationship than previously thought.

Hannan (1998b) developed a statistical model to predict in-hospital mortality using age, admission status, and several conditions found to be associated with higher than average mortality. Risk adjusted in-hospital mortality ranged from 1.96% (95% CI 1.47-2.57) for patients having surgeons with extremely low annual volumes (<5) of carotid endarterectomy (CE) in hospitals with low annual CE volumes (≤100) to 0.94% (95% CI 0.73-1.19) for patients having surgeons with annual volumes of ≥5 in hospitals with annual CE volumes of >100. These two rates were significantly different.

In the *Gordon (1999b)* study, hospitals were classified into one of four groups based on the average number of study procedures per year: minimal (ten or fewer procedures), low (11 to 20 procedures), medium (21 to 50 procedures), and high volume groups (201 or more procedures). The study population averaged 61.6 years of age, was 55% male, 71% Caucasian, and had predominantly Medicare as a payment source. Poisson regression was used to assess the relationship between in-hospital mortality and hospital volume after case-mix adjustment. The main findings were:

- After case-mix adjustment, patients who underwent complex gastrointestinal surgical procedures at the medium, low, and minimal volume provider groups had a 2.1, 3.3, and 3.2 times greater risk of in-hospital death, respectively, than patients at the high volume provider ($p < 0.001$ for all comparisons);
- After case-mix adjustment, patients at the low, medium, and minimal volume provider groups had longer lengths of stay (16.1, 15.7, and 15.5 days respectively) than patients at the high volume provider (14.0 days) ($p < 0.001$ for all comparisons).
- Adjusted charges at the high volume provider were, on average, 14% less than those of the low volume group, which had the next lowest charges.
- Although mortality rates differed by procedure type, for each procedure, mortality increased as provider volume decreased, following the pattern found in the aggregate analysis.
- After case-mix adjustment, the risk of in-hospital death for patients with malignant diagnoses was significantly higher for the medium, low, and minimal volume groups compared with patients at the high volume provider, relative risk of 3.1, 4.0, and 4.2, respectively, ($p < 0.001$ for all comparisons).
- In conclusion, increased hospital experience is associated with a marked decrease in hospital mortality. The decreased mortality at the high volume provider was also associated with shorter lengths of stay.

Johnston (2000) analysed data on 2,623 unruptured and 9,534 ruptured aneurysms treated at 70 centres during 1994-1997, adjusted for age, sex, race, admission source, and admission status. The key results were:

- Hospital treatment volume was not independently associated with in-hospital death.
- Patients treated at institutions that more frequently used coil embolization were less likely to die in hospital (relative risk [RR] for every 10% of endovascular treated cases: RR =0.91; 95% CI 0.86-0.96, p =0.001 (ruptured aneurysms); RR =0.84; 95% CI 0.78-0.91, p <0.001 (unruptured aneurysms)).

- Patients treated at institutions that used angioplasty for vasospasm had a 16% reduction in risk of in-hospital death compared with patients treated at other institutions (RR =0.84; 95% CI 0.71-0.98, p =0.03).
- Whether this is due to improved outcomes with endovascular therapy or is a marker for other aspects of multidisciplinary care cannot be answered from these data.

Hillner (2000) reported the following findings:

- The review found numerous methodological limitations in almost all of the studies identified:
    - All but one of the studies reviewed used retrospective data.
    - The vast majority were not hypothesis driven.
    - Almost all used convenience samples and had no pre-planned statistical power estimates.
    - Co-morbidity was usually inferred from administrative claims rather than from specific clinical indices or databases.
- Studies of hospital characteristics usually considered only hospital size, residency status and type, characteristics which describe an entire hospital and not just its cancer program. Additional factors which have been suggested to be associated with or indicators of superior cancer care, such as nurse staff levels, multidisciplinary teams, internal quality programs, and the range of services and technology, were generally not considered in the studies reviewed.
- An extensive, consistent literature that supported a volume-outcome relationship was found for cancers treated with technologically complex surgical procedures (e.g. most intra-abdominal and lung cancers). These studies predominantly measured in-hospital or 30 day mortality and used the hospital as the unit of analysis.
- For cancer primarily treated with low risk surgery, there were fewer published studies identified.
- An association with hospital volume and surgeon specialization in colon cancer varied with the volume threshold used. Beneficial effects were observed predominantly for rectal cancer.
- For breast cancer, British studies found that physician specialty and volume were associated with improved long term outcomes, and the single American report showed an association between hospital volume of initial surgery and better five year survival.
- Studies of non-surgical cancers, principally lymphomas and testicular cancer, were few but consistently showed better long term outcomes associated with larger hospital volume or specialty focus.
- Studies in recurrent or metastatic cancer were absent.
- Across studies, the absolute benefit from care at high volume centres exceeds the benefit from breakthrough treatments.
- The authors conclude that although these reports are all retrospective, rely on registries with dated data, rarely have predefined hypotheses, inadequately control for case-mix differences, and may have publication and self interest biases, most support a positive volume-outcome relationship in initial cancer treatment.

- The authors point out that the direction of a causal relationship between volume and outcome may be difficult to determine: does higher volume lead to better outcome or do better units and clinicians attract more patients?
- The authors conclude that further study of the relationship between volume and outcome is required to define the relevant processes of care which produce the better outcome.

Ellison (2000) found that:

- Overall in-hospital mortality after radical prostatectomy adjusted for patient characteristics was relatively low (0.25%).
- Patients at low volume centres were 78% more likely to have in-hospital mortality than those at high volume centres (adjusted OR =1.78, 95% CI 1.7-2.6).
- Overall length of stay decreased at all hospitals between 1989 and 1995.
- Average length of stay was longer was higher at low than at high volume centres (7.3 vs. 6.1 days, $p < 0.0001$).
- Further study is necessary to examine the association of hospital volume with other important outcomes, including incontinence, impotence and long term patient survival after radical prostatectomy.

The key findings from *Harmon (1999)* were:

- During the five year period, 9,739 resections were performed by 812 surgeons at 50 hospitals. The majority of surgeons (81%) and hospitals (58%) were in the low volume group.
- Higher surgeon volume was associated with significant improvement in in-hospital death, length of stay, and cost.
- Medium volume surgeons achieved excellent outcomes similar to high volume surgeons when they operated in high or medium volume hospitals, but not in low volume hospitals.
- The results of low volume surgeons improved with increasing hospital volume but never equalled those of the high volume surgeons.

Begg (1998) undertook a retrospective cohort study of 5,013 older cancer patients to examine whether hospital volume for relatively complex surgical procedures was inversely associated with 30 day mortality, after adjusting for co-morbidity, patient age, and cancer stage. Findings included:

- Higher volume was linked with lower mortality for pancreatectomy ($p = 0.004$), oesophagectomy ($p < 0.001$), liver resection ($p = 0.04$), and pelvic exenteration ($p = 0.04$), but not for pneumonectomy ($p = 0.32$).
- The most striking results were for oesophagectomy, for which the operative mortality rose to 17.3% in low volume hospitals, compared with 3.4% in high volume hospitals, and for pancreatectomy, for which the corresponding rates were 12.9% vs. 5.8%.
- Adjustments for case-mix and other patient factors did not change the finding that low volume was strongly associated with excess mortality.

- These data support the hypothesis that when complex surgical oncologic procedures are provided by surgical teams in hospitals with specialty expertise, mortality rates are lower.

Sosa (1998) reported that:

- Increased hospital volume was associated with markedly decreased in-hospital mortality rates and a decreased or similar length of stay for all three types of procedures, and with decreased or similar hospital charges for resections and stents.
- After adjustment for case-mix differences, the relative risk (RR) of in-hospital death after pancreatic resection was 19.3 and 8 at the low and medium volume hospitals, respectively, versus the high volume hospital.
- After bypasses, the RR of death was 2.7 and 1.9, at the low and medium volume hospitals, respectively, versus the high volume hospital.
- After stents, the RR was 4.3 and 4.8, at the low and medium volume hospitals, respectively, versus the high volume hospital.

Birkmeyer (1999) divided the study population into approximate quartiles based on annual hospital volume of pancreaticoduodenectomies in Medicare patients: very low (<1/y), low (1-2/y), medium (2-5/y), and high (5+/y). Using multivariate logistic regression to account for potentially confounding patient characteristics, the association between institutional volume and in-hospital mortality was examined. The key findings were:

- More than 50% of Medicare patients undergoing pancreaticoduodenectomy received care at hospitals performing fewer than two such procedures per year.
- Adjusted in-hospital mortality rates at these low and very low volume hospitals were threefold to fourfold higher than at high volume hospitals (12% and 16%, respectively, vs. 4%, $p < 0.001$).
- Within the high volume quartile, the ten hospitals with the nation's highest volumes had lower mortality rates than the remaining high volume centres (2.1% vs. 6.2%, $p < 0.01$).
- The strong association between institutional volume and mortality could not be attributed to patient case-mix differences or referral bias.
- Although volume-outcome relationships have been reported for many complex surgical procedures, hospital experience is particularly important with pancreaticoduodenectomy.

In a literature review, *Dudley (2000)* identified the highest quality published study assessing the mortality-volume relationship for each given condition and used these to calculate ORs for in-hospital mortality for low volume hospitals (LVHs) versus high volume hospitals (HVHs). These ORs were then applied to the 1997 California database of hospital discharges to estimate potentially avoidable deaths. Mortality was significantly lower at HVHs for the following conditions:

- elective abdominal aortic aneurysm repair
- carotid endarterectomy
- lower extremity arterial bypass surgery

- CABG
- coronary angioplasty
- heart transplantation
- paediatric cardiac surgery
- pancreatic cancer surgery
- oesophageal cancer surgery
- cerebral aneurysm surgery.

The authors estimated that 602 deaths (95% CI 304-830) at LVHs could be attributed to their low volume, and concluded that additional study is needed to determine the extent to which selective referral to HVHs of patients with the above conditions is feasible.

Khuri (1999) found no statistically significant associations between procedure or specialty volume and 30 day mortality rate for the eight major surgical procedures of intermediate complexity which included non-ruptured abdominal aortic aneurysmectomy, vascular infra-inguinal reconstruction, carotid endarterectomy (CEA), lung lobectomy/pneumonectomy, open and laparoscopic cholecystectomy, partial colectomy, and total hip arthroplasty.

Dardik (1999) categorised the 46 hospitals according to the volume of AAAs performed: there were seven high volume (>100 cases), nine moderate volume (50 to 99 cases), and 30 low volume (<50 cases) institutions. The key results were:

- The operative mortality rate was inversely correlated with hospital volume (4.3% in low volume hospitals, 4.2% in moderate volume hospitals, and 2.5% in high volume hospitals, p =0.08), although no differences were noted in the mean ages or co-morbidity levels of patients who underwent operations in these three hospital populations.
- The operative mortality rate was inversely correlated with the experience of the individual surgeon: one case, 9.9%; two to nine cases, 4.9%; ten to 49 cases, 2.8%; 50 to 99 cases, 2.9%; and more than 100 cases, 3.8% (p =0.01).
- Multivariate analysis results identified patient age (p =0.002), low hospital volume (p =0.039), and very low surgeon volume (p =0.01) as independent predictors of operative mortality.
- The mean length of stay decreased with increasing surgeon volume: one case, 22.7 days; two to nine cases, 10.6 days; ten to 49 cases, 10.0 days; 50 to 99 cases, 10.9 days; and more than 100 cases, 9.6 days (p <0.0001).

Glasgow (1999) reported that:

- In-hospital mortality was 14.8% overall.
- Three quarters of patients were treated at hospitals that average three or fewer hepatic resections for HCC per year. These low volume providers represent 97.1% of all hospitals treating patients with HCC state-wide.
- Significant reductions in risk adjusted operative mortality rates (22.7% to 9.4%, p =0.002, multiple logistic regression) and risk adjusted length of stay (14.3 to 11.3 days, p =0.03, multiple linear regression) were observed as hospital volume increased.

Gordon (1998) used regression analyses to examine whether hospital share of pancreaticoduodenectomies was a significant predictor of the in-hospital mortality rate, adjusting for study year and patient characteristics. The key findings were:

- During the study period, one institution increased its yearly share of pancreaticoduodenectomies from 20.7% to 58.5%, and the state-wide in-hospital mortality rate for the procedure decreased from 17.2% to 4.9%.
- After adjustment for patient characteristics and study year, hospital share remained a significant predictor of mortality: an estimated 61% of the decline in the state-wide in-hospital mortality rate was attributable to the increase in share of discharges at the high volume provider, demonstrating the effectiveness of regionalization for high risk surgery.

Patti (1998) reported the following findings:

- 88% of hospitals performed an average of two or fewer resections annually, accounting for 50% of all patients treated.
- The mortality rate in hospitals with >30 oesophagectomies for the five year period was 4.8%, compared with 16% for hospitals with <30 oesophagectomies. This could not be accounted for by other health variables affecting the patients' risk for surgery.
- This striking correlation between a hospital's frequency of oesophagectomy and hospital mortality supports the proposition that high risk general surgical procedures should be restricted to hospitals that can exceed a yearly minimum procedure volume.

Silvestri (1998) reported that:

- Half of all lobectomies and nearly 60% of pneumonectomies were performed by GSs. Patients were similar in age, sex, race, and the proportion in each severity of illness sub-class.
- Mortality was significantly higher in patients who underwent lobectomy by GSs vs. TSs (5.3% vs. 3.0%, $p < 0.05$) and in patients with extreme co-morbidities (43.6% vs. 25.4%, $p = 0.03$) or age >65 years (7.4% vs. 3.5%, $p < 0.05$).
- Seventy percent of TSs performed >10 cases in the series, whereas 75% of GSs performed <10 ($p = 0.05$).
- Logistic regression analysis failed to identify any significant variable that might explain the mortality differences between TSs and GSs.

Pearce (1999) used multiple logistic regression to test the influence of the following factors on outcome: surgeon volume, certification for added qualifications in general vascular surgery, hospital size, hospital volume, patient age, and gender. The key findings were:

- The in-hospital mortality rate increased with age.
- A doubling of surgeon volume was associated with a 4% reduction in risk for adverse outcome for CEA ($p = 0.006$), an 8% reduction for LEAB, and an 11% reduction for AAA ($p = 0.0002$).

- Although hospital volume was significant in predicting better outcomes for CEA and AAA procedures, it was not associated with better outcomes for LEAB.
- Certification for added qualifications in general vascular surgery was a significant predictor of better outcomes for CEA (15% lower risk rate of death or complications, p =0.002) and AAA (24% lower risk rate of death or complications, p =0.009).

STUDIES TO ASSESS THE RISK EFFECT OF LENGTH OF STAY ON IN-PATIENT MORTALITY

If hospitals prematurely discharge patients, length of stays significantly lower than expected might be considered indicative of poor quality care. On the other hand, if poor quality of care causes complications, it would tend to extend length of stay. Under this assumption, longer than expected length of stays could be viewed as indicative of poor quality care.

Studies carried out to review the effect of length of stay on in-patient mortality can be classified by patient group as those involving:

- general population of patients
- patients with cardiovascular disease
- patients with other medical conditions
- surgical patients.

General population of patients

The studies done on a general population of patients that have been reviewed are:

- *Yuan (2000)* examined the association between hospital type and mortality and length of stay (LOS) in a retrospective cohort analysis of data on 16.9 million hospitalised Medicare beneficiaries admitted for ten common medical conditions and ten common surgical procedures between 1984 and 1993.
- *Louis (1999)* examined the effects of the introduction of a hospital financing reform system in Italy in 1995 (which aimed to make hospitals more accountable for their productivity) on LOS, re-admission rates and mortality rates for nine common medical and surgical conditions.

Yuan (2000) grouped a total of 5,127 acute care US hospitals into six mutually exclusive hospital types based on teaching status and financial structure. The findings that emerged included:

- Patients at teaching not for profit [TNFP] hospitals had significantly lower risk adjusted 30 day mortality rates than patients at other hospital types when all diagnoses or procedures were combined.
- The results were mostly consistent when diagnoses and procedures were examined separately.
- After adjustment for patient characteristics, patients at other hospital types had 10% to 20% shorter lengths of stay than patients at TNFP hospitals for most diagnoses and procedures studied.

In the *Louis (1999)* study, mean LOS decreased from 9.1 days to 8.8 days ($p < 0.001$) between 1993 and 1996 while there was little or no change in mortality or re-admission rates.

Patients with cardiovascular disease

The studies done on patients with cardiovascular disease that have been reviewed are:

- *Philbin (1997a)* studied the relationship between LOS and the rate of death in a retrospective, observational study of 3,914 patients hospitalised with congestive heart failure (CHF) in 15 acute care community hospitals in New York.

- *Philbin (1997b)* evaluated the clinical relevance of LOS in congestive heart failure (CHF) by studying its relationship to outcomes among 1,402 predominantly elderly patients with moderately severe or severe CHF admitted to ten acute care community hospitals in New York.
- *Ni (1999)* used multi-variable methods to assess trends in LOS and in-hospital mortality for congestive heart failure (CHF), while controlling for confounding factors, such as age, sex, and co-morbidity.
- *Polanczyk (2000)* assessed trends over ten years in mortality, LOS, and other factors among 6,676 patients with a primary discharge diagnosis of congestive heart failure (CHF) admitted between 1986 and 1996.

Philbin (1997a) calculated an index of severity of illness and a severity weighted expected LOS for each patient. The reported findings were:

- There was significant variability in mean total LOS (7.6 to 12.7 days), mean number of acute care days (acute LOS) (7.1 to 10.3 days), and death rates (4.3 to 12.0%) among the centres.
- Minimal variation in mean expected LOS (5.2 to 6.1 days) and mean severity score (2.8 to 3.3) was observed.
- Mean total LOS ($r = 0.14$, $p = 0.61$) and acute LOS ($r = 0.11$, $p = 0.69$) were not related significantly to death rate for the 15 centres.
- When the hospitals were separated into tertiles based on rank order of total LOS and acute LOS, no differences among the sub-groups were noted in the number of cases per hospital, deaths per hospital, death rates, expected LOS, and severity scores.
- Inter-hospital variation in total LOS was partially explained by the care of patients who did not require acute hospitalisation.

Philbin (1997b) followed patients over six months after discharge and found that:

- Longer than average LOS had a neutral or negative association with patient outcomes.
- Longer LOS was linked to a higher adjusted mortality rate during the index hospitalisation, as well as a greater adjusted risk of death during the post-discharge period.

Ni (1999) reported the following findings:

- The hospital admission rates for CHF were stable over time in all age groups.
- The annual hospital re-admission rate remained constant over time, with an average rate of 15.3%.
- The average length of hospital stay decreased from 5.01 days in 1991 to 3.95 days in 1995.
- The in-hospital mortality rate decreased from 6.9% in 1991 to 4.7% in 1995, independent of LOS.

Polanczyk (2000) adjusted hospital mortality and LOS for socio-demographic characteristics, co-morbidities, invasive procedures, hospital disposition, and LOS where appropriate. It emerged that:

- There was a significant increasing trend in heart failure severity as assessed by a CHF-specific risk adjustment index.

- The proportion of patients who underwent invasive procedures was significantly higher in the 1994-1996 period.
- The standardised mortality ratio (observed mortality/predicted mortality) progressively fell during the study period. Compared with patients admitted before 1991, those admitted after 1991 had a 24% lower observed than predicted mortality.
- Adjusted LOS exhibited a downward trend (7.7 days in 1986-1987 vs. 5.6 days in 1994-1996, $p < 0.001$).

Patients with other medical conditions

The studies done on patients with other medical conditions that have been reviewed are:

- *Benenson (1999)* conducted a retrospective chart review for three cohorts of pneumonia patients admitted via the emergency department to evaluate the effects of a clinical pathway on time to antibiotic treatment, mortality, and average LOS three years after pathway implementation.
- *McCormick (1999)* examined the association between LOS and mortality in a prospective cohort study of 1,188 adult patients with community acquired pneumonia admitted to one of four hospitals.

Benenson (1999) found that following implementation of the clinical pathway:

- The mean time to antibiotic administration decreased from 315 minutes pre-pathway to approximately 171 minutes at three years ($p < 0.0001$).
- The percentage of patients who received antibiotics in the ED increased from 58% pre-pathway to 97% at three years ($p < 0.0001$).
- LOS decreased from 9.7 pre-pathway to 6.4 days at three years ($p < 0.0001$).
- There was a non-significant decrease in in-hospital mortality (9.6% pre-pathway to 4.9%).

McCormick (1999) reported the following:

- Substantial inter-hospital variation was observed in the average lengths of stay.
- After adjustment for baseline differences in case-mix, mortality rates during the first 14 days after discharge were similar in patients admitted to the hospital with the shortest length of stay and those admitted to hospitals with longer mean lengths of stay (relative risk =0.7; 95% CI 0.3-1.7).
- No differences were seen in other health outcomes including hospital re-admission within 14 days, return to usual activities, return to work, and pneumonia related symptoms.

