

**Relationship Between Volume and Quality of
Health Care:
A Review of the Literature**

July 1995

NHS CENTRE FOR REVIEWS AND DISSEMINATION

The NHS Centre for Reviews and Dissemination (CRD) is a facility commissioned by the NHS Research and Development Division. Its aim is to identify and review the results of good quality health research and to disseminate actively the findings to key decision makers in the NHS and to consumers of health care services. In this way health care professionals and managers can ensure their practice reflects the best available research evidence.

The reviews will cover

- the effectiveness of care for particular conditions;
- the effectiveness of health technologies; and
- evidence on efficient methods of organising and delivering particular types of health care.

Further Information

If you would like further information please contact:

General Enquiries:	01904 433634
Information Service	01904 433707
Publications:	01904 433648
Fax:	01904 433661
Email:	revdis@york.ac.uk

Publications

CRD Report 1	Which Way Forward for the Care of Critically Ill Children?	£7.50
CRD Report 2	Relationship Between Volume and Quality of Health Care: A Review of the Literature	£5.00

This report is a review of the literature which has examined the relationship between hospital volume and outcome in different procedures and diagnoses. It was carried out by the NHS Centre for Reviews and Dissemination at the University of York. The aim of this report is to provide information to health care decision makers which can be taken into account when making policy. In particular, it is hoped that it will contribute to the debate over the regionalisation of services in the United Kingdom. This review was carried out by Amanda Sowden, Jon Deeks, Ian Watt and Trevor Sheldon.

ACKNOWLEDGEMENTS

We would like to thank the following for useful comments: Dr Iain Chalmers, UK Cochrane Centre; Dr Mike Chambers, Department of Epidemiology, London Hospital Medical College; Nick Freemantle, Centre for Health Economics, University of York; Professor Alan Maynard, Centre for Health Economics, University of York; Dr Nick Phinn, United Health, Humberside; Dr John Posnett, York Health Economics Consortium, University of York; Dr Charles Stiller, Childhood Cancer Research Group, Department of Paediatrics, University of Oxford and Professor Alan Williams, Centre for Health Economics, University of York

We would also like to thank the following for their contributions to the review: Dr Nick Black, Health Services Research Unit, London School of Hygiene and Tropical Medicine; Dr D Field, Department of Child Health, University of Leicester; Dr Maggie Harding, Department of Public Health Sciences, St George's Hospital Medical School; Dr Jeremy Jones, Nuffield Community Care Studies Unit, University of Leicester; Professor Harold Luft, Institute for Health Policy Studies, University of California; Ministry of Health, New Zealand; Dr Kathy Rowan, Intensive Care National Audit and Research Centre and Dr William Tarnow-Mordi, Department of Child Health, Ninewells Hospital and Medical School, Dundee.

We would also like to thank The Yorkshire Collaborating Centre for Health Services Research for providing support for part of this work and to E.B. Sarawitz for assistance in editing.

CONTENTS

Abstract	5
Introduction	8
Methods of the review	9
Reviewing the evidence	10
Reviews of specific diseases and procedures	11
Problems of interpreting the evidence	13
Direction of causality	14
Adjustment for patient case-mix	15
Evidence for a volume-outcome in CABG surgery	15
Analysis of the volume-outcome data	17
Statistical methods	18
Results	21
The importance of adjustment in observational data	23
Conclusions, research and policy implications	24
References	28

Abstract

Objectives

To carry out a systematic review of the literature in order to examine the evidence for a relationship between the volume of activity of a hospital and patient outcomes in a range of procedures.

Data sources

Search of MEDLINE (from 1985 to 1994) and of the Science Citation Index on the Bath Information and Data Service (BIDS) (from 1993 to 1994). Both letters and editorials were excluded. The reference lists of identified articles were also searched. Key relevant journals were also hand-searched: Medical Care from 1971 to 1994 and Health Policy from 1986 to 1994. Researchers in the field in the United Kingdom and the United States were consulted to identify published and unpublished evidence.

Study selection

Studies were included if they empirically assessed a relationship between the frequency with which procedures were carried out, or patients with a particular diagnosis were treated, and health outcomes such as mortality or morbidity. Individual studies were judged to be relevant by one reviewer. Data extraction was checked by another reviewer.

Data extraction

For studies of coronary artery bypass graft surgery (CABG) which examined the effect of volume on outcome, details of the source of data and the date of data collection, number of hospitals and patients, volume cut-off points used, statistical adjustments made in the analyses, together with the relationship found, were recorded.

Data synthesis

A narrative overview of most of the evidence was carried out. In the case of CABG surgery a meta-analysis by regression was used. The studies were pooled in order to assess the degree to which the estimates of the effect of volume on outcome were affected by the degree of adjustment for case-mix. Factors adjusted for included age, sex, severity of diagnoses and comorbidity. In addition the analysis investigated the relationship between the estimates of volume effect and year of data.

Results

There is a large literature, primarily from North America which has examined the relationship between volume and outcome for a wide range of (mainly surgical) procedures. Almost all of the research in this area has used an observational study design

with hospital mortality rates as the measure of quality. The methods used in each study differed considerably and ranged from simple comparisons of outcomes by hospital volume to complex simultaneous-equation models. A wide range of levels of volume of activity have been used to categorise hospitals as high or low volume.

A positive relationship between volume and outcome (where higher volumes were associated with lower mortality rates) has been reported for most procedures: abdominal aortic aneurysm, vascular surgery, biliary tract surgery, cardiac catheterization and angiography, CABG surgery, total hip replacement, prostatectomy, intestinal operations, hysterectomy, acute myocardial infarction, perinatal illness, hernia and trauma care. In contrast, studies of treatment for fracture of the femur, appendicectomy, stomach operations and burn care have been reported to have either no relationship or a negative relationship (where higher volumes were associated with higher mortality rates) between volume and outcome.

Few of the studies adjusted for the effects of case-mix differences between high and low volume hospitals. The 15 identified studies which examined CABG surgery differed with respect to the extent of adjustment for confounding. The analysis carried out suggests that the greater the adjustment for the effects of case-mix, the smaller the size of the estimate of benefit associated with increased volumes of activity.

The available studies did not shed much light on the possible causes of any volume-outcome relationship and the level (e.g. hospital, ward, clinician) at which it might operate. Thus it was difficult to disentangle any direction of cause i.e. whether increased volume may have generated better outcomes or if better units attracted more patients.

Conclusions

Whilst most studies report a positive relationship between hospital volume and outcome for several elective procedures, they may have biased estimates of the size of the effect of volume because of poor adjustment for the effect of differences in case-mix between high and low volume hospitals. There is also a paucity of information as to why high volume might be associated with better outcomes. For example, there is little evidence whether clinician-related elements such as better judgement about high and low risk patients may be related to volume or whether high volume clinicians keep up with the literature and develop an effective protocol of practice. If the latter were true then the adoption of clinical guidelines by other clinicians may improve quality of care and close the gap between smaller and larger units. Alternatively, there may be institution related elements such as the level of facilities which produce better outcomes. There is also little evidence as to whether merging hospitals to create larger units will result in a change in outcomes

over time. Because of the uncertainty in both the size and the interpretation of any effect of volume reported, caution should be exercised in using this research literature to justify policies of reorganisation of health care delivery. Therefore, the main recommendation is that policy-makers should be cautious when invoking the assumed improvements in outcome achieved by volume as a key argument for centralisation of services.

