THE UNIVERSITY of York

CENTRE FOR HEALTH ECONOMICS
YORK HEALTH ECONOMICS CONSORTIUM
NHS CENTRE FOR REVIEWS & DISSEMINATION

Concentration and Choice in the Provision of Hospital Services

The Relationship Between Hospital Volume and Quality of Health Outcomes

CRD REPORT 8 (Part I)

Concentration and Choice in the Provision of Hospital Services

The Relationship Between Hospital Volume and Quality of Health Outcomes

Amanda J. Sowden
NHS Centre for Reviews and Dissemination

Roberto Grilli
Istituto di Ricerche Farmacologiche 'Mario Negri', Milan

Nigel Rice
Centre for Health Economics

CRD Report 8 (Part I)
January 1997

ISBN 1 900640 05 8

**This report can be ordered from: Publications Office, NHS Centre for Reviews and Dissemination, University of York, York YO1 5DD. Telephone 01904 433648; Facsimile: 01904 433661: email: crdpub@york.ac.uk
Price £12.50**

The NHS Centre for Reviews and Dissemination is funded by the NHS Executive and the Health Departments of Scotland, Wales and Northern Ireland; a contribution to the Centre is also made by the University of York. The views expressed in this publication are those of the authors and not necessarily those of the NHS Executive or the Health Departments of Scotland, Wales or Northern Ireland.

Printed by York Publishing Services Ltd.

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 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 Reports

1	Which Way Forward for the Care of Critically Ill Children? (1995)	£7.50
2	Relationship Between Volume and Quality of Health Care: A Review of the Literature (1995)	£5.00
3	Review of the Research on the Effectiveness of Health Service Interventions to Reduce Variations in Health (1995)	£12.50
4	Undertaking Systematic Reviews of Research on Effectiveness. CRD Guidelines for those Carrying Out or Commissioning Reviews (1996)	£7.50
5	Ethnicity and Health: Reviews of Literature and Guidance for Purchasers in the Areas of Cardiovascular Disease, Mental Health and Haemoglobinopathies (1996)	£12.50
6	Making Cost-Effectiveness Information Accessible: The NHS Economic Evaluation Database Project. CRD Guidance for Reporting Critical Summaries of Economic Evaluations (1996)	£7.50
7	A Pilot Study of 'Informed Choice' Leaflets on Positions in Labour and Routine Ultrasound	£7.50

Effectiveness Matters

Vol.1	(1)	Aspirin and Myocardial Infarction (1995)
	(2)	Helicobacter Pylori and Peptic Ulcer (1995)
Vol.2	(1)	Influenza Vaccination and Older People (1996)
	(2)	Screening for Prostate Cancer (1997)

ACKNOWLEDGMENTS

This paper is part of a report on '*Concentration and Choice in the Provision of Hospital Services*' which was prepared by a team or researchers directed by:

Brian Ferguson, Deputy Director, Centre for Health Economics University of York
John Posnett, Director, York Health Economics Consortium, University of York
Trevor Sheldon, Director, NHS Centre for Reviews and Dissemination, University of York

We acknowledge the contribution of the NHS CRD Information Service in developing the search strategies for this study.

****

Other titles in the *Concentration and Choice in the Provision of Hospital Services* series are:

Summary Report

II. *The Relationship Between Volume and the Scope of Activity and Hospital Costs*

III. *The Relationship Between Concentration, Patient Accessibility and Utilisation of Services*

TABLE OF CONTENTS

1. BACKGROUND

The concentration of hospital services in larger units has been proposed as a means of improving the quality of health care. It is suggested that improved benefits to patients can be expected from the quality of care offered by experienced clinicians working in high volume hospitals or centres.[1] A recent review of the recommendations of the medical Royal Colleges and specialist associations in England found that a number of colleges and associations were proposing volume thresholds based on an assumed relationship between the number of procedures carried out and quality of care.[2] These assumptions to an extent informed the proposal for a three-tiered regional network for cancer services in the UK.[3]

In the Netherlands some operations have been regionalised by regulation; open heart surgery for example, can only be performed in licensed hospitals, in which a minimum of 600 procedures must be carried out each year.[4] Similarly, 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.[5] In a report of the Joint International Society and Federation of Cardiology/World Health Organisation Task Force on Coronary Angioplasty, it was recommended that a physician performing percutaneous transluminal coronary angioplasty should carry out an adequate number of procedures to maximise patient safety and the efficiency of the laboratory. A minimum caseload for a single physician was estimated to be about one case per week.[6]

The evidence used to suggest improved quality of care with concentration of services comes from a large international literature which has examined the relationship between volume of hospital and/or physician activity and clinical outcomes. One of the first studies to assess this relationship was carried out by Luft et al in 1979.[7] Over the past 15 years, publication of other studies in this area has been prolific. Studies have focused on particular groups of patients undergoing a specific procedure or with a specified diagnosis and measured their outcome, usually in terms of hospital mortality rates. A number of reviews of this literature have been published over the past 8 years; of which some have reviewed several procedures and/or conditions[8-11] or specific diseases or conditions such as coronary artery bypass graft surgery (CABG),[14] solid organ transplantation,[12] cancer [13, 14] or speciality.[10]

The hospital mortality rate for many procedures has been repeatedly shown to be related to the frequency with which they are carried out at individual hospitals and the notion of a volume - outcome relationship has generally become accepted as common sense.

Against this background a systematic review of the research was carried out to assess the evidence for a volume-outcome relationship. The three main questions addressed in this review were:

i) what is the evidence of a relationship between increased hospital or physician volume and patient outcomes?

ii) to what degree are the results common across procedures and conditions?

iii) to what extent are any reported differences in outcome associated with volume really attributable to volume, or to other factors such as patient case-mix?

2. METHODS OF THE REVIEW

2.1 Identification of studies

The following searches were carried out:

i) Medline was searched using both key words and Mesh headings (1980 - 1996) as were Embase, Health Planning and Administration, Dissertation Abstracts and Entis (Research report database). (See Appendix for search strategies)

ii) hand searches of key relevant journals (Medical Care)

iii) the references of identified studies were checked

iv) experts in the field and other Health Technology Assessment bodies in the UK and internationally have been contacted to help identify published and unpublished studies (see Appendix for individuals contacted)

2.2 Inclusion criteria

To be included in the review, studies had to be empirical and satisfy the following criteria of relevance, outcome and design:

2.2.1 Relevance

- Studies where a comparison was made between the outcomes of patients treated in a) hospitals/centres with different volume levels or b) actual death rates in units of high or low volume are compared with expected death rates derived from reliable data from units of another volume or c) where outcomes are compared before and after concentration of services.
- English language and foreign language papers.

2.2.2 Outcome

The outcomes of interest were one (or more) of the following:

- mortality (in-hospital or other)
- morbidity (e.g. infection rates)
- psycho-social (e.g. satisfaction)
- quality of life

2.2.3 Design

The following types of study design were included:

- RCT
- controlled trial
- before/after
- prospective cohort
- retrospective cohort
- case control
- cross sectional

2.3 Exclusion criteria

Articles which did not present empirical findings (e.g. policy statements or editorials) were not included in the review. Studies using the same data sets for the same period of time (e.g. duplicate publications) were excluded; where different authors analysed the same data, the study with the best methods was included. Also, studies where process rather than outcome measures were the main focus were excluded.

2.4 Methodological quality assessment

Each study satisfying the inclusion criteria was assessed according to the following criteria:

a) type of study design, (experimental studies are given more weight than observational studies and within observational studies more weight is given to prospective as opposed to retrospective designs)

b) process of patient identification (e.g. from medical records, from routinely collected case abstracts provided by the hospitals, or from insurance claims data)

c) avoidance of selection bias (number of eligible patients included at the start of the study)

d) the degree to which patient case-mix has been adjusted for, any differences found in outcomes could be attributed to differences in patient case-mix between hospitals. A hierarchy of patient case-mix adjustment has been developed and is outlined below.

Adjustment score	Criteria
0	no case-mix adjustment
I	adjustment for demographic variables
II	adjustment for demographic variables and comorbidity
III	adjustment for demographic variables, comorbidity and stage or severity of illness

Each of the studies has been given a score from 0 to III dependent upon the extent to which adjustment for patient case-mix has been made: the more comprehensive the adjustment the higher the score. Any studies which used a randomised design were given a score of III. The box outlines the scoring system used.

2.5 Data extraction

a) For each relevant study, data was extracted in a systematic way, so as to highlight the type of procedure or condition, the setting, health professionals involved, the methods used and the results found (see Appendix for data extraction sheet).

2.6 Data synthesis

a) Studies are grouped according to the procedure or condition, and within this, studies have been ranked according to the extent of adjustment for patient case-mix (i.e. their case-mix adjustment score).

b) A qualitative overview is presented, taking into account the methodological rigour of each individual study. Details of individual studies are given in table format. Where studies are similar enough (e.g. procedure, volume measure, patient type and outcomes measured) formal pooling of the data has been attempted.

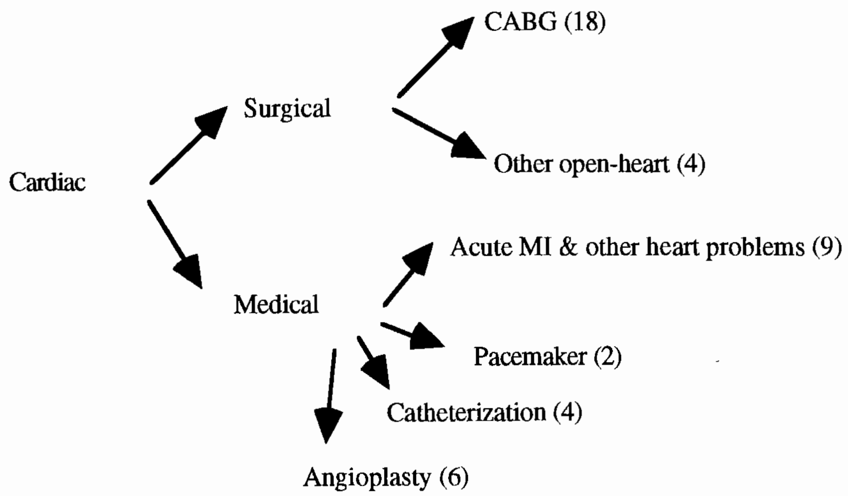
The relevance and quality of each individual study was assessed by one reviewer, and data extraction was carried out by one reviewer. Patient case-mix adjustment scores were allocated by two reviewers.

3. RESULTS

The different procedures and diagnoses for which the relationship between volume and outcome has been assessed are presented, along with the number of studies for each category (some studies assessed the relationship in more than one procedure/diagnosis). A total of 221 evaluations are included in the following categories: cardio-vascular surgery, respiratory medicine, abdominal procedures, orthopaedic surgery, intensive care, urology/gynaecology, trauma care, AIDS, cataract surgery, cancer and miscellaneous (for example patients with cirrhosis).

3.1 Cardiovascular

Cardiovascular surgery was divided into cardiac and vascular procedures, of which cardiac was further sub-divided into surgical and medical. The number of studies which have investigated the volume - outcome relationship for cardiac procedures was:



3.1.1 Coronary artery bypass graft (CABG) surgery

Twenty-four studies have examined the relationship between volume and outcome for CABG surgery, however six[15-20] were excluded since these duplicated analysis of the same data source for the same time period. The 18 remaining studies included were all carried out in the US, between the years 1979 and 1996.[7, 21-37] These studies have mainly used the in-hospital mortality rate as their outcome measure and adjustment for patient case-mix ranged from no adjustments to specific clinical risk factors for cardiac surgery (case-mix adjustment scores ranged from 0 to III).

In one of the five studies with grade III adjustment for patient case-mix the ratios of the adjusted mortality rates for hospitals performing less than 50 procedures per year compared with hospitals performing more than 151 operations decreased from 1.89 in 1990 to 1.36 in 1992.[34] From the same data set, surgeons performing less than 50 procedures annually had higher mortality rates than surgeons performing more than 50 operations.[27] Data from a much earlier time period (1977) was used to examine in-hospital mortality in relation to hospital and surgeon volume. The volume of procedures performed in hospital (mean was 356 per year) was found to be negatively associated with mortality but surgeon volume (mean was 109) was not significantly associated with mortality.[22]

In a more recent study, when 30-day mortality was not adjusted for differences in patient case-mix or adjusted for age only, a significant relationship between hospital volume and mortality was found. One hundred patients per year was the threshold found for maximal statistical significance. However, after risk adjustment (grade III) the significant relationship disappeared and the 100 cases break point was not supported.[37] In a very recent analysis of over 120,370 patients, using data from 1991 to 1993 no thresholds of statistical significance were found except for the lowest volume practice (<100) which had the highest observed and expected mortality (O/E ratio 1.6 to 1.7%). Overall, very weak inverse correlations were found between volume and mortality. The variability of outcome was significant in practices with less than 600 cases per year and varied little at more than 600 cases per year.[36] In the other studies with grade II or lower adjustment for case-mix the majority reported improved outcomes with higher volumes (both hospital and surgeon).[7, 21, 23-26, 28-33, 35] (See Table 1).

In order to further examine the evidence for a volume-outcome relationship in CABG surgery the degree to which the reported relationship reflects differences in patient case-mix was assessed. Data from each individual study were extracted by using the cutoff point closest to 200 procedures per year to define high and low volume hospitals. This particular figure was used as it was the cut off point common to most studies and therefore allowed comparison. Additionally, a number of studies have reported that there is a threshold of around 200 procedures per year. Data from studies in which volume had been analysed as a continuous rather than categorical variable were not included in this pooling as it was not possible to extract the necessary data.[22, 23, 25,37] Studies in which the same data source and time period had been used were only included once (excluded studies were).[15-20,27,35] Two other studies were also excluded from the pooling exercise: the outcome assessed in one study was not mortality[26] and in the other study all hospitals were high volume.[24]

The estimates of benefit associated with higher volume for each study were plotted against the case-mix adjustment scores given to each study, (See Appendix for a discussion of the statistical methods used). Patient variables that have been shown to be significantly associated with increased mortality include: age, sex, previous heart operations, ejection fraction (heart pumping capacity), diabetes, previous myocardial infarction, dialysis dependence, cardiac catheterisation crash, unstable 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, reoperation and mitral valve insufficiency.[20, 23]

In total, ten studies were included in the analysis and the prognostic variables controlled for varied from age and sex to clinical risk factors (see tables 2 and 3 for included and excluded studies). The studies differed in the number of hospitals and patients included and in their volume categories. One study presented data on a 20% sample of elderly Medicare beneficiaries[33] and as it was unclear how the volumes of patients related to hospital volumes the results of this study were only included as part of a sensitivity analysis.

Most of the studies reported a positive relation between volume and outcome. Figure 1 shows the estimate of the benefit (odds ratio of mortality) associated with carrying out more than 200 procedures per year compared with less than 200 procedures per year for each study plotted against the four point case-mix adjustment scale. Studies with more adequate adjustment for case-mix have odds ratios closer to 1 and so lower estimates of the benefit of high volume. Table 4 presents results from the statistical modelling of these trends. Model B shows that the interaction term between volume and the degree of adjustment is significant and greater than one. This means that as the degree of adjustment for case-mix increases, the estimate of the advantage of increased volume is significantly reduced. Figure 2 shows the estimated effect on mortality of high volume hospitals compared with low volume hospitals by year of data collection. The year-volume interaction was not significant.