Surgical patients

The studies done on surgical patients that have been reviewed are:

- *Velasco (1996)* examined the effects of introducing a critical pathway for elective CABG on LOS, post-operative mortality, re-admission rates and costs.

- *Jacobs (1999)* reports on the effects on the mortality rate and average LOS among CABG patients following the setting up of an intensivist directed cardiothoracic intensive care unit (CTU) at a tertiary care teaching hospital in New York State.
- *Cowper (1997)* examined the impact of early hospital discharge on short term clinical outcomes among 83,347 non-health maintenance organisation (HMO) Medicare patients who underwent CABG in the United States in 1992.
- *Edwards (1991)* carried out a population based study of all 11,199 carotid endarterectomies (CE) performed by all surgeons in a single state between 1979 and 1988 to examine relationships between mortality, LOS, physician volume, and hospital volume.
- *Munoz (1990)* tested the hypothesis that hospital mortality would differ for each individual surgical oncologist by the volume of patients treated, by examining data on all 2,627 elective surgical oncologic admissions to one hospital between 1985 and 1987.
- *Rios (1999)* compared outcomes of pancreatoduodenectomy (PD) for the management of benign and malignant pancreatic disease among patients undergoing PD between 1983 and 1996 (Group A, n=98) and those undergoing the procedure between 1956 and 1982 (Group B, n=55).
- *Ho (2000b)* examined the effect of wait time for hip fracture surgery on post-surgery length of stay (LOS) and in-patient mortality in two Canadian provinces and in two U.S. states.

Velasco (1996) examined the effects of introducing a critical pathway for elective CABG which had been developed through a co-operative effort involving surgeons, anaesthesiologists, nurses, social workers, physical therapists, nutritionists, and patient case managers) at a New York hospital. Results revealed that the pathway significantly reduced hospital LOS and direct costs without increasing post-operative mortality and re-admission rates.

Jacobs (1999) reported that following the introduction of new management strategies to improve patient care on the cardiac surgery service with specific reference to post-operative care, the CABG mortality rate decreased significantly while the overall LOS for all cardiac surgical patients in the CTU decreased.

In the study by *Cowper (1997)* early discharge was defined as a post-operative LOS of less than or equal to five days. The following results were reported:

- In 1992, 6% of Medicare patients undergoing CABG were discharged within 5 days of the operation.
- Patients discharged early tended to be younger and male and have fewer co-morbid illnesses.
- Risk adjusted rates of (a) death within 60 days and (b) death or cardiovascular re-admission within 60 days were lowest among patients discharged early.
- These results suggest that physicians were able to identify low risk candidates for early discharge.

Edwards (1991) reported that:

- Mortality rate from CE was 2.1% and the post-operative stroke rate was 3.7%.

- High physician volume decreased the mortality rate (p <0.05) and stroke rate (p <0.01) by 50% and significantly (p <0.001) reduced hospital cost and LOS independent of patient complexity.

Munoz (1990) reported that patients of the 57 low volume surgeons had a longer average LOS and had a higher mortality compared with patients of the 17 high volume surgeons. This was due, in part, to a greater severity of illness.

Rios (1999) found that:

- Pre-operative biliary drainage was performed in 62% of Group A and 3% of Group B patients (p <0.001).
- Post-operative complications were comparable in both groups.
- Both the peri-operative mortality rate and the mean post-operative LOS were significantly different between the two groups: 1% mortality in Group A versus 16% in Group B (p <0.001), and 17 days in Group A (22 days for benign disease) versus 25 days in Group B (p <0.01).

Ho (2000b) analysed discharge abstracts for patients admitted to an acute care hospital with a primary diagnosis of hip fracture who received hip fracture surgery, were admitted from home or the emergency room, were age 45 or older, stayed in the hospital for one year or less, and were not trauma patients. The main findings were:

- Canadians had longer wait times for surgery, longer post-surgery LOS and higher in-patient mortality than the Americans.
- Significant differences in mortality across the geographical regions studied were not attributable to the wait time for surgery.
- The model indicated that the effect of wait time for surgery on post-surgery LOS was small.

STUDIES TO EXAMINE THE RELATIONSHIP OF CASE FATALITY RATES TO THE QUALITY OF IN-PATIENT CARE

Quality of care consists of the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.

A number of studies have been done to identify:

- whether deaths associated with admission are preventable
- the extent to which case fatality rates are influenced by the quality of in-patient care.

Studies of this nature which have been reviewed can be categorised by the following patient groups:

- general population of patients
- patients with cardiovascular disease
- patients with other medical conditions
- paediatric patients
- surgical patients.

General population of patients: Thomas (1998) review

Thomas (1998) conducted a major review of articles published up to 1997 that provided information on the validity of hospital mortality rates as a measure of performance quality; a total of 18 articles published over 30 years were reviewed.

The 18 articles reviewed were grouped into three categories as follows:

- Articles that examine validity empirically, focusing on hospital level relationships between risk adjusted mortality rates and process based measures of quality (eight studies).
- Articles that examine conditions affecting validity. These include those that:
  - provide evidence about population level relationships between quality of care and patients' mortality risk (three studies)
  - focus on potential systematic measurement errors associated with inadequate risk adjustment (systematic error variance reflects the influence on the measure of factors that are unrelated to the concept being studied)
  - quantify the potential for random measurement error when using mortality rates to evaluate hospital quality performance (random error variance represents the portion of total variance that remains after accounting for valid variance (variance related to the concept being studied) and systematic error variance; it is the degree of unreliability of the measure).
- Articles that infer validity from observed behavioural responses to public release of hospital mortality rate data.

Of the 18 papers reviewed, eight carried out empirical investigations of validity. Of these, three examined correlational validity (also called construct validity), while the other five examined predictive validity. In examinations of correlational validity, the measure under study is compared with one or more measures thought, but not known, to be highly valid. The question of interest is therefore: how likely is the candidate measure to yield information that is consistent with other accepted measures of the concept?

All three of the studies that assessed correlational (construct) validity examined relationships between hospitals' risk adjusted mortality rates and a process based measure of quality performance. These studies are reviewed below:

Williams (1979) performed a hospital level regression analysis of quality related factors on perinatal mortality rate. By linking birth certificate records of 3.44 million babies delivered in Californian hospitals in 1960 and 1965-1973 with death certificates for the 39,000 fetal and neonatal deaths during this period, an observed to expected perinatal mortality ratio for each of 504 hospitals was developed. Results from stepwise regression analysis suggested that:

- Risk adjusted perinatal mortality was lower in hospitals with higher specialist to generalist physician ratios and hospitals that routinely measure and recorded infants' Apgar scores.
- Rates were found to relate to annual volume of deliveries in a U-shaped relationship, declining with increases in volume up to 2,859 annual births, and then increasing with greater volumes.

Thomas (1991) examined relationships between hospital mortality rates and quality of care by correlating providers' condition-specific observed to expected mortality ratio with their condition-specific quality problem rates (the percentage of reviewed cases at each hospital identified as involving one or more quality problems) for 50,000 cases in ten clinical conditions in 42 hospitals.

- Earlier population level analyses had indicated valid relationships between mortality risks and quality of care for five of ten clinical conditions studied: pneumonia, cardiac surgery, AMI, cardiac arrhythmia, and femur/pelvic fractures.
- Empirical investigation of validity for these five conditions, found significant ($p < 0.05$) positive correlations, suggestive of valid hospital level relationships between mortality rate and process measures, for cardiac surgery ($r=0.55$), AMI ($r=0.57$), and pneumonia ($r=0.31$), but not cardiac arrhythmia and femur/pelvic fractures.
- Thomas (1998) suggests that the partially negative findings in this study could be due to the reliability of Peer Review Organisation peer review judgements, shown in some studies to be problematic.

Hartz (1993) examined relationships between peer review judged quality of care problems and hospitals' risk adjusted mortality rates for 4,132 hospitals in 38 states, covering two million Medicare admissions.

- They found significant ($p < 0.05$) positive correlations ($r=0.19$) between hospital quality problem rates and risk adjusted mortality rates within 14 states, but not in the other 28 states studied (correlations were either non-significant or negative).

- Thomas (1998) suggests that the partially negative findings in this study could be due to the reliability of Peer Review Organisation peer review judgements, shown in some studies to be problematic.

In the case of the five studies reviewed by Thomas (1998) that examined predictive validity, the ability of a hospital mortality rate measure to predict hospital quality performance was assessed. A measure with high predictive validity should perform well in predicting the occurrence of an attribute e.g. good quality care. In these five studies, sets of superior quality hospitals and poor quality ones were identified on the basis of risk adjusted mortality rates and then these quality predictions were evaluated by comparing processes of care between the two types of outliers.

Knaus (1986) performed clinical management audits of 13 hospital intensive care units (ICU) after ranking them on the basis of risk adjusted mortality rates. Then comparisons were made between low outlier ICUs and high outlier ICUs on quality factors determined through clinical management audit.

- The ICU having the lowest risk adjusted mortality rate (O/E =0.59) was found to have several characteristics associated with good quality – for example, a full-time ICU director and 24-hour in-unit physician coverage.
- The ICU with the highest O/E ratio (1.58) had poor staff coverage and had very poor communication between nursing staff and physicians.
- Thomas (1998) suggests that although this study might be considered to support the validity of risk adjusted mortality rate as a measure of quality, it represents only weak evidence, both because of the small number of hospitals included in the study and because of the post-hoc nature of the clinical management audits.

With data from a chain of 93 proprietary hospitals, *DuBois (1987)* developed a hospital level model to predict the number of in-hospital deaths as a function of four case-mix indicators. Comparing the hospitals' observed mortality rates with rates predicted by the model, this study identified nine low outliers (presumed good quality) and 11 high outliers (presumed poor quality). Samples of medical records for stroke, AMI and pneumonia obtained from six hospitals in each of the groups were audited for quality using condition-specific criteria and also using physicians' implicit judgements about preventability of death. The conclusions were:

- Based on compliance with condition-specific criteria, results showed no differences between high and low outlier hospitals for any of the conditions studied.
- When quality of care was judged implicitly, no differences were observed between low and high outlier hospitals for stroke and AMI.
- For pneumonia patients, preventable death rates were judged to be higher in the group of high outlier hospitals than in the group of low outlier hospitals when quality of care was judged implicitly; however, low levels of inter-rater reliability existed for physicians performing implicit reviews: kappa scores were as low as 0.11 for pneumonia cases reviewed.
- Thomas (1998) points out that subsequent research has shown that hospital level risk adjustment models can produce severely biased

predictions, and that mortality risk adjustments instead must be based on patient level models.

Park (1990) sought to determine whether hospitals identified as high mortality rate outliers were actually delivering poorer quality care than other hospitals. Mean explicit criteria quality scores (degree of compliance with process criteria) were compared between high outlier and non-outlier hospitals for 1,126 congestive heart failure (CHF) and 1,150 AMI cases from hospitals in four states. The initial risk model used only adjusted for age, sex and ethnicity at the patient level. Key findings were:

- With demographic risk adjustment models, no differences were found in quality scores by outlier status for either CHF or AMI; however, for both CHF or AMI patients, mean quality scores in high outlier hospitals were as high or higher than those in non-outlier hospitals.
- 80% of differences in adjusted mortality rates between high outlier and non-outlier hospitals were not attributable to quality but to residual severity differences and random variation in the outcome.
- Although death rates in high outlier hospitals were 5.0 to 10.9 higher per 100 admissions than in non-outlier hospitals, 56% to 82% of the excess could have resulted from purely random variation.

For a subsequent set of analyses with better risk models, it was found that:

- the average quality of care score for CHF cases treated in high outlier hospitals was poorer than the average in low outlier hospitals
- For AMI patients, quality score differences between groups of high outlier and low outlier hospitals were not significant, but patterns were considered suggestive of better quality in low outlier hospitals.

Hannan (1990) used data from a cardiac surgery registry to model patients' CABG mortality risks. For the 28 hospital programs evaluated, crude CABG mortality rates ranged from 2.2% to 14.3%; three programs were found to have significantly fewer deaths than expected and four were found to have significantly higher mortality rates than expected. Mean generic process of care criteria quality scores were compared for 63 CABG deaths sampled from three low outlier and four high outlier hospitals. Reviewers found that 45% of deaths in high outlier hospitals had associated quality of care problems, while only 4.4% (one case) experienced questionable quality of care in the low mortality hospitals.

Researchers from the *Californian Office of Statewide Health Planning and Development [COSHPD] (1996)* stratified 228 hospitals on the basis of risk adjusted AMI mortality rates into three categories – better than expected mortality, expected mortality and worse than expected mortality – and selected ten hospitals from each stratum. About 325 AMI admissions were sampled from each stratum of hospitals (1,005 cases in total) and medical records were reviewed to assess the degree of compliance with hypothesised good processes for the treatment of AMI. The key findings were:

- Although percentages of patients receiving aspirin therapy did not differ among strata, patients in low mortality hospitals were found more likely than those in other hospitals to receive aspirin therapy within six hours of presentation; they were also more likely to receive

heparin, and to undergo revascularisation, coronary angiography, and pulmonary artery catheterisation.
- Patients in high mortality hospitals were more likely to receive beta-blockers than those in other hospitals.

These findings were considered to support the validity of risk adjusted AMI mortality rate as a measure of hospital quality performance.

Three of the studies reviewed by Thomas (1998) looked for evidence of valid variance (the portion of total variance that is related to the concept being studied e.g. quality differences among hospitals) by examining population level relationships between mortality risks and quality of care.

The study by *Kahn (1990)* was highlighted by Thomas (1998) as one of the most carefully designed investigations available in the literature of relationships between patients' mortality risks and quality of medical care processes. In that study risk models for 30 day mortality were developed for each of five medical conditions using physiologic and other variables abstracted from patients' medical records. Quality reviews were performed on 14,000 cases using explicit, disease-specific process of care criteria, and on a sub-sample (n=1,197) of these cases using a structured implicit review methodology. On the basis of explicit quality review the following results were obtained:

- CHF patients judged as having received poor quality care were found to be 1.74 times more likely to die within 30 days than patients whose care was acceptable.
- AMI patients with poor care were 25% more likely to die than other AMI patients.
- For stroke and pneumonia cases, the mortality ORs were both 1.36.
- Only for patients with hip fractures were the authors unable to show a significant relationship between quality of care and mortality risk (OR= 0.90).

When *Rubenstein (1990)* carried out quality of care assessments based on implicit review rather than on explicit criteria on a sample of 1,197 records from the study by Kahn et al (1990), the evidence of a significant relationship between mortality risk and quality of care was even stronger. Cases judged as having received poor quality care were found 2.08 times more likely to die within 30 days of hospitalisation than cases whose care was considered acceptable. In summary, this study provides strong empirical evidence that poor quality care increases patients' risks of death.

Thomas (1991) related Medicare peer review quality judgements to mortality risks for 50,000 cases in ten clinical conditions in 42 hospitals. First, condition-specific models for predicting patients' in-hospital mortality risks were developed using administrative data covering all hospital admissions in the Twin Cities, Minnesota during 1987 and 1988. These models were then applied to estimate the mortality risks for nearly 50,000 Medicare admissions that had been reviewed for quality of care. The authors then compared the O/E ratio for patients whose care had been judged of poor quality with the O/E ratio for all patients whose care had been judged as acceptable. The results were:

- Analyses revealed significant relationships between quality and mortality risk for patients admitted for cardiac surgery, for AMI, or for cardiac arrhythmia.
- Somewhat weaker but still significant relationships were noted between quality of care and mortality risks for pneumonia patients and femur/pelvic fracture patients.
- However, O/E ratios were found not to differ as a function of measured quality for stroke, heart failure, bowel procedure, prostatic disorder, and septicaemia cases.
- Findings were consistent with those reported by Kahn (1990) for the same conditions.

In addition to the 11 studies outlined above, Thomas (1998) reviewed several other articles which provide evidence on systematic-error variance associated with inadequate risk adjustment when using mortality rates to evaluate hospital quality performance, or on random error variance. They also reviewed two validation studies that infer validity from observed behavioural responses to public release of hospital mortality rate data.

From their in-depth review, *Thomas (1998)* concludes that:

- Poor quality care increases patients' risks of mortality (a range of conditions were examined in the studies reviewed).
- On average, quality of care provided in hospitals identified as high mortality rate outliers is poorer than that provided in low mortality rate outlier hospitals.
- However, risk adjusted mortality rates are seriously inaccurate when used as a measure of quality for individual hospitals.
- Publication of hospital mortality rates may seriously misinform the public about hospital quality.

General population of patients: other studies

Other studies done on a general population of patients that have been reviewed are:

- *Mitchell (1997)* reviewed 81 research papers examining relations among organisational structures or processes in care delivery systems and morbidity and mortality outcomes.
- *Best (1994)* determined whether medical records of 111 patients who died in Veterans Affairs (VA) hospitals with high overall ratios of observed to expected mortality rates show greater evidence that life might have been meaningfully prolonged with more expert care than the records of matched patients who died in VA hospitals with low O/E ratios.

The main conclusions from the review by *Mitchell (1997)* were:

- Most research relating mortality and other adverse outcomes to organisational variables has been conducted in acute care hospitals since 1990, with these outcome indicators linked more frequently to organisational structures than to organisational or clinical processes.

- There is support in some studies, but not in others, that the following factors distinguish hospitals with lower rates of mortality and complications from those with higher rates:
    - nursing surveillance
    - quality of working environment
    - quality of interaction with other professionals.
- Risk adjustment methods suggests that variations in mortality and complications are influenced by patient variables more than by organisational variables.
- Adverse events may be a more sensitive marker than mortality of differences in organisational quality in acute care hospitals and long term care.
- As severity adjustment methods become more refined, many of the small differences currently seen in mortality and complications may disappear.
- Overall, the acute care studies are not conclusive regarding the extent to which the organisational features of care delivery systems positively influence outcomes such as mortality.
- Researchers need to refine and better define adverse events that are logically related to the co-ordinative organisational processes among caregivers.
- Researchers need to go much beyond mortality, morbidity, and adverse events in evaluating the linkage between the organisation of care and outcomes.

In the study by *Best (1994)*, patient matching criteria included age, diagnosis responsible for length of stay, type of VA hospital, and the mortality probability predicted by logistic regression. Expert physicians blindly reviewed pairs of medical records, assessing comparative care on a symmetrical, nine alternative visual analogue scale. The following results emerged:

- A slight shift in distribution toward better care in low ratio hospitals was not statistically significant.
- Results of an additional analysis, not dependent on pairing, showed that preventability of death was more strongly related to the following factors than to hospital O/E mortality ratios:
    - physicians' estimates of mortality risk at admission
    - whether the patient had been transferred from a nursing home
    - do not resuscitate status
    - accuracy of discharge coding.

Patients with cardiovascular disease

The studies done on patients with cardiovascular disease that have been reviewed are:

- *Thomas (1993)* used a computerised discharge database to model mortality risks for patients hospitalised in 1987 for three different conditions: cardiac disease (ischaemic heart disease, coronary artery disease, angina, and left ventricular aneurysm), AMI, and septicaemia. A database of peer review quality findings was then used to determine whether the ratio of observed to expected deaths in each of these conditions related validly to quality.

- *Carlisle (1992)* compared the quality of process of care and mortality among patients (aged 65 years and older) hospitalised with AMI from three health maintenance organisations (HMOs) with those among a fee for service (FFS) national sample of AMI patients.
- *Scott (2000)* evaluated the effects of a quality improvement intervention in a before-after time series study of 649 consecutive patients suffering AMI in a health district between March 1996 and August 1998.
- *Allison (2000)* examined the association of hospital teaching status with quality of care and mortality for 114,411 Medicare patients from 4,361 hospitals (22,354 patients from 439 major teaching hospitals, 22,493 patients from 455 minor teaching hospitals, and 69,564 patients from 3,467 non-teaching hospitals) who had AMI between February 1994 and July 1995.
- *O'Connor (1998)* performed a regional prospective study of 8,641 consecutive patients undergoing isolated CABG by all of the 23 cardiothoracic surgeons practising in northern New England during the study period.

Thomas (1993) sought to develop separate mortality risk models for ten different conditions including cerebrovascular disorders, pneumonia, cardiac disease, AMI, heart failure and shock, arrhythmia and conduction disorders, bowel procedures, femur/hip fractures, prostatic disorders, and septicaemia. Only the results for three conditions (cardiac disease, AMI, and septicaemia) were reported. In assessing the validity of each of the mortality models, the study focused sequentially on three questions:

- Does the model yield similar mortality risk estimates when applied to a different data set than that used for model estimation?
- At the patient level, using the model to control for differences in risk, are cases that suffer poor quality care more likely to die than cases that receive care of acceptable quality?
- At the provider level, are quality inferences based on risk adjusted mortality rates likely to be consistent with peer review judgements about quality?