Introduction

A policy of regionalisation – the concentration of hospital services in larger units – has been proposed as a means of reducing the cost of treatment while improving the quality of health care. Benefits to patients, it is suggested, can be expected both from economies of scale and from the quality of care offered by highly experienced specialists in ‘centres of excellence’.¹ These assumptions for example, underpin the recent proposal for a new three-tiered regional network for cancer services in the UK.² Similarly, in the Netherlands some operations have been regionalised by regulation.³ There, for example, coronary artery bypass graft (CABG) surgery can only be performed in licensed hospitals, in which a minimum of 600 CABG procedures must be carried out each year. In the United States, the American College of Surgeons has recommended that open-heart surgery teams perform at least 150 operations per year so that the skills required for such complicated procedures can be maintained, developed and improved.⁴

However, the number of patients treated – hospital volume – is only one element likely to affect quality of care in hospitals. Other factors may include affiliation with a medical school, staffing levels, equipment and expenditure per patient day.⁵ Factors relating to the clinician – skills, expertise – and the patient population, most notably case-mix (eg severity of illness), are also of fundamental importance in determining outcome. These influences confound estimates of the relationship between volume and outcome,⁶ but they are given little if any consideration in most of the literature. The hypotheses tested mainly focus on the gross effects of the number of cases treated, on quality of care, with some adjustment for patient case-mix, usually age and sex, plus concomitant illness. Few studies have taken into consideration the clinical risk factor of the patient’s condition at the time of treatment. A recent review that examined ‘the condition of the literature on differences in hospital mortality’ found that, out of 16 studies, only three used a severity of illness measure to adjust for differences among patients.⁷

Most of the research in this area has measured quality of care provided by units treating different numbers of patients by means of hospital mortality rates – usually extracted (in the USA) from hospital discharge abstracts. Thus, it is assumed that units with higher death rates provide poorer care.⁸ However, without clinical data – factors such as severity of the primary diagnosis, number and complexity of concomitant conditions, age distribution and demographic characteristics – it is difficult to attribute differences in mortality to variations in the quality of care^{8–10}. Patients who are younger or less severely ill can be expected to do better regardless of the size of the unit.

Such confounding is a problem with observational studies where patients are not randomly allocated to hospitals, clinicians or treatments. Statistical adjustment in order to control the effects of confounding is frequently used. However, it is impossible to be sure how adequately the effects of confounding factors have been taken into account.^{11 12}

One of the main objectives of this paper is to assess how well each volume–outcome study was controlled for patient case-mix. Without adequate adjustment, comparisons of hospital performance have little meaning and any conclusions may therefore be misleading. The methods used in the review are summarised in the box.

Methods of the review

The review was based on a search of MEDLINE (1985–1994) and of the Science Citation Index on the Bath Information and Data Service (BIDS) (1993–1994). Letters and editorials were excluded.

The reference lists of identified articles were also searched. Key relevant journals were also hand-searched: Medical Care (1971–1994) and Health Policy (1986–1994). Other health services researchers both in the United Kingdom and in the United States were consulted to identify published and unpublished evidence.

The review of the literature covers aspects of care – procedures and diagnoses – where there is a significant literature, and only includes studies that measured health outcomes such as mortality or morbidity. Given the constraints of resources and time, and the large literature, the authors have chosen to concentrate particularly on studies of a single procedure, coronary artery bypass graft surgery (CABG).

Relevance of individual studies was assessed by one reviewer. Data extraction was checked by another reviewer.

The methods used in the studies ranged from simple comparisons of outcomes by hospital volume to complex simultaneous-equation models. A hierarchy of adjustment for case-mix was developed to explore the relationship between the degree of adjustment for confounding factors and the estimate of the relationship between volume and outcome.

Most of the research was carried out in North America. In-patient mortality is generally used as the measure of quality, although other factors – post-hospitalisation mortality

within a fixed time period, morbidity, length of hospital stay – are occasionally used. For most procedures, there is general agreement throughout the literature that outcome (mortality) is consistently better in higher-volume hospitals.^{3,13-37}

Reviewing the evidence

In 1988 more than 100 papers were used to examine the relationship between volume and outcome.¹⁵ Only 26 of these were considered to be methodologically sound (on the basis that they included a sufficient number of hospitals and cases to offer statistically valid results or that their main objective was to examine the volume–outcome relationship). In 1990 an updated and expanded version of this review was published.¹⁶ While some of the papers included in the later review focused on only one procedure, the majority considered a number of procedures.

A consistent relationship between volume and outcome, with worse outcomes at lower volumes, was reported for: abdominal aortic aneurysm (weakening of the aorta wall), vascular surgery, biliary tract surgery, appendicectomy, cardiac catheterisation and angiography, CABG surgery, total hip replacement, prostatectomy, intestinal operations, hysterectomy, treatment for acute myocardial infarction (heart attack), neonatal intensive care and hernia repair. In contrast, studies of treatment for fracture of femur and stomach operations suggested a negative relationship between volume and outcome.

Another review, also published in 1990,¹⁷ considered six main areas: neonatal intensive care, surgery, coronary care (acute myocardial infarction), cardiac catheterisation, trauma care and burns care. No clear association between volume and outcome was found for coronary care, neonatal intensive care or burns care. However, a positive relationship was reported for some surgical procedures, cardiac catheterisation and care of the severely injured. Methodological issues, such as patient case-mix and the difficulty in interpreting the data, were discussed. Despite such problems the authors concluded that the evidence was sufficient to support the concentration of services for some surgery, cardiac catheterisation and specialised trauma care.

One recent review¹⁸ of outcomes following surgery included studies where factors such as hospital characteristics, population profiles and referral preferences were examined in addition to individual surgeon volume and hospital volume. Overall, hospital volume favoured outcome whereas surgeon volume was more ambiguous. The review concluded that there is evidence of differences in the outcome of some surgical procedures, but that the cause of these differences is not clear.

A review published in 1994,¹⁹ was carried out as part of an advisory process for achieving optimum provision of tertiary services in New Zealand. Areas specified included: cardiac surgery, major trauma, neurosurgery, care of acute spinal cord injury, general intensive care, renal transplantation, bone marrow transplantation and neonatal intensive care. The authors reported positive, significant benefits of concentration of services for cardiac surgery, neurosurgery, oncology (cancer) services, transplantation and general intensive care. However, consideration of the effects of potential confounding factors was very limited.

Reviews of specific diseases and procedures

Another approach has been to focus on reported outcomes by treatment.^{3,13,14} Eight studies of CABG surgery, carried out in the United States^{20–22,26,27,29,31,32} between 1979 and 1989, were reviewed in the Netherlands³ to examine the relationship between volume and outcome for this procedure. The authors reported a positive relationship between the number of procedures carried out in a single centre and outcome, although they acknowledged that few of the studies included clinical risk factors, such as patient case-mix or severity of illness, and their likely effect on mortality. These findings led to a proposal for the centralisation of CABG services .

Solid organ – kidney, heart, liver, pancreas – transplantation has also been the subject of a review of the relationship between the number of operations performed and patient survival.¹³ However, the few studies that were available were methodologically very weak, in particular lacking any information on how patient case-mix was dealt with. Despite this, reports of improved outcomes at higher volumes were seen as confirming the 'practice makes perfect' hypothesis (increased experience results in improved outcomes): a policy to regionalise transplantation services was recommended.