Figure 1 **How estimates of benefits increased volume (>200) of CABG surgery vary by adequacy of adjustment for case-mix**

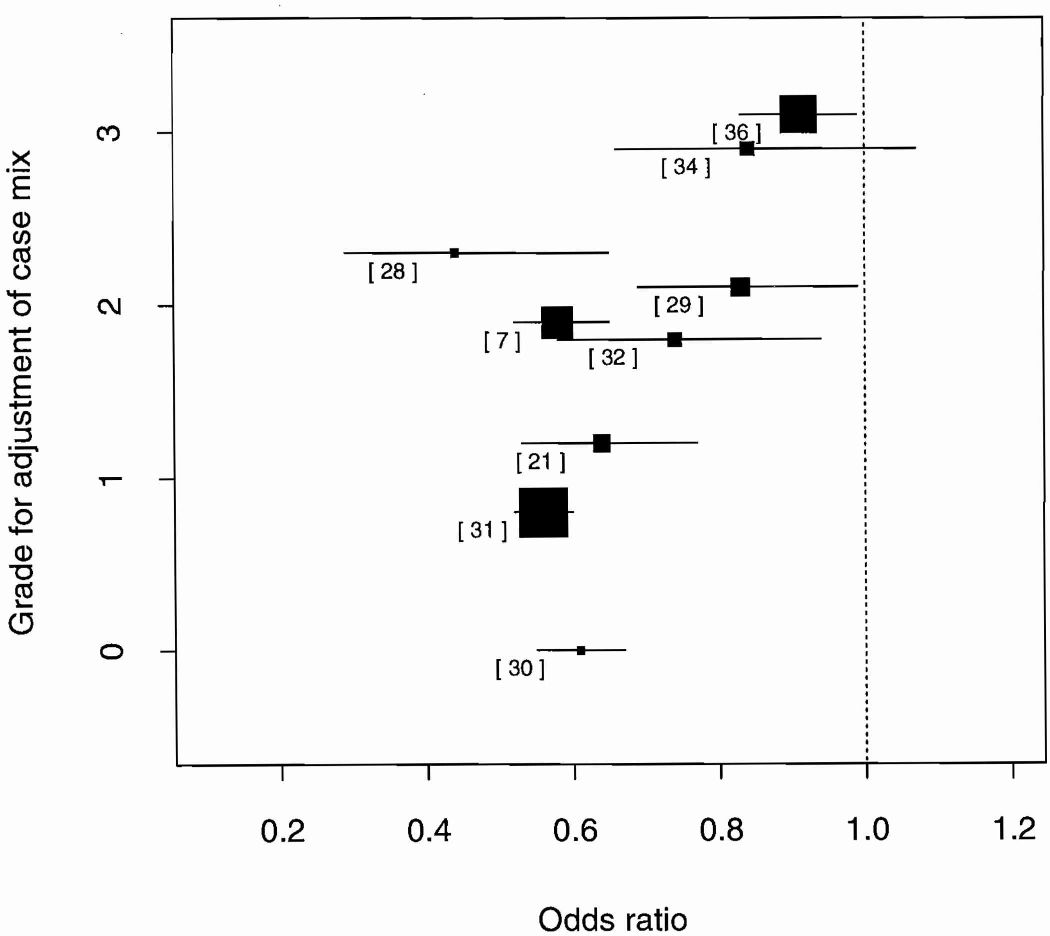


Figure 2 Odds ratio against year

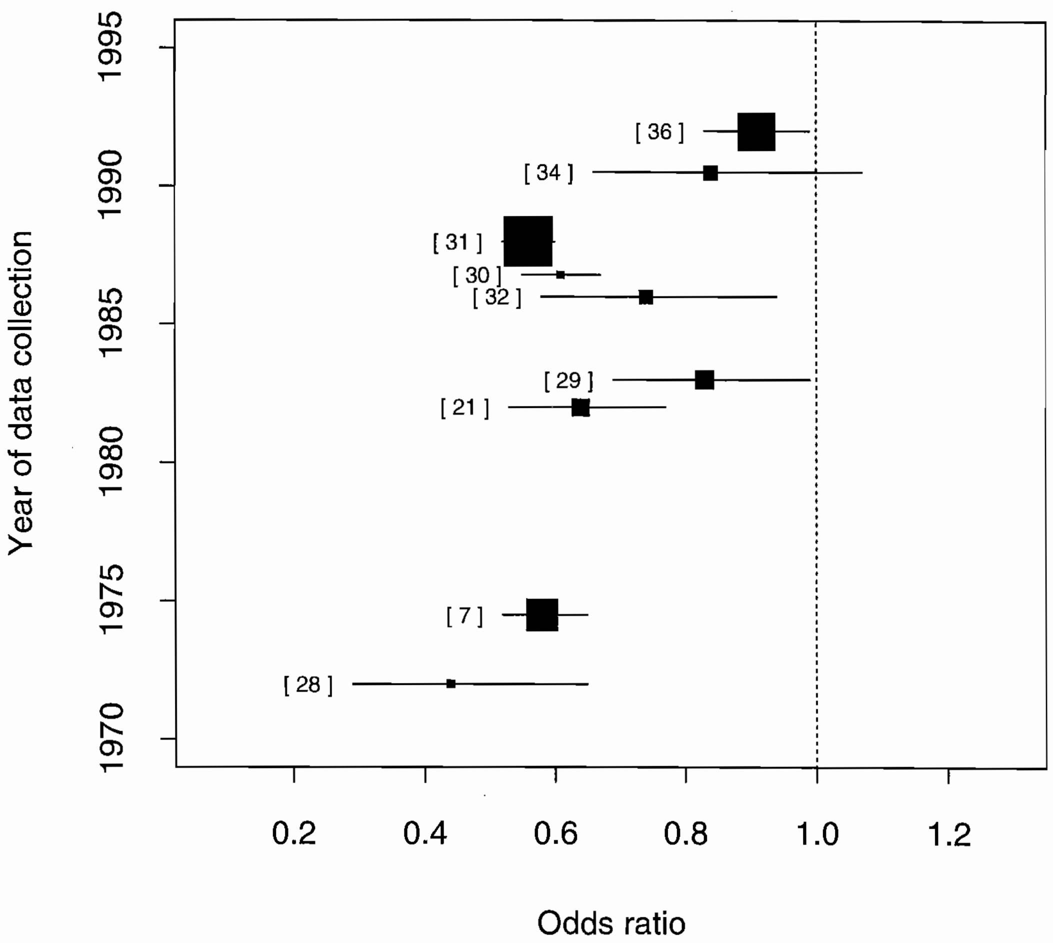


Table 4: Volume and adjustment parameter estimates

Model	Odds ratio (95% CI)	Significance
Model A		
Volume	0.66 (0.57, 0.77)	P<0.001
Model B		
Volume (when adjustment = 0)	0.44 (0.36, 0.53)	P<0.001
Adjustment-Volume interaction	1.25 (1.14, 1.38)	p<0.001
Model C*		
Volume (when year=1972)	0.52 (0.36, 0.74)	P<0.001
Year-Volume interaction	1.02 (0.99, 1.04)	n.s.

* Omits study by Johnson as non-adjusted results were used.[30]

3.1.2 Open heart

Four studies from the United States have investigated the volume-outcome relationship in other types of open heart surgery.[7, 38-40] One study based on a retrospective analysis of discharge abstract data examined the relationship in children undergoing surgery for congenital heart disease.[38] In-hospital mortality and length of stay were the outcomes measured and patients were classified into four categories depending on the complexity of their procedure. Age, sex, race and transfer status were also used to adjust for differences in case-mix (case-mix adjustment score = III). Both unadjusted and adjusted mortality were found to be significantly lower in high volume hospitals. In hospitals dealing with less than 10 cases per year adjusted death rates were around 18.5% and in high volume hospitals (over 300 cases per year) about 3.0%. Unadjusted length of hospital stay did not differ significantly between hospitals, but after adjustment significant differences were found, with a longer length of stay of around 1.6 days in hospitals with less than 10 cases per year. In a study[7] with grade II adjustment for patient case-mix, mortality decreased in patients undergoing open-heart surgery with increasing volume (results in this study were presented graphically).

Two studies have examined the relation between volume and mortality in patients undergoing cardiac transplantation.[39,40] In one study with grade II adjustment for patient case-mix there was a 40% increased risk of post-operative mortality and a 33% increased risk of 12 month mortality in centres performing less than nine transplants per year compared with centres performing more than nine transplants per year.[39] In a second study with grade I adjustment for case-mix total transplantation volume and mortality showed no significant correlation and when the analysis was carried out for sub-groups of patients there was still no significant correlation.[40] (See Table 5).

3.1.3 Myocardial infarction and other heart problems

Nine studies have examined the relationship between volume and outcome for patients with acute myocardial infarction or other heart problems.[22, 23, 25, 28, 41-43] Seven studies were carried out in the United States, one from Australia and one pooled the data from the International Tissue Plasminogen Activator/Stroptokinase Mortality Trial which included 13 countries.[41] Most of the studies used retrospective study designs utilising discharge abstract data to examine the relationship between hospital or physician volume and in-hospital mortality and/or length of stay. The patient variables adjusted for in the analyses were mainly age, sex and the presence of comorbidity (case-mix adjustment scores ranged from I to III).

Seven of the studies focused on the outcomes of patients with acute or suspected myocardial infarction and the findings of these studies were inconsistent. The Australian study used a prospective cohort design to examine the effectiveness of coronary care in relation to level of care provided by individual hospitals. Patients were interviewed to determine their medical history and a severity score was calculated for each individual (case-mix adjustment score = III). Logistic regression failed to detect any evidence of a benefit with more specialised care for any sub-group based on severity of disease. Average mortality in coronary care units was 17% and this did not differ significantly from the mortality rates at other hospitals where facilities were not as elaborate (no information was given on volume levels).[42]

In one study[41] which included data from 13 countries, the outcomes of patients with acute myocardial infarction were compared in relation to the size and comprehensiveness of cardiovascular services. The patient sample was likely to be more homogenous than in most studies reviewed, as all patients met criteria for inclusion in the tissue plasminogen randomised controlled trial. The outcomes of a number of physiological measures (collected at baseline) were included in the adjustment for patient case-mix (case-mix adjustment score = III). Whilst no significant differences in in-patient mortality between the different hospital categories were found, at 6 months post-treatment, mortality was greater in patients treated at smaller centres (less than 300 beds compared with more than 300 beds). Hospital length of stay was also significantly longer for patients treated in small hospitals by about one day, and larger centres had the lowest rate of haemorrhage.

In a third study where case-mix adjustment was III, a significant negative relationship between physician volume and mortality was found, but no relationship between hospital volume and mortality.[22] In the studies with less rigorous adjustment for patient case-mix the results were inconsistent.[23,25,28,43] The two other diagnoses included in this section were congestive heart failure and atrial fibrillation. Case-mix adjustment scores were II and I. Both analyses showed no statistically significant differences between high and low volume hospitals or physicians.[25] (See Table 6)

3.1.4 Pacemaker implantation

Two recent studies have investigated hospital volume and patient outcome after pacemaker implantation, one in the UK[44] and one in the US.[45] Both studies adjusted for age, sex, and comorbidity (case-mix adjustment score = II). The UK study reviewed pacemaker practice in one district general hospital over a 79 month period and compared complication rates with national UK data and data from two large specialist centres.[44] The pacemaker records of 201 transplant patients were retrospectively reviewed and found to have similar age and sex patterns as the 1992 national data. The presenting symptoms and ECG indications for the hospital studied appeared to differ from the 1992 data (no statistical testing reported). The complication rate of 2.5% was similar to the rates achieved at the two large specialist centres (2% & 2.8%). In the US study outcomes were compared both between and within hospitals over time.[45] No significant differences in mortality were detected between hospitals and although volume increased significantly over time the risk adjusted mortality rate did not change significantly. (See Table 7).

3.1.5 Cardiac catheterization/angiography

Four US studies examined the relationship between hospital volume, surgeon volume and patient outcomes in patients undergoing cardiac catheterization/angiography.[22,28,35,46] Three were based on retrospective analyses of patient discharge abstracts from different administrative databases or patient discharge abstracts.[22, 28,35] Patients' age, sex, number of diagnoses and in one study disease stage were used to control for case-mix differences between hospitals. One study which made no adjustment for case-mix, reported a nationwide survey of all institutions with an open-heart surgical team to determine the complication and mortality rate.[46]

In the study with Grade III adjustment for patient case-mix an association between hospital volume and mortality was found but no association between physician volume and mortality (mean hospital volume was 399 cases annually and mean physician volume was

97 cases).[22] A decrease in hospital mortality with increased volume was also found in the less adequately adjusted studies.[28, 35, 46] (See Table 8)

3.1.6 Percutaneous transluminal coronary angioplasty (PTCA)

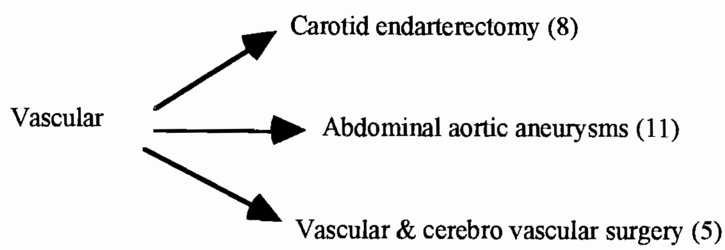
The seven studies that have examined the relationship between volume and outcome for PTCA were all carried out in the US and were based mainly on retrospective analyses of large databases, two of which were clinical registries[47, 48] and three of which were administrative databases.[49-51] One study used a randomised controlled trial design to determine the relationship between equipment size, operator experience and PTCA outcomes[52] and one retrospectively reviewed patients' case records.[53] One was excluded as it used data that was already included.[51] The outcomes measured in these studies ranged from in-hospital mortality to major complications such as emergency bypass surgery and acute myocardial infarction.

The adjustments for patient case-mix ranged from age, sex, race and secondary diagnoses to several clinical predictors (the case-mix adjustment scores ranged from 0 to III). The importance of adequately adjusting for patient case-mix has been demonstrated in an analysis of detailed clinical data from the medical records of 2086 patients undergoing angioplasty.[54] The variables that were found to significantly predict adverse events included age over 75 years, female sex, severe cardiac dysfunction, enlarged heart on x-ray and unstable angina. The ACC/AHA Task Force also reported that patient related factors such as age, sex and clinical variables (such as history of hypertension, diabetes, prior myocardial infarction, prior bypass surgery and impairment of left ventricular function) are associated with procedural mortality.[6]

In the three studies with grade III adjustment for patient case-mix the results indicated that when patients were randomised to physicians who had carried out more than 500 procedures or to physicians who had performed less than 50 procedures, no statistically significant differences in angiographic or clinical success were found.[52] A decrease in major complications was found when hospital volume was greater than 400 procedures per year but no differences in mortality.[48] Physician volume did not significantly affect mortality but more complications, emergency CABG and a longer hospital stay were reported for patients treated by physicians carrying out less than 50 procedures per year.[53]

Studies with grade II adjustment and below all showed positive relationships between increased volume and a range of outcomes.[47,49,50] (See Table 9)

The number of studies which have investigated the volume - outcome relationship for vascular procedures was:



3.1.7 Carotid endarterectomy (CE)

Eight studies have investigated the relationship between volume of CE procedures and outcome.[45,55-61] All studies were from the US and were based mainly on retrospective analyses of administrative databases, except for one study which used data collected prospectively from The Cleveland Vascular Society registry[56] and one study which reviewed patient records from one community hospital.[60] The outcomes measured were in-hospital mortality, in-hospital stroke and length of stay.

None of the studies scored the maximum III points for patient case-mix adjustment (adjustment scores ranged from 0 to II). Three studies had grade II adjustment for case-mix and all of these reported no differences in outcomes between high and low volume hospitals.[55,57,45] In one of these studies mortality and stroke both increased by 50% in patients treated by physicians performing less than 12 procedures per year and length of hospital stay increased by 4 days.[55] In another study however, annual physician caseloads of <12, 12 - 50 and >50 were not associated with differences in mortality or stroke.[57] The other five studies had either grade I or no adjustment for patient case-mix. [56,58-61] (See Table 10).

3.1.8 Abdominal aortic aneurysms

Eleven studies have investigated the relationship between volume and outcome for patients with abdominal aortic aneurysm repair (resection and replacement of the aneurysm with an artificial graft) [7,25,32,56,62-68] Most were carried out in the US or Canada and were based mainly on retrospective analyses of administrative databases. One was carried out in Norway using a prospective multicentre cohort design. All studies used mortality rates as the outcome measure and one also used length of hospital stay. The degree to which patient case-mix was considered varied between studies (case-mix adjustment scores ranged from I to III).

Of great importance in assessing the outcomes of aneurysm surgery is classification according to whether surgery is elective (to prevent acute rupture of the aneurysm) or emergency (rupture or expanding aneurysm) as the associated risk of mortality is dependent upon the condition of the aneurysm. For example, mortality rates were found to range from around 5 to 11% for elective procedures and from 42 to 56% for emergency operations.[62] In another study patients with rupture or fistulas were 12% more likely to die than patients with uncomplicated aneurysm.[65]

In addition, the presence of other medical conditions has also been shown to be predictive of increased mortality in patients undergoing aneurysm surgery. A recent study used 10 years of data from one US community to determine the variables that were independently predictive of mortality after ruptured abdominal aortic aneurysms. The most significant determinants of perioperative mortality were found to be a history of chronic obstructive lung disease and chronic renal insufficiency. Patients with chronic obstructive lung disease had an operative mortality of 78% and patients with chronic renal insufficiency had a mortality rate of 80%.[69] Patients with three or four additional diagnoses were found to have a 4 to 7% greater probability of experiencing a poor outcome after aneurysm surgery compared with patients with only 1 diagnosis.[65]

Outcomes were found to be consistently better at higher volume hospitals, across the five studies with grade III adjustment for case-mix (and the hospital volumes compared across studies were similar). In one study the SMR was 30% higher in hospitals with greater than 14 cases per year.[63] Similarly, a 48% higher mortality rate was found for physicians working in hospitals performing less than five procedures annually compared with hospitals carrying out more than 36 procedures.[32] When the average annual hospital volume was 23 cases per year, for every additional four operations performed per hospital the risk of mortality was reduced by 1%.[65] In hospitals performing less than 21 operations per year there was a 2% increased odds of dying compared with hospitals performing more than 21 procedures per year.[64] When annual hospital volume was below six cases, mortality was 12% compared with 5% when volume was greater than 38 cases per year.[67] When physician volume was examined, one study found no significant differences between high and low volume physicians (mean =10 cases per year),[65] whilst in another mortality was 9% in surgeons performing less than 6 operations per year compared with 4% for those performing more than 26 per year.[67] The results were mixed in studies with a lower grade of adjustment[7,25,56,62,66,68] (See Table 11).

3.1.9 Vascular (and cerebro-vascular) surgery

Four studies from the US were identified which evaluated the volume-outcome relationship in vascular (cerebro-vascular) surgery [7,25,56,63] and 1 study focused on patients with stroke.[28] All were based on retrospective analyses of administrative databases, except for one, where data from The Cleveland Vascular Society computer registry were used.[56] The prognostic factors included in the adjustment for patient case-mix were mainly age, sex and the number of secondary diagnoses (the case-mix adjustment scores ranged from I to III). The outcomes measured were mortality, length of stay and in one study morbidity (defined as major amputation).