The main findings were:

- Overall, the results provided strong support for the validity of one of the mortality indicators, weak support for another, and no support for the third.
- Using the cardiac disease mortality model, cases receiving care of acceptable quality died at approximately the rate predicted by the model, while cases receiving poor quality care died at more than twice the expected rate. This, concludes the authors, provides strong support for the validity of risk adjusted mortality rate as an indicator of provider quality of care for cardiac disease.
- For the AMI disease group, the model did not allow inferences to be made about quality on the basis of risk adjusted mortality.
- For the septicaemia condition group, using the authors' model, risk adjusted mortality was not an indicator of quality of care.

- Results reported for these three conditions are representative of the range of findings obtained for the seven other disease groups studied.
- The strength of relationships between quality of care and risk adjusted mortality in these analyses was not related to the proportion of outcome variance explained by risk models
- It was concluded that before inferences about hospital quality are made using any risk adjusted mortality indicator, the validity of the quality-outcome relationship must be established explicitly for that particular measure and risk model.

Carlisle (1992), in a comparison of the quality of process of care and mortality among older patients hospitalised with AMI from three HMOs with those among a fee for service (FFS) sample of AMI patients, found that:

- After adjustment for sickness at admission, there were no significant mortality differences between the HMO and FFS groups at either 30 (23.2% vs. 23.5%) or 180 days (34.4% vs. 34.5%) after admission.
- Compliance with process criteria was higher for the HMO group as a whole ($p < 0.05$). The HMOs had greater compliance with three of five scales measuring different aspects of care for patients with AMI.

In the study by *Scott (2000)*, after a six month baseline period, clinical practice guidelines were issued to providers followed by sequential feedback of clinical indicator data over a one year period. Resultant changes in practice were then evaluated during a 12 month post-intervention period. The findings were:

- The proportion of eligible patients receiving early thrombolysis, lipid-lowering drugs and cardiac rehabilitation increased, respectively, from 30.8 to 70.0% ($p = 0.001$), from 23.4 to 56.4% ($p = 0.003$), and from 23.6 to 54.3% ($p = 0.003$).
- In-hospital death rate decreased from 15.8 to 8.6% ($p = 0.02$).
- Incidence of post-infarct angina decreased from 30.1 to 14.3% ($p < 0.001$).
- Mean length of stay decreased from 7.4 to 6.3 days ($p = 0.001$).
- Despite the absence of control groups, the present study suggests that clinical guidelines combined with feedback of clinical indicators are useful in improving quality of care.

Allison (2000) reported that:

- Among major teaching, minor teaching, and non-teaching hospitals, respectively, for patients meeting strict inclusion criteria, administration rates for:
  - aspirin during hospitalisation were 91.2%, 86.4%, and 81.4% ($p < 0.001$)
  - angiotensin-converting enzyme inhibitors at discharge were 63.7%, 60.0%, and 58.0% ($p < 0.001$)
  - beta-blockers at discharge were 48.8%, 40.3%, and 36.4% ($p < 0.001$)
  - reperfusion therapy at admission were 55.5%, 58.9%, and 55.2% ($p = 0.29$).

- Differences in unadjusted 30 day, 60 day, 90 day, and two year mortality rates among hospitals were significant at $p < 0.001$ for all time periods, with a gradient of increasing mortality from major teaching to minor teaching to non-teaching hospitals.
- Mortality differences were attenuated by adjustment for patient characteristics and were almost eliminated by additional adjustment for receipt of therapy.
- In conclusion, in this study of elderly patients with AMI, admission to a teaching hospital was associated with better quality of care based on three of four quality indicators and lower mortality.

O'Connor (1998) ranked surgeons according to risk adjusted mortality rates and grouped them into tertiles. Findings were:

- The mortality rate was 3.3% in the lowest surgeon mortality tertile and 5.8% in the highest tertile.
- Fatal heart failure accounted for 80.0% of the difference in aggregate mortality rates, ranging from 1.9% in lowest surgeon mortality tertile to 4.0% in the highest tertile ($p < 0.001$).
- Rates of other causes did not differ significantly across surgeon mortality tertiles.
- Differences in rates of fatal heart failure could not be explained by differences in pre-operative left ventricular dysfunction or other patient characteristics.
- In conclusion, in this study most of the difference in observed mortality rates between cardiothoracic surgeons was attributable to differences in rates of heart failure.

Patients with other medical conditions

The studies done on patients with other medical conditions that have been reviewed are:

- *Hayward (1993)* evaluated six potential generic screens (including in-hospital death) for poor quality in-patient care by reviewing 675 general medicine admissions at a university teaching hospital.
- *McGloin (1999)* undertook a six month hospital audit to determine whether any unexpected deaths occurring on general wards were potentially avoidable, and to assess whether the quality of care on general wards prior to unexpected ward death or admission to intensive care affected subsequent outcome.

Hayward (1993) reviewed a stratified random sample of 675 general medicine admissions to evaluate the usefulness in-hospital death and 28 day re-admission as indicators of poor quality in-patient care. The main findings were:

- Patients who died in hospital were substantially more likely than those who were discharged alive to be rated as having had sub-standard care (30% vs. 10%, $p < 0.001$).
- In contrast, cases who had subsequent early re-admissions did not have poorer quality ratings.

McGloin (1999) reviewed the medical notes of patients either dying unexpectedly (i.e. not having a prior 'do not resuscitate' order) or requiring intensive care unit (ICU) admission. Findings included:

- In the six month study period, 317 of the 477 hospital deaths occurred on the general wards, of which 20 (6%) followed failed attempts at resuscitation.
- Thirteen of these unexpected deaths were considered potentially avoidable: gradual deterioration was observed in physiological and/or biochemical variables, but appropriate action was not taken.
- In the same period, 86 hospital in-patients were admitted on 98 occasions to the ICU, 31 of whom received sub-optimal care pre-ICU admission due either to non-recognition of (the severity of) the problem or to inappropriate treatment. Both ICU (52% vs. 35%) and hospital (65% vs. 42%) mortality was significantly higher in these patients compared to well managed patients ($p < 0.0001$).
- Patients with obvious clinical indicators of acute deterioration can be overlooked or poorly managed on the ward which may lead to potentially avoidable unexpected deaths or to a poorer eventual outcome following ICU admission.

Paediatric patients

The studies done on paediatric patients that have been reviewed are:

- *Salinas (1997)* assessed and quantified the effect of quality of care on death preventability, independent of social and biological variables by comparing 181 avoidable perinatal deaths (cases) to 341 non-avoidable ones (controls).
- *Esposito (1999)* reports on the results of a retrospective multidisciplinary consensus panel review of 138 deaths attributed to mechanical trauma in children (aged 18 years or less) occurring in the rural state of Montana between 1989 and 1992 to determine the rates of preventable mortality and inappropriate care (appropriateness was defined according to nationally accepted guidelines), as well as the nature of treatment.

In the study by *Salinas (1997)*, judgement criteria on death preventability were based predominantly on compliance with explicit hospital medical care standards, determined by peer review. The main findings were:

- The overall perinatal mortality rate was 24.8 per 1000 births and could be reduced by 35% if all avoidable perinatal deaths were prevented.
- Sixteen per cent of the deaths presented structural and 31.2% process deficiencies; both predominated among avoidable perinatal deaths (3.5.4% vs. 5.3%, $p < 0.000$; and 79.3% vs. 5.9%, $p < 0.000$, respectively).
- Structural deficiencies increased the risk of an avoidable perinatal death 11-fold (95% CI 4.1-26.9, $p < 0.001$) and process deficiencies 88-fold (95% CI 37.2-204.5, $p < 0.001$), after controlling for confounders.
- The strength of the association between quality of care and preventable perinatal mortality was estimated.

Esposito (1999) reported that:

- Less than 1% (n=1) of deaths were judged frankly preventable.
- 11 deaths (8%) were judged possibly preventable, giving a total preventability rate of 9% for all cases reviewed.
- Considering only in-hospital deaths (n=77), the total preventability rate was 16%.
- The rate of inappropriate care rendered for all deaths, regardless of preventability, was 36% (16% in the pre-hospital phase and 74% for in-hospital deaths).
- Inappropriate care for the paediatric group was more prevalent in patients less than or equal to 14 years old.
- The nature of inappropriate care was most frequently associated with the management of respiratory problems, including airway control and management of chest trauma.

Surgical patients

The studies done on surgical patients that have been reviewed are:

- *Silber (1997a)* assessed whether the correlation between two outcome measures (hospital mortality and complication rates after surgery) which are intended to measure the same underlying quantity (hospital quality of care), increases as data quality improves.
- *Pronovost (1999)* undertook an observational study of patients discharged from 39 non-federal acute care hospitals in Maryland with a principal procedure code for abdominal aortic surgery between January 1994 and December 1996 (n=2,987) to determine whether organisational characteristics of intensive care units (ICUs) are related to clinical outcomes for such patients.
- *Silber (1992)* investigated whether factors that predict overall mortality following two common surgical procedures are different from those that predict adverse occurrences (complications) during the hospitalisation or death after an adverse occurrence by examining data on 5,972 Medicare patients undergoing elective cholecystectomy or transurethral prostatectomy.
- *Daley (1997a)* determined the validity of risk adjusted surgical mortality rates as measures of quality of care by assessing the process and structure of care in 20 surgical services (in Veterans Affairs Medical Centres) with higher than expected and lower than expected risk adjusted 30 day mortality rates.

Silber (1997a) used three predictive models varying in adjustment for severity of illness. Findings and conclusions were:

- Two hospital rankings, based on each of the two outcomes, were well correlated when not adjusted for severity; however, as clinical data were added to the models, the correlation tended to disappear.
- Before claims of construct validity can be made, investigators must show that correlations between outcomes purporting to measure quality of care are sustained after adequate correction for severity.
- Most importantly, it should be recognised that inadequately controlled confounding variables may lead to a spurious high correlation between an accepted and a new outcome measure, and a false sense of adequate construct validity.

Pronovost (1999) found that:

- In-hospital mortality varied among hospitals from 0% to 66% for patients undergoing abdominal aortic surgery.
- In multivariate analysis adjusted for patient demographics, co-morbid disease, severity of illness, hospital and surgeon volume, and hospital characteristics, not having daily rounds by an ICU physician was associated with a threefold increase in in-hospital mortality (OR =3.0; 95% CI 1.9-4.9).
- Not having daily rounds by an ICU physician was also associated with an increased risk of:
    - cardiac arrest (OR =2.9; 95% CI 1.2-7.0)
    - acute renal failure (OR =2.2; 95% CI 1.3-3.9)
    - septicaemia (OR =1.8; 95% CI 1.2-2.6)
    - platelet transfusion (OR =6.4; 95% CI 3.2-12.4)
    - re-intubation (OR =2.0; 95% CI 1.0-4.1).

Silber (1992) reported that:

- The death rate was associated with both hospital and patient characteristics.
- The rate of adverse occurrences was associated primarily with patient characteristics.
- The failure to rescue rate (number of deaths in patients who developed an adverse occurrence/number of patients with an adverse occurrence) was associated more with hospital characteristics, and was less influenced by patient admission severity of illness.
- Therefore, factors associated with hospital failure to rescue are different from factors associated with adverse occurrences or death.

In the study by *Daley (1997a)*, several dimensions of process and structure of care were assessed including technology and equipment, technical competence of staff, leadership, relationship with other services, monitoring of quality of care, co-ordination of work, relationship with affiliated institutions, and overall quality of care. It was found that:

- Surgical services with lower than expected risk adjusted surgical morbidity and mortality rates had significantly more equipment available in surgical intensive care units than did services with higher than expected outcomes (4.3 vs. 2.9, p <0.05).
- Ratings by site visitors of overall quality of care were significantly higher for surgical services with lower than expected morbidity and mortality rates (6.1 vs. 4.5 for high outliers, p <0.05). Technology and equipment were rated significantly better among low outlier services (7.1 vs. 4.8 for high outliers, p <0.001).
- Masked site visitor teams correctly predicted the outlier status (high versus low) of 17 of the 20 surgical services visited (p <0.001).

STUDIES TO ASSESS THE USE OF CASE FATALITY RATES IN COMPARING HOSPITAL PERFORMANCE

A number of studies have specifically assessed the use of case fatality rates as comparative indicators of hospital performance. These can be classified by patient group as those involving:

- general population of patients
- patients with cardiovascular disease
- patients with other medical conditions
- paediatric patients
- surgical patients.

General population of patients

The studies done on a general population of patients that have been reviewed are:

- *Thomas (1999)* examined whether risk adjusted mortality rates are valid indicators of hospital quality using an analytical model to estimate the sensitivity and predictive error of hospital mortality rate performance indicators.
- *Jarman (1999)* analysed eight million discharges from NHS hospitals over a four year period (where the primary diagnosis was one of the diagnoses accounting for 80% of in-patient deaths) to determine which factors best explain variation in standardised hospital in-patient mortality ratios.
- *Zalkind (1997)* used Monte Carlo simulation to examine the relationship between outlier status based on adjusted mortality rates and theoretical underlying quality of care in hospitals.
- *Flanders (1994)* used data for 169 acute care hospitals on 17 conditions and procedures to evaluate whether observed 30 day mortality rates differ significantly from model predicted 30 day mortality rates (adjusted for demographic characteristics, principal diagnosis, and co-morbidities) for situations when the number of hospital deaths is small.
- *McKee (1995)* examined methodological issues related to the publication of hospital mortality league tables, with particular reference to severity adjustment and sample size, by analysing data on admissions to 22 hospitals for a range of principal diagnoses (aortic aneurysm, carcinoma of the colon, cervical cancer, cholecystectomy, fractured neck of femur, head injury, ischaemic heart disease, and peptic ulcer). In-hospital mortality rates were adjusted by disease severity and calculated on the basis of both admissions and episodes.
- *Treurniet (1999)* studied the association between regional variations in 'avoidable' mortality and variations in disease incidence, and analysed whether the proportion of in-hospital deaths could explain the regional variations in incidence adjusted mortality for a selection of conditions.
- *Bond (1999)* evaluated associations among hospital characteristics, staffing levels of health care professionals, and mortality rates in 3,763 United States hospitals.
- *Ballard (1994)* evaluated the potential confounding effect of illness severity factors associated with referral selection and hospital mortality on comparisons of risk adjusted hospital mortality based on the Health Care Financing Administration (HCFA) Medicare hospital mortality model.
- *Sirio (1999)* performed a retrospective cohort study of 116,340 severely ill patients admitted between 1991 and 1995 to a community based sample of 28

hospitals to assess variations in severity adjusted hospital mortality among intensive care unit (ICU) patients across a major metropolitan region.

- *Goldman (1994)* reported on a study undertaken by the Department of Veterans Affairs (VA) in 1991 to measure hospital-specific mortality rates for VA facilities and to use these rates as a screening tool for focused medical record reviews.
- *Rosenthal (1998)* determined whether standardised hospital mortality rates varied for six common medical diagnoses that are likely to be managed by similar practitioners (AMI, congestive heart failure, pneumonia, stroke, obstructive lung disease, and gastrointestinal haemorrhage) in a retrospective cohort study of 89,851 patients (aged 18 years or over) discharged from 30 hospitals between 1991 and 1993.

In the *Thomas (1999)* study, the following six parameters were shown to determine accuracy of hospital mortality rate performance indicators:

- mortality risks of patients who receive good quality care
- mortality risks of patients who receive poor quality care
- proportion of patients (across all hospitals) who receive poor quality care
- proportion of hospitals considered to be 'poor quality'
- patients' relative risk of receiving poor quality care in 'good quality' and in 'poor quality' hospitals
- number of patients treated per hospital.

Key findings from analyses were:

- Using best available values for model parameters, analyses demonstrated that in nearly all situations, even with perfect risk adjustment, identifying poor quality hospitals on the basis of mortality rate performance is highly inaccurate.
- Of hospitals that delivered poor quality care, fewer than 12% were identified as high mortality rate outliers, and more than 60% of outliers were actually good quality hospitals.
- Under virtually all realistic assumptions for model parameter values, sensitivity was less than 20% and predictive error was greater than 50%.
- The authors conclude that the best candidate measures for reporting hospital performance are likely to be process based indicators such as the degree of compliance with explicit condition-specific criteria. In addition, hospitals find process based measures, unlike outcome statistics, 'actionable' i.e. helpful in determining how quality of care can be improved.

Other points to note in relation to this study are:

- Mortality following CABG surgery was identified as the single exception to the above findings. Using their Texas Foundation for Medical Care database, it was calculated that for CABG surgery, patients in the worst performing 10% of hospitals were 5.4 times as likely to receive poor quality care as patients in the other 90% of hospitals. Using these parameter values, they found that with the median hospital volume (n=530) in a CABG surgery mortality report for New York State, 80% of facilities identified as high mortality rate outliers were in fact poor quality hospitals.

- Key parameter values utilised in their analyses were derived from published research on Medicare patients hospitalised for serious medical conditions. The authors concede that for other conditions and types of patients, the mortality risks associated with good quality care and poor quality care might differ from values used in their analyses.
- The authors choice of 200 as the patient sample size in their analyses could be criticised as being unfairly small. However, this sample is not atypical for samples appearing in published hospital performance reports. Furthermore, although the accuracy of mortality rates in the identification of poor quality hospitals improves as sample sizes increase, this study indicates that even with sample sizes as large as 900 patients, one third of high outlier hospitals could be false positives.

The main findings from the study by *Jarman (1999)* were:

- The four year crude death rates varied across hospitals from 3.4% to 13.6% (average for England 8.5%).
- Standardised hospital mortality ratios ranged from 53 to 137 (average for England 100).
- The percentage of cases that were emergency admissions (60% of total hospital admissions) was the most powerful predictor of this variation in mortality, with the ratio of hospital doctors to beds and general practitioners to head of population the next best predictors.
- When analyses were restricted to emergency admissions (which covered 93% of all patient deaths analysed) number of doctors per bed was the best predictor.

Hennell (1999) has put forward a possible explanation for the observation by Jarman (1999) that ‘more doctors means fewer deaths’. For any given population the standardised admissions rate is positively correlated with the standardised death rate but is inversely correlated with the standardised hospital death rate (defined as any death within 30 days of a hospital admission). Where a population is admitted to hospital fairly frequently, a higher proportion of admissions will not be associated with subsequent death; hence there will be a lower apparent hospital mortality. Although Jarman (1999) included a calculation of the standardised admission ratio (though not separated into ‘all cases’ and ‘emergencies’), unfortunately, the figures used were the aggregates for the health authority of hospital location rather than individual figures for each hospital's emergency catchment area (which is often very different). Even in this form, standardised admissions entered significantly into both the all cases and the emergency multiple regression models. It is not impossible that a fully specified version of this variable might displace some of the doctor variables; the observed relation could simply be ‘More doctors means more admissions’.

Bunker (1999) has pointed out that another limitation of the study by Jarman (1999) is that the severity of the primary illness could not be estimated and that co-morbidity was limited to a count of sub-diagnoses for each patient in hospital episode statistics. The inadequacy of such routine data has been well documented. The dangers of using inadequate information has been demonstrated in earlier research which reported the counterintuitive finding that many sub-diagnoses, for example, adult onset diabetes mellitus, essential hypertension, previous AMI, and angina, were associated with lower death rates. Bunker emphasises that severity of illness and co-morbidity are best judged prospectively, preferably in consultation with physicians. The second best is retrospective review of individual patients' records comparing condition-specific diagnoses and procedures case by case, while the least satisfactory approach is to rely on routinely collected undifferentiated data.

Zalkind (1997) simulated mortality experience for a hypothetical set of hospitals, with perfect risk adjustment and with prior perfect knowledge of the identity of hospitals providing poor quality care. A key question addressed was could random variation noise obscure the signal of differences in underlying rates of quality of care problems in the absence of case-mix differences. Classification of hospitals as 'outliers' was done compared with 'true' hospital quality, based on underlying rates for quality of care problems in mortality cases. Key findings were:

- Predictive error rates with respect to 'quality' for both 'outlier' and 'non-outlier' hospitals are substantial under a variety of patient load and cut-off point choices for determining outlier status.
- With a volume of 200 patients per hospital, sensitivity averages 10% to 20% for identifying poor quality hospitals, while predictive error averages 60%.
- Using overall death rates as an indicator of underlying quality of care problems may lead to substantial predictive error rates, even when adjustment for case-mix is excellent.
- The authors conclude that outlier status should only be used as a screening tool and not as the information provided to the public to make informed choices about hospitals.

Flanders (1994) found that if the number of deaths in some hospitals is small, p-values calculated using the nominal chi-square distribution can be misleading. The Monte Carlo simulation is a more appropriate approach to the analysis of hospital mortality or small area analysis for situations in which the number of deaths is small.

The key findings from *McKee (1995)* were:

- The numbers of deaths from specific conditions were often small and the corresponding confidence intervals wide.
- Rankings of hospitals by death rate are sensitive to adjustment for severity of disease. There are some differences that cannot be explained using routine data.
- Comparison of crude death rates may be misleading.
- Some adjustment for differences in severity is possible, but current systems are unsatisfactory.
- Differences in death rates should be studied, but because of the scope for manipulating data, this should be undertaken in a collaborative rather than a confrontational way.
- Any decision to publish league tables of death rates will be on political rather than scientific grounds.