A major recent review of oncology services in the UK¹⁴ used observational data to examine whether centralised treatment and/or entry to a clinical trial improved survival. The following sites were included: stomach (2 studies), colorectal (5 studies), lung (3 studies), breast (5 studies), cervix (4 studies), ovarian (7 studies), prostate (3 studies), testis (3 studies), Hodgkin's disease (2 studies), miscellaneous sites (6 studies) and childhood cancers (13 studies). Only eight of the 53 studies presented information on the number of cancer patients treated at the centre or hospital. The review concluded that patients treated at major centres dealing with larger numbers of cases, or at teaching hospitals and other specialist centres, had improved survival rates and that there was no evidence that centralised referral led to increased mortality.

However, it is possible that the differences in patient outcomes between centres may reflect variations in prognostic factors such as the severity of illness of patients treated. One unit may have had a high proportion of elderly patients with more advanced disease who were admitted as emergencies, whereas another centre may have dealt mainly with younger patients with less advanced disease. The effects of clinical variables are illustrated in a recent study³⁸ which used a prospective design to assess post-operative complications and mortality and survival rates achieved by surgeons in patients with colorectal cancer. Post-operative mortality rates ranged from 0% to 20%. However, other factors – age, emergency admission, pre-existing heart or respiratory disease – were also identified as being associated with post-operative mortality. After statistical adjustment was made for these elements the relative performance of surgeons changed. While some of the variations in crude measures of patient survival reflected differences in surgeons' performance, they also reflected differences in the patient population. Importantly, surgeons who undertook fewer procedures tended to deal with a higher proportion of older patients who had been admitted as emergencies and had more advanced disease. In contrast, surgeons who dealt with larger numbers of patients, tended to have a higher proportion of younger, fitter patients, who had smaller tumours removed during an elective procedure.

Overall, this suggests that the results of studies which are inadequately controlled for clinical factors may be biased. On the basis of the evidence cited, it would be incautious to argue for the existence of a clear-cut positive volume–outcome relationship in cancer services.

Estimates of the size of the positive relationship between volume and outcome differed in relative and absolute terms. For example, for open-heart surgery, vascular surgery, transurethral resection of the prostate (TURP) and CABG surgery, hospitals in which at least 200 operations were carried out annually had adjusted death rates 25% to 41% lower than hospitals where less than 200 procedures were carried out each year.²⁰ In hospitals performing less than 100 CABG procedures per year, the adjusted death rate was 5.3% compared with 3.5% for hospitals with over 200 procedures per year. This suggests that 3.6 lives may be saved per 200 operations if patients are treated in high-volume hospitals.²¹ In one study an increase in CABG surgery volume from 20 to 80 procedures per year was estimated to reduce the predicted death rate from 6.7% to 5.6%.²⁷

Using the results of observational studies, 33% of all in-patient deaths for total hip replacement could have been prevented if the operation had been carried out in hospitals doing more than 50 procedures per year (actual death rate was 19% for low volume and 11% for high volume), a saving of 80 lives per 1000 operations.²⁰ For TURP an increase

in the number of operations per year from 20 to 100 was associated with a decrease in the predicted death rate from 2.7% to 2.3%, a 14.8% decrease. This decline is not very substantial in absolute terms (4 deaths prevented per 1000 operations) because of the relatively low mortality associated with the procedure.²⁷

Problems of interpreting the research

Despite the impression given by the literature that mortality rates – outcomes – may be improved as more operations are performed in a single centre, the results of the studies reviewed are difficult to interpret for a number of reasons.

Information on in-patient deaths is readily available in the USA from either hospital discharge abstracts or claims data. However, mortality can only ever be a partial measure of quality. While it may be the most accessible outcome measure available, it is unlikely to be the most appropriate one. Morbidity and quality of life can be equally important outcomes.^{39,40} and increased survival may sometimes be achieved at high cost to the patient. Health status is more difficult to quantify than mortality and data on quality of life are not routinely collected and recorded.

Volume levels are not absolute: whereas one study categorised hospitals carrying out more than 200 procedures per year as high volume,¹⁹ another study adopted a cut-off point of 650 procedures per year.³³ Equally, cut-off points vary between procedures.^{23,27,31} This makes it very difficult to translate such findings into meaningful recommendations for practice.

Research has tended to concentrate on the number of procedures carried out in a hospital rather than on the number performed by each clinician. This is due in part to the lack of data on clinician activity and outcomes.⁴¹ In the CABG studies included in this review, for example, only 3 tested the correlation between volume and outcome for the surgeon as well as for the hospital.^{22,32} One study found a positive relationship²² and 2 found no relationship^{32, 51} between volume and outcome.

It is difficult to attribute an apparently positive relationship between volume and outcome to a particular factor associated with the institution, clinician or some combination of the two. Other variables, such as hospital size, affiliation with a medical school, quality of facilities, availability of specialist staff and differences in the effectiveness of individual clinical skills, may mask the true relationship between volume and outcome. Rarely is information about the entire process of care considered in detail. Rather the in-patient episode is treated as a 'black-box'.^{18,42}

Direction of causality

The positive relationship between high volume and outcome could be used to support the 'practice makes perfect' hypothesis.³⁴ It seems plausible, as higher volume means greater experience with a particular procedure. Alternatively, in some health-care systems the same results may support a 'selective referral' hypothesis,⁴³ in which hospitals with good outcomes attract more patients. It may also be the case that higher-volume hospitals attract better clinicians. However, these effects, if real, are not generalisable to situations in which all hospitals increase their volume or smaller hospitals merge to create larger units.

When clinician volume is considered, it may be found that highly skilled doctors attract more patients, or that clinicians performing larger numbers of a particular procedure progress more rapidly along a learning curve. Direction of causality is ambiguous. Since the majority of studies use data from a cross-section of hospitals observed at specific points in time, rather than from a cohort of hospitals over a time continuum, and because they do not use experimental designs, it is difficult to interpret the direction of the reported relationship. These data do not provide the evidence needed to show whether quality would improve if volume were increased in smaller hospitals or, indeed, if quality could be sustained if volume were increased in larger units. To do this, studies need to monitor change over time or, preferably, use experimental designs.

Adjustment for patient case-mix

Studies of hospital mortality rates need to distinguish between the effects of differences in severity of illness and differences in quality of care. Higher mortality rates in low-volume hospitals, for example, may be due to a higher proportion of emergency or urgent cases, whereas lower mortality rates in high-volume hospitals may reflect the better results obtained from a greater number of elective procedures and a lower-risk patient population.²⁷

Routine hospital mortality data in the USA (derived from hospital claims forms or patient abstracts) include: patients' ages, sex, race, principal diagnosis and up to five secondary diagnoses, the principal procedure and up to three secondary procedures plus type of admission (emergency, urgent or planned). Although researchers use one or more of these administrative variables to adjust for mortality risk, routine claims data provide little relevant information about the patient's condition. Clinical data such as the results of

physical examinations, laboratory tests or radiological procedures, are better predictors of risk.¹⁰

One reliable way to determine the relationship between mortality rates and the number of patients treated is to assign large numbers of comparable patients randomly to hospitals with different volume levels. Though difficult to set up, such randomised controlled trials would eliminate confounding. None have been carried out to date. The main alternative is to attempt to control for the effects of confounding factors by using statistical methods.