In the one study with grade III adjustment for patient case-mix, mortality rates were 16% higher in patients undergoing amputation of the lower limb (no trauma) treated in hospitals with below the average annual volume (mean = 11 cases per year).[63] In the studies with grade II, grade I and no adjustment for patient case-mix the results were conflicting.[7,25,28,56] (See Table 12)

3.1.10 Respiratory

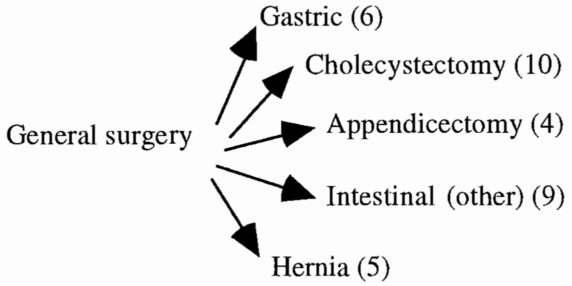
Only one study was found in this area, which examined four types of respiratory problem: pneumonia, chronic pulmonary disease, respiratory failure and respiratory infection.[25]

Mortality and length of hospital stay were the outcomes measured and the analysis was based on patient discharge data. The patient risk factors adjusted for in the analyses were age and sex and in the analysis of patients with pneumonia, comorbidity was also included (case-mix adjustment score = I or II).

The analysis of patients with pneumonia had the best adjustment for patient case-mix (case-mix adjustment score = II) found that higher volume hospitals had significantly greater rates of mortality than lower volume hospitals (mean = 211 cases). Mortality between physicians did not differ significantly (mean = 8 cases). In the analyses with grade I adjustment for patient case-mix the results were inconsistent.[25] (See Table 13)

3.2 Abdominal procedures

Abdominal procedures were divided into two categories: general surgery and medical. The number of studies which have investigated the volume - outcome relationship for general surgery was:



3.2.1 Gastric operations

Six retrospective studies from the US have investigated the relationship between volume and patient outcomes after gastric operations.[7,32,35,63,65] The categories of operations performed ranged from all stomach operations to specific procedures such as vagotomy. All six studies used discharge abstracts as their data source and measured outcome in terms of in-hospital mortality and in one study length of hospital stay was also included.

The patient risk factors adjusted for were age, sex, race, secondary diagnoses, clinical characteristics and stage of illness (case-mix adjustment scores were II or III).

Three studies scored the maximum III points for patient case-mix adjustment. Two of these studies focused on patients undergoing operations for ulcers; one found a significant effect of volume, where every additional 17 operations performed decreased the probability of death by 1%. However, physician volume in the same study did not affect mortality.[65] The other study found no statistically significant differences in outcome between hospitals of different volumes (average procedures performed = 24).[63] In patients undergoing partial gastrectomy there was an increased risk of death when treated by surgeons performing one or less operations on average per year. When low volume surgeons operated in low volume hospitals (less than 5 cases annually) mortality was higher than if they operated in higher volume hospitals.[32] The three remaining analyses which had grade II adjustment for patient case-mix had mixed results.[7, 35] (See Table 14)

3.2.2 Cholecystectomy

Ten studies carried out in the US have examined the relationship between volume and outcome in patients undergoing cholecystectomy.[7,32,33,35,45,63,70,71] All of the studies except for one, were based on retrospective analyses of administrative databases; the remaining study used prospective data from the Virginia Statewide Infection Control Programme to measure rates of post-operative wound infection.[70] Patient age, sex, race, admission status, secondary diagnoses and procedures and severity of illness were used to adjust for differences in case-mix between hospitals (case-mix adjustment scores ranged from 0 to III).

Only two of the ten studies scored the maximum III for patient case-mix adjustment. In both studies the overall mortality rates were low (around 1%) and they reported decreased mortality in higher volume hospitals.[32,63] In hospitals with less than 168 procedures per year mortality was 1.5% as compared with 1.2% in hospitals with more than 168 procedures per year.[32] Physician volume did not affect mortality.[32]

The eight studies where adjustment for patient case-mix was grade II or below had mixed results.[7,33,35,70,71] In the one study where longitudinal as well as cross-sectional analyses were carried out, volume increased over time but mortality did not change significantly (grade II for case-mix adjustment).[45] (See Table 15).

3.2.3 Appendicectomy

Four studies have measured patient outcome after appendicectomy in relation to hospital volume.[28,35,43,70] All were carried out in the US and three were based on retrospective analyses of discharge abstracts from the Commission on Professional and Hospital Activities (CPHA) for the years 1972, 1973, and 1982 and one on data collected prospectively by The Virginia Statewide Infection Control Programme. Adjustments for patient risk factors were age, sex and secondary diagnoses (case-mix adjustment scores ranged from 0 to II). Outcomes of appendicectomy are dependent upon the condition of the appendix at the time of operation (ruptured, inflamed or normal). Poorer outcomes would be expected from the removal of a ruptured appendix and therefore it is important that the condition of the appendix is taken into account when comparing outcomes.

Both grade II studies found better outcomes with increased volume.[28,35] In a grade I study which examined the rate of normal tissue removed - as a function of surgeon volume (mean number of cases per surgeon was 3) - no statistically significant differences in normal tissue removed were found (mean number of patients per physician was 3).[43] (See Table 16).

3.2.4 Intestinal

The nine studies which have compared outcomes in patients undergoing intestinal operations at different volume hospitals or by different volume physicians were all carried out in the US and were mainly based on retrospective analyses of discharge abstract data.[7,25,28,32,33,35,70,72] The outcomes measured were in-hospital mortality, length of hospital stay and in one study the rate of post-operative wound infection. Adjustment for patient case-mix included age, sex, comorbidity and severity of illness (case-mix adjustment scores ranged from 0 to III).

The study with grade III adjustment for patient case-mix found that hospitals with an annual volume of 40 cases or above had a mortality rate for colectomies of 5.9%

compared with 8.3% for hospitals treating less than 40 patients per year. Similarly, surgeons performing more than 8 colectomies per year had better mortality rates than surgeons performing less than 8 per year. The surgeon relationship remained significant independent of the hospital volume.[32]

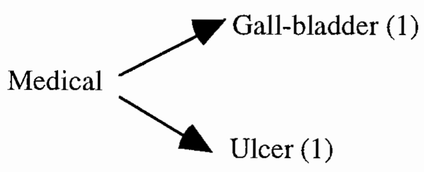
Four of the studies with grade II adjustment for patient case-mix reported improved outcomes at higher volume hospitals (e.g. above 20, above 50, above 130 and a mean of 40 cases per year).[7,28,35,72] One grade II study reported no hospital volume relationship for patients undergoing large bowel resection[25] and for patients with gastro-intestinal bleeding hospitals with higher volume had greater mortality.[25] The 2 studies with poor (or no) adjustment for patient case-mix both found better outcomes (post-operative wound infection and mortality) at higher volume hospitals.[33,70] (See Table 17)

3.2.5 Hernia repair

Five US studies have examined patient outcomes after hernia repair in relation to hospital and surgeon volume.[23,28,33,35,70] Four studies were based on retrospective analyses of discharge abstracts and one which measured rates of post-operative wound infection used prospective data collected by The Virginia Statewide Infection Control Programme. Adjustments for case-mix included age, sex, secondary diagnoses and disease stage (case-mix adjustment scores were 0 or II).

In the three studies with grade II adjustment for patient case-mix, two reported improved outcomes at higher volumes above 380 cases per year[28] and 105 cases per year.[35] In one study comparing mortality in 330 general hospitals, no differences in outcome were reported between hospitals, however, when hospitals were compared over time a significant relationship was found between increased volume and adjusted mortality.[23] In the two studies with either poor or no adjustment for patient case-mix, one found no relationship between volume and either in-hospital or 60-day mortality[33] and the other found that hospital volume was a significant predictor of post-operative wound infection.[70] (See Table 18)

Two studies have investigated the volume - outcome relationship for abdominal medical diagnoses:



3.2.6 Gall-bladder

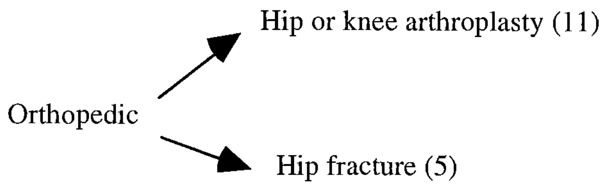
One study has examined the relationship between volume and outcome in patients with a non-surgical gall-bladder diagnosis.[63] This US study based on a retrospective analysis of data from the Commission of Professional Hospital Activities and the Professional Activities study examined outcomes for 88839 cases. Mortality was adjusted for a number of risk factors, including laboratory values (case-mix adjustment score = III). Significantly more deaths than expected were found in high volume hospitals (mean volume was 73 cases per year). (See Table 19)

3.2.7 Ulcer

One study has examined the relationship between volume and outcome in the medical treatment of patients with an ulcer diagnosis.[63]. This US study based on a retrospective analysis of data from the Commission of Professional Hospital Activities and the Professional Activities study examined outcomes for 138268 cases. Mortality was adjusted for a number of risk factors, including laboratory values (case-mix adjustment ore = III) and was found to be non-significantly higher in low volume hospitals (mean annual volume was 114 cases). (See Table 20)

3.3 Orthopaedic surgery

The number of studies identified which have investigated the volume - outcome relationship for orthopaedic surgery was:



3.3.1 Hip or knee arthroplasty

There were eleven studies which have examined the outcome of patients undergoing hip or knee arthroplasty in relation to hospital and surgeon volume. [7,23,25,28,33,35,45,73-75] All eleven studies were carried out in the US and were mainly based on retrospective analyses of discharge abstract data (case-mix adjustment scores ranged from 0 to III).

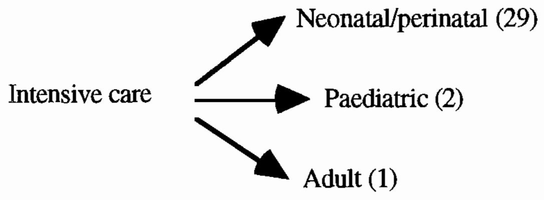
One study with grade III adjustment for patient case-mix assessed the relationship between hospital volume, length of stay and post-operative complications using data collected as part of the Knee Replacement Patient Outcomes Research Team (PORT).[73] The risk of developing post-operative complications was found to be less in higher volume hospitals and also length of hospital stay was shorter (where the average N of treatments was 3.5). Five studies had an adjustment score of II for patient case-mix. [7,23,28,35,45] In two of these studies both cross sectional and longitudinal comparisons were made. In both of the cross sectional analyses, outcomes did not differ significantly between hospitals. Similarly, no effect of volume on mortality over time was found, even when volume increased over time.[23,45] No statistically significant differences in mortality were found according to surgeon volume (where the median number of procedures was 3 per year).[35] In the five studies with poor or no adjustment for patient case-mix findings were mixed.[25,33,74,75] (See Table 21)

3.3.2 Hip fracture

Four studies have examined the relationship between the outcome of patients with a hip fracture in relation to hospital and physician volume at the place of treatment [25,45,63,76] and one in patients with fracture of the femur.[28] All studies were carried out in the US and were based on retrospective analyses of patient discharge abstracts. Mortality and length of stay were the outcomes measured and adjustment for case-mix included age, sex, comorbidity and clinical risk factors (case-mix adjustment scores were II or III). One study with grade III adjustment for patient case-mix found that there was no effect of volume on mortality, where the annual average number of cases was 45.[63] Three studies with grade II adjustment for patient case-mix had mixed results.[25,28,35] In another grade II study where both cross sectional and longitudinal analyses were reported, higher volume hospitals were found to have lower rates of mortality when hospitals were compared (mean number of cases was 3 per year). In the longitudinal analysis however, although volume increased over time, mortality rates remained constant.[45] (See Table 22)

3.4 Intensive care

The number of studies identified which have investigated the volume - outcome relationship for intensive care was:



3.4.1 Neonatal/perinatal intensive care

Nine studies[23,28,77-83] have investigated the relationship between volume and outcome in neonatal/perinatal intensive care, 14 studies have compared outcomes according to the level of care provided[84-97] and 6 studies have evaluated the impact of regionalisation of neonatal/perinatal services on outcomes.[98-103] Most studies were carried out in the United States and were mainly based on retrospective analyses of administrative databases or linked birth and death certificates. The outcome most frequently measured was the in-hospital mortality rate and most studies adjusted for birthweight as a possible confounder (case-mix adjustment scores ranged from 0 to III).

Of the nine studies that assessed the relationship between volume and outcome directly only one scored the maximum grade III in its adjustment for case-mix.[77] This study was based in the UK and used a prospective cohort design to measure survival to discharge in infants treated in intensive care units carrying out a minimum of 500 days of ventilation annually compared with special care units performing less than 500 days of ventilation annually. Infants of less than 28 weeks gestation had significantly better survival rates when treated in intensive care units compared with special care units. However, differences in survival were not significant between more mature infants.

Seven of the remaining studies had grade II adjustment for case-mix and one had no adjustment for case-mix. The results of these studies were mixed; some reported improved outcomes at higher volumes,[23,28,82] one reported no differences in outcome between physicians of different volumes,[79] and some reported improved outcomes for low birth weight babies at higher volume hospitals, but no differences for normal weight babies or in fact poorer outcomes at larger volume hospitals.[78,80,81] In one study with no adjustments for case-mix higher volume hospitals had a higher mortality rate.[83]

In those studies where outcomes were assessed according to hospital level rather than volume or according to regionalisation of services, the findings were mixed. Some studies reported no differences in outcome between hospitals or centres of different levels and others reported better outcomes at larger, level III hospitals, while others reported poorer outcomes at level II or III hospitals compared with level I. (See Table 23)

3.4.2 Paediatric intensive care

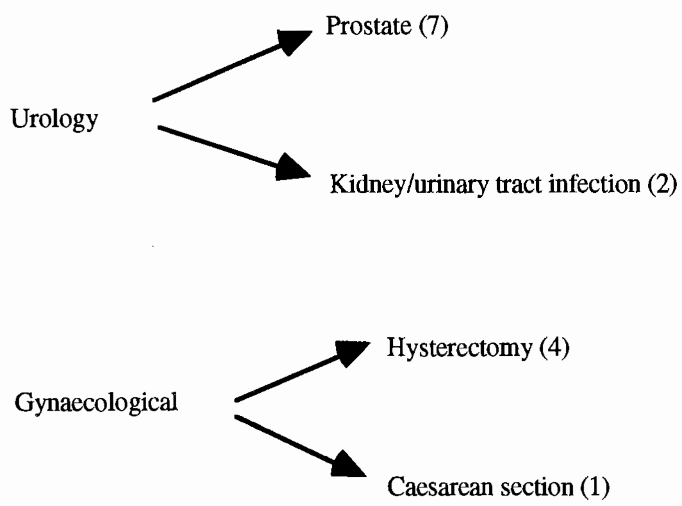
One grade III study based in the US used a prospective cohort design to investigate the relationship between volume and outcome in paediatric intensive care.[104] Mortality was compared (using medical records, after controlling for physiological risk factors using the PRISM score) in 16 different intensive care units of different volumes (monthly volumes ranged from 13 to 63 cases). No statistically significant association between volume per month and mortality was found. An earlier study also with grade III adjustment for case-mix found that adjusted mortality was lower in tertiary centres than in non-tertiary centres.[105] (See Table 24).

3.4.3 Adult intensive care

One study with grade III adjustment for case-mix was based on data derived from the Intensive Care Society's UK APACHE II study and examined the association between the volume of patients admitted to intensive care units and subsequent mortality.[106] 8796 patients aged 15 years and above were included in 26 intensive care units in the UK. Average volume levels ranged from 8.3 to 37.7 cases per month. APACHE II scores based on 14 physiologic variables plus weightings for age and chronic ill-health were calculated for each patient. No significant association between the percentage of patients dying and monthly unit volume was found after adjustment for APACHE II scores. (See Table 25).

3.5 Urology/gynaecology

The number of studies which have investigated the volume - outcome relationship for urology was:



3.5.1 Prostate

The outcomes of patients undergoing prostate operations have been assessed in relation to hospital (and surgeon) volume in seven studies carried out in the US.[7,28,33,35,71,107,108] In-hospital mortality, length of stay and post-operative complications were the outcomes measured. The hospital volumes compared varied dramatically across studies (e.g. cut off volumes of 40, 100 and 200 procedures per year was used to define high and low volume). Adjustments for case-mix included age, sex, comorbidity and medical history (case-mix adjustment scores ranged from I to III).

Only one study scored the maximum III for patient case-mix adjustment.[108] Death within 90-days was measured in 4570 patients undergoing prostatectomy. Volumes compared were less than 40 cases per year, 40 to 90 and more than 90 cases annually. No statistically significant differences were found in mortality between the different volume hospitals.

Five studies had grade II adjustment for patient case-mix.[7,28,35,71,107] Two studies using the same data source and time period reported on the effects of hospital[71] or surgeon volume on mortality[107] and post-operative complications.[71] With the exception of these two studies all reported improved outcomes at higher volumes. A result which was consistent with the study with poor adjustment for patient case-mix.[33] (See Table 26)

3.5.2 Kidney/urinary tract infection and urology

One US study compared mortality and hospital length of stay in patients with kidney/urinary tract infection treated at different volume hospitals.[25] The case-mix adjustment score was II. No statistically significant differences were detected in either mortality or length of stay according to hospital (mean was 99 cases per year) or physician volume (mean was 4 cases per year). One other study with grade II adjustment for patient case-mix measured mortality in patients undergoing urological operations.[110] When surgeon volume was greater than 8 cases (over a 3 year period) mortality was 4% lower compared with surgeon volume of less than five cases for non-emergency patients. In contrast, volume did not significantly affect mortality in emergency patients. (See Table 27)

3.5.3 Hysterectomy

Four studies based in the US have examined outcomes in relation to hospital and surgeon volume in women undergoing hysterectomy.[28,35,70,71] Mortality, post-operative wound infection, complications, readmission and hospital length of stay were the outcomes measured. Adjustments for patient case-mix were either no adjustment (score = 0) or adjustment for age, sex, and comorbidity (score = II). In the three studies with grade II adjustment for patient case-mix, improved outcomes were reported with increasing volume, (where the volumes were above and below 784 cases per year,[28] above and below 100 cases per year [71] and a mean of 142 cases per year[35]). (See Table 28)

3.5.4 Caesarean section

One study based in the US examined the post-operative wound infection rate in 3478 women undergoing caesarean section and found that volume was not a significant predictor of infection.[70] However, no adjustment for possible differences in case-mix were included in this analysis (case-mix adjustment grade = 0). (See Table 29)

3.6 Trauma care

Twenty-four studies (either directly or indirectly) have assessed the relationship between volume and outcome in trauma care.[111-134] Most of these have used a retrospective analysis of hospital records or autopsy reports as their data source. Comparisons have been made between geographical areas where trauma systems are in operation and areas without designated trauma centres, between trauma centres and facilities without trauma centres and also between areas before and after the introduction of a trauma system. Most studies have attempted to control for patient case-mix, usually with the well validated Trauma Score and/or the Injury Severity Score.