Treurniet (1999) used regression analysis to examine the relationship between mortality and incidence on the one hand, and between incidence adjusted mortality and in-hospital mortality on the other. The results were:

- Significant regional mortality variations were found for cervical cancer, cancer of the testis, hypertensive and cerebrovascular disease (CVA), influenza/pneumonia, cholecystitis/lithiasis, perinatal causes and congenital cardiovascular anomalies.
- Regional mortality differences in general were only partly accounted for by incidence variations; the only exception was cervical cancer, which no longer showed significant variations after adjustment for incidence.
- The contribution of in-hospital mortality variations to total cause-specific mortality variations varied between conditions: the highest percentage of explained variance was found for mortality from CVA (60.1%) and appendicitis (29.2%).
- Incidence adjusted mortality rates are likely to be more sensitive for quality-of-care variations than the 'crude' mortality variations.

The key findings from the study by *Bond (1999)* were:

- Controlling for severity of illness, hospital characteristics associated with lower mortality were occupancy rate and private non-profit and private for profit ownership.
- Mortality rates decreased as staffing level per occupied bed increased for medical residents, registered nurses, registered pharmacists, medical technologists, and total hospital personnel.
- Mortality rates increased as staffing level per occupied bed increased for hospital administrators and licensed practical vocational nurses.

Ballard (1994) analysed data from the 1988 Medicare hospitalisation data file on patients treated at nine hospitals and evaluated the relationship between distance from patient residence to the admitting hospital and risk adjusted hospital mortality. The following findings emerged:

- Hospital mortality for referral patients was substantially lower than predicted (5.0% vs. 7.5%, $p < 0.001$).
- Controlling for other factors, distance from patient residence to the hospitals was independently associated with mortality among each of the hospitals examined.
- The authors warn that the HCFA Medicare hospital mortality model should be used with extreme caution to evaluate hospital quality of care for national referral centres because of residual confounding due to severity of illness factors associated with geographic referral.

Findings from *Sirio (1999)* include:

- Hospital standardised mortality ratios (SMR) (actual/predicted mortality) ranged from 0.85 to 1.21 (less than a 1.5-fold difference), and tended to be somewhat lower in the major teaching hospitals.
- The considerable overlap in the 99% confidence intervals among the various hospitals, implies that it is difficult to distinguish the best

hospitals from the average hospitals as well as the average from the less than average hospitals, on the basis of mortality experience.

The authors highlight some methodologic limitations of the study:

- Analysis was limited to a single outcome measure whereas the quality of care encompasses multiple dimensions, including the appropriateness and process of care, functional outcomes, patient satisfaction, and long term mortality.
- Factors not controlled for in this study and which may contribute to variations in mortality include functional status, mental health, and social support.
- No information was collected regarding the use of do not resuscitate orders, a practice that may differ across institutions.

Goldman (1994) compared 1986 mortality rates for each VA medical centre with aggregated VA data, adjusting for patient characteristics related to mortality, and conducted medical record reviews of cases for hospitals with elevated levels of mortality. It was reported that:

- The mortality rate analysis identified 44 facilities with rates that were significantly elevated overall or within specific patient categories.
- Both the mortality rate analysis and the medical record review suggested that quality concerns were more common in predominantly psychiatric medical centres than in other facilities.
- The findings of this study suggest that there is value in performing intensive review of outlier providers in outcome analyses.
- However, serious questions remain concerning the validity of the use of administrative databases to assess hospital mortality.

Rosenthal (1998) categorised hospitals into quintiles on the basis of standardised mortality ratios (SMRs). Correlations between SMRs and agreement between quintile rankings were determined for each pair of diagnoses. The key findings were:

- Correlations between hospital SMRs for individual diagnoses were generally weak. For the 15 possible pairs of diagnoses, Pearson coefficients ranged from -0.10 to 0.43; only six were 0.30 or greater.
- Agreement between hospital quintile rankings was also generally low, with weighted kappa values ranging from -0.12 to 0.42. Three of 15 kappa values were less than 0 (i.e. agreement lower than chance), and only four exceeded 0.20, the threshold for ‘fair’ agreement.
- Although simulated analyses found that random variation and relatively low hospital volumes accounted for some of the difference in standardised mortality ratios for diagnoses, a large proportion of the difference remained unexplained.
- Although variability may be decreased by restricting analyses to hospitals with large volumes, the findings indicate that for many hospitals, diagnosis-specific mortality rates may be an inconsistent measure of hospital quality, even when data are aggregated for multiple years.

Patients with cardiovascular disease

The studies done on patients with cardiovascular disease that have been reviewed are:

- *Leyland (1998)* used multilevel modelling to examine the effects of patient characteristics, hospital, and area of residence, on all cause mortality within 30 days of discharge after hospital admission for AMI and on death from AMI at any time during the study period.
- *Iezzoni (1996a)* compared expected versus observed AMI death rates for 100 acute care hospitals using data on 11,880 adults admitted for AMI in 1991, and compared the effects of using ten different severity measures to adjust for patient severity of illness.
- *Ghali (1998)* examined observed and risk adjusted death rates for 23 Canadian hospitals performing an overall total of 50,357 CABG procedures between 1992 and 1996, and also examined changes over time in hospital-specific death rates.
- *Marshall (1998)* described a new analytic strategy of examining risk adjusted mortality outcome indicators with time to overcome the problems of their instability and their inability to provide reliable comparisons of small volume hospitals, and applied this technique to data on 24,029 patients undergoing coronary artery bypass surgery.
- *Hartz (1992)* presents a method to compare quality of care between providers of revascularization (CABG and angioplasty) procedures based on complication rates adjusted for patient characteristics using detailed clinical data.
- *Poloniecki (1998)* carried out an observational study of 3,983 patients (aged 16 years or more) undergoing open heart surgery in a regional cardiothoracic unit between January 1992 and August 1995, to detect changes in 30 day in-hospital mortality using a cumulative risk adjusted mortality chart.
- *Iezzoni (1995a)* compared measured mortality risks from 11 widely used severity systems for 9,407 medically treated stroke patients from 94 hospitals, with 916 (9.7%) in-hospital deaths.
- *Iezzoni (1997b)* examined the use of risk adjusted mortality comparisons to assess hospital performance in a study comparing commercial severity measures applied to data on patients admitted for AMI, CABG surgery, pneumonia, or stroke.
- *Landon (1996)* applied 14 common severity measures to data on 7,765 patients undergoing CABG surgery at 38 hospitals to determine whether judgements about risk adjusted hospital death rates are sensitive to the specific severity method.
- *Luft (1993)* used data from hospital discharge abstracts for 132,750 patients (aged 18 years or over) undergoing CABG in 115 California hospitals between 1983 and 1989 to assess whether differences in risk adjusted mortality rates for hospitals primarily reflect chance variation.
- *Localio (1997)*, using the public reports of the Pennsylvania Health Care Cost Containment Council on CABG surgery for 1990 to 1992 as a case study, assessed the sensitivity of results to the choice of data and statistical methodology.
- *Hannan (1997b)* undertook a study to determine whether performing CABG surgery on high risk patients adversely affects the risk adjusted mortality rates for patients of surgeons and hospitals in New York State compared with the impact of performing surgery on more routine patients.
- *Hannan (1994)* examined changes in the risk adjusted mortality associated with CABG procedures performed in New York State during the first four years of New York's Cardiac Surgery Reporting System (1989 to 1992).

- *Chen (1999a)* used data from the Cooperative Cardiovascular Project on 149,177 elderly Medicare patients with AMI in 1994-1995 to examine the care and outcomes of patients admitted to three types of hospitals: those ranked high in cardiology (top ranked hospitals) in the influential annual published list of 'America's Best Hospitals'; hospitals not in the top rank that had on-site facilities for cardiac catheterization, coronary angioplasty, and bypass surgery (similarly equipped hospitals); and the remaining hospitals (non-similarly equipped hospitals).
- *Chen (1999b)* assessed whether Medicare patients with AMI admitted to one of HCIA-Mercer's '100 top hospitals' received better care or had better outcomes than patients treated in other hospitals.

Leyland (1998) analysed Scottish Morbidity Records from December 1992 to November 1993 for 14,359 episodes of AMI, the death records of those who died, and 9,391 death records for individuals who died after AMI but who had not been in hospital in the 30 days before death. Key findings were:

- They found significant differences between hospitals by age, sex, and medical history. The ORs for death ranged from 0.62 (95% CI 0.50-0.80) to 1.28 (1.07-1.59), relative to the average performance for Scotland as a whole.
- Analysis including area of residence, deaths occurring out of hospital, and more detailed information about patients showed no significant differences between hospitals for patients aged 70 years.
- The random variability of deaths from AMI before admission to hospital between areas was much greater than that for deaths in hospital or within 30 days of discharge, but the two were strongly correlated, so that patients from areas with high death rates before admission also tended to have higher rates of death in hospital or within 30 days of discharge.
- For death after admission, it was estimated that the relative contribution of area effects and hospital effects above that expected by chance, was 41% and 59% of the observed variance, respectively.
- The authors found clear influences of area effect, but these are complex: rates of AMI vary greatly and only 26% of deaths occur in hospital or shortly after discharge. Therefore the proportions of patients reaching hospital and those at risk of death will also vary and simple variables such as distance from the hospital and admission criteria become relevant.

The following results emerged from the study by *Iezzoni (1996a)*:

- Unadjusted mortality rates ranged from 5–26% for the 100 hospitals.
- Severity measure agreement ranged from only fair to good on which hospitals had the highest (more deaths than expected) or lowest (fewer deaths than expected) mortality rates.
- Severity measures based on medical records frequently disagreed with measures based on discharge abstracts.
- Thirty two hospitals had significantly better or worse death rates than expected using one or more, but not all ten, severity measures.

- Thus, whether individual hospitals were identified as especially good or poor frequently depended on the particular severity measure used.
- These findings raise important questions for report card efforts to judge hospital performance by using severity adjusted death rates.

Findings from *Ghali (1998)* included the following:

- The overall death rate following CABG was 3.6%.
- Inter-hospital comparisons showed that average severity of illness varied considerably across hospitals.
- Despite risk adjustment accounting for this variable severity, there was considerable variation in adjusted death rates across the 23 hospitals, from 1.95% to 5.76% ($p < 0.001$).
- For some hospitals, death rates decreased between 1992/93 and 1995/96, whereas for others the rates were stable or increased.
- There may be differences in quality of care across hospitals.

Marshall (1998) applied a new analytic strategy for examining mortality indicators (Time Series Monitors of Outcome) to data on 24,029 patients undergoing CABG between April 1987 and September 1992 in forty three Department of Veterans Affairs centres. The ratio of observed to expected operative mortality (adjusted for patient-specific risk factors) was calculated for each hospital for each of the 11 six month periods. The results were:

- This method identified four high and one low outlier hospitals based on significant deviations from the overall mean and three upward and seven downward trending outlier hospitals based on significant deviations in trend with time.
- This method should help reduce misclassification of outliers due to random variation in outcomes as well as provide more reliable comparative information from which to evaluate hospital performance.

Findings from *Hartz (1992)* include:

- When sample sizes are limited, adverse outcome rates may be a more sensitive measure of quality of care than mortality rates.
- Complication rates differed significantly and substantially among the hospitals.
- Clinical adjustment changed the hospital rankings for the bypass surgery hospitals, but not for the angioplasty hospitals.
- Adjustment for co-morbidities did not affect hospital rankings for either procedure.

Poloniecki (1998) examined the number of operations required for statistical power to detect a doubling of mortality, and determined the control limits (at level of $p = 0.01$) for the detection of an adverse trend. Key findings were:

- There was a highly significant variation in annual case-mix (Parsonnet scores 8.7-10.6, $p < 0.0001$).
- There was no significant variation in mortality after adjustment for case-mix (OR 1-1.5, $p = 0.18$) with monitoring by calendar year.
- With continuous monitoring, however, nominal 99% control limits based on 16 expected deaths were crossed on two occasions.

- Hospital league tables for mortality from heart surgery will be of limited value because year to year differences in death rate can be large (OR =1.5) even when the underlying risk or case-mix does not change.
- Statistical quality control of a single series with adjustment for case-mix is the only way to take into account recent performance when informing a patient of the risk of surgery at a particular hospital.
- If there is an increase in the number of deaths the chances of the next patient surviving surgery can be calculated from the last 16 deaths.

The key findings from the study by *Iezzoni (1995a)* were:

- Unadjusted hospital mortality rates ranged from 0 to 24.4%.
- For 27 hospitals, observed mortality rates differed significantly from expected rates when judged by one or more, but not all 11, severity methods.
- The agreement between pairs of severity methods for identifying the worst 10% or best 50% of hospitals was fair to good.
- In conclusion, efforts to evaluate hospital performance based on severity adjusted in-hospital death rates for stroke patients are likely to be sensitive to how severity is measured.

Iezzoni (1997b) used logistic regression to predict risk of in-hospital death using severity scores for patients admitted for AMI (n=11,880), CABG surgery (n=7,765), pneumonia (n=18,016), and stroke (n=9,407). Odds ratios for in-hospital death were compared across pairs of severity measures. The following results emerged:

- For AMI, the severity measure called Disease Staging performed best in discriminating between patients who lived and those who died (c statistic=0.86).
- The measure called All Patient Refined Diagnosis Related Groups was the best discriminator for CABG patients (c statistic=0.83).
- The measure, MedisGroups, had the highest c statistics for pneumonia (c=0.85) and stroke (c=0.87).
- Different severity measures predicted different probabilities of death for many patients.
- Severity measures frequently disagreed about which hospitals had particularly low or high z scores (which compared actual and expected death rates).
- Agreement in identifying low and high mortality hospitals between severity adjusted and unadjusted death rates was often better than agreement between different severity measures.
- Different severity measures frequently produce different impressions about relative hospital performance.
- Severity adjusted mortality rates alone are unlikely to isolate quality differences across hospitals.

Landon (1996) found that unadjusted death rates ranged from 0% to 11.2% across hospitals. Hospital mortality rate performance rankings were generally consistent among severity methods for 33 hospitals, but not for five others.

Luft (1993) examined observed and expected semi-annual mortality rates for each hospital to identify consistent patterns over time. Data on in-patient mortality was adjusted for age, gender, chronic co-morbidities, timing of surgery, and presence of additional procedures. Using the quartile of patients with the highest predicted risk (average mortality, 10%), high and low outlier hospitals were identified from two consecutive years of pooled data and outcomes two years later were examined. Each hospital year observation was also examined individually to identify outliers and to assess differences for observed and expected death rates, patient volume, and other characteristics. The main findings to emerge were:

- Some hospitals showed consistently lower than expected in-patient mortality.
- Some hospitals had periods of significantly higher than expected mortality followed by a correction.
- High outlier hospitals that were selected based on two years of data had mortality rates two years later that averaged 31% above expected, while low outlier hospitals had rates two years later that averaged 28% below expected.
- On a contemporaneous basis, high outliers had proportionately more transfers to other acute care hospitals and longer post-operative stays among survivors.
- Risk adjusted outcomes for CABG patients derived from administrative data exhibit substantial patterns of consistency.
- Such patterns cannot be detected for low risk patients but are evident for the top quartile of patients stratified by risk.
- Even with reporting lags and changes in hospital outcomes over time, a policy of channelling high risk patients away from high outlier hospitals and towards low outlier hospitals could lower the overall risk adjusted mortality rate by 54%.

Localio (1997) reported that:

- Statistical power calculations demonstrated that the annual volume of bypass surgery for many hospitals and for most surgeons is too small for meaningful mortality comparisons.
- The number of hospitals and physicians designated as mortality "outliers" in the Council's reports results in part from a failure to adjust critical P values for multiple comparisons. Hierarchical statistical models implemented by mixed effects logistic regression, by contrast, can detect true differences in performance without producing false outliers.
- Mortality analyses are sensitive to the choice of co-morbidities used for severity adjustment of a mortality model.
- Small area analyses indicate large differences in the rates of bypass surgery across Pennsylvania, with lower population based rates of surgery associated with higher population based in-patient mortality.
- Analyses of mortality by operative procedure, rather than by patient diagnosis, should consider the potential for selection bias caused by the decision to elect surgery.

- The clinical and statistical issues of operative mortality are sufficiently complex to merit review by independent experts before public release of hospital and physician performance measures.

Hannan (1997b) calculated risk adjusted mortality rates for 31 hospitals and 87 surgeons for high risk (a predicted mortality rate of at least 7.5%) and low risk patients during the time period 1990 to 1992. It was reported that:

- The risk adjusted mortality for all high risk patients was lower (2.94%) than the risk adjusted mortality for other patients (3.02%).
- Fifteen of the 31 hospitals had a lower risk adjusted mortality for all patients than they did for low risk patients only, and no differences in either direction were statistically significant.
- Forty-one of 87 surgeons (47%) had risk adjusted mortality for all patients that was at least as low as the risk adjusted mortality for low risk patients.
- In general, hospitals and surgeons with the lowest risk adjusted mortality for all cases also had the lowest risk adjusted mortality for high risk cases.
- The authors conclude that there is no systematic bias against operating on high risk CABG patients in the risk adjusted performance system in New York.

Hannan (1994) subdivided surgeons and hospitals into three groups on the basis of their risk adjusted mortality associated with CABG procedures in 1989. The risk adjusted mortality for each of the three groups was computed for 1992 and compared with their 1989 mortality. The results indicate that:

- All groups of providers exhibited large reductions in the risk adjusted mortalities, with the groups that showed the highest initial mortalities manifesting the most improvement.
- However, the group rankings remained the same in 1992 as they were in 1989. For example, when the hospital groups were based on the tertiles of risk adjusted mortality observed in 1989, the risk adjusted mortality decreased from 2.72% to 2.19% for group 1, from 4.24% to 2.51% for group 2, and from 7.12% to 2.77% for group 3.
- Notably, the risk adjusted mortalities of the three groups were all significantly different from one another in 1989, but were not significantly different from one another in 1992.
- The volume of operations performed by the various provider groups did not change substantially in the four year period.

The key results from the study by *Chen (1999a)* were:

- Admission to a top ranked hospital was associated with lower adjusted 30 day mortality (OR =0.87; 95% CI 0.76-1.00, p =0.05 for top ranked hospitals vs. the others).
- Among patients without contraindications to therapy, top ranked hospitals had significantly higher rates of use of aspirin (96.2% as compared with 88.6% for similarly equipped hospitals and 83.4% for non-similarly equipped hospitals, p <0.01) and beta-blockers (75.0%

vs. 61.8% and 58.7%, $p < 0.01$), but lower rates of reperfusion therapy (61.0% vs. 70.7% and 65.6%, $p = 0.03$).
- The survival advantage associated with admission to top ranked hospitals was less strong after adjustments were made for factors including the use of aspirin and beta-blockers (OR =0.94; 95% CI 0.82-1.08, $p = 0.38$).

Chen (1999b) found that among four hospital peer groups, the top 100 hospitals had similar thirty day mortality and use of aspirin, beta-blockers, and reperfusion compared with their peers, but lower lengths of stay and in-hospital costs, with similar or lower re-admission rates. This, suggests the authors, that HCIA-Mercer's ‘100 Top Hospitals’ study may be better suited for identifying hospitals with higher performance on financial and operating measures than superior clinical performance in treating elderly AMI patients.

Patients with other medical conditions

The studies done on patients with other medical conditions that have been reviewed are:
- *Rockall (1995)* carried out a population based, prospective study of the management and outcome of unselected cases of acute upper gastrointestinal haemorrhage in 5,565 patients (aged 16 years and over) in 74 acute hospitals in four health regions in the UK to assess the effect of risk standardisation on the variation observed in hospital mortality for this condition.
- *Hofer (1996)* undertook a study in Michigan to evaluate the feasibility of using mortality rates for medical diagnoses to distinguish 172 average quality hospitals from 19 poor quality ones (5% vs. 25% of deaths being preventable, respectively).
- *Iezzoni (1996b)* examined whether judgements about an individual hospital's risk adjusted mortality is affected by the severity adjustment method using data from 105 acute care hospitals on 18,016 adults hospitalised in 1991 for pneumonia.
- *Hand (1997)* looked at variations in hospital mortality and average length of stay (LOS) among 16,249 Medicare patients with pneumonia from 20 Illinois hospitals with the largest number of discharges for pneumonia for the years 1989 through 1992.

Rockall (1995) found that:
- Crude mortality in individual hospitals ranged from 0% to 29%.
- When hospitals were ranked in order of increasing mortality, risk standardisation for age, shock, and co-morbidity resulted in 21 of the 74 hospitals changing ranks by ten or more places.
- After further standardisation for diagnosis, endoscopic stigmata of recent haemorrhage, and rebleeding, 32 hospitals moved ten or more places from their original rank; one hospital moved 45 places.

The authors concluded that:
- Risk standardisation to correct for variation in case-mix results in apparently significant differences in mortality rates becoming non-significant.

- The current state of routine data collection does not allow for anything but the most basic case-mix adjustment to be made.
- Simple league tables of crude mortality are misleading in this disorder and cannot be regarded as a reflection of the quality of health care.