The more statistical adjustment takes into account patient factors which influence outcome and are associated with hospital volume, the more likely it is to get an accurate assessment of the influence of volume. Systems which adjust for risk of dying based on detailed *clinical data* seem to be the most valid.⁴⁴ A recent study in the USA compared CABG surgery data from four sites participating in the Ischemic Heart Disease Patient Outcomes Research Team. Two data sets contained clinical data and two administrative data. The study found that data on chronic or asymptomatic conditions, such as mitral valve insufficiency (malfunction), cardiomegaly (enlargement of the heart), previous myocardial infarction, tobacco use and hyperlipidaemia, were less likely to be available from administrative than from clinical data sources.⁴⁵ Yet each of these factors places the patient at greater risk of dying. Similarly, another study of CABG surgery to evaluate whether the source of data affected the assessment of quality of care outcome found that the risk factors derived from administrative data were not sufficiently accurate or precise to yield quality measures comparable to those derived from the CABG surgery abstracts.⁴⁶

The clinical variables which best predict surgical mortality are parameters which indicate the patient's general physiological status. These tend to be similar across conditions (although about one-fifth of predictors were thought to be condition-specific). The physiological variables identified in a recent study as good predictors of mortality were: heart rate, arterial oxygenation and pH (acid-base balance), blood pressure, measures of consciousness level and respiratory (breathing) rate.⁴⁷ Even with a set of clinical predictors like these it is difficult to ensure that all the important effects of confounding have been taken into account. Other factors may arise.

Evidence for a volume–outcome relationship: CABG

The following detailed examination of the research on the volume–outcome relationship for a single procedure illustrates the importance of being able to adjust adequately for the effects produced by differences in patient case-mix.⁴⁹ CABG surgery was chosen since it

has been studied extensively. There is also wide variability in the degree of adjustment made for case-mix, and an independent literature identifying the patient factors which accurately predict survival.

The significant risk factors – independent of the procedure – which affect post-operative mortality in CABG surgery are set out in Table 1.^{23,48} However, adjustment for variables in the patient population is limited by the information contained in the studies. As most of the datasets are based on discharge abstract data, very few of the key prognostic factors indicated in Table 1 are included. Early studies rarely adjusted for more than the age and sex of patients. In contrast, later studies based on clinical data have used more comprehensive adjustment, taking into account specific concomitant illness, medical history and recent cardiac procedures known to be associated with outcome.

Table 1 Significant independent risk factors which affect post-operative mortality in CABG surgery

Age
Sex
Previous heart operations
Ejection fraction (heart pumping capacity)
Diabetes
Previous myocardial infarction
Dialysis dependence
Disasters (e.g. massive damage to the heart)
Cardiac catheterisation crash
Unstable (progressive) angina
Intractable congestive heart failure
Emergency procedure
Creatine (heart enzyme) levels > 168 mmol/l
Severe left ventricular disease
Pre-operative haematocrit (red blood cell volume) < 0.34
Chronic pulmonary (lung) disease
Prior vascular surgery
Re-operation
Mitral valve insufficiency

Analysis of the volume–outcome data

The volume–outcome relationship in 15 studies using observational data from the USA (see Tables 3 and 4) was analysed to investigate the relationship between the estimate of the effect of volume on mortality rates and the degree to which potential differences in patient case-mix had been taken into account (see box, ‘Statistical methods’).⁴⁹

The cut-off point between high- and low-volume hospitals was defined as more or less than 200 CABG procedures carried out per year, the only cut-off point that was common to all studies, thus allowing comparison of results to be made. In addition, several authors have suggested that there is a threshold of about 200 CABG procedures per year.⁵⁰

Patient numbers and adjusted mortality rates were extracted from each study along with the variables used to adjust for patient case-mix. Some studies presented the expected rather than adjusted death rates. In these instances the low-volume mortality rate was adjusted by multiplying by the ratio of the expected death rates in the high-volume to the low-volume group.

Each study was given a score from 0 to 3 (Table 2) indicating the adequacy of adjustment based upon the evidence of prognostic factors discussed above. The assessor was blind to the results of each study.

Table 2 Gradings of adjustment for case-mix in CABG studies

Adjustment score	Criteria
0	No case-mix adjustment
1	Adjustment for age, sex and whether patients had concomitant illness
2	Adjustment for age, sex, and the nature of other heart and medical conditions as described in discharge abstracts
3	Adjustment for age, sex, and the nature of other heart and medical conditions as described in clinical databases

Statistical methods

The estimates of benefit associated with higher volume (odds ratio) for each study is plotted against the degree of adjustment used in the study on the four-point classification

A statistical model was developed in order to investigate whether there is a systematic change in the estimates of the volume effect as the degree of adjustment for patient case-mix is improved. Logistic regression was used to model the reported risks of death in high- (>200) and low-(≤ 200) volume hospitals in each study. A covariate indicating high and low volume was included to estimate the effect of volume on mortality. All models also included a covariate for each study, so that the volume effects were estimated on the basis of pooled within-study comparisons.⁵¹

The model presented in this report included interaction terms which measured the modification of any volume effect according to the degree of patient-mix adjustment, and the modification of any volume effect related to the year of data collection. It is these interaction terms which are of primary interest in the analysis.

The statistical models were initially fitted to data from the six studies with the high-low volume cut-off point near 200 cases per year. Data from the study using cut-off points which could not be directly linked to hospital volumes were included in a sensitivity analysis. Overdispersion (residual heterogeneity) was accounted for in the models by appropriately rescaling the standard errors. The significance of the variables was assessed from z-scores calculated as the ratio of the effect sizes to the rescaled standard errors.⁵²

Table 3 Studies of coronary artery bypass surgery included in the review

Study (Reference)	Data Source	Years Used	Cut-off point used to define low volume	Adjusted* mortality rate in low volume hospitals	Adjustments made for confounding (Adjustment score)	Number of hospitals	Adjusted* Number of Deaths / Number of Patients	Adjusted* odds ratio comparing mortality in high to low hospitals
<i>Studies included in the main analysis</i>								
Miaerki 1986 (28)	CPHA discharge abstracts	1972	215	5.5% (5.2%, 5.7%)	age, sex and single or multiple diagnosis, admission blood pressure (1)	Low volume: 109 High volume: 5	Low volume: 246/3619 High volume: 28/905	0.44 (0.29, 0.65)
Luft 1979 (20)	CPHA discharge abstracts	1974-5	200	7.0% (6.4%, 7.7%)	age, sex and single or multiple diagnosis (1)	Low volume: 157 High volume: 25	Low volume: 1089/19098 High volume: 518/15232	0.58 (0.52, 0.65)
Rosenfeld 1987** (30)	CPHA discharge abstracts	1982	150	4.9% (4.6%, 5.4%)	no adjustments made (0)	Low volume: 109 High volume: 5	Low volume: 125/2508 High volume: 875/26980	0.64 (0.53, 0.77)
Showstack 1987 (21)	CHFC discharge abstracts	1983	200	4.8% (4.4%, 5.2%)	sex, age, ethnic group and presence of acute myocardial infarction, congestive heart failure, cardiac catheterization, and coronary angioplasty (2)	Low volume: 35 High volume: 42	Low volume: 157/3797 High volume: 539/15189	0.83 (0.69, 0.99)
Hannan 1989 (22)	New York State discharge abstracts	1986	223	6.0% (5.5%, 6.7%)	age, sex, whether the admission was scheduled or unscheduled, disease condition/stage, number of secondary diagnoses, co-diagnosis of cancer, and whether valve or aneurysm surgery was also performed (2)	Low volume: 10 High volume: 17	Low volume: 86/1530 High volume: 347/8244	0.74 (0.58, 0.94)
Hannan 1994 (25)	New York State CSRS prospective clinical data	1989-92	200	4.8% (4.4%, 5.1%)	age, sex, coronary heart disease, high grade stenosis, ischemia unstable angina, ejection fraction, previous myocardial infarction, preoperative intra-aortic balloon pump, congestive heart failure, "disasters", diabetes, morbid obesity, chronic obstructive pulmonary disease, dialysis dependent, previous open heart surgery. (3)	Low volume: 4 High volume: 26	Low volume: 71/1850 High volume: 1779/54742	0.84 (0.66, 1.07)
<i>Studies included in a sensitivity analysis</i>								
Riley 1985 (27)	20% Sample of Elderly Medicare Beneficiaries	1979-80	19 (not inflated from the 20% sample)	6.5% (5.7%, 7.2%)	no adjustments made	Total: 909	Low volume: 313/4658 High volume: 85/1499	0.83 (0.65, 1.07)