The Trauma Score is essentially a field scoring system based on four physiological measures - systolic blood pressure, capillary refill, respiratory rate and respiratory expansion - combined with the Glasgow Coma Scale. Patients' probability of survival has been shown to be associated with Trauma Score values; the Trauma Score ranges from 16 with a related probability of survival of around 99% to 1 with a probability of survival of

0%. The Injury Severity Score is a method for measuring the severity of injuries received in motor vehicle accidents and defines major trauma based on anatomic injury. It has been found to correlate with mortality and an ISS of 16 is predictive of a 10% chance of mortality. Patients' age has also been found to be an important predictor of survival: in the Major Trauma Outcome Study an age greater than 55 years (in the 23,000 patients studied) was found to be significantly associated with increased mortality. Quantitating injury severity is essential in evaluating the quality of trauma care.[135]

Five studies had grade III adjustment for patient case-mix.[115,117,123,130,131] However, only two of these studies measured volume directly in relation to outcome.[123,131] One of these studies[123] evaluated the changes in the performance of an experimental trauma system compared with changes in two comparator regions where a trauma system was not in operation in the UK. Data were collected prospectively and the outcomes measured were mortality and (in survivors) morbidity (follow-up was for six months post incident). Standardised rates of mortality and morbidity did not differ significantly between the experimental and control regions and in further analyses of these data no statistically significant difference in mortality between high and low volume departments (volumes ranged from less than 10 to more than 90 cases per year) in the three regions were found.[123] Similarly, in the second study, no differences were found in mortality in a tertiary trauma unit for patients with mainly blunt injuries as it doubled in volume over a 4-year period.[131] Two studies with grade II adjustment for case-mix also assessed the impact of volume on mortality.[119,128] A significant inverse relationship between volume and unadjusted mortality was reported where high volume was greater than 200 cases over the 22 month study period and low volume was less than 140 cases. However, after adjustment this relationship was no longer significant. In a separate analysis (to predict mortality) the relationship between volume and mortality was shown to be significant, accounting for around 30 to 40% of the observed variation in mortality rates.[128] Institutional volume (annual volume ranged from 60 to 844 cases) was not a significant predictor of mortality, nor was trauma surgeon volume for adult patients with penetrating injuries or for paediatric patients. However, for adult blunt injured patients surgeon volume was a significant predictor of survival.[119]

The 20 remaining studies have assessed outcomes in relation to the introduction of a trauma system or in geographical areas with and without trauma centres. As these studies do not directly measure volume in relation to outcome their results are presented separately from those studies that directly assessed the effect of volume on outcome.

Two studies with grade III adjustment for case-mix focused on outcomes in trauma patients before and after the introduction of special trauma facilities[117,130] and one compared the outcomes of trauma care with national standards.[115] In one study, overall mortality fell from 42% to 26% (although this was not statistically significant) during the period of regionalisation.[117] Mortality in patients with acute spinal injury compared before and after the introduction of an acute spinal cord injury unit decreased significantly after its introduction (145 versus 7.5%). However, when mortality was classified by injury severity score differences were non-significant (e.g 10% versus 8% for before and after).[130] No statistically significant differences in survival were found between national standards and either the outcomes of patients treated in a trauma or a non-trauma centre.[115] In those studies where case-mix adjustment was grade II or below, the results were fairly consistent with reports of improved outcomes after the introduction of a trauma system or in areas with trauma systems. (See Table 30)

3.7 Patients with AIDS

Two American studies have examined the relationship between experience and outcome for people with AIDS (one for all AIDS related diagnoses and one for AIDS patients with pneumocystis carinii pneumonia).[136,137] One study was based on a retrospective analysis of administrative data[136] and one was based on a retrospective analysis of a population based disease surveillance registry.[137] A number of patient risk factors were controlled for, including age, sex, concomitant infection and in one study[137] the Severity Classification for AIDS hospitalisations and the Justice Stage Assessment (which used laboratory and clinical data to classify patients into stages of disease severity) was used. One study[137] had grade III adjustment for patient case-mix and one had grade II.[136] The variables significantly associated with increased mortality in patients with AIDS have been reported in a separate study[138] and included a longer duration of symptoms prior to admission, high respiratory rate, presence of rales and lower haemoglobin. Both studies had similar volume categories (under and over 30 and under and over 42 cases per year) and both found statistically significantly higher rates of mortality in lower volume hospitals. (See Table 31)

3.8 Cataract surgery

One US study used a prospective cohort design to examine the relationship between surgical technique, patient and surgeon characteristics and clinical outcomes (adverse events and visual acuity) in cataract surgery.[139] Around 800 patients and 75 ophthalmologists were included in this study. Detailed information was collected for each patient and entered into a multiple linear logistic regression (case-mix adjustment score = III). High (201 - 399) and very high (400 plus) volume surgeons were found to have a greater rate of adverse events than lower volume (51 - 200) surgeons. The most common adverse event was posterior capsular opacification and the odds ratio for surgeons carrying out more than 200 operations annually for posterior capsular opacification was 2.5. However, no statistically significant association between volume and visual acuity was subsequently found. (See Table 32)

3.9 Cancer

Eighteen studies were included which examined hospital and physician characteristics in relation to outcomes of care for patients with cancer.[17,25,63,65,140-152] The cancer sites included breast, colorectal, pancreatic, malignant teratoma, oesphageal, stomach, lung and childhood cancers. Most studies used population based cancer registries as their data source and measured either in-hospital mortality or survival at 1 to 5 years; one study also measured the incidence of pancreatic fistula. The patient risk factors adjusted for varied between the studies and case-mix adjustment scores ranged from 0 to III. A study scored III if age, sex (where appropriate), stage and other clinical variables had been adjusted for in the analysis. Eight studies were given scores of III.[63,65,141,143,145,148,150]

Two breast cancer studies (one from the US[17] and one from Italy[140]) compared survival at different volume hospitals. One UK study compared surgeons with different volumes.[145] The UK study with grade III adjustment for case-mix found that treatment by a high volume surgeon (over 29 cases per year) improved 5 year survival.[145] The two other studies with grade II adjustment or below had inconsistent results.[17,140] (See Table 34)

Seven studies examined patient outcomes in relation to colorectal volume.[25,63,65,150-152,172] Four studies scored the maximum III points for case-mix adjustment.[63,65,150,172] One of these studies reported increased in-hospital mortality in low volume hospitals (mean annual volume was 17 cases)[63] and one reported no associations between mortality and hospital (mean annual volume was 50.4 cases) or surgeon (mean was 8.4) volume.[65] In the study of patients undergoing laparotomy with colorectal resection, unadjusted mortality and morbidity differed between surgeons with volumes ranging from 44 to 110 cases per year. However, after adjustment using the POSSUM system, no statistically significant differences in mortality or morbidity between surgeons could be detected.[150] Significant variation was found in overall post-operative mortality, local recurrence, anstomotic leak and survival at 10 years in patients treated by one of thirteen surgeons with volumes that ranged from 21 to 98 procedures over the six year study period. However, these variations were not shown to be correlated with volume.[172] In the three studies with grade II or below adjustment for case-mix, the results were inconsistent.[25,151,152]

Two US based studies have focused on patients with pancreatic cancer.[142,148] One study with grade III adjustment for case-mix found that surgeon volume was related to the rate of pancreatic fistula. Patients treated by surgeons with the highest volume (76 cases in 20 months) had the lowest risk of fistula compared to lower volume surgeons in the same hospitals.[148] In a second study where case-mix adjustment was grade II, both higher volume hospitals (over 50 cases during the study period) and surgeons (over 9 cases during the study period) were related to lower in-hospital mortality.[142]

One UK based study examined 5 year risk adjusted mortality (case-mix adjustment score was III) in 454 patients with malignant teratoma.[141] Five-year mortality was found to be 60% lower in patients treated at a cancer unit which treated over 50% of patients with this cancer in the area.

In one UK study examining mortality in 1143 patients with oesophageal cancer, where case-mix adjustment was grade III, there was a 17% lower rate or operative mortality in surgeons performing more than three operations annually. There was a 4% reduction in five year mortality with surgeons treating more than six new cases per year.[143]

One study with grade III adjustment for case-mix examined in-hospital mortality in 341 patients with stomach cancer. No statistically significant association between mortality and either hospital or surgeon volume was found.[65]

One US study focused on 12439 patients undergoing lung cancer resection (case-mix adjustment score = II) and found that the risk of death was significantly related to volume (OR 0.7, 95% CI: 0.6 to 1.0 for 9 to 16 cases per year and OR 0.6, 95% CI: 0.4 to 0.8 for more than 24 procedures per year).[144]

Two UK based studies have examined five year mortality in childhood cancers, both with case-mix adjustment scores of below II.[146,147] One study found no differences in survival rates[147] and one found that children treated in centres with an average of at least six patients per year had lower mortality than those treated in lower volume centres.[146]

One US based study examined mortality and length of hospital stay in 2627 patients undergoing elective oncologic procedures in one medical centre, where adjustment for case-mix was grade III.[149] Lower volume surgeons (5 cases or less over a 3 year period) had higher mortality and longer lengths of stay than higher volume surgeons (8 patients or more over a 3 year period). (See Table 33)

3.10 Miscellaneous

Six studies were classified as miscellaneous and focused on patients with dehydration,[25] cirrhosis,[28] on patients undergoing neurosurgery[153] on children with acute diarrhoeal disease,[154] on patients undergoing laminectomy[70] and on patients undergoing elective general surgical operations.[155] Case-mix adjustment scores ranged from 0 to II. In the four studies with grade II adjustment for patient case-mix the following results were found. For patients with dehydration, being treated by high volume physicians (mean volume = 5 cases per year) increased the chance of dying, whereas hospital volume had no effect. In patients with cirrhosis, being treated at higher volume hospitals improved mortality (low/high volume level was 41 cases per year). Mortality and length of stay for emergency neurosurgical patients treated by low (less than 5 patients over a 3 year period) and high (more than 8 patients over a 3 year period) volume surgeons did not differ significantly. Similarly, for non-emergency patients no significant differences were found in mortality between patients of low and high volume surgeons. However, length of stay was significantly longer for patients treated by low volume surgeons (4.43 days versus 6.88 days respectively). In children with acute diarrhoeal disease a significant weight gain was found in those hospital wards who treated a higher percentage of acute diarrhoeal cases.

In the two studies with no adjustment for patient case-mix, wound infection was significantly increased in patients undergoing laminectomy in lower volume hospitals. In patients having elective surgery, being treated in either a district general hospital or a community hospital did not affect mortality, wound infection rate or length of hospital stay. (See Table 34)

4. DISCUSSION

The literature on links between volume of activity and clinical outcomes suggests that for some procedures or diagnoses there may be some quality gains as volume increases. In other areas the literature suggests an absence of significant volume gains. Therefore, generalisation is not possible. Other reviews of the literature have reported similar findings and have tended to conclude that the evidence was sufficient to support the concentration of services for some procedures.[4, 9,11,156] In contrast, one recent review examining the outcome of patients with solid cancers found that the literature did not support the idea that centralisation of treatment leads to improved results.[14] One UK review examining outcomes in patients undergoing surgical procedures was fairly sceptical about the relationship between volume and outcome and concluded that even though differences in outcome exist between hospitals, the volume-outcome literature does not prove a causal link.[10]

In the procedures or specialities where volume-quality links have been reported (from the best quality studies) the results are still difficult to interpret for a number of reasons.

4.1 Are the outcomes measured appropriate?

Information on in-patient deaths is readily available in the USA from either hospital discharge abstracts or claims data. However, the use of mortality in these studies is problematic for two reasons. Firstly, the mortality rates recorded usually relate to inpatient deaths only or at best 30 day mortality. This may not be a good proxy for longer term survival. Indeed differences in short term survival may simply reflect different discharge policies between hospitals. A hospital that discharges patients at higher risk of dying would as a consequence register a lower inpatient mortality rate. A study by the Stanford Centre for Health Care Research found that a substantial proportion of deaths (for particular conditions) occurred outside the hospital.[157]

Second, and more fundamental, mortality, over whatever time period, 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 and increased survival may sometimes be achieved at high cost to the patient. In hip replacement for example, more appropriate outcome measures are likely to be technical success, morbidity and patient satisfaction,[109] yet in most of the studies reviewed the outcome measured was mortality. Health status is more difficult to quantify than mortality and data on quality of life are not routinely collected and recorded. Length of hospital stay is sometimes measured as a proxy for complication rates, as a longer length of stay might be suggestive of post-operative complications, however, this has rarely been validated. Relatively few of the studies reviewed measured outcomes other than mortality. Although, the studies evaluating outcomes in trauma care tended to focus on avoidable deaths often determined by review of coroners reports, rather than just mortality rates. Mortality would seem to be particularly inappropriate as an outcome measure for low risk procedures. To accurately measure quality of care for such procedures more sensitive measures are needed.

The use of effectiveness rates (such as mortality) alone for measuring quality have also been questioned. It has been suggested that the appropriate use of a medical procedure also needs to be considered as this is an integral component of the quality of care delivered to patients.[158]

4.2 What is high and low volume?

In the majority of studies reviewed volume was measured as the number of patients or procedures per year, although studies have used other periods of time. Some studies have included volume as a continuous measure in a regression analysis to assess whether volume is related to outcome. Others compare the outcome above and below the mean volume for a particular group of hospitals or clinicians, while others choose some threshold or cut-off point to represent high and low volume. In many of the studies where volume is specified as a categorical variable the rationale behind the cut points used to define the categories chosen is not clear. What is considered high and low volume varies enormously between procedures, for example where a hospital performing 50 hip replacements per year might be considered a high volume facility, a hospital performing 200 CABG surgeries per year might be considered low volume. Volumes within procedures also vary considerably (for example in CABG surgery, one study used above and below 200 procedures per year to define high and low volume while a second study adopted a cut-off point of 650 procedures per year). These inter-study variations make it very difficult to summarise the research evidence and translate these findings into useful policy recommendations.

4.3 Hospital or physician volume?

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 and the smaller numbers of events per clinician. Any observed relationship between volume and outcome at the hospital level may in fact be related to the volume or experience of the surgeon carrying out the procedure. Alternatively it may be related to a whole host of variables such as operating room staff, surgical techniques used, etc. In one study for example, where characteristics of high and low volume hospitals were compared, high volume hospitals were found to be larger, more urban, more expensive, and more likely to be teaching hospitals or affiliated with medical schools.[65]

Some studies have examined the relationship between mortality and physician volume. However, it is not clear whether high volume physicians work in higher volume hospitals (or whether the relationship between higher volume hospitals and mortality is really an effect of having high volume physicians). In other words the studies which simply use the physician volume variable, alone or with hospital volume alone or with the hospital volume variable, are not able to distinguish between any hospital or physician effects. In order to be able to sensibly consider the potential contribution of each of these effects together and separately the analysis needs to examine the interaction between hospital and physician volume. One study came close to looking at this.[35] Here the hospital volume and the proportion of patients in the hospital operated on by low volume physicians were examined. However, no interaction terms were included. This study found that independent of hospital volume, the lower the proportion of patients treated by low volume physicians, the lower the mortality rate in a number of surgical procedures. However, physician volume was not as significantly associated with mortality as was hospital volume. Ideally multilevel modelling techniques should be employed which allows the patient, clinician and hospital level effects to be analysed appropriately.[159]

All the physician volume studies are affected by problems with the definition of volume where, for example, the consultant physician may not have been the person who did the operation. Also, it is not clear whether a physician's total experience with other similar operations is taken into account. For example in colorectal cancer surgery the number of similar operations performed may affect outcome.

One study assessed hospitals' total experience with heart transplants to determine whether the outcomes in low volume centres were a function of low volume directly or a function of overall experience. Aggregate outcomes in newly established low volume centres were compared in their initial year to outcomes in the second year. Mortality rates increased significantly in the second year of operation, suggesting that experience assessed as time doing the operation is not a proxy for volume.[39]

If a volume-outcome relationship does exist for clinicians it may be that some of the quality gains are due to higher volume clinicians keeping up with the literature and the use of the most appropriate management practices available as opposed to the acquisition of purely personal (manual) skills. If this is the case then quality improvement may be achieved in other settings by clinicians with lower volume adopting guidelines based upon best practice. Thus where outcomes of a procedure are poor, there may be a number of measures other than attempts to manipulate volume that could possibly improve results.[34]

It is likely that the true effects will be the result of a complex interaction between physician and hospital volume and future research should investigate the correlation between hospital and physician volume in order to better understand the nature of any associations.