Findings from the study by *Hofer (1996)* suggested that:

- For individual diagnosis related groups (DRGs), mortality rates were a poor measure of quality, even using the optimistic assumption of perfect case-mix adjustment.
- For AMI, high mortality rate outlier status (using two years of data and a 0.05 probability cut-off) had a positive predictive value of only 24%, thus, more than three fourths of those labelled poor quality hospitals (high mortality rate outliers) actually would have average quality.
- Aggregating all medical DRGs and assuming very large quality differences and perfect case-mix adjustment, the sensitivity for detecting poor quality hospitals is 35% and the positive predictive value is 52%.

The authors conclude that although DRG-specific hospital mortality rates may be useful for some surgical diagnoses, they probably cannot accurately detect poor quality outliers for medical diagnoses. Even collapsing to all medical DRGs, hospital mortality rates seem unlikely to be accurate predictors of poor quality.

Iezzoni (1996b) calculated observed to expected hospital death rates following admission for pneumonia for each severity adjustment method. The key findings from were:

- Unadjusted mortality rates for the 105 hospitals ranged from 1.4% to 19.6%; the overall in-hospital morality rate was 9.6%.
- After adjusting for age, sex, DRG, and severity, 73 facilities had observed mortality rates that did not differ significantly from expected rates according to all 14 severity methods; two had rates significantly higher than expected for all 14 severity methods.
- For 30 out of 105 hospitals, observed mortality rates differed significantly from expected rates when judged by one or more but not all 14 severity methods.
- Kappa analysis showed fair to excellent agreement between severity methods.

Hand (1997) reported that:

- The distributions showed trends toward lower mortality and shorter stays over the four years. Correlation of performance from year to year at each hospital for mortality was low with none of the calculated correlation coefficients significant at $p <0.05$, while correlations for LOS were higher (all coefficients significant at $p <0.01$).
- For LOS, the correlation between 1991 and 1992 was 0.766 ($p <0.00005$, $r(2)=0.587$), showing that nearly 60% of variance was caused by differences in performance.

- In contrast, for mortality in 1991 and 1992, the correlation was 0.301 (p =0.0986, r(2) =0.091), showing that less than 10% of variance between hospitals was caused by differences in performance.
- Similar results were obtained when the 20 hospitals were ranked and their rank correlations calculated.
- Based on these findings the authors conclude that for pneumonia in Medicare patients, differences in length of stay between hospitals are caused by differences in performance, while differences in mortality are random.

Paediatric patients

The study done on paediatric patients that has been reviewed is:

Parry (1998) carried out a longitudinal study of mortality in nine neonatal intensive care units (ICUs) in the United Kingdom to assess whether crude league tables of mortality and league tables of risk adjusted mortality accurately reflect the performance of hospitals. Data on 2,671 very low birth weight or pre-term infants admitted to neonatal ICUs between 1988 and 1994 were analysed. Hospital mortality rates were adjusted using the clinical risk index for babies (CRIB) score. The findings were:

- Hospitals had wide and overlapping confidence intervals when ranked by mortality in annual league tables; this made it impossible to discriminate between hospitals reliably.
- In most years there was no significant difference between hospitals, only random variation. The apparent performance of individual hospitals fluctuated substantially from year to year.
- On the basis of their findings, the authors conclude that annual league tables are not reliable indicators of performance or best practice as they do not reflect consistent differences between hospitals.

Surgical patients

The studies done on surgical patients that have been reviewed are:

- *Roos (1990)* examined correlations among hospital rankings for various outcome indicators following five common surgical procedures.
- *Khuri (1997)* examined the effect of case-mix adjustment on the relative ranking of 44 hospitals by post-operative all cause 30 day mortality rates following major non-cardiac surgery. A cohort study design was used to analyse the outcomes of 87,078 major non-cardiac operations performed between October 1991 and December 1993 at 44 Veterans Affairs Medical Centres.
- *Daley (1997b)* used a cohort design to analyse 30 day post-operative mortality and morbidity rates for 87,078 major non-cardiac operations at 44 Veterans Affairs Medical Centres and to assess correlation between hospital rankings by risk adjusted mortality and morbidity rates.
- *Ansari (1999)* undertook a comparison of crude and adjusted in-hospital mortality rates after prostatectomy in 21 hospitals using routinely collected hospital discharge data for the period 1987 to 1994 to illustrate the value and limitations of using comparative mortality rates as a surrogate measure of quality of care. Variables used to control for severity of illness including age, specific co-morbid

conditions, length of stay, emergency admission, and hospital location, were identified using ICD9-CM coded Victorian hospital morbidity data.

- *Copeland (1995)* conducted a one year prospective analysis of 3,004 patients undergoing non-day case general surgery in a district general hospital and estimated mortality rates by surgeon adjusted for patient severity using the Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM) scoring system.
- *Mozes (1999)* undertook a cross-sectional study of 303 patients undergoing surgical repair of traumatic femoral neck fracture in three medical centres to assess the merit of multi-outcome measurements in the evaluation of quality of care.
- *Silber (1997b)* analysed data on 74,647 patients who underwent general surgical procedures to assess the correlation among adjusted hospital outcome rankings for three different measures of quality of care: in-hospital death, complication and failure to rescue (in-hospital death following a complication).
- *Khuri (1998)* carried out a prospective assessment of pre-surgical risk factors, process of care during surgery, and outcomes 30 days after surgery on veterans undergoing major surgery in 123 Veterans Affairs Medical Centers (VAMCs), as part of the National VA Surgical Quality Improvement Program, and then identified high and low outlier facilities on the basis of observed to expected mortality ratios.

Roos (1990) found that the ranking of hospitals according to in-hospital mortality was highly correlated with their ranking according to 30 day post-surgical mortality.

Key findings from the study by *Khuri (1997)* were:

- patient risk factors predictive of post-operative mortality included:
  - serum albumin level
  - American Society of Anesthesia class
  - emergency operation
  - 31 additional pre-operative variables.
- Considerable variability in unadjusted mortality rates for all operations was observed across the 44 hospitals (1.2-5.4%); after risk adjustment, observed to expected ratios ranged from 0.49 to 1.53.
- Rank order correlation of the hospitals by unadjusted and risk adjusted mortality rates for all operations was 0.64.
- 93% of the hospitals changed rank after risk adjustment, 50% by more than five and 25% by more than ten places.
- Risk adjustment had an appreciable impact on the rank ordering of the hospitals and provided a means for monitoring and potentially improving the quality of surgical care.

Daley (1997b) examined rates of 21 post-operative adverse events (morbidities) and mortality occurring within 30 days after major non-cardiac operations. The main findings were:

- Wide variation in the unadjusted rates of one or more post-operative morbidities for all operations was observed across the 44 hospitals (7.4-28.4%); risk adjusted observed to expected ratios ranged from 0.49 to 1.46.

- The Spearman rank order correlation between the ranking of the hospitals based on unadjusted morbidity rates and risk adjusted observed to expected ratios for all operations was 0.87.
- There was little or no correlation between the rank order of the hospitals by risk adjusted morbidity and risk adjusted mortality.

Ansari (1999) reported that:

- The overall crude mortality rates between hospitals achieved borderline significance (p =0.06); these differences were no longer significant after adjustment (p =0.21).
- On crude analysis of mortality rates, four hospitals were initially identified as 'low' outlier hospitals; after adjustment, none of these remained outside the 95% CI, whereas a new hospital emerged as a 'high' outlier (OR =4.56, p =0.05).
- Regression adjustment of routinely collected data on prostatectomy from the Victorian In-patient Minimum Database reduced variance associated with age and correlates of illness severity.

Copeland (1995) estimated the ratio of observed to expected numbers of deaths and complications (O:E ratio) for each surgeon, both overall and within specialty zones. Results of this study include:

- Mortality rates varied from 1.0 to 4.9 per cent whereas O:E ratios ranged from 0.83 to 1.06; morbidity rates varied from 5.3 to 12.6 per cent with O:E ratios 0.86-1.02.
- On the basis of their findings the authors highlight the serious limitation in using 'raw' uncorrected mortality and morbidity statistics to compare surgeon or hospital performance.

In the study by *Mozes (1999)*, the outcomes examined included mortality, functional capacity, post-operative complications, and length of stay; the explanatory variables included were sociodemographic details, co-morbidity indices, pre-operative functional capacity, depression, and cognition. Results indicated that:

- Centre A was a ‘good’ outlier for mortality rate but, in contrast, was a ‘bad’ outlier for complication rate and length of stay.
- Centre B was a ‘bad’ outlier for functional capacity but a ‘good’ outlier for length of stay.

The authors concluded that outcome studies for quality assurance programs should include all relevant outcomes, as the assumption that one major outcome may be representative for quality of care assessment may be misleading.

In the analysis of outcomes among general surgical patients, *Silber (1997b)* adjusted outcomes for severity and case-mix. The findings included:

- For 142 hospitals, the correlation between hospital rankings based on the death rate and those ranked by the complication rate was only 0.208 (p =0.013).
- A similarly low correlation was present between the complication and failure rate rankings, r = -0.090 (p =0.287).

- A higher correlation was observed between the death and failure rate rankings, r =0.90 (p <0.001).

Khuri (1998), using data for the financial year 1997, compared hospital outlier status based on risk adjusted observed to expected mortality ratios with that based on unadjusted mortality rates. The findings were:

- 11 VAMCs were low outliers for risk adjusted observed to expected mortality ratios, while 13 VAMCs were high outliers.
- Identification of high and low outliers by unadjusted mortality rates would have ascribed an outlier status incorrectly to 25 of 39 hospitals, an error rate of 64%.

STUDIES TO EXAMINE TECHNICAL ISSUES IN COMPILING CASE FATALITY RATES

The main technical issues that need to be addressed in compiling case fatality rates are:

- methods of linking data about hospital episodes and mortality
- type of hospital admission and diagnostic specificity
- time cut-off point and inclusion of all deaths within period or only those which occur in hospital
- method of death diagnosis recording used
- dealing with patient transfer
- adjusting for severity of condition and other potential confounders
- accuracy and completeness of data
- statistical power.

Methods of linking data about hospital episodes and mortality

The technical issues concerned with linking hospital episode data have been studied in the Oxford Record Linkage Study for many years. Record linkage is the process of bringing together related records, or abstracts of records, which have been compiled separately. It may be used to bring together different data about the same medical event (e.g. a hospital admissions record, a discharge record, clinical data, laboratory data), data on different episodes of illness for the same person, or data from different sources relating to the same person (e.g. death certificate and hospital episode data). The crux of the actual process of linkage is the comparison of two records and the decision as to whether they refer to the same person. Linked medical records lend themselves, in particular, to the calculation of re-admission rates as outcome measures.

Gill (1993) and *Kendrick (1993)* have outlined the four main steps involved in the linkage of records:

- file blocking
- matching
- linkage
- validation and checking.

The first step is *file blocking* which involves putting records in an order (e.g. alphabetical) to make searching more efficient for the purpose of making comparisons between records. With very large data sets, comparing every record in a system with every other record may be impossible, and blocking is used to minimise the number of comparisons needed. It is desirable to use at least two different methods of blocking to effect matches by one method that may have been missed by the other (e.g. blocking on dates of birth followed independently by blocking on names). Thus records will fail to be compared only if they differ on an item from both sets of blocking criteria.

The next step is *matching* which is the process of comparing records to determine whether they do or do not relate to the same person. When records created at different times and in different places are to be linked, it should be possible, in principle, to link these using a personal identification number such as the NHS number.

Currently, partial identifiers such as forenames, surname, sex, date of birth, and postcode of residence, are used to identify different records relating to the same individual. When using partial identifiers, such as forenames and surname, character by character matching is not recommended if precision of matching is required. ‘Failure to match’ records belonging to the same person on this basis occurs partly because there are fairly high levels of error in spelling and recording names, and particularly because the recording of names may vary. For each of the identifying items used to link records, there may be a discrepancy rate of up to 3% in pairs of records belonging to the same person. Thus exact character by character matching could miss up to 15% of true links if five separate identifying items are used (the sum of the item-specific discrepancies).

In computerised linkage of medical records, probability matching, using an array of identifiers, achieves much higher levels of correct matching than is generally achievable by exact character by character comparisons. A composite score or match weight can be calculated for each pair of records indicating the probability that they relate to the same person, using calculations based on the discriminating power of each identifying item used in the matching process. In the Scottish Record Linkage System which employs probability matching (using name, initial, sex, and date of birth) clerical checking shows that on a pair-wise basis, both the false positive rate (the proportion of linked pairs of records which do not in fact refer to the same individual) and the false negative rate (the proportion of truly matched pairs which the system fails to link) are around 1%.

After records have been matched, *linkage* is undertaken, which is the process of assembling correctly matched records, identified as relating to the same person, into a time-sequenced composite record for the individual.

Finally, *validation checks and corrections* are carried out in which any inconsistencies between different records for the same person are identified and corrected. This cycle may, in practice, account for a significant part of the resource required to match and link records in an established linkage system.

Other relevant studies that have been reviewed are:

Henderson (1992) reported on the extent to which death certificates which specify that death occurred in hospital can be matched and linked with routine hospital in-patient information systems. Data on patients (aged 65 years and over) who died in hospital between 1979 and 1985 in six health districts in southern England covered by medical record linkage were used in the analyses. The main results were:

- 98.2% of hospital record abstracts which specified that death occurred in hospital were linked by standard computer based techniques to death certificates.
- Conversely, however, only 94.4% of death certificates which specified that death occurred in hospital could be linked to the abstracts of corresponding hospital in-patient records. A major factor contributing to the latter failures may be a difference of definition of what

constitutes a death 'following hospital admission' in patients who die shortly after arrival at hospital.
- Hospital in-patient record abstracts corresponding to death certificates for deaths in hospital may not invariably exist when death occurs shortly after the arrival of the patient at hospital.

Type of hospital admission (emergency or planned) and diagnostic specificity

Jarman (1999) analysed 8 million discharges from NHS hospitals over a four year period (where the primary diagnosis was one of the diagnoses accounting for 80% of in-patient deaths) to determine which factors best explain variation in standardised hospital in-patient mortality ratios. The authors found that the percentage of cases that were emergency admissions (60% of total hospital admissions, accounting for 93% of all patient deaths analysed) was the most powerful predictor of this variation in mortality.

Rosenthal (1997b) examined the consistency of hospital mortality rates across different diagnoses using standardised mortality ratios for patients discharged in 1991 from US hospitals. Results revealed that:
- Correlations between standardised mortality ratios for different diagnoses were relatively weak, ranging from 0.03 to 0.34.
- Agreement between hospital rankings (based on standardised mortality ratios), as measured by the weighted kappa statistic, was also weak.
- In conclusion, hospital mortality rates for individual diagnoses are weakly associated. Thus, it is not valid to generalise conclusions about hospital performance from a single diagnosis.

Method of death diagnosis recording

The calculation of case fatality rates after an admission can in England use one of three ways of recording the death diagnosis:
- death from any cause
- death due to a specific condition recorded as the underlying cause on a death certificate
- death due to a specific condition recorded anywhere on the death certificate.

Goldacre (1993) undertook a retrospective cohort study using linked hospital records and death certificates for people who died shortly after hospital admission in six UK district health authorities to examine the extent to which individual diseases, when recorded as being present shortly before death, were certified as causes of death. Three broad patterns of certification were distinguished:
- Firstly, there were diseases that were usually recorded on death certificates when death occurred within four weeks of hospital care for them. Examples included lung cancer (on 91% of such death certificates), breast cancer (92%), leukaemia and lymphoma (90%), anterior horn cell disease (89%), multiple sclerosis (89%), AMI (90%), stroke (93%), aortic aneurysm (87%), and spina bifida (89%). These diseases were also usually certified as the underlying cause of death.

- Secondly, there were diseases which, when present within four weeks of death, were commonly recorded on death certificates but often not as the underlying cause of death. Examples included tuberculosis (on 76% of such certificates; underlying cause on 54%), thyroid disease (49%; 21%), diabetes mellitus (69%; 30%) and hypertension (43%; 22%).
- Thirdly, there were conditions which, when death occurred within four weeks of their treatment, were recorded on the death certificate in a minority of cases only. Examples of these included fractured neck of femur (on 25% of such certificates), asthma (37%), and anaemia (22%).
- There was 'convergence' in certification practice towards the common cardiovascular and respiratory causes of death.
- There was also evidence that conditions regarded as avoidable causes of death may not have been certified when present at death in some patients.

Time cut-off point and inclusion of all deaths within period or only those which occur in hospital

Many of the studies noted in previous chapters that use in-hospital mortality as a performance indicator while disregarding deaths occurring soon after discharge, may give misleading information. A number of studies have examined the place of death of patients recently admitted to hospital, while few have compared the use of different time cut-off points for calculating case fatality rates.

Rosenthal (2000) examined the relationship of in-hospital and 30 day (post-admission) mortality rates and the association between in-hospital mortality and hospital discharge practices in a retrospective cohort analysis of data on 13,834 patients with congestive heart failure admitted to 30 hospitals in Ohio in 1992-1994. Standardised mortality ratios (SMRs) were determined for in-hospital and 30 day mortality by dividing observed death rates by predicted death rates. It was found that:

- In-hospital SMRs ranged from 0.54 to 1.42, and six hospitals were classified as statistical outliers ($p < 0.05$); 30 day SMRs ranged from 0.63 to 1.73, and seven hospitals were outliers.
- Although the correlation between in-hospital SMRs and 30 day SMRs was substantial ($R=0.78$, $p < 0.001$), outlier status changed for seven of the 30 hospitals. Nonetheless, changes in outlier status reflected relatively small differences between in-hospital and 30 day SMRs.
- Rates of discharge to nursing homes or other in-patient facilities varied from 5.4% to 34.2% across hospitals. However, relationships between discharge rates to such facilities and in-hospital SMRs ($R=0.08$, $p=0.65$) and early post-discharge mortality rates ($R=0.23$, $p=0.21$) were not significant, suggesting that in-hospital SMRs were not biased by differences in post-discharge mortality or discharge practices.

Smith (1993) highlighted some of the methodological difficulties of identifying patients who die within 30 days of an operation using the Scottish Hospital Discharge (SMR1) file and Registrar General's Death Registration file. Key findings included:

- Patients who died prior to discharge or transfer accounted for only 52% of the total peri-operative deaths.
- 32% of deaths occurred following transfer to a different specialty or hospital.
- 8% of deaths occurred on re-admission following discharge home.
- The remaining 8% of the total peri-operative deaths were shown to have taken place outside hospital.

Higginson (1998) analysed trends in place of death of cancer patients, by region and by diagnosis, within England for the period 1985-1994. It was found that:

- The percentage of all cancer deaths which occurred in a NHS hospital or nursing home fell gradually from 58% (1985) to 47.3% (1994), while the percentage who died in non-NHS hospitals, nursing homes, hospices and communal establishments increased.
- Although hospital is still the most common place of death from cancer, the percentage of cancer patients who die in hospital is reducing.

Brameld (1998) measured the trend and pattern of time spent in hospital during the last year of life and identified trends in the place of death among 68,875 persons (aged 65 years and over) who died between 1985 and 1994 in Western Australia. The findings were:

- Increasing proportions of all age groups (65-74, 75-84, and 85+ years) were admitted to hospital at least once in the year before death during 1985-94, but the chance of admission decreased with age.
- There was a trend towards a greater number of shorter admissions per person; however, the average total bed days per person showed no significant increase, except at ages 65-74 years.

Pollock (1995) reported on a study which investigated factors associated with the high proportions of 'death certificate only' registrations (DCOs) for all cancers registered in the 56 districts in south-east England between 1987 and 1989. 'Death certificate only' (DCO) registration ratios (standardised for age and sex) as a proportion of all registrations, were derived for each of the 56 districts in the Thames Regions. The findings were:

- Around 24% of all patients deaths were registered as DCOs by the Thames Cancer Registry between 1987 and 1989.
- Of these, 40.9% died in an acute NHS hospital setting, 37.1% died at home, 10.4% died in hospices and 3.4% died in non-NHS hospitals.
- Increasing age, decreasing survival, district of residence and place of death were positively associated with DCOs.
- Changing patterns of treatment and terminal care may make case ascertainment and registration more difficult for registry staff in the future, although the minimum contract data set should assist in this.
- The current trends to shorten lengths of stay and increase day case and out-patient treatment could adversely affect registration and case ascertainment, especially if fewer people die in hospital.

Wennberg (1989) compared rates of hospital use and mortality in 1985 among Medicare enrollees in two US states (Boston and New Haven, Connecticut). It was found that:

- Adjusted rates of discharge, re-admission, and length of stay, were 47, 29, and 15% higher, respectively, in Boston.
- 40% of Boston's deaths occurred in hospitals as compared with 32% of New Haven's.
- High variation medical conditions (those for which there is little consensus about the need for hospitalisation) accounted for most of these differences.
- By contrast, discharge rates for low variation medical conditions (which tend to reflect the incidence of disease) were similar.
- In-patient case fatality rates were lower in Boston than in New Haven (relative risk [RR] =0.85; 95% CI 0.78-0.92), but when all deaths (regardless of place of death) were measured, the mortality rates in Boston and New Haven were nearly identical (RR =0.99; 95% CI 0.93-1.05).
- The authors conclude that population based as well as hospital based statistics are needed to evaluate differences in hospital mortality rates for high variation medical conditions.