Table 4 Studies of coronary artery bypass surgery excluded from the review

Study (Reference)	Data Source	Years Used	Reason for Exclusion
<i>Studies not included</i>			
Rosenfeld 1987** (30)	CPHA	1972	Same data source as Maerki 1986
Luft 1987 (31)	CPHA	1972	Volume categorisations are not clear. Same data source as Maerki 1986
Luft 1980 (26)	CPHA	1974-5	Analysis presented as a regression model. Same data source as Luft 1979
Sloan 1986 (29)	CPHA	1972-81	Volume categorisations are not clear. Same data source as Maerki 1986 and Luft 1979
Hughes 1987 (32)	CPHA	1982	Analysis presented as a regression model. Same data source as Rosenfeld 1987
Zelen 1991 (33)	New York State	1986	Same data source as Hannan 1989
Hannan 1991 (23)	New York State	1989	Same data source as Hannan 1994
Farley 1992 (24)	HCUP	1980-87	Analysis presented as a regression model
Williams (1991) (45)	Philadelphia's medical schools	1985-1987	All hospitals were high volume (>200 cases per year)

* Different adjustments were made between the studies as indicated in column 5.

**Rosenfeld 1987 presented data for both 1972 (excluded) and 1982 (included).

CPHA = Commission on Professional and Hospital Activities, Ann Arbor, Michigan; CHFC = California Health Facilities Commission; HCUP = Hospital Cost and Utilisation Project, Division of Provider Studies in the Agency for Health Care Policy and Research; CSRS - Cardiac Surgery Reporting System, New York

Results

Fifteen studies using observational data from the USA were identified which have examined the relationship between volume and outcome in CABG surgery (Tables 3 and 4).^{9–11,19–29,50} However, where data were duplicated (Table 3), they were only included once for each set of studies.^{9,10,19,21,22,26,27} One further study was excluded as all hospitals performed more than 200 procedures per year.⁵⁰

The studies mainly obtained their data from hospital discharge abstracts. The prognostic variables controlled for in the 7 studies varied from age and sex to some clinical risk factors (Tables 3 and 4). One study presented data on a 20% sample of elderly Medicare beneficiaries.²¹ As it was unclear how these patient volumes related to hospital volumes the results of this study were confined to the sensitivity analysis.

All the studies included in the analysis found a positive relationship between volume and outcome with five of the seven showing this result as statistically significant. Of the three studies which included clinician volume one found a positive relationship between volume and outcome²² and two did not.^{32,50}

Table 5 gives details of the significance of the statistical modelling of these trends. Model B indicates that, as the degree of adjustment for case-mix increases, the estimate of the advantage of increased volume is significantly diminished. Baseline level of risk was added as an explanatory variable but did not materially alter the estimate of the adjustment–volume interaction. The degree of adjustment for case-mix has improved with time and so year and adjustment variables are highly correlated (Spearman's rank correlation coefficient =0.79, n=6, p < 0.1).

Table 5 Volume and adjustment parameter estimates

Model	Odds ratio (95% CI)	Significance
Model A		
Volume	0.66 (0.57, 0.75)	P < 0.001
Model B		
Volume (when adjustment=0)	0.54 (0.44, 0.67)	P < 0.001
Adjustment–volume interaction	1.15 (1.01, 1.32)	P = 0.04

According to these results, a positive relationship between volume and outcome suggested by most of the studies might be confounded by differences in patient case-mix between high- and low-volume hospitals. As the degree of adjustment is increased the estimated net beneficial effect of increased volume is reduced (odds ratio moves nearer to 1). (see Figure 1).

This relationship could not be explained by looking at differences between studies in the baseline levels of risk in low volume units. It is unclear whether more extensive and detailed adjustment for the effects of patient case-mix would further reduce or, indeed, increase this effect, since such data are not available. The analysis also showed that the size of the estimated benefit of high volume diminished over time.⁴⁹

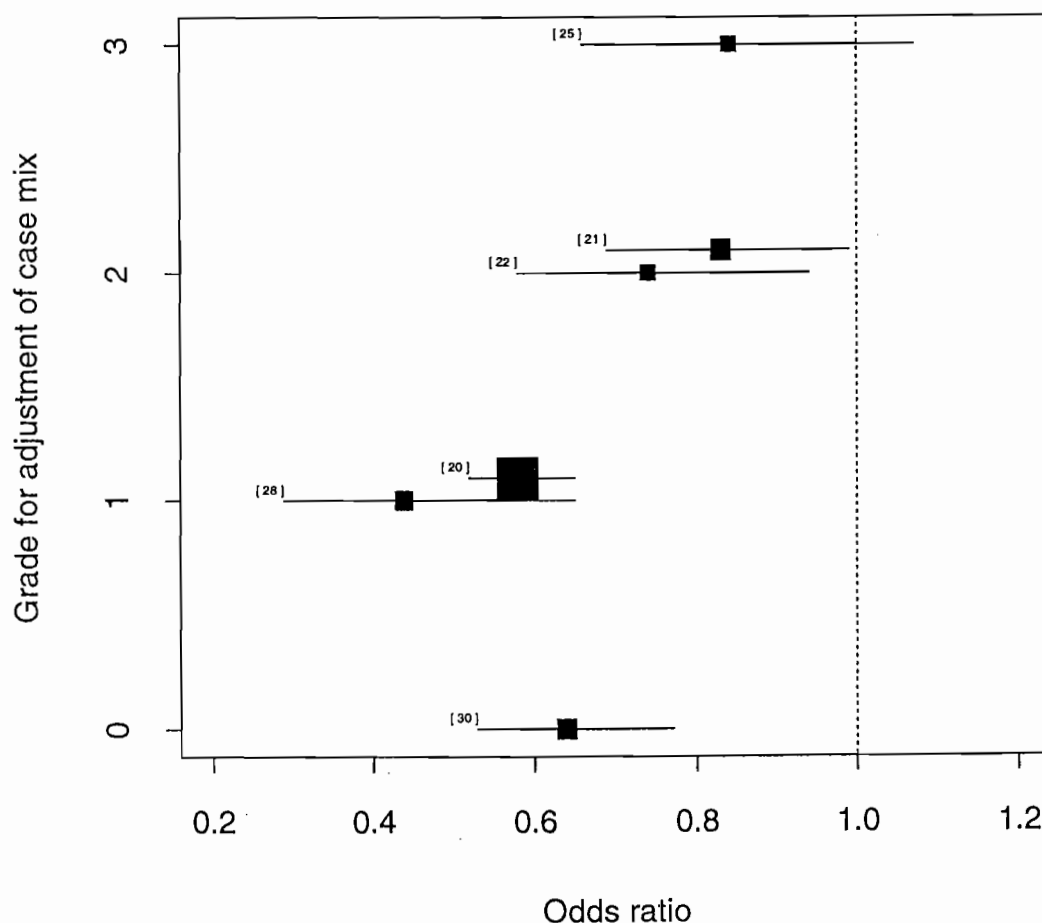


Figure 1 Estimated effect on mortality of high volume hospital compared to low volume hospital by degree of adjustment for case-mix