4.4 Direction of causality?

Positive relationships between high volume and outcome could be used to support the 'practice makes perfect' hypothesis. 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, in which hospitals with good outcomes attract more patients. It may also be the case that higher-volume hospitals attract better clinicians. The direction of causality is ambiguous since the majority of studies used data from a cross-section of hospitals observed at specific points in time, rather than from a cohort of hospitals over a period of time. Such data are unable to provide the evidence needed to show whether quality would improve if smaller hospitals increased their volume. Changes in volume over time need to be monitored in order to provide such evidence. Only a small number of the studies reviewed used a longitudinal design to assess the effects of changes in volume over time. Where units grew in size over time no improvements were found as hospitals increased their volume even when cross sectional associations had been found.[45, 131]

4.5 Precise definition of procedure

The procedure whose outcome is being measured should be precisely defined because complication rates for similar procedures may differ. Thus any differences detected in outcome may be due to risks associated with particular procedures rather than to the volume of procedures performed. An example is in total hip replacements, where the outcomes for the two main types (cemented and uncemented) differ. Outcomes for non-cemented prostheses have tended to be poorer than for cemented prostheses.[160] Therefore, any comparison in outcomes between hospitals needs to take into account the likely effect of differences in the type of procedure performed.

4.6 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.

Routine hospital mortality data in the USA (derived from hospital claims forms or patient abstracts) include: patients' age, sex, race, principal diagnosis and several secondary diagnoses, the principal procedure and several secondary procedures plus type of admission (emergency or elective). Although researchers tend to 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. Administrative data tends to use fairly simple categories such as whether a disease is or is not present, rather than the severity of the condition. Clinical data such as the results of physical examinations, laboratory tests or radiological procedures, are better predictors of risk. In a review examining the literature on differences in hospital mortality it was reported that out of 16 studies only three used a severity of illness measure to adjust for differences among patients.[161]

The effects of confounding factors may be controlled for using statistical methods. The more statistical adjustment takes into account patient factors which influence outcome, the more likely it is to obtain an unbiased estimate of the influence of volume on outcome. Systems which adjust for risk of dying based on detailed clinical data seem to be the most valid. 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.[162] 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. An example of the importance of adequate adjustment for case-mix can be seen in a study of treatment of benign prostatic hyperplasia. Analysis of claims data suggested that transurethal 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.[163] However, a study which used better adjustment for case-mix such as severity of concomitant illness, found that mortality for the two procedures was similar.[164] The apparent higher mortality rate associated with TURP is likely to have been the result of patient selection; with more frail patients undergoing TURP than open resection (as the former is less traumatic).[165]

In coronary artery bypass graft surgery, hospitals performing less than 100 procedures per year were found to have significantly greater rates of mortality than hospitals performing more than 100 procedures per year. However, when mortality was risk adjusted the differences disappeared.[37] Similarly, a study where the outcomes for stroke patients were compared before and after the introduction of a stroke unit found that crude data suggested that patients identified after the introduction of the stroke service were significantly more likely to be alive. After adjustment for age and sex this relationship remained significant, however, after adjustment for several prognostic indicators the differences between the two groups were non-significant. Thus suggesting that the improvements found after the introduction of the stroke service may have been due to differences in case-mix between the two cohorts rather than the new stroke service.[166]

Even databases which contain prospectively collected clinical data are not free from problems. For example a recent critique of the Cardiac Surgery Reporting System (CSRS), which contains detailed information about all open heart procedures in New York has raised questions about the data collection methods. There were significant increases in the prevalence of several risk factors (such as renal failure, low ejection fraction) which were more likely to reflect changes in coding practices than in patients' characteristics. This spurious increase in risk factors accounted for 66% of the increase in predicted mortality and thus for 41% of the total reduction in statewide risk-adjusted mortality.[167]

The one study in which the volume-outcome relationship was evaluated as part of a randomised controlled trial is a good example of a study design which ensures a fairly homogeneous group of patients from which to compare outcomes.[52] Patients entered into a randomised controlled trial have to meet certain criteria which ensures a degree of homogeneity more than would be expected in observational studies. Therefore, differences detected between the outcomes of patients treated at different volume hospitals or by different volume surgeons are more likely to be attributable to differences in quality of care rather than differences in patient case-mix. In this rather small study no statistically significant differences in outcome were detected between the two groups receiving angioplasty. Future research could well benefit from such an approach to evaluation of the volume-outcome relationship.

Patient case-mix is also important in trying to apply the results in a practical setting. Most studies report the results for the whole patient sample, not taking into account any possible differences in the volume effect for different types of patients. This ignores the possibility that the volume effect, if it exists at all, may only operate on subgroups of patients with particular characteristics, such as baseline risk. It might be the case, for example, that the benefits of being treated in a larger unit, by more experienced clinicians are confined to those patients who have complications or are at high risk of mortality. Some studies in the perinatal area have examined this issue. For example, one study found that the advantages of Level III hospitals were confined to babies weighing less than 5lb.[78] Another study found that although very low birthweight babies had better survival when born in Level III facilities, normal and low birthweight babies did better when born in Level II hospitals, though case mix adjustment grade was only II in this study.[95] Future research therefore, needs to look at within hospital heterogeneity of case mix.

4.7 Random events

Outcomes of care are a function of not only patient characteristics and quality of care but also of random events. When outcomes differ between hospitals or clinicians it is important to consider how much of the variation could have occurred by chance. A study where a re-analysis (using more appropriate statistical methods) was performed on computerised discharge data for all adult patients with pneumonia in central Pennsylvania found that variation which was explainable as random had been incorrectly attributed to variations in hospital quality.[168]

5. CONCLUSIONS

Overall, the literature on links between volume of activity and clinical outcomes suggests that for some procedures or specialities there may be some quality gains as hospital or clinician volume increases. In other areas the research suggests an absence of statistically significant volume gains. However, any association found may be confounded by other variables such as differences in patient case-mix between high and low volume hospitals (or clinicians). The bulk of the research, because it does not sufficiently take into account case-mix differences, probably overestimates the size of the impact of volume on the quality of care. In the few cases where volume quality links have been suggested by more reliable studies, the thresholds indicated in some studies are relatively low and could be reached through specialisation of tasks within a hospital rather than an increase in the size of the provider. However, where volume is associated with quality, the direction of causation is not established. It is difficult to use findings of a positive relationship between volume and outcome across hospitals or clinicians to infer what would happen to healthcare outcomes if existing low volume units expanded.

It is likely that outcomes are dependent upon a combination of factors reflecting hospital, clinician and indeed patient characteristics. Future research needs to move away from simple analysis of statistical linear associations and explore the complex range of factors which may influence health care outcomes.

TABLES AND APPENDICES

Table 1 Coronary Artery Bypass Graft (CABG)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Clark et al (1996)[36] USA To determine the relationship between average surgeon volume and outcome in CABG surgery	Patients undergoing CABG surgery = 120,377 Practices = 180	Retrospective review of the National Cardiac Database 1991 - 1993	Operative mortality	Annual caseload: mean number of cases per month x 12 = ≤100 101-150 151-200 201-300 301-400 401-500 501-600 601-700 701-800 801-900 >900	Risk model based on a number of demographic, comorbidity and physiological factors Correlation tests, t-tests, linear regression and logistic regression	No thresholds of statistical significance were found except for the lowest volume group which had the highest observed and expected mortality (O/E ratio1.6 to 1.7%) Weak to very weak inverse correlations were found between volume and mortality (coefficient = -0.0003, odds ratio of 1.000) Variability of outcome was significant in lower volume practices (<600) and varied little at more than 600 cases per year	Case-mix adjustment score = 3

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hannan et al (1994)[34] and Hannan et al (1995)[27] USA To assess changes in outcomes of coronary artery bypass graft surgery (CABG) in New York since 1989. To examine the longitudinal relationship between surgeon volume and in-hospital mortality for coronary artery bypass graft surgery (CABG) in New York State.	Patients undergoing CABG surgery = 57187 Hospitals = 30 Physicians: 528 (over the 4 year period)	Retrospective analysis of the Cardiac Surgery Reporting System	In-hospital mortality	Annual volumes were: 1989 = 12269 1990 = 13946 1991 = 14944 1992 = 16028 Surgeon volume: <50, n=143, patients = 3577 51-100, n=135, patients = 10371 101-150, n=106, patients = 10514 >151, n=144, patients = 29957	42 potential risk factors, including demographic data, admission status, preoperative complications Logistic regression	Adjusted in-hospital mortality decreased over the 4 year period, 4.17% to 2.45% and the average severity of illness increased for all volume categories (2.62% to 3.54%). The risk adjusted mortality rate in 1989 for the lowest volume hospital was 5.45 v 3.77 for the highest volume hospital and in 1992 the rates for the same hospitals were 1.44 v 1.86 respectively. Low volume surgeons had higher risk adjusted mortality rates but the relationship became weaker over the 4 years. Surgeons performing <50 operations in 1989 had significantly higher mortality rates than the other 3 groups (7.94% vs 3.57%, for >151 procedures). In 1991 and 1992 there were no significant differences in adjusted mortality rates among any of the volume groups. The ratios of the adjusted mortality rates for <50 operations to the adjusted rates of >151 operations decreased to 5.72/3.03 = 1.89 in 1990 and 1.39 and 1.36 in 1991 and 1992 respectively. In 1989 low volume surgeons had a more severely ill mix of patients than the other surgeon groups (2.76% vs 2.69%, 2.45% and 2.71%). In 1990 and 1991 the lowest volume surgeons had the least risky case mix.	Case-mix adjustment score = 3

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kelly & Hellinger (1987)[22] USA To examine the effects of selected characteristics of hospitals and physicians on the mortality rates of heart patients	Patients undergoing CABG surgery = 3883 Short term hospitals across the US = 26 Physicians = 99	Retrospective analysis of patient discharge abstracts from the Hospital Cost and Utilisation Project 1977	In-hospital mortality	Physician and hospital volume were expressed as continuous variables	Age, sex, number of diagnoses and disease stage (extent of complications) Multivariate regression	The volume of procedures performed was negatively associated with adjusted in-hospital mortality (coefficient = -0.005, $p = 0.05$). Physician volume was not significantly associated with in-hospital mortality Age was significantly associated with mortality ($p = 0.05$) with patients aged 60 and above being more likely to die. Above the age of 70 mortality increased dramatically, patients were 5.7% more likely to die than patients aged between 18 - 45 years The presence of additional diagnoses significantly increased the chance of mortality as did disease stage ($p = 0.05$)	Case-mix adjustment score = 3

Table 1 Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Shroyer et al (1996)[37] USA To determine whether risk-adjusted CABG mortality rates for patients in VA medical centres are significantly related to volume	Patients undergoing CABG surgery = 23986 VA hospitals = 44	Retrospective analysis of data from the VA Continuous Improvement in Cardiac Surgery Study 1987 - 1992	30-day mortality	ANOVA found an annual threshold of: <100 >100	Previous heart operation, surgical priority, New York Heart Association functional class, age, pulmonary rates, peripheral vascular disease, cerebral vascular disease ANOVA and Poisson regression models (5-year analysis)	When there was no adjustment for case-mix, or adjustment for age only, a significant relationship between hospital volume and mortality was found (p<0.02) 100 patients per year was the threshold found for maximal statistical significance (mean O/E 1.26 v 0.95, p<0.03) Low volume hospitals had greater numbers of patient with risk factors for operative death The Poisson Additive model analysis showed no systematic relationship between risk-adjusted mortality and hospital volume. O/E ratio did not change with volume and 100 cases break point was not supported	7% of records were excluded due to missing data Case-mix adjustment score = 3

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991) [25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients undergoing CABG surgery = 2565 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 243.17 (s.d. 122.73) Physician volume = 21.25 (s.d. 30.41)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	High volume hospitals had significantly lower mortality than low volume hospitals (coefficient = -0.489, $p<0.05$), as did low volume physicians (coefficient = -0.302, $p<0.001$) High volume hospitals and physicians had significantly shorter LOS than low volume hospitals and physicians (coefficient = -0.026, $p<0.05$ and -0.025, $p<0.05$ respectively)	Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farley et al (1992)[23] USA To examine how patient outcomes for several procedures respond to changes in volume over time	Patients undergoing CABG surgery = 146890 Short term general hospitals = 62	Retrospective analysis of data collected as part of the Hospital Cost and Utilisation Project (discharge abstract data) 1980 - 1987	Mortality	Volume was measured as a continuous variable Both within and between hospital differences were measured	Age, disease staging Multivariate regression	A significant relationship between increased volume over time and adjusted mortality was found (coefficient = -3.772 p<0.001) A significant effect of volume on adjusted mortality was also found in the cross sectional analysis (coefficient = -1.263, p<0.001)	Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hannan et al (1989)[32] USA To test the combined relationship of hospital and physician volume with in-hospital mortality rates for several procedures	Patients undergoing CABG surgery = 9774 hospitals = 27 physicians = 353	Retrospective analysis of New York State discharge abstracts from the Statewide Planning and Research Cooperation System 1986	In-hospital mortality	A continuous volume measure was used which logistic regression defined as: hospital volume: > 223, n = 10, cases = 1530 224 - 309, n = 5, cases = 1355 310 - 650, n = 8, cases = 3661 651 - 1081, n = 4, cases = 3228 physician volume > 116, n = 355, cases = 4834 >116, n = 28, cases = 4940 (low and high physician volume was also defined within each of the hospital volume categories)	Age, sex, race, admission status, upto 4 secondary diagnoses and procedures, and severity of illness (measured in terms of diagnosis). Logistic regression	The overall unadjusted mortality rate was 4.5% Age ($p<0.001$), sex ($p<0.01$), admission status ($p<0.01$), number of secondary diagnoses($p<0.001$) and disease stage ($p<0.001$) were significantly associated with mortality in the regression analysis as was physician volume Physician volume had a significant impact on risk adjusted mortality: 4.92% vs 4.04% for low and high volume respectively ($p = 0.05$). Hospitals performing less than 650 procedures per year had a risk adjusted mortality rate of 5.28% vs 3.82% for hospitals performing more than 650 procedures per year; the ratio of these percentages was1.38 ($p = 0.05$) Higher volume physicians operated on lower risk patients	There were originally 16 procedures included in this study, but only those 5 which had statistically significant volume-mortality relationships were presented Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing CABG surgery = 29503 Hospitals = 120 Surgeons = 800	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Surgeon volume: median number of procedures per year = 12, used to differentiate less experieinced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Low surgeon volume was positively related to better outcomes (-4.2107, $p<0.10$)	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Only results for surgeon volume presented. Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Johnson (1988)[30] USA To empirically examine whether the volume of bypass surgey performed in 1986 was an indicator of superior outcomes and quality.	Patients undergoing CABG surgery = 2034 Metropolitan hospitals in the Twin Cities = 12 Surgeons = 180	Retrospective analysis of the Council of Hospital Corportations Medical record abstract database 1986	In-hospital mortality and length of stay	Hospital volume: < 200 > 200	Age, sex and patient severity (secondary diagnoses and/or procedures) Multiple regression	A simple correlation between hospital volume and mortality (-0.40) was found which disappeared when patient factors were taken into account	273 (12%) eligible patients were excluded due to involvement of other unspecified bypass surgery or because they were single bypass surgery only Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Leape et al (1993)[26] USA To determine the appropriateness of use of CABG surgery in New York State (volume was included as a hospital characteristic)	Patients undergoing CABG surgery = 1338 New York State hospitals = 15	Random selection of hospitals stratified according to New York upstate/ downstate location and volume in 1989, 15 out of the 30 were then selected A random selection of 1426 medical records were retrospectively reviewed	30 day in-hospital mortality and complications	Low volume < 325 High volume > 325	Age, risk (using the Parsonnet score), angiographic disease category and emergency status. Logistic regression	Mortality was significantly higher for patients aged 75 and over than for patients under 75 (RR 4.2, 95% CI: 1.1 - 13.6), who had cardiogenic shock (30%), PTCA complications (15%) and congestive heart failure (11%), (RR for all 3 10.1% 95% CI: 3.8 - 22.9) compared with all other patients Mortality was significantly lower with chronic stable angina (0.5%) and in asymptomatic patients (0.0%) (RR for both 0.1 95% CI: 0.0 - 0.5) Mortality rates and complications were related to operative risk as indicated by the Parsonnet score. Patients in the high risk category were more likely to die (4.5% v 0.2%) RR 21, 95% CI: 5 - 74) or to have complications (32% v 11%) OR 3.0 95% CI: 2.6 - 3.5) man were patients in the low risk category. After adjustment differences in operative mortality were non-significant, but complication rates varied significantly from 9% - 26% (p = 0.009) There were no significant differences in mortality between high and low volume hospitals (10%, 95% CI: 7 - 13 v 13%, 95% CI: 9 - 16 respectively) in high risk patients	8% of eligible cases were excluded to missing data, unable to locate record, miscoded or because another major procedure was also being performed Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing CABG surgery: n = 17165 Hospitals = 182	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: > 200 = 157 hospitals, 9549 patients < 200 = 25 hospitals, 157 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume (presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing CABG surgery = 5172 Hospitals = 114	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 215 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.020$, $p<0.01$)	Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Showstack et al (1987)[29] USA To address the question of whether a volume- outcome relationship exists for patients undergoing CABG surgery with comorbidity	Patients undergoing CABG surgery = 18986, of which 11505 were scheduled and 7491 were non-scheduled California hospitals = 77	Retrospective analysis of patient discharge abstracts from the California Health Facilities Commission 1983	In-hospital mortality and poor outcome (total of in-hospital death and long post-operative stay for survivors)	20 - 100: 12 hospitals, 759 patients 101 - 200: 22 hospitals, 3038 patients 201 - 350: 26 hospitals, 6835 patients > 350: 17 hospitals, 8354 patients	Age, sex, ethnic group, presence of acute MI, congestive heart failure, angina, coronary catheterization and angioplasty	Patients undergoing non-scheduled surgery were more likely to have had a current MI (22% v 9%, $p<0.001$) and.or congestive heart failure (6% v 2%, $p<0.001$) and/or cardiac catheeterization (72% v 23%, $p<0.001$) Sex, age, percent with acute MI, congestive heart failure, angina, cardiac catheterization and angioplasty were all associated with poor outcome or mortality Adjusted mortality was significantly higher in low volume hospitals than in high volume hospitals (5.2% v 3.1%, $p<0.01$), for non-scheduled this was also true (7.7% v 4.6%, $p<0.01$) but for scheduled patients the difference was not statistically significant Poor outcomes adjusted for risk were significantly higher in low volume hospitals than in high volume hospitals (21.7% v 12.0%, $p<0.001$) and remained true for scheduled (16.1% v 9.3%, $p<0.001$) and non-scheduled procedures (27.9% v 16.3%, $p<0.001$)	Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Williams et al (1991[24] USA To measure hospital and surgeon specific mortality rates for patients with CABG surgery and to examine possible reasons for any differences	Patients undergoing CABG surgery = 4613 Medical schools in Philadelphia = 5	Retrospective analysis of the discharge abstracts and itemised bills of consecutive patients in the 5 hospitals July 1985 - December 1987	In-hospital mortality	Volume categories were not specified, the number of procedures and outcomes achieved were presented for each hospital individually	Age and disease stage (illness severity was based on a disease staging system that classified patients into 1 of 3 stages) Univariate analyses and logistic regression	Patient age and disease severity were significantly associated with mortality ($p<0.0001$). Each 10 year age increase was associated with an increase in the mortality odds ratio of 1.5 and the mortality rate almost doubled for each increase in disease stage (3.6 in stage 1, 6.7% in stage 2 and 14.0 in stage 3) and differed across the 5 hospitals) There were no significant differences in mortality rates between hospitals for CABG only procedures. For CABG procedures with catheterization significant differences occurred; patients in hospitals with either 1493 or 1097 procedures per year had lower mortality rates than patients in hospitals with either 1123, 503 or 397 procedures per year ($p<0.0004$ [rates not given]) No consistent differences between surgeon volume and mortality were found, for example the highest volume surgeon (588 procedures) and the lowest volume surgeon (138 procedures) had similar adjusted mortality rates (5.9% v 7.4%) No statistically significant association was found between hospital volume and mortality	Case-mix adjustment score = 2