Fisher (2000) undertook a cross-sectional analysis to explore whether geographic variations in Medicare hospital utilisation rates are due to differences in local hospital capacity, after controlling for socioeconomic status and disease burden, and to determine whether greater hospital capacity is associated with lower Medicare mortality rates among a 20% sample of 1989 Medicare enrollees. The following results emerged:

- The per capita supply of hospital beds varied by more than twofold across U.S. regions.
- Residents of areas with more beds were up to 30% more likely to be hospitalised, controlling for area differences in sociodemographic characteristics and disease burden.
- A greater proportion of the population was hospitalised at least once during the year in areas with more beds.
- In areas with more beds, death was more likely to take place in an in-patient setting.
- Residence in areas with greater levels of hospital resources was not associated with a decreased risk of death.

Tolle (1999) investigated the reasons for Oregon’s low in-hospital death rate, the state that currently has the lowest in-hospital death rate in the US (31%). The principal factors determining location of death in the US are the use and availability of beds in acute care hospitals. The authors reported that, in Oregon:

- Decisions to avoid hospital admission are far more common than discharge of the actively dying.
- Do not resuscitate orders were reported for 91% of nursing home residents in one study, and living wills were reported for 67% of a random sample of adult Oregon decedents in a second study.

- In the second study, decisions not to start treatment were far more common than decisions to stop treatment (79% compared with 21%).

Sirio (1999) performed a retrospective cohort study of 116,340 severely ill patients admitted between 1991 and 1995 to 28 hospitals to assess variations in severity adjusted hospital mortality. Findings included:

- While standardised mortality ratios tended to decline over the four years of the study, excluding the increasing ($p < 0.001$) number of patients discharged to skilled nursing facilities (these patients had markedly higher severity of illness than those discharged home) attenuated much of the decline in standardised mortality over time.
- In-hospital mortality may be significantly influenced by discharge timing, especially among terminal patients who are discharged to die outside hospital. If hospitals know that they are being evaluated on the number of in-patient deaths, severely ill patients may be transferred to other facilities.
- As more patients are transferred to nursing facilities, the length of life from hospital admission to death (wherever that may occur) may be a more appropriate outcome indicator.

Mullins (1998) carried out a retrospective one year follow-up study of 90,048 injured patients discharged from acute care hospitals in Washington between 1991 and 1993 with at least one diagnosis indicating trauma, to evaluate the adequacy of hospital death rates as an outcome measure following trauma and to determine the frequency of post-discharge death. It was found that:

- Among 1,912 injured patients who died in hospital, 825 death certificates (43%) listed a non-injury cause of death.
- The overall mortality rate at hospital discharge was 21.2 per 100,000 hospitalised injured patients (12.1 per 100,000 for trauma deaths and 9.1 per 100,000 for those designated as non-trauma deaths).
- Patients with trauma related death designations were younger (mean age, 51.5 years vs. 77.9 years), had shorter lengths of stay (median stay, 2 days vs. 5 days), and sustained more severe injures ($p < 0.001$).
- When the 1,273 deaths that occurred within 30 days of hospital discharge were also included, rates for trauma-designated deaths increased to 14.1 per 100,000 and those for non-trauma-designated deaths increased to 21.3 per 100,000.

The authors conclude that:

- Hospital discharge death rates are incomplete measures of death frequency for injured patients. Designation of the cause of death, especially among older, hospitalised, injured patients often reflects pre-existing medical conditions.
- Adequate assessment of mortality following trauma requires measurement of the frequency of death following hospital discharge.

Garnick (1995) compared 30 day and 180 day post-admission hospital mortality rates for all Medicare patients and those in three categories of cardiac care: CABG, AMI, and congestive heart failure. Using data for 1989, hospitals were ranked on a measure of actual and predicted mortality, and rankings were compared for measures of mortality within 0-30 days, 31-180 days, and 0-180 days post-admission. The reported findings were:

- For the conditions studied, when hospitals were ranked using the 30 day data; some hospitals had much lower than predicted 30 day mortality rates, while others had much higher than predicted mortality rates.
- Data from the time period 31-180 days post-admission yielded results that corroborate the 0-30 day post-admission data.
- Hospital performance on one condition was related to performance on the other conditions, but the correlation was much weaker in the 31-180 day interval than in the 0-30 day period.
- The 30 day data discriminated the top and bottom fifths of the 180 day data extremely well, especially for AMI outcomes.
- The authors concluded that using data on cumulative hospital mortality from 180 days post-admission does not yield a different perspective from using data from 30 days post admission for the conditions studied.

Dealing with patient transfer

Hospitals that receive a significant number of transfers from other institutions may have as a consequence patients with more severe disease.

Gordon (1996) carried out a retrospective cohort study to determine whether a widely implemented method of severity adjustment underestimated the risk of death and other outcomes among patients transferred from other acute care hospitals and to examine the impact of this potential bias on hospital outcomes profiles. The analyses were based on data on 40,820 adult medical and surgical patients from 1988 to 1991, of whom 38,946 were direct admissions and 1,874 were inter-hospital transfers. The reported findings were:

- Admission severity of illness was directly related ($p < 0.001$) to rates of in-hospital death and length of stay.
- 49% of inter-hospital transfers had moderate to high severity, compared with 35% of direct admissions ($p < 0.001$).
- However, adjusting for severity and other co-variates, the risk of in-hospital death was higher in transfers than in direct admissions (multivariable OR =1.99; 95% CI 1.64-2.42).
- Results were consistent in medical and surgical admissions, when examined separately, and among individual diagnostic categories.
- The authors estimate that, independent of quality of care, severity adjusted mortality and length of stay would appear 17% and 8% higher, respectively, for hospitals in which 20% of patients were inter-hospital transfers than for hospitals in which 2% of patients were transfers.
- Failure to account for transfer status may produce biased performance profiles and, therefore, may create disincentives for hospitals to accept transfers from other acute case facilities.

Bernard (1996) compared in-hospital mortality and length of stay among 8,740 patients transferred from other hospitals and 76,047 non-transferred patients in an academic medical centre between July 1989 and December 1993. Key findings were:

- Hospitalisations of the transfer patients were more complex and resource-use intensive.
- Transferred patients were more likely (p <0.0000) to suffer in-hospital death (9.4% vs. 2.5%) and were more likely to be length of stay outliers (28% vs. 10%).
- After case-mix adjustment and exclusion of length of stay outliers, transferred patients on the three services (surgery, medicine, and paediatrics) remained in the hospital 1.62, 1.15, and 0.84 days longer (p <0.0001) than non-transferred patients.

Adjusting for severity of condition and other potential confounders

A number of studies, mainly from the USA, have shown that severity of condition may be an important factor in interpreting case fatality rates.

Romano (2000) analysed hospital discharge abstracts for 116,174 adult patients with AMI and the in-patient medical records for a sample of 974 patients with AMI admitted to 30 California hospitals to determine if a widely used proprietary risk adjustment system, APR-DRGs, misadjusts for severity of illness and misclassifies provider performance. Key findings were:

- The APR-DRG Risk of Mortality class was a strong predictor of death within 30 days (c =0.831-0.847), but was further enhanced by adding age and sex.
- The APR-DRG Risk of Mortality system is a powerful risk adjustment tool, largely because it includes all relevant diagnoses, regardless of timing.
- Although some late diagnoses may not be preventable, APR-DRGs appear suitable only if one assumes that none is preventable.

Iezzoni (1995b) undertook a retrospective cohort study of 11,880 patients managed medically for AMI in 100 hospitals to determine whether assessments of illness severity, defined as risk for in-hospital death, varied across four severity measures. For each of four common severity measures, patients were ranked according to probability of death as predicted by the severity measure, age, and sex; rankings were then compared across the different severity measures. The presence or absence of each of six clinical findings considered to indicate poor prognosis in patients with AMI was determined for patients ranked differently by different severity measures.

- The key finding was that some pairs of severity measures assigned very different severity levels to more than 20% of patients.
- The authors concluded that evaluations of patient outcomes need to be sensitive to the severity measures used for risk adjustment.

Iezzoni (1996c) studied 9,407 stroke patients from 94 hospitals to see whether severity adjusted predictions of the likelihood of in-hospital death for stroke patients differed among severity measures. For each stroke patient, the study compared the probabilities of death as predicted by the five different severity measures combined with data on age and sex. The frequencies of seven clinical indicators of poor prognosis in stroke were examined for patients with very different odds of death predicted by different severity measures. Odds ratios were considered very different when the odds of death predicted by one severity measure was less than 0.5 or greater than 2.0 of that predicted by a second measure. The reported findings were:

- There were 916 (9.7%) in-hospital deaths among the study sample.
- Some pairs of severity measures ranked over 60% of patients very differently by predicted probability of death.
- The authors conclude that studies of severity adjusted stroke outcomes may produce different results depending on which severity measure is used for risk adjustment.

Iezzoni (1996d) carried out a retrospective cohort study of 18,016 persons (aged 18 years and older) managed medically for pneumonia, to see whether predictions of patients' likelihood of dying in-hospital differed among severity methods. The probability of death was calculated for each patient using logistic regression with age, age squared, sex, and each of five severity measures as independent variables. Patients were ranked by calculated probability of death and rankings were compared across the five different severity methods. Frequencies of 14 clinical findings considered poor prognostic indicators in pneumonia were examined for patients ranked differently by different methods. The reported findings were:

- There were 1,732 (9.6%) in-hospital deaths among the study sample.
- The three severity methods based on discharge abstract data rated more than 25% of patients differently by predicted likelihood of death when compared with the rankings of the two clinical data based methods.
- The authors conclude that results of outcomes studies may vary depending on which severity method is used for risk adjustment.

Iezzoni (1998) examined whether five severity adjustment measures rated severity differently for 7,764 CABG patients from 38 hospitals. Two of the severity measures were based on clinical data taken from the first two hospital days; the other three used diagnoses and other information coded on standard, computerised hospital discharge abstracts.The odds of death predicted from each of the different severity scores along with data on age, age squared, and sex, were compared. The following findings were reported:

- The overall in-hospital death rate was 3.2%.
- Code based severity measures predicted very different odds of dying than clinical measures for more than 30% of patients.
- Although code based measures had superior predictive power, this may reflect their reliance on diagnosis codes for life threatening conditions occurring late in the hospitalisation, possibly as complications of care.
- The latter compromises the utility of code based severity measures for drawing inferences about quality of care based on severity adjusted CABG death rates.

Davis (1995) attempted to establish what information is needed to adjust for risk when using outcomes to evaluate the quality of health care. The study used retrospective analyses of existing, computerised clinical databases containing laboratory test results, information on chronic co-existing conditions, and nursing evaluations of functional status, to predict in-hospital mortality for stroke and pneumonia patients admitted to one tertiary teaching hospital between 1987 and 1992. The predictive values of the nursing functional status assessments and the clinical laboratory data were compared. The following results emerged:

- For both conditions, the functional status data had as much prognostic information as the laboratory data.
- Specifically, a nurse's report that a patient required total assistance for bathing was the best single predictor of in-hospital mortality in the models for patients with either cerebrovascular disease or pneumonia.
- If hospitals admit patients with different levels of functional impairment, it is important to account for these differences before comparing outcomes across facilities.
- Assessments of functional status are a simple, inexpensive measure that may have considerable value.

Poses (2000), in a prospective cohort study of patients presenting in the emergency department with acute congestive heart failure (CHF), found that hospital rankings of 30 day and one year adjusted mortality rates for patients with CHF depended on the method used to adjust for severity.

Elixhauser (1998) developed a comprehensive set of co-morbidity measures for use with large administrative in-patient datasets by analysing data on in-hospital mortality and length of stay for all adult, non-maternal in-patients from 438 acute care hospitals in California in 1992 (n = 1,779,167). It was found that

- The set of 30 co-morbidity measures developed were associated with substantial increases in length of stay and mortality both for heterogeneous and homogeneous disease groups.
- Several co-morbidities including mental disorders, drug and alcohol abuse, obesity, coagulopathy, weight loss, and fluid and electrolyte disorders, are important predictors of outcomes, yet commonly are not measured.
- The authors suggest that different co-morbidities have independent effects on outcomes and probably should not be simplified as an index because they affect outcomes differently among different patient groups.

Accuracy and completeness of data

Accuracy and completeness of data are essential in the compilation of indicators based on case fatality rates. This issue relates to all indicators calculated from hospital episode and mortality data and is not addressed specifically in this report.

Statistical power

Thomas (1999) examined whether risk adjusted mortality rates are valid indicators of hospital quality using an analytical model to estimate the sensitivity and predictive error of hospital mortality rate performance indicators. The number of patients treated per hospital was one of the factors shown to determine accuracy of hospital mortality rate performance indicators.

SUMMARY OF FINDINGS FROM THE LITERATURE REVIEW

In reviewing the literature we attempted to address the following questions:

- Do properly compiled case fatality rates reflect the quality of in-patient care (see Chapter 5)?
- Are case fatality rates a useful way of comparing hospital performance (see Chapter 6)?
- Does the volume of cases influence case fatality rates (see Chapter 3)?
- Does the length of stay influence case fatality rates (see Chapter 4)?
- What other factors influence case fatality rates (see Chapter 2)?
- How should case fatality rates be compiled if they are to reflect the quality of in-patient care (see Chapter 7)?

Case fatality rates and quality of in-patient care

The main conclusions from the major review by Thomas (1998) were:

- Poor quality care increases patients’ risks of mortality.
- On average, the quality of care provided in ‘high mortality’ outliers is poorer than that provided in ‘low mortality’ outliers.

Studies of individual diseases show that the validity of the relationship between quality of care and mortality must be established explicitly for each condition. Thus, there is an established relationship for cardiac surgery, chronic cardiac disease, pneumonia and perinatal care. There is no proven relationship for septicaemia and stroke. The results for acute myocardial infarction are equivocal with some studies showing a weak relationship, others none at all.

Variations in mortality seem to be more influenced by patient variables than organisational features of the care delivery system.

Use of case fatality rates in comparing hospital performance

Most of the relevant studies have been done in the USA. All researchers warn against using even sophisticated case fatality measures for comparing hospital performance. The main conclusions from the review and study by Thomas (1998, 1999) were:

- Risk-adjusted mortality rates are seriously inaccurate when used as a measure of quality for individual hospitals.
- Publication of hospital mortality rates may seriously misinform the public about hospital quality.
- The best measures for reporting hospital performance are likely to be process based indicators such as the degree of compliance with explicit criteria for a specific condition.

A number of studies have compared severity adjusted case fatality measures but have found that hospital rankings varied considerably depending on the way that severity had been estimated.

Condition-specific case fatality rates do not seem to be useful for medical conditions but may be for surgical ones, in particular coronary artery bypass grafting, for which they have been shown to be a good measure of hospital performance.

A British study of neonatal paediatric intensive care units concluded that case fatality rates were unreliable comparative performance indicators. Hospitals had confidence intervals that were wide and overlapping, the performance of individual hospitals fluctuated substantially from year to year, and in most years there was no significant difference between hospitals, only random variation

Other researchers have noted that death following admission for a specific condition may be a rare event and that there are issues of inadequate statistical power in interpreting comparative hospital information calculated on this basis.

Relationship between case volumes and case fatality rates

The main conclusions from the reviews by Sowden (1997, 1998) were:

- Much of the research into the relationship between volume and case fatality is of poor quality.
- Most studies report a reduction in case fatality as volumes increase.
- There is no general relationship between volume and quality.
- Research is difficult to interpret because of confounding due to case-mix, limited measurement of outcomes and lack of clarity about the direction of cause and effect.

Hillner (2000) reviewed the effect in patients with cancer and concluded that:

- There were numerous methodological limitations in almost of all the studies.
- Relationship between volume and outcome was found consistently for cancers treated with technologically complex surgical procedures.

Surgical procedures for conditions other than cancer for which an association between volume and case fatality rate has been found include coronary artery bypass grafting and paediatric cardiac surgery. Equivocal results have been reported for elective abdominal aneurysm repair and carotid endarterectomy. No relationship has been found for pneumonectomy, cholecystectomy and hip arthroplasty.

There have been far fewer studies on medical conditions and no consistent conclusions can be drawn.

Relationship between length of stay and case fatality rates

There have been few studies of the relationship between length of stay and case fatality rate. Most have been done on patients with cardiovascular disease and pneumonia and no relationship has been found.

Relationship between other factors and case fatality rates

A number of studies have been done to identify predictors and causes of case fatality. Not surprisingly, in the studies on general patients the main factors included age and diagnoses of cancer and cardiovascular disease.

Organisational factors during the in-patient stay that may be of importance include total hours of nursing and other professional care and the use of specialized units. After discharge, factors such as the quality of a nursing home setting, follow up care, distance from the hospital and social support mechanisms may be important.

For almost all the single diseases that were studied the severity of the condition was a major influence on case fatality rates.

For cardiovascular conditions, on which most work has been done, additional factors include a variety of clinical complications. Women are reported to have significantly worse case fatality rates after coronary bypass surgery.

For other surgical conditions the most important additional risk factors were co-morbidities and emergency rather than elective admission.

Compilation of case fatality rates

Very little work has been done to date on the merits of different methods of compiling case fatality rates. The work currently being done by NCHOD will be a major contribution to the subject.

Because of the different geographical patterns of where people die it is essential that case fatality rates are calculated for a specific period after admission and not as in-hospital rates.

Issues relevant to the compilation of condition-specific case fatality rates that need further work are:

- Include all admissions or emergency admissions only.
- Use deaths from any cause or use condition-specific deaths that are either the underlying cause or that appear anywhere on a death certificate.
- Include deaths up to 30, 90 or 365 days post-admission or operation.
- Include adjustments for age and sex.
- The most appropriate severity adjustment measures to use.

Because deaths after admission for many conditions are rare events, the calculation of comparative case fatality rates for these conditions also raises issues of statistical power.

REFERENCES

Abrahamsson, P., Dellborg, M., Rosengren, A., and Wilhelmsen, L. (1998). Improved long-term prognosis after myocardial infarction 1984-1991. *Eur. Heart J.* **19**, 1512-1517.

Allison, J. J., Kiefe, C. I., Weissman, N. W., Person, S. D., Rousculp, M., Canto, J. G., Bae, S., Williams, O. D., Farmer, R., and Centor, R. M. (2000). Relationship of hospital teaching status with quality of care and mortality for Medicare patients with acute MI. *J. Am. Med. Assoc.* **284**, 1256-1262.

Ansari, M. Z., Ackland, M. J., Jolley, D. J., Carson, N., and McDonald, I. G. (1999). Inter-hospital comparison of mortality rates. *International Journal for Quality in Health Care* **11** , 29-35.

Ansari, M. Z., Costello, A. J., Ackland, M. J., Carson, N., and McDonald, I. G. (2000b). In-hospital mortality after transurethral resection of the prostate in Victorian public hospitals. *Aust. N.Z. J. Surg.* **70**, 204-208.

Ansari, M. Z., Collopy, B. T., Hart, W. G., Carson, N. J., and Chandraraj, E. J. (2000a). In-hospital mortality and associated complications after bowel surgery in Victorian public hospitals. *Aust. N.Z. J. Surg.* **70**, 6-10.

Arboix, A., Garcia, E. L., Massons, J., Oliveres, M., and Targa, C. (2000). Acute stroke in very old people: Clinical features and predictors of in- hospital mortality. *J. Am. Geriatr. Soc.* **48**, 36-41.

Ballard, D. J., Bryant, S. C., O'Brien, P. C., Smith, D. W., Pine, M. B., and Cortese, D. A. (1994). Referral selection bias in the Medicare hospital mortality prediction model: are centers of referral for Medicare beneficiaries necessarily centers of excellence? *Health Serv. Res.* **28**, 771-784.

Begg, C. B., Cramer, L. D., Hoskins, W. J., and Brennan, M. F. (1998). Impact of hospital volume on operative mortality for major cancer surgery. *Journal of the American Medical Association* **280**, 1747-1751.

Beguin, C., France, F. R., and Ninane, J. (1997). Systematic analysis of in-patients' circumstances and causes of death: A tool to improve quality of care. *Int. J. Qual. Health Care* **9**, 427-433.

Benenson, R., Magalski, A., Cavanaugh, S., and Williams, E. (1999). Effects of a pneumonia clinical pathway on time to antibiotic treatment, length of stay, and mortality. *Acad. Emerg. Med.* **6**, 1243-1248.

Bernard, A. M., Hayward, R. A., Rosevear, J., Chun, H., and McMahon, L. F. (1996). Comparing the hospitalizations of transfer and non-transfer patients in an academic medical center. *Acad. Med.* **71**, 262-266.