The importance of adjustment in observational data

This example highlights the importance of adequate adjustment for confounding when observational data are used to compare the outcomes of different units. The importance of adjustment for confounding has also been documented in other areas of clinical care. The majority of studies of neonatal intensive care compare outcomes by using birthweight to measure initial neonatal risk;^{24,41,53–56} differences in mortality rates have been reported at hospitals or units of varying sizes. However, other, unmeasured risk factors may well have accounted for the differences. Indeed, three recent studies have shown that the correlation of birthweight and mortality may not be adequate for comparing the performance of different units^{57–59} because initial disease severity is not necessarily related to birthweight.⁵⁷

In 1993 the scoring system Clinical Risk Index for Babies (CRIB) was developed for comparing the performance of neonatal intensive care units (NICUs) in the UK.⁵⁷ It is based on routine data on physiological status recorded within 12 hours of birth. Table 6 shows the CRIB score components.

Table 6 Components of the Clinical Risk Index for Babies (CRIB)⁵⁷

Birthweight

Gestational age

Presence of congenital malformations

Maximum and minimum appropriate FiO_2 (blood oxygen concentration) in first 12 hours

Maximum base excess in first 12 hours (indicator of alkalosis, bicarbonate mmol/l)

CRIB was found to be substantially more accurate than birthweight alone in predicting hospital death. The authors concluded that comparisons of mortality rates without adequate adjustment can be misleading. Outcome between NICUs can only be compared if there is an accurate measure of disease severity.⁵⁷

The authors of an illness severity index (Score for Neonatal Acute Physiology, SNAP) developed for neonatal intensive care in the USA⁵⁹ found that illness severity and birthweight were both powerful independent predictors of neonatal mortality. Neither could be excluded if meaningful comparisons were to be made between NICUs.

Examples from other clinical areas also highlight the risk of drawing conclusions about effectiveness without adequate adjustment. Analysis of claims data for treatment of benign prostatic hyperplasia (BPH), suggested that transurethral resection of the prostate

(TURP), a minimally invasive procedure, was associated with twice the re-operation rate and carried a higher mortality risk than open surgery, even after adjusting for age and concomitant illness.⁶⁰ However, a study which used better adjustment for prognostic factors, such as the severity of concomitant illnesses, found that mortality for the two procedures was similar.⁶¹ The apparent higher death rate associated with TURP may be the result of selection criteria: more frail patients underwent TURP than open resection since it is less traumatic.

Another example comes from a recent study of patient outcomes in 26 intensive care units (ICUs) in the United Kingdom.⁶² Crude hospital mortality rates produced a positive correlation between volume and outcome. However, when a validated measure of case severity - the Acute Physiology and Clinical Health Evaluation Severity Score (APACHE II)⁶³ - was applied, high-volume ICUs were found to admit a greater proportion of less severely ill patients⁶⁴ and hence on average had better outcomes. After adjustment for differences in patient case-mix the association between volume and in-hospital mortality disappeared.⁶²

The case study of CABG and the examples above illustrate the problem of relying on observational data. Any conclusions about variations in the quality of care provided by different hospitals based on this type of analysis *alone* should be viewed with caution and should take into account the degree to which the effects of patient case mix and other confounding have been taken into account.

Conclusions, research and policy implications

Most researchers argue for an association between volume and mortality rates which seems to have found wide acceptance in the policy community. For example The Office of Technology Assessment stated in a recent report that: 'Although the exact nature of the relationship between volume and outcomes remains murky, the research overall has tended to reinforce the idea that simultaneously reducing costs (through improved efficiency at high-volume institutions) while improving the quality of care (through better care outcomes) is an achievable goal. This association however, is not well supported by the evidence. Nor can existing research provide the information needed to answer a number of important questions concerning the nature of any volume–outcome relationship that might exist.

There is little evidence to show why high volume might be associated with better outcomes and little attempt has been made to 'unpack' this concept. Knowing how well something works is different from knowing how it works.⁶⁵ For example, if there is a benefit due to higher volumes, is this caused by institution-related elements such as

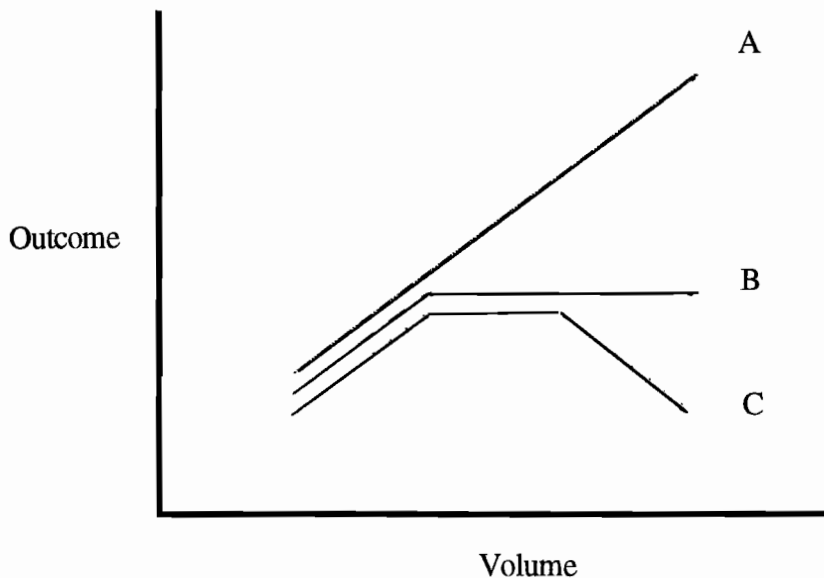
specially designed facilities for specific workloads? Alternatively, any relationship may be attributable to clinician-related elements such as the learning or experience effects in individual doctors.¹⁶ For example, there may be substantial differences in the performance of a surgeon carrying out a procedure twice a year, compared with a surgeon doing the same procedure 15 times a year, but little difference between performing an operation 40 or 55 times a year. There may also be a point over and above which quality is no longer improved by increased volume and may start to decrease as doctors become bored, rushed or pass on work to less experienced junior staff.

Clinicians carrying out procedures many times a year may achieve some of their quality gains because they keep up with the literature and use the most appropriate management practices available – that is, they develop effective (implicit or explicit) protocols – and not due solely to the acquisition of purely personal (e.g. manual) skills. If this is the case, then such quality improvement may be achieved in other settings by adopting guidelines based upon the best practice. Clinical guidelines are becoming increasingly common in health care. Future research could investigate whether clinicians or hospitals with lower volumes could use such guidelines to achieve the outcomes of larger specialist centres. Thus, where outcomes of a procedure are poor, there may be a number of measures other than attempts to manipulate volume that could improve results.²⁵

Where a volume–outcome relationship has been identified the direction of causality remains uncertain. Both the 'practice makes perfect' hypothesis (greater experience at high-volume hospitals) and the 'selective referral' hypothesis (higher level of referral at hospitals with better outcomes) have been suggested as explanations. However, as the data do not provide any information on how outcomes change as volume changes over time it is difficult to interpret causality. The fact that a high-volume hospital has good outcomes for a particular procedure does not automatically imply that increasing volume at another hospital will improve its outcomes. There is a need for prospective studies (with adequate controls) to monitor the effects of changing volumes on quality of care.