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Grumbach et al (1995)[31] USA To determine how regionalization of facilities for CABG surgery affects geographic access to CABG and surgical outcomes	Patients undegoing CABG = 116593 Hospitals = 157	Retrospective analysis of computerised hospital discharge abstracts for New York, and California and provincial health plans in Ontario, British Columbia and Manitoba 1987 to 1989	In-patient mortality within 14 days of CABG	Hospital volume: <100: n = 37, patients = 5877 100-199: n = 42, patients = 16247 200-499: n = 53, patients = 46008 > 500: n = 25, patients = 48464	Age and sex Adjustments made using the overall population of patients receiving CABG in California hospitals from 1987 to 1989 as the referent. The expected number of deaths at each hospital was compared with the observed number and an adjusted rate was then calculated Regression models were used on the California data to predict expected, calculate adjusted mortality using 22 predictor variables	Unadjusted (crude) mortality rates for hospitals < 100 cases per year in New York were 3.5% vs 2.5% for >500 cases, and the adjusted rates were 4.1% vs 2.6% (p<0.001). For Canadian hospitals performing 100 - 199 procedures unadjusted mortality was 4.1% vs 2.9% for > 500 and adjusted mortality was 4.3% vs 3.3% (n.s.) The California adjusted mortality rates using predictor variables were 4.7% among hospitals with volumes < 100 cases per year and 2.4% for hospitals with > 500 cases per year (p<0.001)	No Canadian hospital performed less than 100 operations Comparisons were made between hospitals <100 vs > 500 cases per year for adjusted mortality rates only Case-mix adjustment score = 1

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged Medicare beneficiaries	Aged Medicare patients undergoing CABG surgery = 6157 Short stay hospital = 990	Retrospective analysis of the Medicare Statistical System of the Health Care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 7, 1697 patients 7 - 11, 1524 patients 12 -19, 1437 patients > 19, 1499 patients	Age and sex Multiple regression analysis	Age was significantly associated with mortality ($p<0.001$) High surgical volume was significantly associated with lower 60 day mortality (coefficient = -0.13, $p<0.03$) and lower in-hospital mortality (coefficient -0.15, $p<0.03$)	This was a 20% probability sample Case-mix adjustment score = 1

Table 1: Coronary artery bypass graft surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Rosenfeld et al (1987)[21] USA To explore reasons for improved in-hospital survival rate for CABG surgery between 1972 and 1982	Patients undergoing CABG surgery in 1972: = 6469 and in 1982, = 29488 Hospitals = 111 in 1972 and 107 in 1982	Retrospective analysis of case abstract data from the Commission on Professional Activities 1972 and 1982	In-hospital mortality	Volume in 1972: < 90, 2390 patients 90 - 149, 1371 patients > 150, 2708 patients Volume in 1982: < 90, 754 patients 90 - 149, 1754 patients > 150, 26980 patients	Age and sex were adjusted for by dividing the 1982 proportion of patients in each age - sex cell by the 1972 cell specific death rates and summing over all cells	Unadjusted mortality in 1972 was 6.0% v 3.4% in 1982. Adjusted mortality in 1972 was 7.7% v 6.0% in 1982. Increased age and female sex were associated with mortality. In 1982 the lowest volume group had an actual rate of 6.4% and an expected rate of 7.6 and hospitals with > 150 cases had an actual rate of 3.2% and an expected rate of 4.9	The 1982 data is a two-thirds sample of eligible hospitals and cases Only results for 1982 presented. Case-mix adjustment score = 1

Table 2 Studies Included in the CABG analysis

Study (Reference)	Data Source	Years Used	Cut-off point used to define low volume hospitals	Adjusted* crude 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 i high to low hospitals
Studies included in the main analysis								
Maerki 1986 (28)	CPHA discharge abstracts	1972	215	6.8% (6.0% , 7.6%)	age, sex and single or multiple diagnosis, admission blood pressure (2)	Low volume: 109 High volume: 5	Low volume: 246/3619 High volume: 28/905	0.44 (0.29, 0.65)
Luft 1979 (7)	CPHA discharge abstracts	1974-5	200	5.7% (5.4%, 6.0%)	age, sex and single or multiple diagnosis (2)	Low volume: 157 High volume: 25	Low volume: 1089/19098 High volume: 518/15232	0.58 (0.52, 0.65)
Rosenfeld 1987** (21)	CPHA discharge abstracts	1982	150	5.0% (4.1%, 5.8%)	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 (29)	CHFC discharge abstracts	1983	200	4.1% (3.5%, 4.7%)	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)
Johnson 1988 (30)	Minneapolis/ St. Paul CHC medical abstract records	1986	200	6.6% (4.6%, 8.6%)	no adjustment (0)†	Low volume: 8 High volume: 4	Low volume: 39/592 High volume: 58/1442	0.59 (0.39, 0.89)
Grumbach 1995 (31)	New York, California, Canada discharge abstracts	1987-9	200	5.3% (5.0%, 5.6%)	Age and sex (1)	Low volume: 79 High volume: 78	Low volume: 1182/22124 High volume: 2817/94472	0.55 (0.51, 0.58)
Hannan 1989 (32)	New York State discharge abstracts	1986	223	5.6% (4.4%, 6.8%)	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 (34)	New York State CSRS prospective clinical data	1989-92	200	3.8% (2.9%, 4.7%)	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)
Clark 1996 (36)	Society of Thoracic Surgery National Cardiac Database	1991-93	200	3.0% (2.8%, 3.2%)	STS non stratification system. Score based on clinical information, co-morbidity and socio-economic charactiersitcs of the patient (3)	Data not provided	Low volume: 581/19381 High volume: 2766/100996	0.91 (0.83, 0.99)
Studies included in a sensitivity analysis								
Riley 1985 (33)	20% Sample of Elderly Medicare Beneficiaries	1979-80	19 (not inflated from the 20% sample)	6.5% (5.7%, 7.2%)	Age and sex (1)	Total: 909	Low volume: 313/4658 High volume: 85/1499	0.83 (0.65, 1.07)

* Different adjustments were made between the studies as indicated in column 6.
** Unadjusted mortality rates are shown: Results of controlling for case-mix were presented as Tobit regressions - no volume effect was found
† Adjustment for case mix was used in this study. However, the results were presented in such a way that we were able only to use unadjusted death rates.

Table 3 Studies excluded from the CABG analysis

Study (Reference)	Data Source	Years Used	Reason for Exclusion
Studies not included			
Rosenfeld 1987** (21)	CPHA	1972	Same data source as Maerki 1986
Luft 1987 (15)	CPHA	1972	Volume categorisations are not clear. Same data source as Maerki 1986
Luft 1980 (16)	CPHA	1974-1975	Analysis presented as a regression model. Same data source as Luft 1979
Kelly 1987 (22)	HCUP	1977	Analysis presented as a regression model
Sloan 1986 (17)	CPHA	1972-1981	Volume categorisations are not clear. Same data source as Maerki 1986 and Luft 1979
Hughes 1987 (35)	CPHA	1982	Analysis presented as a regression model. Same data source as Rosenfeld 1987
Freeland 1987 (18)	CHFC	1983	Same data source as Showstack 1987
Zelen 1991 (19)	New York State	1986	Same data source as Hannan 1989
Hannan 1991 (20)	New York State	1989	Same data source as Hannan 1994
Farley 1992 (23)	HCUP	1980-1987	Analysis presented as a regression model
Williams 1991 (24)	Philadelphia's medical schools	1985-1987	All hospitals were high volume (>200 cases per year)
Burns 1991 (25)	SDHS	1988	Analysis presented as a regression model
Leape 1993 (26)	New York State	1990	Outcome used: Appropriateness of treatment
Hannan 1995 (27)	New York State	1989-1992	Same data as Hannan 1994
Shroyer, 1996 (37)	VA CICSS	1987-1992	Analysis presented as a regression model

* Different adjustments were made between the studies as indicated in column 6.
** Rosenfeld 1987 presented data for both 1972 and 1982. Only the 1982 data were included as the 1972 data are already included in the study by Maerki, 1986.

Abbreviations used in Table 2a and 2b
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 State.
VA CICSS: VA Continuous Improvement in Cardiac Surgery Study.

Table 4 Volume and adjustment parameter estimates - see p.11

Table 5 Open heart surgery

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Jenkins et al (1995)[38] USA To examine the impact of hospital case-load on in-hospital mortality for paediatric congenital heart surgery in California and Massachusetts	Children undergoing surgery for congenital heart disease = 2833 Acute care hospitals = 37	Retrospective analysis of statewide discharge abstract data from the California Office of Statewide Health Planning and Development and the Rate Setting Commission 1988 (California) 1989 (Massachusetts)	In-hospital mortality and length of stay (LOS)	Annual volume <10: 16 hospitals, 26 patients 10 - 100: 10 Hospitals, 583 patients 101 - 300: 9 hospitals, 1261 patients >300: 2 hospitals, 963 patients	Age, sex, race, transfer status, presence of cardiopulmonary bypass code and patients were classified according to the complexity of the procedure (4 categories) X^2 test or ANOVA were used to test for differences across groups of patients and multivariable regression was used to adjust for the intra-institutional correlation among patients	Overall, in-hospital mortality was 7.7% and varied significantly among complexity categories (p<0.001) as did LOS (p = 0.0001). Patients in category 4 (greatest complexity) had greater mortality and longer LOS than those in category 1 (15.9% vs 4.8% and 16.6 vs 11.0 days respectively) Patient characteristics varied according to volume; high volume hospitals had a higher percentage of complex patients (p = 0.0001) Unadjusted mortality was lowest at high volume hospitals (p = 0.05). Adjusted death rates were significantly lower at high volume hospitals (18.5, 95% CI: 4.7 - 50.9 for hospitals <10 cases vs 3.0, 95% CI: 2.1 - 4.3 for hospitals with >300 cases) Unadjusted LOS did not differ significantly by volume, although after adjustment it was longer by 1.6 days at hospitals with <10 cases (p = 0.08), longer by 3.6 days (p = 0.001) at hospitals with 10 - 100 cases and by 3.3 days (p = 0.0001) at hospitals with 101 - 300 cases in comparison with facilities treating >300 cases In multivariate analyses independent risk factors for mortality were complexity (p = 0.001), procedure for cardiopulmonary bypass (p<0.0001) and admission from another acute care facility (p<0.001) and compared with hospitals >300 cases the estimated odds ratio for mortality was 7.7 (95% CI: 1.6 - 37.8) for patients at hospitals with <10 cases (p = 0.1), 2.9 (95% CI: 1.6 - 5.3) for patients at hospitals with 10 to 100 cases (p<0.0005) and 3.0 (95% CI: 1.8 - 4.9) for patients at hospitals with 101 - 300 cases	Case-mix adjustment score = 3

Table 5 Open heart surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hosenpud et al (1994)[39] USA To determine the effect of small transplantation volume on transplant outcome	Patients undergoing cardiac transplan-tation = 7893 Cardiac transplant centres = 150	Retrospective analysis of the Scientific Registry of the United Network for Organ Sharing 1987 - 1991	Operative and 12 month mortality	Annual volume: <9 (80 centres) >9 (70 centres) (Experience was also measured in 13 centres)	First or repeat transplantation, medical condition and presence of congenital heart disease T-test, X^2 test and regression analysis	Risk of mortality decreased to a basal level in centres performing between 8 and 10 transplants per year Increased risk of operative mortality was 40.3% and at 12 months 33% in centres performing less than 9 transplants per year ($p<0.001$) Experience: mortality increased in the second year of operation (experience assessed as time in operation does not substitute for volume)	Case-mix adjustment score = 2 Volume and overall experience were investigated in separate analyses

Table 5 Open heart surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Laffel et al (1992)[40] USA To determine the relation between experience with heart transplantation and mortality	Patients undergoing heart transplantation = 1123 Heart transplantation centres = 56	Retrospective analysis of the registry of the International Society for Heart and Lung Transplantation 1984 - 1986	90-day mortality	Experience was the main focus: new centres were examined Volume was assessed as the correlation between total number of transplantations and mortality for each centre	Age, sex, co-existing disease, rare indication for transplant, ejection fraction, mechanical hemodynamic support, triple drug antirejection therapy, ischemic time, for the volume analysis subgroups were analysed separately Univariate, bivariate and regression analysis	Total transplantation volume and mortality showed no significant correlation. Analysis of subgroups also showed no significant correlations between volume and mortality Patient's who received one of the centres first 5 transplants had higher mortality than patients who received subsequent transplants. New centres staffed by experienced cardiologists had lower mortality rates. In contrast previous training of performing surgeon was not related to mortality	Case-mix adjustment score = 2

Table 5 Open heart surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing open heart surgery = 27471 Hospitals = 587	Retrospective analysis of discharge abstract data from the Commission on Professional and hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: < 200 = 541 hospitals, 11997 patients > 200 = 46 hospitals, 15474 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score =2

Table 6 Acute myocardial infarction and other heart problems

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Barbash et al (1994)[41] 13 countries To evaluate the outcome of patients with myocardial infarction (MI) in relation to the size and comprehensiveness of cardiovascular services	Patients with acute MI receiving thrombolytic therapy = 8206 Hospitals n = 438	Retrospective analysis of the multinational database of the International Tissue Plasminogen Activator/Strepto-kinase Mortality Trial October 1988 to August 1989	Mortality, reinfarction, bleeding complication, stroke and hospital length of stay	The hospitals were classified as having coronary revascularization (n = 82) and no coronary revascularization (n = 356) and further divided into < 300 beds (n = 254) and > 300 beds (n = 184) (within each classification)	Sex, age, antecedent angina pectoris, previous myocardial infarction, diabetes, history of hypertension, smoking, compromised hemodynamic state at admission Logistic regression	There were no significant differences in in-hospital mortality between the hospitals, however mortality at 6 months was significantly higher in patients admitted to small centres with revascularization facilities (17% v 11.8 to 12.3, p = 0.03) Patients admitted to small hospitals stayed significantly longer than those admitted to large hospitals (11.2 & 11.6 v 9.8 & 10.3 days, p<0.0001). There were no significant differences in the length of stay between centres with or without revascularization services The rate of reinfarction was lowest in small centres with revascularization services (1.7 v 3.6, 3.9, 4.9, p = 0.01) and the rate of haemorrhage was lowest in large centres without revascularization services (4.9 v 7.1, 6.0, 7.1, p = 0.01). Stroke rate did not differ significantly between centres	Significant differences were found between hospitals in the distribution of antecedent angina and time to treatment at baseline Case-mix adjustment score = 3

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kelly & Hellinger (1987)[22] USA To examine the effects of selected characteristics of hospitals and physicians on the mortality rates of heart patients	Patients with a diagnosis of acute myocardial infarction with no surgical procedure = 11033 Short term hospitals across the US = 146 Physicians = 926	Retrospective analysis of patient discharge abstracts from the Hospital Cost and Utilisation Project 1977	In-hospital mortality	Physician and hospital volume were expressed as continuous variables. Mean hospital volume = 146 (s.d. = 96) Mean physician volume = 30 (s.d.= 26)	Age, sex, number of diagnoses and disease stage. Multivariate regression	There was a significant negative relationship between adjusted in-hospital mortality and physician volume (coefficient = -0.049, p = 0.05). No statistically significant relationship between hospital volume and adjusted mortality was found Patients with higher disease stages were significantly more likely to die than patients with lower disease stages (p = 0.05) as were older patients compared with younger (p = 0.05)	Case-mix adjustment score = 3