Best, W. R. and Cowper, D. C. (1994). The ratio of observed-to-expected mortality as a quality of care indicator in non-surgical VA patients. *Med. Care* **32**, 390-400.

Birkmeyer, J. D., Finlayson, S. R. G., Tosteson, A. N. A., Sharp, S. M., Warshaw, A. L., and Fisher, E. S. (1999). Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. *Surgery* **125**, 250-256.

Bond, C. A., Raehl, C. L., Pitterle, M. E., and Franke, T. (1999). Health care professional staffing, hospital characteristics, and hospital mortality rates. *Pharmacotherapy* **19**, 130-138.

Brameld, K. J., Holman, C. D. 'A., Bass, A. J., Codde, J. P., and Rouse, I. L. (1998). Hospitalisation of the elderly during the last year of life: An application of record linkage in Western Australia 1985-1994. *Journal of Epidemiology and Community Health* **52**, 740-744.

Braun, B. I. (1991). The effect of nursing home quality on patient outcome. *J. Am. Geriatr. Soc.* **39**, 329-338.

Bunker, J. P. and Black, N. (1999). Differences in death rates in English hospitals. Data are inadequate basis for drawing conclusion of paper [letter]. *BMJ* **319**, 854-855.

California Office of Statewide Health Planning and Development (1996). Report of the California Hospital Outcomes Project: Hospital Specific Detailed Statistical Tables. Vol. 2. Technical Appendix. Sacramento, California: Office of Statewide Health Planning and Development.

Canto, J. G., Every, N. R., Magid, D. J., Rogers, W. J., Malmgren, J. A., Frederick, P. D., French, W. J., Tiefenbrunn, A. J., Misra, V. K., Kiefe, C. I., and Barron, H. V. (2000). The volume of primary angioplasty procedures and survival after acute myocardial infarction. National Registry of Myocardial Infarction 2 Investigators. *N. Engl. J. Med.* **342**, 1573-1580.

Carlisle, D. M., Siu, A. L., Keeler, E. B., McGlynn, E. A., Kahn, K. L., Rubenstein, L. V., and Brook, R. H. (1992). HMO vs fee-for-service care of older persons with acute myocardial infarction. *Am. J. Public Health* **82**, 1626-1630.

Casale, P. N., Jones, J. L., Wolf, F. E., Pei, Y., and Eby, L. M. (1998). Patients treated by cardiologists have a lower in-hospital mortality for acute myocardial infarction. *Journal of the American College of Cardiology* **32**, 885-889.

Chen, J., Radford, M. J., Wang, Y., Marciniak, T. A., and Krumholz, H. M. (1999b). Performance of the '100 top hospitals': what does the report card report? *Health Aff. Millwood.* **18**, 53-68.

Chen, J., Radford, M. J., Wang, Y., Marciniak, T. A., and Krumholz, H. M. (1999a). Do "America's Best Hospitals" perform better for acute myocardial infarction? . *N. Engl. J. Med.* **340**, 286-292.

Cooper, A., Hannan, E. L., Bessey, P. Q., Farrell, L. S., Cayten, C. G., Mottley, L., Lucas, C., Mullins, R. J., and Demetriades, D. (2000). An examination of the volume-mortality relationship for New York State trauma centers. *J. Trauma Inj. Infect. Crit. Care* **48**, 16-24.

Copeland, G. P., Sagar, P., Brennan, J., Roberts, G., Ward, J., Cornford, P., Millar, A., and Harris, C. (1995). Risk-adjusted analysis of surgeon performance: a 1-year study . *Br. J. Surg.* **82**, 408-411.

Cowper, P. A., Peterson, E. D., DeLong, E. R., Jollis, J. G., Muhlbaier, L. H., and Mark, D. B. (1997). Impact of early discharge after coronary artery bypass graft surgery on rates of hospital readmission and death. The Ischemic Heart Disease (IHD) Patient Outcomes Research Team (PORT) Investigators. *J. Am. Coll. Cardiol.* **30**, 908-913.

Czaplinski, C. and Diers, D. (1998). The effect of staff nursing on length of stay and mortality. *Med. Care* **36**, 1626-1638.

Daley, J., Khuri, S. F., Henderson, W., Hur, K., Gibbs, J. O., Barbour, G., Demakis, J., Irvin, G., III, Stremple, J. F., Grover, F., McDonald, G., Passaro, E., Fabri, P. J., Spencer, J., Hammermeister, K., Aust, J. B., and Oprian, C. (1997b). Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. *J. Am. Coll. Surg.* **185**, 328-340.

Daley, J., Forbes, M. G., Young, G. J., Charns, M. P., Gibbs, J. O., Hur, K., Henderson, W., and Khuri, S. F. (1997a). Validating risk-adjusted surgical outcomes: site visit assessment of process and structure. National VA Surgical Risk Study. *J. Am. Coll. Surg.* **185**, 341-351.

Dardik, A., Lin, J. W., Gordon, T. A., Williams, G. M., Perler, B. A., Panetta, T. F., Ohki, T., O'Donnell-TF, J., Ascher, E., and Sidawy, A. N. (1999). Results of elective abdominal aortic aneurysm repair in the 1990s: A population-based analysis of 2335 cases. *Journal of Vascular Surgery* **30**, 985-995.

Davis, R. B., Iezzoni, L. I., Phillips, R. S., Reiley, P., Coffman, G. A., and Safran, C. (1995). Predicting in-hospital mortality. The importance of functional status information. *Med. Care* **33**, 906-921.

Dubois, R. W., Rogers, W. H., Moxley, J. H., Draper, D., and Brook, R. H. (1987). Hospital inpatient mortality. Is it a predictor of quality? *N. Engl. J. Med.* **317**, 1674-1680.

Dudley, R. A., Johansen, K. L., Brand, R., Rennie, D. J., and Milstein, A. (2000). Selective referral to high-volume hospitals: Estimating potentially avoidable deaths. *J. Am. Med. Assoc.* **283**, 1159-1166.

Edwards, W. H., Morris-JA, J., Jenkins, J. M., Bass, S. M., and MacKenzie, E. J. (1991). Evaluating quality, cost-effective health care. Vascular database predicated on hospital discharge abstracts. *Ann. Surg.* **213**, 433-438.

Elixhauser, A., Steiner, C., Harris, D. R., and Coffey, R. M. (1998). Comorbidity measures for use with administrative data . *Med. Care* **36**, 8-27.

Ellison, L. M., Heaney, J. A., and Birkmeyer, J. D. (2000). The effect of hospital volume on mortality and resource use after radical prostatectomy. *J. Urol.* **163**, 867-869.

Esposito, T. J., Sanddal, N. D., Dean, J. M., Hansen, J. D., Reynolds, S. A., and Battan, K. (1999). Analysis of preventable pediatric trauma deaths and inappropriate trauma care in Montana. *J. Trauma* **47**, 243-251.

Fisher, E. S., Wennberg, J. E., Stukel, T. A., Skinner, J. S., Sharp, S. M., Freeman, J. L., and Gittelsohn, A. M. (2000). Associations among hospital capacity, utilization, and mortality of U.S. Medicare beneficiaries, controlling for sociodemographic factors. *Health Serv. Res.* **34**, 1351-1362.

Flanders, W. D., Shipp, C. C., FitzGerald, D. M., and Lin, L. S. (1994). Analysis of variations in mortality rates with small numbers. *Health Serv. Res.* **29**, 461-471.

Fox, K. M., Magaziner, J., Hebel, J. R., Kenzora, J. E., and Kashner, T. M. (1999). Intertrochanteric versus femoral neck hip fractures: differential characteristics, treatment, and sequelae. *J. Gerontol. A. Biol. Sci. Med. Sci.* **54**, M635-M640.

Frances, C. D., Go, A. S., Dauterman, K. W., Deosaransingh, K., Jung, D. L., Gettner, S., Newman, J. M., Massie, B. M., and Browner, W. S. (1999). Outcome following acute myocardial infarction: are differences among physician specialties the result of quality of care or case mix? *Arch. Intern. Med.* **159**, 1429-1436.

Garnick, D. W., DeLong, E. R., and Luft, H. S. (1995). Measuring hospital mortality rates: Are 30-day data enough? *Health Serv. Res.* **29**, 679-695.

Ghali, W. A., Quan, H., and Brant, R. (1998). Coronary artery bypass grafting in Canada: Hospital mortality rates, 1992-1995. *Canadian Medical Association Journal* **159**, 926-930.

Gill, L., Goldacre, M., Simmons, H., Bettley, G., and Griffith, M. (1993). Computerised linking of medical records: methodological guidelines. *J. Epidemiol. Community Health* **47**, 316-319.

Glasgow, R. E., Showstack, J. A., Katz, P. P., Corvera, C. U., Warren, R. S., and Mulvihill, S. J. (1999). The relationship between hospital volume and outcomes of hepatic resection for hepatocellular carcinoma. *Archives of Surgery* **134**, 30-35.

Glenn, L. L. and Jijon, C. R. (1999). Risk-adjusted in-hospital death rates for peer hospitals in rural and urban regions. *J. Rural. Health* **15**, 94-107.

Goldacre, M. J. (1993). Cause-specific mortality: understanding uncertain tips of the disease iceberg. *J. Epidemiol. Community Health* **47**, 491-496.

Goldberg, R. J., McCormick, D., Gurwitz, J. H., Yarzebski, J., Lessard, D., and Gore, J. M. (1998). Age-related trends in short- and long-term survival after acute myocardial infarction: a 20-year population-based perspective (1975-1995). *Am. J. Cardiol.* **82**, 1311-1317.

Goldman, R. L. and Thomas, T. L. (1994). Using mortality rates as a screening tool - The experience of the Department of Veterans Affairs. *Joint Commission Journal on Quality Improvement* **20**, 511-522.

Goldsmith, I., Lip, G. Y. H., Tsang, G., and Patel, R. L. (1999). Comparison of primary coronary artery bypass surgery in a British Indo-Asian and white Caucasian population. *European Heart Journal* **20**, 1094-1100.

Gordon, H. S. and Rosenthal, G. E. (1996). Impact of interhospital transfers on outcomes in an academic medical center. Implications for profiling hospital quality. *Med. Care* **34**, 295-309.

Gordon, H. S. and Rosenthal, G. E. (1999a). The relationship of gender and in-hospital death: increased risk of death in men. *Med. Care* **37**, 318-324.

Gordon, T. A., Bowman, H. M., Bass, E. B., Lillemoe, K. D., Yeo, C. J., Heitmiller, R. F., Choti, M. A., Burleyson, G. P., Hsieh, G., and Cameron, J. L. (1999b). Complex gastrointestinal surgery: Impact of provider experience on clinical and economic outcomes. *Journal of the American College of Surgeons* **189**, 46-56.

Gordon, T. A., Bowman, H. M., Tielsch, J. M., Bass, E. B., Burleyson, G. P., and Cameron, J. L. (1998). Statewide regionalization of pancreaticoduodenectomy and its effect on in-hospital mortality. *Ann. Surg.* **228**, 71-78.

Hand, R., Klemka, W. L., and Inczauskis, D. (1997). Mortality and length of stay as performance indicators for pneumonia in the elderly. *J. Investig. Med.* **45**, 183-190.

Hannan, E. L., Kilburn, H., O'Donnell, J. F., Lukacik, G., and Shields, E. P. (1990). Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. *JAMA* **264**, 2768-2774.

Hannan, E. L., Kumar, D., Racz, M., Siu, A. L., and Chassin, M. R. (1994). New York State's Cardiac Surgery Reporting System: Four years later. *Ann. Thorac. Surg.* **58**, 1852-1857.

Hannan, E. L., Siu, A. L., Kumar, D., Racz, M., Pryor, D. B., and Chassin, M. R. (1997b). Assessment of coronary artery bypass graft surgery performance in New York. Is there a bias against taking high-risk patients? *Med. Care* **35**, 49-56.

Hannan, E. L., Racz, M., Ryan, T. J., McCallister, B. D., Johnson, L. W., Arani, D. T., Guerci, A. D., Sosa, J., and Topol, E. J. (1997a). Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. *JAMA* **277**, 892-898.

Hannan, E. L., Popp, A. J., Tranmer, B., Fuestel, P., Waldman, J., and Shah, D. (1998b). Relationship between provider volume and mortality for carotid endarterectomies in New York State. *Stroke* **29**, 2292-2297.

Hannan, E. L., Racz, M., Kavey, R. E., Quaegebeur, J. M., and Williams, R. (1998a). Pediatric cardiac surgery: The effect of hospital and surgeon volume on in-hospital mortality. *Pediatrics* **101**, 963-969.

Harmon, J. W., Tang, D. G., Gordon, T. A., Bowman, H. M., Choti, M. A., Kaufman, H. S., Bender, J. S., Duncan, M. D., Magnuson, T. H., Lillemoe, K. D., Cameron, J. L., Brennan, M. F., Fischer, J. F., Zinner, M. J., Rhoads, J. E., Stremple, J. F., Traverso, L. W., and Lewis, F. R. (1999). Hospital volume can serve as a surrogate for surgeon volume for achieving excellent outcomes in colorectal resection. *Ann. Surg.* **230**, 404-413.

Hartz, A. J., Kuhn, E. M., Kayser, K. L., Pryor, D. P., Green, R., and Rimm, A. A. (1992). Assessing providers of coronary revascularization: a method for peer review organizations. *Am. J. Public Health* **82**, 1631-1640.

Hartz, A. J., Gottlieb, M. S., Kuhn, E. M., and Rimm, A. A. (1993). The relationship between adjusted hospital mortality and the results of peer review. *Health Serv. Res.* **27**, 765-777.

Hartz, A. J., Pulido, J. S., and Kuhn, E. M. (1997a). Are the best coronary artery bypass surgeons identified by physician surveys? *Am. J. Public Health* **87**, 1645-1648.

Hartz, A. J., Kuhn, E. M., and Krakauer, H. (1997b). The relationship of the value of outcome comparisons to the number of patients per provider. *Int. J. Qual. Health Care* **9**, 247-254.

Hayward, R. A., Bernard, A. M., Rosevear, J. S., Anderson, J. E., and McMahon-LF, J. (1993). An evaluation of generic screens for poor quality of hospital care on a general medicine service . *Med. Care* **31**, 394-402.

Henderson, J., Goldacre, M. J., Griffith, M., and Simmons, H. (1992). Recording of deaths in hospital information systems: implications for audit and outcome studies. *J. Epidemiol. Community Health* **46**, 297-299.

Hennell, T. (1999). Differences in death rates in English hospitals. Effects of admission rates may have been understated [letter]. *BMJ* **319**, 854.

Hermansson, M., Stael, v. H., and Zilling, T. (1999). Surgical approach and prognostic factors after peptic ulcer perforation. *European Journal of Surgery* **165**, 566-572.

Higginson, I. J., Astin, P., and Dolan, S. (1998). Where do cancer patients die? Ten-year trends in the place of death of cancer patients in England. *Palliative Medicine* **12**, 353-363.

Hillner, B. E., Smith, T. J., and Desch, C. E. (2000). Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. *J. Clin. Oncol.* **18**, 2327-2340.

Ho, V. (2000a). Evolution of the volume-outcome relation for hospitals performing coronary angioplasty. *Circulation* **101**, 1806-1811.

Ho, V., Hamilton, B. H., and Roos, L. L. (2000b). Multiple approaches to assessing the effects of delays for hip fracture patients in the United States and Canada. *Health Serv. Res.* **34**, 1499-1518.

Hofer, T. P. and Hayward, R. A. (1996). Identifying poor-quality hospitals. Can hospital mortality rates detect quality problems for medical diagnoses? *Med. Care* **34**, 737-753.

Horner, R. D., Matchar, D. B., Divine, G. W., and Feussner, J. R. (1995). Relationship between physician specialty and the selection and outcome of ischemic stroke patients. *Health Serv. Res.* **30**, 275-287.

Iezzoni, L. I., Shwartz, M., Ash, A. S., Hughes, J. S., Daley, J., and Mackiernan, Y. D. (1995a). Using severity-adjusted stroke mortality rates to judge hospitals. *Int. J. Qual. Health Care* **7**, 81-94.

Iezzoni, L. I., Ash, A. S., Shwartz, M., Daley, J., Hughes, J. S., and Mackiernan, Y. D. (1995b). Predicting who dies depends on how severity is measured: implications for evaluating patient outcomes. *Ann. Intern. Med.* **123**, 763-770.

Iezzoni, L. I., Ash, A. S., Shwartz, M., Daley, J., Hughes, J. S., and Mackiernan, Y. D. (1996a). Judging hospitals by severity-adjusted mortality rates: the influence of the severity-adjustment method. *Am. J. Public Health* **86**, 1379-1387.

Iezzoni, L. I., Shwartz, M., Ash, A. S., and Mackiernan, Y. D. (1996c). Predicting in-hospital mortality for stroke patients: results differ across severity-measurement methods. *Med. Decis. Making* **16**, 348-356.

Iezzoni, L. I., Shwartz, M., Ash, A. S., Hughes, J. S., Daley, J., and Mackiernan, Y. D. (1996b). Severity measurement methods and judging hospital death rates for pneumonia. *Med. Care* **34**, 11-28.

Iezzoni, L. I., Shwartz, M., Ash, A. S., and Mackiernan, Y. D. (1996d). Using severity measures to predict the likelihood of death for pneumonia inpatients. *J. Gen. Intern. Med.* **11**, 23-31.

Iezzoni, L. I., Ash, A. S., Shwartz, M., and Mackiernan, Y. D. (1997a). Differences in procedure use, in-hospital mortality, and illness severity by gender for acute myocardial infarction patients: are answers affected by data source and severity measure? *Med. Care* **35**, 158-171.

Iezzoni, L. I. (1997b). The risks of risk adjustment . *JAMA* **278**, 1600-1607.

Iezzoni, L. I., Ash, A. S., Shwartz, M., Landon, B. E., and Mackiernan, Y. D. (1998). Predicting in-hospital deaths from coronary artery bypass graft surgery. Do different severity measures give different predictions? *Med. Care* **36**, 28-39.

Jacobs, M. C., Hussain, E., Hall, M. H., and Cohn, E. (1999). Stranger in a strange land: Internists in cardiothoracic intensive care. *New Horizons: Science and Practice of Acute Medicine* **7**, 562-568.

Jarman, B., Gault, S., Alves, B., Hider, A., Dolan, S., Cook, A., Hurwitz, B., and Iezzoni, L. I. (1999). Explaining differences in English hospital death rates using routinely collected data. *BMJ* **318**, 1515-1520.

Jenkins, K. J., Newburger, J. W., Lock, J. E., Davis, R. B., Coffman, G. A., and Iezzoni, L. I. (1995). In-hospital mortality for surgical repair of congenital heart defects: preliminary observations of variation by hospital caseload. *Pediatrics* **95**, 323-330.

Johnston, S. C. (2000). Effect of endovascular services and hospital volume on cerebral aneurysm treatment outcomes. *Stroke* **31**, 111-117.

Jones, R. H., Hannan, E. L., Hammermeister, K. E., DeLong, E. R., O'Connor, G. T., Luepker, R. V., Parsonnet, V., and Pryor, D. B. (1996). Identification of preoperative variables needed for risk adjustment of short-term mortality after coronary artery bypass graft surgery. *J. Am. Coll. Cardiol.* **28**, 1478-1487.

Kahn, K. L., Rogers, W. H., Rubenstein, L. V., Sherwood, M. J., Reinisch, E. J., Keeler, E. B., Draper, D., Kosecoff, J., and Brook, R. H. (1990). Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA* **264**, 1969-1973.

Keeler, E. B., Rubenstein, L. V., Kahn, K. L., Draper, D., Harrison, E. R., McGinty, M. J., Rogers, W. H., and Brook, R. H. (1992). Hospital characteristics and quality of care. *JAMA* **268**, 1709-1714.

Kendrick, S. and Clarke, J. (1993). The Scottish Record Linkage System. *Health Bull. Edinb.* **51**, 72-79.

Khuri, S. F., Daley, J., Henderson, W. *et al.* (1998). The Department of Veterans Affairs' NSQIP: The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. *Ann. Surg.* **228**, 491-507.

Khuri, S. F., Daley, J., Henderson, W., Hur, K., Gibbs, J. O., Barbour, G., Demakis, J., Irvin, G. III, Stremple, J. F., Grover, F., McDonald, G., Passaro, E., Fabri, P. J., Spencer, J., Hammermeister, K., and Aust, J. B. (1997). Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. *J. Am. Coll. Surg.* **185**, 315-327.

Khuri, S. F., Daley, J., Henderson, W., Hur, K., Hossain, M., Soybel, D., Kizer, K. W., Aust, J. B., Bell-RH, J., Chong, V., Demakis, J., Fabri, P. J., Gibbs, J. O., Grover, F., Hammermeister, K., McDonald, G., Passaro, E., Phillips, L., Scamman, F., Spencer, J., and Stremple, J. F. (1999). Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. *Ann. Surg.* **230**, 414-429.