The shape of any volume–benefit curve has yet to be determined. Research is needed to establish whether outcomes improve continuously with increasing volume (Figure 2A), or whether there is a threshold (Figure 2B), or even a negative effect at very high volumes (Figure 2C). This could be used to help planners decide whether to aim for the highest-volume hospitals or whether to avoid hospitals performing fewer than a certain number of given procedures per year. Past research has used very different volume categories within the same procedure, so that high volume in one study equals low volume in another study. This disparity makes it extremely difficult to translate the research into more meaningful recommendations for policy-makers.

Figure 2A, B, C Practice makes perfect hypothesis: possible shapes of the volume–outcome curve



Policy discussions must also focus on the likely impact of centralisation on hospitals, clinicians and patients. For example, if a hospital loses specific services in a process of centralisation, the viability of the remaining services will be affected. In turn, hospitals which increase their volume need to expand a wide range of services to be able to cope with the extra patient load. The trade-off involved in concentrating emergency services is contentious, since travel time and distance - as well as availability of specialist services - are of crucial importance. Patients with emergency problems may not benefit from the services of a major trauma centre if prolonged travel time greatly increases the risk of dying. Policy-makers also need to take into account the transportation costs associated with centralisation of services. Until more evidence about outcomes becomes available from countries where centralisation has been adopted, it is difficult to draw any conclusions about the likely effects on patient outcomes in the United Kingdom.

When services are concentrated or shifted by moving a contract from a low- to a high-volume unit, changes should be carefully monitored. Information should be collected about the pattern of referrals (in terms of differences between a local district general hospital and a new centralised service) to monitor equity of care. This is important in the light of evidence that utilisation rates may decline as distance from a health facility

increases.^{66,67} Although evidence is scant, barriers to access may significantly affect outcomes, particularly among those with limited access to adequate transportation, such as the elderly, women with pre-school children, low-income families, the disabled and adolescents, all of whom are likely to have a greater need for health care.⁶⁸

This document does not review the arguments for or against concentration of services and the findings should not be interpreted to mean that such a policy is unjustified. There are other legitimate reasons which could justify some centralisation. Quality in health care has been defined as consisting of the dimensions of: effectiveness (for individual outcome), efficiency and economy, equity, social acceptability, relevance to need (appropriate for a specific population) and access to the service.⁶⁹ However, it is rarely possible to maximise all these dimensions simultaneously and trade-offs have to be made. Policy-makers may need to be more explicit about the relative weights given to other dimensions of quality, given the uncertain effect of volume on clinical effectiveness, in particular the extent to which health policy is being driven by assumptions about changes in efficiency.*

This review has shown that most of the *research* reports a positive relationship between volume and outcome of care but that the *evidence* is more uncertain. The validity of some of the research findings is suspect because of problems in adjusting for patient-mix and the relationships reported are difficult to interpret. Therefore, the main recommendation is that policy-makers should be cautious when invoking the assumed improvements in outcome achieved by volume as a key argument for centralisation of services.

* The NHS Centre for Reviews and Dissemination will be examining the research evidence on the existence and size of economies of scale at the end of 1995.

References

- 1 Bunker JP, Luft HS, Enthoven A. Should surgery be regionalized? Surgical Clinics of North America 1982; 62: 657–668.
- 2 Consultative Document: A Policy Framework for Commissioning Cancer Services. London. Department of Health, 1994.
- 3 Banta D & Bos M. The relation between quantity and quality with coronary artery bypass graft (CABG) surgery. Health Policy 1991; 18: 1–10
- 4 American College of Surgeons. Guidelines for minimal standards in cardiac surgery. American College of Surgeons Bulletin. 1984 January 67–69.
- 5 Flood AB, Scott R, Ewy W. Does practice make perfect? Part II the relation between volume and outcomes and other hospital characteristics. Medical Care 1984; 22: 115-125
- 6 Holland WW, Detels R, Knox G. (ed.s) Oxford Textbook of Public Health. 1991; 2: (p. 70)
- 7 Fink A, Yano EM, Brook RH. The condition of the literature on differences in hospital mortality. Medical Care 1989; 27: 315-336.
- 8 Des Harnais SILF, McMahon LF, Wroblewski RT, Hogan AJ. Measuring hospital performance: the development and validation of risk adjusted indexes of mortality, readmissions and complications. Medical Care 1990; 28: 1127-1141.
- 9 Smith DW, Pine M, Bailey CR, Jones B, Brewster A, Krakauer H. Using clinical variables to estimate the risk of patient mortality. Medical Care 1991; 29: 1108-1119.
- 10 Orchard C. Comparing health care outcomes. BMJ 1994; 308: 1493-1496.
- 11 Davey Smith G, Phillips A, Neaton JD. Smoking as 'independent' risk factor for suicide: illustration of an artefact from observational epidemiology. Lancet 1992; 340: 709-712.
- 12 Sheldon TA. Please bypass the PORT. BMJ 1994; 309: 142-143.

- 13 Banta HD, Engel GL, Schestén T. Volume and outcome of organ transplantation. International Journal of Technology Assessment in Health Care 1992; 8: 490-505.
- 14 Stiller C. Centralised treatment, entry to trials and survival. British Journal of Cancer 1994; 70: 352-362.
- 15 Office of Technology Assessment. The Quality of Medical Care: Information for Consumers. Washington DC: United States Government Printing Office, 1988.
- 16 Luft HS, Garnick DW, Mark DH, McPhee SJ. Hospital volume, physician volume and patient outcomes. Assessing the evidence. Michigan. Health Administration Press Perspectives. 1990
- 17 Black N, Johnston A. Volume and outcome in hospital care: evidence, explanations and implications. Health Services Management Research 1990; 3: 108-114.
- 18 Houghton A. Variation in outcome of surgical procedures. British Journal of Surgery 1994; 81: 653-660.
- 19 Ministry of Health. Tertiary Services Review. September 1994
- 20 Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. NEJM 1979; 301: 1364-1369.
- 21 Showstack J.A., Rosenfield K.E., Garnick D.W., Luft H., Schaffazrick MD, Fowles J. Association of volume with outcome of coronary artery bypass graft surgery. JAMA 1987; 257: 785-789.
- 22 Hannan EL, O'Donnel JF, Kilburn H, Bernard HR, Yazici A. Investigation of the relationship between volume and mortality for surgical procedures performed in New York state hospitals. JAMA 1989; 262: 503-510.
- 23 Hannan EL, Kilburn H, Bernard H, O'Donnell, Lukacik G, Shields EP. Coronary artery bypass surgery: The relationship between in hospital mortality rate and surgical volume after controlling for clinical risk factors. Medical Care 1991; 29: 1094-1107.