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Reznik et al (1987)[42] Australia To examine the effectiveness of coronary care in relation to level of care in teaching hospitals and small country hospitals	Patients with suspected acute myocardial infarction =2265 Australian public hospitals = 18	Prospective cohort where patients were interviewed to determine history of the attack and past general medical history; hospital records were also reviewed May 1979 - October 1980	Mortality and cardiopulmonary resuscitation	Hospitals were classified into 4 levels (level 1 = most elaborate) and then into either unit or ward care	A severity score was calculated for each patient (age, extent of infarct, admission systolic blood pressure, pulse rate, cardiac failure, shock and cardiac enzyme values) and patients were stratified into 5 risk groups Analysis of variance and covariance were used to compare continuous variables across levels of care and X^2 analysis for categorical variables. Logistic regression was used to adjust mortality rates for severity	Patients did not differ significantly in age or sex across different hospital levels Unadjusted mortality was 21% for level 1 care; 22% for level 2; 21% for level 3 and 19% for level 4 (n.s.) Average mortality in coronary care units was lower (17%) (n.s.) Differences in ward mortality across levels were significant ($X^2 = 14.9$, df = 3, $p<0.01$) Differences in mortality between coronary care unit and ward care was significant at levels 1, 2 and 3 but not at 4. Differences were largest at level 1 and fell to non-significant at level 4 Logistic regression failed to show any benefit at higher levels of care for any sub-group based on severity	Case-mix adjustment score = 3

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with atrial fibrillation = 4287 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 164.83 (s.d. 77.97) Physician volume = 8.81 (s.d. 9.51)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals or physicians There were no significant differences in LOS between high and low volume hospitals or physicians	Case-mix adjustment score = 2

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farley et al (1992)[23] USA To examine how patient outcomes for several procedures respond to changes in volume over time	Patients with a diagnosis of acute myocardial infarction = 974803 Short term general hospitals = 426	Retrospective analysis of data collected as part of the Hospital Cost and Utilisation Project (discharge abstract data) 1980 - 1987	Mortality	Volume was measured as a continuous variable Both within and between hospital differences were measured	Age, disease staging Multivariate regression	A significant relationship between increased volume over time and adjusted mortality was found (coefficient = -2.982, $p<0.001$) (as volume increases mortality declined) A significant effect of volume on adjusted mortality was also found in the cross sectional analysis (coefficient = -0.841, $p<0.001$)	Case-mix adjustment score = 2

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients with acute myocardial infarction = 98066 Hospital - 906	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 291 cases per year	Age, sex, multiple diagnoses, admission blood pressure. Actual and expected death rates were modelled as a function of hospital volume using regression.	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.049$, $p<0.01$).	Case-mix adjustment score = 2

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with myocardial infarction = 5708 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 205.32 (s.d. 90.30) Physician volume = 14.21 (s.d. 15.70)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	High volume hospitals had significantly greater mortality than low volume hospitals (coefficient = 0.391, p<0.001), but there were no significant differences between physicians with varying caseloads LOS was significantly longer for high volume hospitals compared with low volume hospitals (coefficient = 0.147, p<0.001), there were no significant differences between physicians	Case-mix adjustment score = 1

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with congestive heart failure = 5864 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 221.13 (s.d. 119.48) Physician volume = 9.91 (s.d. 10.08)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals or physicians LOS was significantly longer for high volume hospitals compared with low volume hospitals (coefficient = 0.039, $p<0.01$), there were no significant differences between physicians	Case-mix adjustment score = 1

Table 6 Acute myocardial infarction and other heart problems (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Shortell & LoGerfo (1981)[43] USA To examine the relationship among hospital structural characteristics, individual physician characteristics, medical staff organisation characteristics and quality of care	Coronary conditions including myocardial infarction, ischeamic heart disease or disorders of heart rhythm = 51159 East North Central region hospitals = 95	Retrospective analysis of case abstract data from the Commission on Professional and Hospital Activities (CPHA) 1973	In-hospital mortality	Physician volume was expressed as the number of patients in each diagnosis divided by the number of appropriate physicians on staff (mean number of patients per physician = 5.43, s.d. = 4.51)	Patients were stratified based on their systolic blood pressure level and age. Multiple regression	Significant independent effects of age and systolic blood pressure as well as a strong joint effect were found The volume of patients per physician was strongly associated with lower adjusted mortality rates ($p < 0.05$) Hospitals seeing 60 or fewer patients per year had an adjusted mortality ratio of 1.17 compared with 1.00 for those seeing more than 60 patients per year ($p<0.05$)	This was a 35% sample of all hospitals in the East North Central region stratified by bed size, ownership and teaching status 3745 eligible cases were excluded due to missing or miscoded data Case-mix adjustment score = 1

Table 7: Pacemaker implantation

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Boles (1994)[45] USA To examine linkages between hospital volume and outcomes for several procedures	Patients having pacemakers inserted (n not given) Acute care general hospitals = 1751 (not procedure specific)	Retrospective analysis of Medicare Hospital Information from the Health care Financing Administration 1988 and 1990	In-hospital mortality	Mean hospital volume: 2.695 (s.d 0.870) Cross sectional and longitudinal assessments were made	Age, sex, comorbidity, type and source of admission, previous hospitalisations Linear regression	No significant differences in mortality between hospitals. Volume increased over time ($t = 14.98$, $p<0.0001$) but the risk adjusted mortality rate did not change significantly over time	Case-mix adjustment score =2

Table 7: Pacemaker implantation (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Doherty et al (1994)[44] UK To review the pacemaker practice in 1 hospital over a 79 month period and compare it with national data	Patients having a pacemaker implanted for the first time = 201 1 district general hospital	Retrospective analysis of pacemaker records from 1 hospital, which were compared with UK national data (1992) and data from 2 large specialist centres January 1987 - July 1993	Complication rates	1 district general hospital and 2 specialist centres and 1992 national UK data	Age, sex, presenting symptoms and ECG indications were compared Statistical testing was not reported	The mean age of the hospital patients was similar to the mean age of the UK sample (74 years v 71.9 years respectively) as was the percentage of males (47.5% v 56% respectively) Presenting symptoms and ECG indications appeared to differ The overall complication rate for the hospital was 2.48% v 2% and 2.8% for the 2 specialist centres	Case-mix adjustment score = 2

Table 8: Cardiac catheterization/angiography

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kelly & Hellinger (1987)[22] USA To examine the effects of selected characteristics of hospitals and physicians on the mortality rates of heart patients	Patients undergoing cardiac catheterisatio n = 4835 Short term hospitals across the US = 39 Physicians = 145	Retrospective analysis of patient discharge abstracts from the Hospital Cost and Utilisation Project 1977	In-hospital mortality	Physician and hospital volume were expressed as continuous variables Mean hospital volume = 399 (s.d. 401) Mean physician volume = 97 (s.d. 71)	Age, sex, number of diagnoses and disease stage. Multivariate regression	No statistically significant relationship between mortality and physician volume was found. Hospital volume and mortality were associated (coefficient = -0.001) Age ($p = 0.05$) was significantly associated with mortality, with patients aged 70 or above being 2.7% more likely to die in hospital than those aged 18 - 45 years. Disease stage ($p = 0.05$) was also significantly associated with mortality	Case-mix adjustment score =3

Table 8: Cardiac catheterization/angiography (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing cardiac catheterization (CC) = 76584 Hospitals = 150 Surgeons = 2987	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 246.25, s.d. 348.3) Surgeon volume: median number of procedures per year = 11, used to differentiate less experieinced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-2.3021, $p<0.01$) and a lower proportion of patients operated on by low volume surgeons (4.8976, $p<0.01$) were positively related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score =2

Table 8: Cardiac catheterization/angiography (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing angiography/ cardiac catheterization = 26678 Hospitals = 360	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 598 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.542$, $p<0.01$)	Case-mix adjustment score = 2

Table 8: Cardiac catheterization/angiography (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Adams et al (1973)[46] USA To determine the rate of complications due to coronary arteriography in 1970-71	Patients undergoing coronary arterio-graphy = 46904 Hospitals = 373	A nationwide survey to all institutions with an open heart surgical team 1970 - 1971	Compl-ication rate and mortality	Hospital volume over the 2 year period: < 200 200 - 499 500 - 799 > 800	Patients were classified according to technique: femoral or brachial Chi-square test No adjustment for patient risk factors	Overall mortality rate was 0.45% (brachial) and 0.78% (femoral) ($p<0.001$) The mortality rate (for both types of technique) was 8 times higher in hospitals performing less than 200 procedures per 2 years than in hospitals doing more than 800 ($p<0.0001$) The incidence of complications (such as myocardial infarction and cerebral embolism) was significantly higher in lower volume hospitals (e.g for hospitals doing less than 200 procedures myocardial infarction was 1.9% and in hospitals doing more than 800 it was 0.19%, $p<0.001$)	173 hospitals responded Case-mix adjustment score = 0

Table 9: Percutaneous transluminal coronary angioplasty (PTCA)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kimmel et al (1995)[48] USA To assess the relationship between the volume of PTCA in a cardiac catheterization laboratory and major complication after adjusting for case-mix	Patients undergoing a first balloon angioplasty = 19594 Cardiac catheteriz-ation laboratories = 48	Retrospective analysis of the Society for Cardiac Angiography and Interventions Registeries (a clinical registry) 1992 and 1993	In-hospital mortality, emergency bypass surgery, myocardial infarction and major complication (defined as 1 or more of these 3 outcomes)	Volume was defined in 3 ways: 1) < 200 procedures per year, patients = 2010 200 - 399, patients = 6102 400 - 599, patients = 5056 > 600, patients = 6426 2) < 200 > 200 (current ACC/AHA guidelines) 3) volume was examined as a continuous variable	Age, sex, MI within 2 - 14 days, aortic valve disease, congestive heart failure, chronic renal insufficiency, dialysis, diabetes, hypertension, PTCA during diagnostic procedure, emergency PTCA, PTCA attempted on graft, PTCA attempted on left main, lytic therapy prior to PTCA, mitral valve disease, multivessel disease, number of vessels attempted, previous CABG, previous catheterization, previous valve surgery, shock, unstable angina, worst lesion type attempted Patients were stratified by their risk factors and logistic regression was used to examine the effect of risk on outcome	There was a significant reduction in unadjusted mortality (p = 0.04), emergency CABG (p<0.001), acute MI (p = 0.001) and major complication (p<0.001) with increasing volume There were significant differences in case-mix between volume categories (p<0.001, p<0.05). The risk of major complication decreased with increasing volume in both high and low risk patients. Volume was linearly and inversely associated with major complication (p<0.001) Using logistic regression to adjust for risk factors a significant inverse relationship between volume and emergency CABG (p<0.001), acute MI (p<0.001) and major complication (p<0.001) was found. No statistically significant relationship was found between volume and mortality A significant decrease in major complications was observed when volume was greater than 400 procedures per year (adjusted OR 0.66, 95% CI: 0.46 - 0.96) There were no significant differences in mortality (OR 0.71, 95% CI: 0.26 - 1.99), emergency CABG (OR 0.82, 95% CI: 0.49 - 1.39), acute MI (OR 1.06, 95% CI: 0.45 - 2.50) or major complication (OR 0.81, 95% CI: 0.53 - 1.25) between laboratories performing <200 procedures and >200 procedures. Although the rates of each were greater in lower volume laboratories	Case-mix adjustment score = 3

Table 9 Percutaneous transluminal coronary angioplasty (PTCA) (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Shook et al (1996)[53] USA To investigate whether operator volume is associated with PTCA outcomes	Patients undergoing (PTCA) = 2204 (2350 procedures) Operators = 38	Retrospective analysis of case records (computerised database) March 1991 - February 1994	CABG <24 hours after PTCA, morbidity (procedural complica-tions), in-hospital mortality and LOS	Operator volume: High = > 50 cases per year (n=8, 1502 cases) Low = <50 cases per year (n=30, 848 cases)	Age, sex, race, body surface area, procedural priority (elective, urgent, emergent), procedural complexity (1 v multiple vessel PTCA), resting heart rate, blood pressure Chi-square test, unpaired t-test, logistic regression and linear regression	No significant differences in unadjusted mortality or hospital length of stay, between high and low operators. High volume operators had less emergency CABG, fewer procedural complications and lower lengths of stay. Risk adjusted mortality did not differ. Low volume operators required more emergency CABG (risk ratio 2.05, 95% CI: 1.24 to 3.39, p=0.005), had higher complication rates (risk ratio 1.79, 95% CI: 1.32 to 2.43, p<0.001) and a 9% longer length of stay (ratio 1.09, p=0.004) than high volume operators High volume operators had older patients with more complex and more emergent/urgent procedures than low volume operators	Case-mix adjustment score = 3

Table 9 Percutaneous transluminal coronary angioplasty (PTCA) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Talley et al (1995) [52] USA To determine the relationship between equipment size, operator experience, and PTCA cost and patient outcome	Patients undergoing elective PTCA for a single lesion = 50	RCT (concealed allocation) to either a 0.010 or 0.014 PTCA balloon catheter would be used and secondary randomisation determined whether the primary operator	Angio-graphic success(redu-ction of stenosis to < 50% diameter Clinical success angio-graphic success without the occurrence of an inhospital ischemic complication, need for emergency CABG surgery, or repeat PTCA or development of a non-Q wave or Q wave MI) Plus process measures	Randomisation was either to an attending physician who received training in interventional cardiology and had independently carried out more than 500 procedures or to a fellow in interventional cardiology who had performed less than 50 procedures as primary operator (they served as proxies for low volume, inexperienced operators)	Baseline clinical and angiographic characteristics were compared for the 2 groups T-tests or ANOVA, or ordinary least squares regression was used (intention to treat)	There were some baseline differences with low volume operators having more males and significantly more patients with previous MI There were no significant differences in either angiographic (100 v 96%) or clinical success between the low and high volume operators The low volume operators had a significant increase in the time to cross the lesion and fluoroscopic time and a trend in total procedure time	Case-mix adjustment score = 3

Table 9 Percutaneous transluminal coronary angioplasty (PTCA) (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Philips et al (1995)[50] USA To examine whether hospital volumes of PTCA are associated with adverse outcomes	Patients undergoing PTCA = 24856 Acute care hospitals in California = 110	Retrospective analysis of discharge abstracts from the California Office of Statewide Health Planning and Development 1989	In-hospital mortality and CABG surgery after PTCA	< 201, hospitals = 64, patients = 6945 201 - 400, hospitals = 28, patients = 8300 >400, hospitals = 18, patients = 9611	Age, sex, race diabetes, number of vessels dilated and days from admission to PTCA. Regression analysis Patients with acute MI were analysed separately from patients without	Low volume hospitals had significantly higher rates of unadjusted adverse outcomes than higher volume hospitals (mortality: 0.25% v 0.20% for non-AMI and 0.34% v 0.27% for AMI in low v high volume hospitals, p<0.001) Patients with a principal diagnosis of acute MI were concentrated in low and medium volume hospitals and patients without were concentrated in high volume hospitals Low volume hospitals had significantly higher proportion of patients who were women, non-white, diabetic, single vessel PTCAs and a longer period between admission and PTCA (p<0.05) Low volume hospitals had significantly higher than predicted rates of adverse outcomes and high volume hospitals had lower than predicted rates of adverse outcomes (for both MI and non-MI patients) When volume was included as an independent variable, higher volume was associated with significantly fewer adverse outcomes (p<0.001 [rates not given])	Case-mix adjustment score = 2

Table 9 Percutaneous transluminal coronary angioplasty (PTCA) (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Jollis et al (1994)[49] USA To examine the relationship between the volume of angioplasty procedures and mortality	Patients aged 65 years and over undergoing PTCA = 217836 Hospitals = 736 during all 4 years, with a total of 1194 over the 4 years	Retrospective analysis of Medicare Provider Analysis and Review 1987 - 1990	In-hospital mortality, 30-day mortality and rate of CABG after angioplasty	Low volume: <50 procedures per year intermediate volume: 50 -100 high volume: >100 Volume was also specified in 10 groups each containing 10% of the sample according to annual volume 50% of patients were treated in hospitals performing 54 or fewer procedures in Medicare patients per year	Age, sex, race (year of procedure). Logistic regression	Overall unadjusted in-hospital mortality was 2.9%, 30 day mortality was 3.2% and 3.8% also underwent CABG before discharge In-hospital mortality fell from 3.9% in lowest volume hospitals (1 - 46) to 2.5% in highest volume hospitals (>372) and 30 day mortality fell from 4.2% to 2.7% respectively. In the lowest volume hospitals 5.3% underwent CABG surgery versus 2.8% in the highest In-hospital mortality remained significantly associated with volume after controlling for prognostic factors (p<0.001 [rates not given])	15% of the original sample were excluded due to: Medicare eligibility for reasons other than age, incomplete data, residence outside USA or primary illness other than ischemic heart disease Case-mix adjustment score =1