Knaus, W. A., Draper, E. A., Wagner, D. P., and Zimmerman, J. E. (1986). An evaluation of outcome from intensive care in major medical centers. *Ann. Intern. Med.* **104**, 410-418.

Kreder, H. J., Williams, J. I., Jaglal, S., Hu, R., Axcell, T., and Stephen, D. (1998). Are complication rates for elective primary total hip arthroplasty in Ontario related to surgeon and hospital volumes? A preliminary investigation. *Can. J. Surg.* **41**, 431-437.

Lai, S. M., Alter, M., Friday, G., Lai, S. L., and Sobel, E. (1999). Stroke survival after discharge from an acute-care hospital. *Neuroepidemiology* **18**, 210-217.

Landon, B., Iezzoni, L. I., Ash, A. S., Shwartz, M., Daley, J., Hughes, J. S., and Mackiernan, Y. D. (1996). Judging hospitals by severity-adjusted mortality rates: the case of CABG surgery. *Inquiry* **33**, 155-166.

Leyland, A. H. and Boddy, F. A. (1998). League tables and acute myocardial infarction. *Lancet* **351**, 555-558.

Localio, A. R., Hamory, B. H., Fisher, A. C., and TenHave, T. R. (1997). The public release of hospital and physician mortality data in Pennsylvania. A case study. *Med. Care* **35**, 272-286.

Longo, W. E., Virgo, K. S., Johnson, F. E., Oprian, C. A., Vernava, A. M., Wade, T. P., Phelan, M. A., Henderson, W. G., Daley, J., and Khuri, S. F. (2000). Risk factors for morbidity and mortality after colectomy for colon cancer. *Dis. Colon Rectum* **43**, 83-91.

Louis, D. Z., Yuen, E. J., Braga, M., Cicchetti, A., Rabinowitz, C., Laine, C., and Gonnella, J. S. (1999). Impact of a DRG-based hospital financing system on quality and outcomes of care in Italy. *Health Serv. Res.* **34**, 405-415.

Luft, H. S. and Romano, P. S. (1993). Chance, continuity, and change in hospital mortality rates: Coronary artery bypass graft patients in California hospitals, 1983 to 1989. *J. Am. Med. Assoc.* **270**, 331-337.

Mahon, N. G., Codd, M. B., O'Rorke, C., Egan, B., McCann, H. A., and Sugrue, D. D. (1999). Management and outcome of acute myocardial infarction in older patients in the thrombolytic era. *J. Am. Geriatr. Soc.* **47**, 291-294.

Marshall, G., Shroyer, A. L., Grover, F. L., and Hammermeister, K. E. (1998). Time series monitors of outcomes. A new dimension for measuring quality of care. *Med. Care* **36**, 348-356.

McCloskey, J. M. (1998). Nurse staffing and patient outcomes. *Nurs. Outlook* **46**, 199-200.

McCormick, D., Fine, M. J., Coley, C. M., Marrie, T. J., Lave, J. R., Obrosky, D. S., Kapoor, W. N., and Singer, D. E. (1999). Variation in length of hospital stay in patients with community-acquired pneumonia: are shorter stays associated with worse medical outcomes? *Am. J. Med.* **107**, 5-12.

McGloin, H., Adam, S. K., and Singer, M. (1999). Unexpected deaths and referrals to intensive care of patients on general wards. Are some cases potentially avoidable? *J. R. Coll. Physicians Lond.* **33**, 255-259.

McGrath, P. D., Wennberg, D. E., Malenka, D. J., Kellett-MA, J., Ryan-TJ, J., O'Meara, J. R., Bradley, W. A., Hearne, M. J., Hettleman, B., Robb, J. F., Shubrooks, S., Verlee, P., Watkins, M. W., Lucas, F. L., and O'Connor, G. T. (1998). Operator volume and outcomes in 12,988 percutaneous coronary interventions. *J. Am. Coll. Cardiol.* **31**, 570-576.

McKee, M. and Hunter, D. (1995). Mortality league tables: Do they inform or mislead? *Qual. Health Care* **4**, 5-12.

Mehta, R. H., Ruane, T. J., McCargar, P. A., Eagle, K. A., and Stalhandske, E. J. (2000). The treatment of elderly diabetic patients with acute myocardial infarction: Insight from Michigan's cooperative cardiovascular project. *Arch. Intern. Med.* **160**, 1301-1306.

Mitchell, P. H. and Shortell, S. M. (1997). Adverse outcomes and variations in organization of care delivery. *Med. Care* **35**, NS19-NS32.

Mozes, B., Olmer, L., Galai, N., Simchen, E., and Nugent, W. C. (1998). A national study of postoperative mortality associated with coronary artery bypass grafting in Israel. *Annals of Thoracic Surgery* **66**, 1254-1263.

Mozes, B., Maor, Y., Olmer, L., and Shabtai, E. (1999). Comparing medical centers treating hip fractures in the elderly: the importance of multi-outcome measurements. *Am. J. Med. Qual.* **14**, 117-123.

Mullins, R. J., Mann, N. C., Hedges, J. R., Worrall, W., Helfand, M., Zechnich, A. D., and Jurkovich, G. J. (1998). Adequacy of hospital discharge status as a measure of outcome among injured patients. *JAMA* **279**, 1727-1731.

Munoz, E., Mulloy, K., Goldstein, J., Tenenbaum, N., and Wise, L. (1990). Costs, quality, and the volume of surgical oncology procedures. *Arch. Surg.* **125**, 360-363.

Nash, I. S., Corrato, R. R., Dlutowski, M. J., O'Connor, J. P., and Nash, D. B. (1999). Generalist versus specialist care for acute myocardial infarction. *Am. J. Cardiol.* **83**, 650-654.

Ng, T. P., Mak, K. H., Phua, K. H., and Tan, C. H. (1999). Trends in mortality, incidence, hospitalisation, cardiac procedures and outcomes of care for coronary heart disease in Singapore, 1991-1996. *Ann. Acad. Med. Singapore* **28**, 395-401.

Ni, H., Nauman, D. J., and Hershberger, R. E. (1999). Analysis of trends in hospitalizations for heart failure. *J. Card. Fail.* **5**, 79-84.

O'Connor, G. T., Birkmeyer, J. D., Dacey, L. J., Quinton, H. B., Marrin, C. A., Birkmeyer, N. J., Morton, J. R., Leavitt, B. J., Maloney, C. T., Hernandez, F., Clough, R. A., Nugent, W. C., Olmstead, E. M., Charlesworth, D. C., and Plume, S. K. (1998). Results of a regional study of modes of death associated with coronary artery bypass grafting. Northern New England Cardiovascular Disease Study Group. *Ann. Thorac. Surg.* **66**, 1323-1328.

O'Connor, G. T., Malenka, D. J., Quinton, H., Robb, J. F., Kellett-MA, J., Shubrooks, S., Bradley, W. A., Hearne, M. J., Watkins, M. W., Wennberg, D. E., Hettleman, B., O'Rourke, D. J., McGrath, P. D., Ryan, T., and Verlee, P. (1999). Multivariate prediction of in-hospital mortality after percutaneous coronary interventions in 1994-1996. *Journal of the American College of Cardiology* **34**, 681-691.

Park, R. E., Brook, R. H., Kosecoff, J., Keesey, J., Rubenstein, L., Keeler, E., Kahn, K. L., Rogers, W. H., and Chassin, M. R. (1990). Explaining variations in hospital death rates. Randomness, severity of illness, quality of care. *JAMA* **264**, 484-490.

Parry, G. J., Gould, C. R., McCabe, C. J., and Tarnow-Mordi, W. O. (1998). Annual league tables of mortality in neonatal intensive care units: longitudinal study. International Neonatal Network and the Scottish Neonatal Consultants and Nurses Collaborative Study Group. *BMJ* **316**, 1931-1935.

Patti, M. G., Corvera, C. U., Glasgow, R. E., and Way, L. W. (1998). A hospital's annual rate of esophagectomy influences the operative mortality rate. *J. Gastrointest. Surg.* **2**, 186-192.

Pearce, W. H., Parker, M. A., Feinglass, J., Ujiki, M., Manheim, L. M., Sawyer, W., Schneider, J., McCarthy, I. I. I., Blebea, J., and Hertzer, N. (1999). The importance of surgeon volume and training in outcomes for vascular surgical procedures. *Journal of Vascular Surgery* **29**, 768-778.

Pearson, M. L., Kahn, K. L., Harrison, E. R., Rubenstein, L. V., Rogers, W. H., Brook, R. H., and Keeler, E. B. (1992). Differences in quality of care for hospitalized elderly men and women. *JAMA* **268**, 1883-1889.

Pell, J. P., Sirel, J., Marsden, A. K., and Cobbe, S. M. (2000). Sex differences in outcome following community-based cardiopulmonary arrest. *Eur. Heart J.* **21**, 239-244.

Philbin, E. F. and Roerden, J. B. (1997b). Longer hospital length of stay is not related to better clinical outcomes in congestive heart failure. *American Journal of Managed Care* **3**, 1285-1291.

Philbin, E. F., Rogers, V. A., Sheesley, K. A., Lynch, L. J., Andreou, C., and Rocco-TA, J. (1997a). The relationship between hospital length of stay and rate of death in heart failure. *Heart Lung* **26**, 177-186.

Philbin, E. F., Weil, H. F., Erb, T. A., and Jenkins, P. L. (1999). Cardiology or primary care for heart failure in the community setting: process of care and clinical outcomes. *Chest* **116**, 346-354.

Piette, J. D. and Moos, R. H. (1996). The influence of distance on ambulatory care use, death, and readmission following a myocardial infarction. *Health Serv. Res.* **31**, 573-591.

Polanczyk, C. A., Rohde, L. E. P., Dec, G. W., and Disalvo, T. (2000). Ten-year trends in hospital care for congestive heart failure: Improved outcomes and increased use of resources. *Arch. Intern. Med.* **160**, 325-332.

Pollack, M. M., Cuerdon, T. T., Patel, K. M., Ruttimann, U. E., Getson, P. R., and Levetown, M. (1994). Impact of quality-of-care factors on pediatric intensive care unit mortality . *JAMA* **272**, 941-946.

Pollock, A. M. and Vickers, N. (1995). Why are a quarter of all cancer deaths in south-east England registered by death certificate only? Factors related to death certificate only registrations in the Thames Cancer Registry between 1987 and 1989. *British Journal of Cancer* **71**, 637-641.

Poloniecki, J., Valencia, O., and Littlejohns, P. (1998). Cumulative risk adjusted mortality chart for detecting changes in death rate: Observational study of heart surgery. *BMJ* **316**, 1697-1700.

Poses, R. M., McClish, D. K., Smith, W. R., Huber, E. C., Clemo, F. L., Schmitt, B. P., Alexander, D., Racht, E. M., and Colenda, C. C., III (2000). Results of report cards for patients with congestive heart failure depend on the method used to adjust for severity. *Ann. Intern. Med.* **133**, 10-20.

Pronovost, P., Dorman, T., Sadovnikoff, N., Garrett, E., Breslow, M., and Rosenfeld, B. (1999). The association between preoperative patient characteristics and both clinical and economic outcomes after abdominal aortic surgery. *Journal of Cardiothoracic and Vascular Anesthesia* **13**, 549-554.

Pronovost, P. J., Jenckes, M. W., Dorman, T., Garrett, E., Breslow, M. J., Rosenfeld, B. A., Lipsett, P. A., and Bass, E. (1999). Organizational characteristics of intensive care units related to outcomes of abdominal aortic surgery. *JAMA* **281**, 1310-1317.

Rios, G., Conrad, A., Cole, D., Adams, D., Leveen, M., O'Brien, P., and Baron, P. (1999). Trends in indications and outcomes in the whipple procedure over a 40- year period. *American Surgeon* **65**, 889-893.

Rockall, T. A., Logan, R. F., Devlin, H. B., and Northfield, T. C. (1995). Variation in outcome after acute upper gastrointestinal haemorrhage. The National Audit of Acute Upper Gastrointestinal Haemorrhage. *Lancet* **346**, 346-350.

Romano, P. S. and Chan, B. K. (2000). Risk-adjusting acute myocardial infarction mortality: Are APR-DRGs the right tool? *Health Serv. Res.* **34**, 1469-1489.

Roos, L. L. and Brazauskas, R. (1990). Outcomes and quality assurance: facilitating the use of administrative data. *Qual. Assur. Health Care* **2**, 77-88.

Roques, F., Nashef, S. A., Michel, P., Gauducheau, E., de Vincentiis, C., Baudet, E., Cortina, J., David, M., Faichney, A., Gabrielle, F., Gams, E., Harjula, A., Jones, M. T., Pintor, P. P., Salamon, R., and Thulin, L. (1999). Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. *Eur. J. Cardiothorac. Surg.* **15**, 816-822.

Rosenthal, G. E., Harper, D. L., Quinn, L. M., and Cooper, G. S. (1997a). Severity-adjusted mortality and length of stay in teaching and nonteaching hospitals. Results of a regional study. *JAMA* **278**, 485-490.

Rosenthal, G. E. (1997b). Weak associations between hospital mortality rates for individual diagnoses: implications for profiling hospital quality. *Am. J. Public Health* **87**, 429-433.

Rosenthal, G. E., Shah, A., Way, L. E., and Harper, D. L. (1998). Variations in standardized hospital mortality rates for six common medical diagnoses: implications for profiling hospital quality. *Med. Care* **36**, 955-964.

Rosenthal, G. E., Baker, D. W., Norris, D. G., Way, L. E., Harper, D. L., and Snow, R. J. (2000). Relationships between in-hospital and 30-day standardized hospital mortality: implications for profiling hospitals. *Health Serv. Res.* **34**, 1449-1468.

Rubenstein, L. V., Kahn, K. L., Reinisch, E. J., Sherwood, M. J., Rogers, W. H., Kamberg, C., Draper, D., and Brook, R. H. (1990). Changes in quality of care for five diseases measured by implicit review, 1981 to 1986. *JAMA* **264**, 1974-1979.

Salinas, A. M., Coria, I., Reyes, H., and Zambrana, M. (1997). Effect of quality of care on preventable perinatal mortality. *Int. J. Qual. Health Care* **9**, 93-99.

Scott, I. A., Eyeson-Annan, M. L., Huxley, S. L., and West, M. J. (2000). Optimising care of acute myocardial infarction: Results of a regional quality improvement project. *J. Qual. Clin. Pract.* **20**, 12-19.

Shepardson, L. B., Youngner, S. J., Speroff, T., and Rosenthal, G. E. (1999). Increased risk of death in patients with do-not-resuscitate orders. *Med. Care* **37**, 727-737.

Silber, J. H., Williams, S. V., Krakauer, H., and Schwartz, J. S. (1992). Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. *Med. Care* **30**, 615-629.

Silber, J. H. and Rosenbaum, P. R. (1997a). A spurious correlation between hospital mortality and complication rates: the importance of severity adjustment. *Med. Care* **35**, OS77-OS92.

Silber, J. H., Rosenbaum, P. R., Williams, S. V., Ross, R. N., and Schwartz, J. S. (1997b). The relationship between choice of outcome measure and hospital rank in general surgical procedures: implications for quality assessment. *Int. J. Qual. Health Care* **9**, 193-200.

Silvestri, G. A., Handy, J., Lackland, D., Corley, E., and Reed, C. E. (1998). Specialists achieve better outcomes than generalists for lung cancer surgery. *Chest* **114**, 675-680.

Sirio, C. A., Shepardson, L. B., Rotondi, A. J., Cooper, G. S., Angus, D. C., Harper, D. L., and Rosenthal, G. E. (1999). Community-wide assessment of intensive care outcomes using a physiologically based prognostic measure: implications for critical care delivery from Cleveland Health Quality Choice. *Chest* **115**, 793-801.

Smith, T. (1993). Perioperative deaths--detection by linkage of hospital discharge and death registration sub-files. *Health Bull. Edinb.* **51**, 299-309.

Sollano, J. A., Gelijns, A. C., Moskowitz, A. I., Heitjan, D. F., Cullinane, S., Saha, T., Chen, J. M., Roohan, P. J., Reemtsma, K., Shields, E. P., Matloff, J. M., Spencer, F. C., Smith, C. R., and Marrin, C. (1999). Volume-outcome relationships in cardiovascular operations: New York State, 1990-1995. *Journal of Thoracic and Cardiovascular Surgery* **117**, 419-430.

Sosa, J. A., Bowman, H. M., Gordon, T. A., Bass, E. B., Yeo, C. J., Lillemoe, K. D., Pitt, H. A., Tielsch, J. M., Cameron, J. L., Howard, J. M., Warshaw, A. L., and

Jordan-PH, J. (1998). Importance of hospital volume in the overall management of pancreatic cancer. *Ann. Surg.* **228**, 429-438.

Sowden, A., Aletras, V., Place, M., Rice, N., Eastwood, A., Grilli, R., Ferguson, B., Posnett, J., and Sheldon, T. (1997). Volume of clinical activity in hospitals and healthcare outcomes, costs and patient access. *Qual. Health Care* **6**, 109-114.

Sowden, A. J., Deeks, J. J., and Sheldon, T. A. (1995). Volume and outcome in coronary artery bypass graft surgery: true association or artefact? *BMJ* **311**, 151-155.

Sowden, A. J. and Sheldon, T. A. (1998). Does volume really affect outcome? Lessons from the evidence. *Journal of Health Services Research and Policy* **3**, 187-190.

Taylor, D.H., Whellan, D. J., and Sloan, F. A. (1999). Effects of admission to a teaching hospital on the cost and quality of care for Medicare beneficiaries. *N. Engl. J. Med.* **340**, 293-299.

Thiemann, D. R., Coresh, J., Oetgen, W. J., and Powe, N. R. (1999). The association between hospital volume and survival after acute myocardial infarction in elderly patients. *N. Engl. J. Med.* **340**, 1640-1648.

Thomas, J. W. (1991). Validating risk-adjusted outcomes as measures of quality of care in hospitals. Report prepared for the Minnesota Coalition on Health. Ann Arbor, Michigan: University of Michigan, School of Public Health, Department of Health Services Management and Policy.

Thomas, J. W., Holloway, J. J., and Guire, K. E. (1993). Validating risk-adjusted mortality as an indicator for quality of care. *Inquiry* **30**, 6-22.

Thomas, J. W. and Hofer, T. P. (1998). Research evidence on the validity of risk-adjusted mortality rate as a measure of hospital quality of care. *Med. Care Res. Rev.* **55**, 371-404.

Thomas, J. W. and Hofer, T. P. (1999). Accuracy of risk-adjusted mortality rate as a measure of hospital quality of care. *Med. Care* **37**, 83-92.

Tolle, S. W., Rosenfeld, A. G., Tilden, V. P., and Park, Y. (1999). Oregon's low in-hospital death rates: What determines where people die and satisfaction with decisions on place of death? *Annals of Internal Medicine* **130**, 681-685.

Treurniet, H. F., Looman, C. W., van der Maas, P. J., and Mackenbach, J. P. (1999). Variations in 'avoidable' mortality: a reflection of variations in incidence? *Int. J. Epidemiol.* **28**, 225-232.

Urschel, J. D. and Urschel, D. M. (2000). The hospital volume-outcome relationship in general thoracic surgery. Is the surgeon the critical determinant? *J. Cardiovasc. Surg.* **41**, 153-155.

Velasco, F. T., Ko, W., Rosengart, T., Altorki, N., Lang, S., Gold, J. P., Krieger, K. H., and Isom, O. W. (1996). Cost containment in cardiac surgery: results with a critical pathway for coronary bypass surgery at the New York hospital-Cornell Medical Center. *Best Pract. Benchmarking. Healthc.* **1**, 21-28.

Ward, M. M. (1999). Hospital experience and mortality in patients with systemic lupus erythematosus. *Arthritis and Rheumatism* **42**, 891-898.

Wennberg, J. E., Freeman, J. L., Shelton, R. M., and Bubolz, T. A. (1989). Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *N. Engl. J. Med.* **321**, 1168-1173.

Whittle, J., Lin, C. J., Lave, J. R., Fine, M. J., Delaney, K. M., Joyce, D. Z., Young, W. W., and Kapoor, W. N. (1998). Relationship of provider characteristics to outcomes, process, and costs of care for community-acquired pneumonia. *Med. Care* **36**, 977-987.

Williams, R. L. (1979). Measuring the effectiveness of perinatal medical care. *Med. Care* **17**, 95-110.

Wolinsky, F. D., Wyrwich, K. W., and Gurney, J. G. (1999). Gender differences in the sequelae of hospitalization for acute myocardial infarction among older adults. *J. Am. Geriatr. Soc.* **47**, 151-158.

Yuan, Z., Cooper, G. S., Einstadter, D., Cebul, R. D., and Rimm, A. A. (2000). The association between hospital type and mortality and length of stay: a study of 16.9 million hospitalized Medicare beneficiaries. *Med. Care* **38**, 231-245.

Zalkind, D. L. and Eastaugh, S. R. (1997). Mortality rates as an indicator of hospital quality. *Hospital and Health Services Administration* **42**, 3-15.