- 24 Farley DE, Ozminkowski RJ . Volume–outcome relationships and in-hospital mortality: The effect of changes in volume over time. Medical Care 1992; 30: 77-94.
- 25 Hannan EL, Kilburn H, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York state. JAMA 1994; 271: 761-766.
- 26 Luft H. The relation between surgical volume and mortality: an exploration of causal factors and alternative models. Medical Care 1980; 18: 941-959.
- 27 Riley G, Lubitz J. Outcomes of surgery among the Medicare aged: surgical volume and mortality. Health Care Financing Review 1985; 7: 37-47.
- 28 Maerki SC, Luft HS, Hunt SS. Selecting categories of patients for regionalization: Implications of the relationship between volume and outcome. Medical Care 1986; 24: 148-158.
- 29 Sloan FA, Perrin JM, Valvona F. In-hospital mortality of surgical patients: is there an empiric basis for standard setting? Surgery 1986; 99: 446-453.
- 30 Rosenfield K, Luft HS, Garnick DW, McPhee SJ. Changes in patient characteristics and surgical outcomes for coronary artery bypass surgery 1972-82. American Journal of Public Health 1987; 77: 499-500
- 31 Luft HS, Hunt SS, Maerki C. The volume outcome relationship: practice makes perfect or selective referral patterns? Health Services Research 1987; 22: 157-182.
- 32 Hughes RG, Hunt SS, Luft HS. Effects of surgeon volume and hospital volume on quality of care in hospitals. Medical Care 1987; 25: 489-503.
- 33 Zelen J, Bilfinger TV, Anagnostopoulos CE. Coronary artery bypass grafting. The relationship of surgical volume, hospital location and outcome. New York State Journal of Medicine 1991; 91: 290-292.
- 34 Flood AB, Scott WR & Ewy W. Does practice make perfect? Part 1: The relation between hospital volume and outcomes for selected diagnostic categories. Medical Care 1984; 22: 98-114.

- 35 Hughes RG, Garnick DW, Luft HS, McPhee SJ, Hunt SS. Hospital volume and patient outcomes. The case of hip fracture patients. Medical Care 1988; 26: 1057-1067.
- 36 Farber BF, Kaiser DL, Wenzel RP. Relation between surgical volume and incidence of postoperative wound infection. NEJM 1981; 305: 200-204.
- 37 Shortell SM, LoGerfo JP. Hospital medical staff organisation and quality of care: Results for myocardial infarction and appendectomy. Medical Care 1981; 19: 1041-1053.
- 38 McArdle CS, Hole D. Impact of variability among surgeons on postoperative morbidity and mortality and ultimate survival. BMJ 1991; 302: 1501-1505.
- 39 Ashley J, McLachlan G. (Eds.) Mortal or Morbid? A diagnosis of the morbidity factor. London: The Nuffield provincial Hospitals Trust. London, 1985.
- 40 Jenkinson C. (Ed.) Measuring health and medical outcomes. London: UCL Press, 1994.
- 41 Rosenblatt RA, Reinken J, Shoemack P. Is obstetrics safe in small hospitals? Evidence from New Zealand's regionalized perinatal system. Lancet 1985; 2: 429-432.
- 42 Luft H, Garnick DW, Mark DH, Peltzman DJ, Phibbs CS, Lichtenberg E, McPhee SJ. Does quality influence choice of hospital? JAMA 1990; 263: 2899-2906.
- 43 Flood AB, Scott WR, Ewy W. Does practice make perfect? Part I the relation between volume and outcome for selected diagnostic categories. Medical Care 1984; 22: 98-114.
- 44 Iezzoni LI. [editor]. Risk Adjustment for Measuring Health Care Outcomes. Michigan: Health Administration Press, 1994.
- 45 Romano PS, Roos LL, Luft HS, Jollis JG, Doliszny K, The Ischemic Heart Disease Patient Outcomes Research Team. J Clin Epidemiol 1994; 47: 249-260.

- 46 Hartz AJ, Kuhn EM. Comparing hospitals that perform coronary artery bypass surgery: the effect of outcome measures and data sources. American Journal of Public Health 1994; 84: 1609-1614.
- 47 Iezzoni LI, Ash A, Coffman GA, Moskowitz MA. Predicting in-hospital mortality: A comparison of severity measurement approaches. Medical Care 1992; 30: 347-359.
- 48 Higgins TL, Estafanous FG, Loop FD, Beck GJ, Blum JM, Paranandi L. Stratification of morbidity and mortality outcome by preoperative risk factors in CABG patients. A clinical severity score. JAMA 1992; 267: 2344-2348.
- 49 Sowden AJ, Deeks J, Sheldon TA. Volume and outcome in coronary artery bypass graft surgery: true relationship or artefact? BMJ 1995; 311: 151-161.
- 50 Williams SV, Nash DB, Goldfarb N. Differences in mortality from coronary artery bypass graft surgery at five teaching hospitals. JAMA 1991; 266: 810-815.
- 51 Thompson SG. Controversies in meta-analysis: the case of trials of serum cholesterol reduction. Statistical Methods in Medical Research 1993; 2: 173-192.
- 52 STATA Corporation. STATA reference manual: Release 3.1, Vol. 2 6th ed. 449-465 College Station , TX, 1993.
- 53 Williams RL. Measuring the effectiveness of perinatal medical care. Medical Care 1979; 17: 95-110.
- 54 Mayfield JA, Rosenblatt RA, Baldwin LM, Chu J, Logerfo JP. The relation of obstetrical volume and nursery level to perinatal mortality. AJPH 1990; 80: 819-823.
- 55 Field D, Hodges S, Mason E, Burton P. Survival and place of treatment after premature delivery. Archives of Disease in Childhood 1991; 66: 408-411.
- 56 LeFevre ML. Physician volume and obstetric outcome. Medical Care 1992; 30: 866-871.

- 57 The International Neonatal Network. The CRIB (clinical risk index for babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. The Lancet 1993; 342: 193-198.
- 58 Tarnow-Mordi W, Ogston S, Wilkinson AR, Reid E, Gregory J, Saaed M, Wilkie R. Predicting death from initial disease severity in very low birthweight infants: a method for comparing the performance of neonatal units. BMJ 1990; 300: 1611-1614.
- 59 Richardson DK, Phibbs CS, Gray JE, McCormick MC, Workman-Daniels K, Goldman DA. Birth weight and illness severity: Independent predictors of neonatal mortality. Pediatrics 1993; 91: 969-975
- 60 Fisher ES, Wennberg JE. Administrative data in effectiveness studies: the prostatectomy assessment. In Heithoff KA, Lohr KN (editors) Effectiveness and Outcomes in Health Care. Washington DC. National Academy Press. 1990 Ch.13.
- 61 Concato DR, Horwitz RI, Feinstein AR, Elmore JG, Schiff SF. Problems of comorbidity in mortality after prostatectomy. JAMA 1992; 267: 1077-1082.
- 62 Jones J, Rowan K. Is there a relationship between the volume of work carried out in intensive care and its outcome? International Journal of Technology Assessment in Health Care. (In press).
- 63 Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Critical Care Medicine 1985; 13: 818-829.
- 64 Strauss MJ, LoGerfo JP, Yeltatzie JA, Temkin N, Hudson LD. Rationing of intensive care unit services: an everyday occurrence. JAMA 1986; 255: 1143-1146.
- 65 Berwick DM. The double edge of knowledge. JAMA 1991; 266: 841-842.
- 66 Haynes RM, Bentham CG. Accessibility and the use of hospitals in rural areas. Area 1979; 11; 186-191.
- 67 Parkin D. Distance as an influence on demand in general practice. Journal of Epidemiology and Community Health 1979; 33; 96-99.

- 68 Bentham CG, Haynes R. A raw deal in remoter areas? Family Practitioner Services 1986; 13: 84-87.
- 69 Maxwell RJ. Quality assessment in health. BMJ 1984; 288: 1470-1472.