Table 9 Percutaneous transluminal coronary angioplasty (PTCA) (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kelsey et al (1984)[47] USA To examine the effect of investigator experience from the introduction of PTCA through the first 4 years of its application	Patients undergoing PTCA = 3101 Clinical centres = 105	Retrospective analysis of the PTCA registry (a clinical database) September 1977 - September 1981	Success rates (a decrease of at least 20% in luminal diameter narrowing, survival without MI or CABG during hospitalisation)	<10 , centres = 48, patients = 208 10 - 19, centres = 25, patients = 328 20 - 49, centres = 14, patients = 442 50 - 99, centres = 8, patients = 515 100 - 149, centres = 5, patients = 668 150 - 199, centres = 3, patients = 463 >200, centres = 2, patients = 477	Not stated	Centres with fewer than 50 cases had success rates of 55% compared with success rates of 66% for centres with 50 plus cases For investigators carrying out 50 - 99 procedures the success rate was 70%, in 100 -149 procedures the success rate was 72% and in investigators carrying out 150 or more procedures the success rate was 77%	No information is given on patient prognostic factors Case-mix adjustment score = 0

Table 10: Carotid endartectomy

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Boles (1994)[45] USA To examine linkages between hospital volume and outcomes for several procedures	Patients undergoing CE (n not given) Acute care general hospitals = 1751 (not procedure specific)	Retrospective analysis of Medicare Hospital Information from the Health care Financing Administration 1988 and 1990	In-hospital mortality	Mean hospital volume = 2.098 (s.d. 1.106) Cross sectional and longitudinal assessments were made	Age, sex, comorbidity, type and source of admission, previous hospitalisations Linear regression (and regression lines)	No significant differences in mortality between hospitals. Volume decreased over time ($t = -1.922$, $p<0.05$) but the risk adjusted mortality rate did not change significantly over time	Case-mix adjustment score = 2

Table 10 Carotid endartectomy (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Edwards et al (1991)[55] USA To determine the effect of physician, hospital volume and comorbidity on morbidity, mortality and length of stay in CE	Patients undergoing CE = 11199 Acute care hospitals in one state (n not given)	Retrospective analysis of discharge abstract data for one state January 1979 - December 1988	In-hospital mortality, in-hospital stroke and length of stay (LOS)	Hospital volume: low = 1 - 12 cases per year, = (922 patients); medium = 13 - 49, = (5258 patients); high = >50, = (5019 patients) Physician volume: low = 1 - 12, = (5067 patients); medium = 13 - 49, = (3876 patients); high = >50, = (772 patients).	Age, sex and comorbidity (ischemic heart disease, pulmonary disease, diabetes, hypertension) Differences in outcomes by prognostic factors were assessed by chi-square tests	Overall, 48% of patients were defined as complex (had comorbidity). There were no significant differences in mortality, stroke or LOS as a function of complexity Over time complex patients were increasingly referred to high volume physicians Mortality (2.6% low v high 1.2%, p = 0.02) and stroke (4.0% low v high 2.0%, p = 0.008) decreased by 50% as a function of physician volume and LOS also decreased (12.7 low v high 8.1 days, p<0.001) There were no statistically significant differences in mortality or stroke according to hospital volume. Increased hospital volume was significantly associated with decreased LOS (12.9 low v high 10.0 days, p<0.001)	Case-mix adjustment score = 2

Table 10: Carotid endartectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kempezinski et al (1986)[57] USA To examine the impact of operative caseload on outcomes of patients undergoing CE	Patients undergoing CE = 656 (750 operations General medical surgical hospitals = 16 Surgeons = 61	Retrospective review of patients' medical records July 1983 - June 1984	Mortality and post-operative stroke	Average caseloads of surgeons (per years): < 12 12 - 50 > 50 Hospital caseload: < 50 50 - 100 > 100	Patients were classified as either symptomatic or asymptomatic and were compared according to age, sex, carotid bruit, hypertension, smoking, diabetes, myocardial infarct, angina, using the Chi-squared test	No statistically significant differences in either post-operative stroke or mortality were found between surgeons or hospitals, although for surgeons a decreasing trend was noted Surgeons in the high volume group operated on more asymptomatic patients (who were at lower risk of stroke) than the 2 other groups (61% v 44%)	Case - mix adjustment score = 2

Table 10: Carotid endartectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hertzer et al (1984)[56] USA To examine outcomes for 3 types of vascular surgery according to surgeon experience	Patients undergoing CE = 2646 Hospitals in Cleveland-Akron region = 28 Surgeons = 36	Retrospective analysis of The Cleveland Vascular Society computer registry (includes preoperative risks, operative circumstances) 1978 - 1981	Mortality and morbidity (stroke)	Surgeon volume per year: < 10 10 - 25 > 25	Patients were divided into either preoperative neurologic symptoms or no preoperative symptoms Fisher's exact	Overall mortality was 1.2%, post-operative stroke was 2.5% Mortality was 0.8% for surgeons performing 10 - 25 procedures and 2.0% for those performing fewer than 10 (n.s.) and similarly morbidity did not differ significantly	Case-mix adjustment score = 1

Table 10: Carotid endartectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Richardson et al (1989)[61] USA To examine the relationship of post-operative stroke to operative volume	Patients undergoing CE = 705 Hospitals = 41 Surgeons =98	Retrospective review of Medicare records and chart review 1983	Post-operative complica-tions	Annual surgeon volume: <3 (n=47) >12 (n=37) >25	Unclear Chi-square test	Post-operative stroke was 6.1% for surgeons carrying out <3 cases compared with 2.3% for surgeons carrying out >12 ($p<0.01$) Post-operative stroke increased slightly in surgeons carrying out >25 cases, but was non-significant.	The only statistical difference was between 3 and 12 cases per year (therefore not a linear relationship) Case-mix adjustment score = 0

Table 10 Carotid endartectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Ruby et al (1996)[59] USA To define the relationship between operative experience and post-operative morbidity and mortality in CE	Patients undergoing CE = 3997 Hospitals = 32 Surgeons = 226	Retrospective analysis of the Connecticut Hospital Association database (discharge abstract data) 1985 - 1991	Combined stroke and/or death and a prolonged hospital stay (>7 days)	Surgeon volume ranged from: < 1 (43% of surgeons) 2 - 5 (79%) 6 - 10 (12%) > 10 (9%)	No risk adjustment Chi-square tests, chi-square tests for linear trends and odds ratios	Surgeons performing 1 or fewer operations were 2.5 times more likely to have a serious complication than surgeons > 10 cases per year ($p<0.02$) and post-operative length of stay was 3.54 times more likely to be prolonged ($p<0.0001$)	Surgeons' speciality was also assessed Case-mix adjustment score = 0

Table 10: Carotid endartectomy (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Segal et al (1993)[58] USA To study physicians use of CE and explore variations in surgical volume and outcome	Patients undergoing CE =5657 Surgeons = 652	Retrospective analysis of Pennsylvania Medicare data December 1989 - January 1992	In-hospital mortality and re-admission within 181 days for subsequent cerebral occlusion	High surgeon volume > 30: 52 surgeons, 2384 patients low surgeon volume = < 30: 600 surgeons, 3273 patients	No information given	Overall mortality rates were 2% and varied significantly by frequency of procedure ($R^2 = 0.051$). The low volume surgeon mortality rate was 2.6% versus 1.17% for high volume surgeons ($p<0.005$) A mortality rate of 4.6% was found for surgeons carrying out 1 procedure per year Of those patients discharged alive, 33% had at least 1 readmission and 6% subsequently died	No information was presented on the statistical analysis used or about any variables controlled for Case-mix adjustment score = 0

Table 10 Carotid endartectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Slavish et al (1984)[60] USA To examine the safety of CE in a community hospital	Patients undergoing CE = 743 Hospitals = 1 Surgeons=24	Pre-operative review of patient records (1 community hospital) 1977 - 1982	Mortality and morbidity	Annual surgeon volume : <24 >24	None reported Correlation coefficients were presented	There were no statistically significant differences in morbidity or mortality between high and low volume surgeons	Case-mix adjustment score = 0

Table 11 Abdominal aortic aneurysms

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients undergoing intra-abdominal artery operations =9532 Acute care hospitals in the United States = 645	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume Low vol = 14 high vol > 15	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 15.5% The SMR for low v high volume hospitals was 1.20 v 0.91, p<0.001 (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For low, medium and high risk patients SMRs were above 1 in low volume hospitals and below 1 in high volume hospitals (results presented graphically)	Case-mix adjustment score = 3

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hannan et al (1989)[32] USA To test the combined relationship of hospital and physician volume with in-hospital mortality rates for several procedures	Patients undergoing resection of AAA =1635 hospitals = 170 physicians = 508	Retrospective analysis of New York State discharge abstracts from the Statewide Planning and Research Cooperation System 1986	In-hospital mortality	A continuous volume measure was used which logistic regression defined as: hospital volume: < 5, n = 95, cases =252 6 - 10, n = 34, cases = 266 11 - 20, n = 25, cases = 391 21 - 35, n = 11, cases = 342 36 - 72, n = 5, cases = 384 physician volume < 4, n = 412 > 4, n = 96 (low and high physician volume was also defined within each of the hospital volume categories)	Age, sex, race, admission status, upto 4 secondary diagnoses and procedures, and severity of illness. Logistic regression	The overall unadjusted mortality rate was 18% Age (p<0.001), admission status (p<0.001) and disease stage (p<0.001) were significantly associated with mortality Physician volume had a significant impact on risk adjusted mortality: 20.27% vs 14.67% for low and high volume respectively (p<0.05). There was a tendency for low volume physicians to operate on higher risk patients than high volume physicians Among low volume physicians hospital volume did not have a significant effect on mortality. Among high volume physicians risk adjusted mortality was lower for those practising in high volume hospitals. The ratio of mortality rates for low to high volume hospitals was 1.48 (p = 0.05)	There were originally 16 procedures included in this study, but only those 5 which had statistically significant volume-mortality relationships were presented Case-mix adjustment score = 3

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Katz et al (1994)[64] USA To define the operative mortality rates for intact and ruptured AAA in Michigan and to document changes in mortality and to identify risk factors	Patients aged 50 and above with a diagnosis of AAA = 10014 divided into: intact = 8185 and ruptured = 1829 Acute care hospitals = 205 in 1980 176 in 1990	Retrospective analysis of the Michigan Inpatient database 1980 - 1990	Mortality	Volume was defined for intact AAA as: 0 1 - 10 11 - 20 > 21 and for ruptured AAA as: 0 1 - 4 > 5	Age, sex, race, and comorbidity (intact and ruptured specified). Logistic regression was used to control for those confounding variables found by chi-square and t-tests to be significantly related to mortality	Overall mortality was 10.2%. Intact AAA: mortality decreased over time, 13.6% in 1980 vs 5.6% in 1990 (p<0.001) and patients' comorbidity increased over time, 13.6% in 1980 vs 53.2% in 1990 (p<0.001) A significant relationship was found between unadjusted mortality and volume (p = 0.001): hospital with a volume of 21 or more procedures had a mortality rate of 6.2% vs 8.9% for hospitals with a volume of less than 21 procedures Significant predictors of mortality were: female sex (p = 0.001), age (p = 0.001), multiple comorbidities (p = 0.001), kidney failure (p = 0.001) and dysrhythrmia (p = 0.001) Hospital volume under 21 v over 21 procedures was significantly associated with mortality (OR 1.2, p = 0.02) as was year of surgery 1980 compared to 1990) (OR 3.4, p = 0.0001) Ruptured AAA: no trend in mortality over time was identified (43.5% vs 57.5%) In logistic regression significant predictors of mortality were: being female (p = 0.0006), increased age (p = 0.0001), kidny failure (p = 0.0001) and dysrthythmia (p = 0.0001) Low surgical volume was significantly related to increased mortality (53.6% compared with 45.7%, p = 0.002 [odds ratio not given])	Case-mix adjustment score = 3

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kelly et al (1986)[65] USA To examine the effects on post-surgical mortality rates of the volume of several procedures performed by individual physicians and hospitals	Patients undergoing blood vessel surgery for AAA = 999 Short term general hospitals = 77 physicians = 232	Retrospective analysis of data from the Hospital Cost and Utilization Project, based on discharge abstract records 1977	In-hospital mortality	Both hospital and physician volume were expressed as continuous variables Mean hospital volume = 22.82 (s.d. 12.32) Mean physician volume = 9.48 (s.d. 7.11)	Age, sex, number of diagnoses and stage of illness. Logistic regression	Overall morality = 10.2% A significant effect of variation in hospital volume on mortality was found (coefficient = -0.0026, p = 0.05) indicating that patients were more likely to survive when treated in high volume hospitals Each additional 4 operations performed reduced the probability of death by 1% The volume of procedures performed by individual physicians did not have a statistically significant effect on mortality) The number of diagnoses had a significant effect on mortality (p = 0.05); a patient with 3 diagnoses has between a 4 and 7% higher probability of experiencing a negative outcome than a patient with only 1 diagnosis Disease staging variables were significant (p = 0.05) indicating that a patient with rupture or fistulas was 12% more likely to die than a patient with an uncomplicated abdominal aneurysm	Case-mix adjustment score = 3

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Veith et al (1991)[67] USA To assess the quality of care in vascular surgery	Patients undergoing AAA repair (unruptured) = 3570 Surgeons =98	Retrospective review of New York Statewide Planning & Resource Cooperative System database 1985 - 1987	Mortality	Annual hospital volume: 1-5 >38 (average =10.2 per hospital) Annual surgeon volume: 1-5 >26 (average=3.6 per surgeon)	Age, comorbidities and stage of disease Chi-square test	Unadjusted mortality for low volume hospitals was 14% v 5% for high volume hospitals (p<0.0001) and unadjusted mortality for low volume surgeons was 10% v 6% for high volume surgeons (p<0.0001) Adjusted mortality for low volume hospitals was 12% v 5% for high volume hospitals (p<0.001) and for low volume surgeons was 9% v 4% for high volume surgeons (p<0.001)	Case-mix adjustment score = 3

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Amundsen et al (1990)[62] Norway To compare the results of treatment after elective and emergency treatment of AAA related to numbers of operations done, experience and type of hospital	Patients undergoing AAA repair elective = 279 emergency = 165 Hospitals = 26	Prospective multicentre cohort	Mortality	1 - 9 10 - 29 30 - 39 > 40 > 100 was defined as experienced	Patients were divided into: elective and emergency cases, which was further sub-divided into: ruptured and impending rupture and odds ratios for subgroups were computed	Elective: there was a significant difference between hospitals carrying out less than and more than 10 elective procedures per year (p = 0.04, OR 2.7). The mortality for operations undertaken in hospitals with vascular surgical experience was 4.8% compared with 11.3% for units without such experience (p = 0.05, OR 2.6) Emergency: the mortality rate was 55.7% in hospitals carrying out less than 10 procedures per year compared with 42.3% for those carrying out more than 10 procedures (p = 0.10, OR 1.7) The mortality rate for operations undertaken in hospitals with vascular surgical experience was 39% compared with 55% for units without such experience (p = 0.06, OR 1.9)	No information is presented on patients' prognostic variables Case-mix adjustment score =2

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients undergoing aorta repair/ replacement = 1703 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 95.37 (s.d. 53.51) Physician volume = 20.10 (s.d. 33.00)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, but mortality was inverssely correlated with physician vol (coefficient = -0.362, p<0.01) There were no significant differences in LOS between high and low volume hospitals or physicians	Case-mix adjustment score = 2

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing resection and graft - AAA = 4624 Hospitals = 692	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: < 20 = 654 hospitals, 3384 patients > 20 = 38 hospitals, 1240 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume, but flattened out at relatively low volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score =2

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Pilcher et al (1980)[66] USA To examine the relation between aneurysm surgery and mortality in the Vermont community	Patients undergoing resection of AAA = 294 General hospitals = 17	Survey data from individual surgeons on AAA procedures 1970 - 1977	Mortality	High (university medical centre), medium (more than 1 patient per year) and low volume (less than 1 patient per year) Surgeon volume: < 2 2 - 4 > 4	Diabetes, respiratory disease, arteriosclerotic cardiovascular disease, myocardial infarction, hypertension, heart failure, blood loss, anaesthesia duration, aneurysm size recorded (not clear which of these were included in the analysis for elective patients only) Patients were divided into elective and ruptured aneurysm Chi-square and discriminant analysis	Ruptured aneurysms: mortality at the university hospital was 46% v 62 and 86% in the other 2 groups (p<0.05), but mortality was not significantly related to surgeon volume Elective: no significant association was found for hospital volume. However, for surgeon volume mortality was significantly related (7% for high, 15% for medium and 17% for low volume surgeons).	9 hospitals were excluded as they had no patients undergoing this procedure during the study period Case-mix adjustment score = 2

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Wu Wen et al (1996) Canada To determine whether a volume-mortality relationship was observed in Ontario hospitals for AAA	Patients with unruptured AAA: n = 5492 Patients with ruptured AAA: n = 1203 Ontario hospitals = 312	Retrospective analysis of the Hospital Medical Records Institute discharge abstracts 1988 - 1992	In-hospital death and length of hospital stay (LOS)	Five year hospital volume for unruptured aneurysm: < 50 50 - 100 101 - 200 > 200 Five year hospital volume for ruptured aneurysm: < 10 10 - 20 21 - 40 > 40	Age, sex, transfer status, some secondary diagnoses Ruptured and unruptured aneurysms were analysed separately Multiple logistic regression analysis Multiple linear regression	The overall case fatality rate (CFR) for unruptured aneurysms was 3.8% and 40.0% for ruptured aneurysms There was a moderate decline in the crude CFR with rising hospital volume for both ruptured and unruptured (n.s.) and the length of hospital stay relationship was weak and inconsistent (n.s.) In the adjusted analyses each 10 case per year increase in hospital volume was related to a 6% reduction in odds of death (OR 0.94, 95% CI = 0.88 - 0.99) and a 0.29 day reduction in hospital stay (95% CI = -0.22 to -0.35) in unruptured cases. There was no significant relationship between CFR or hospital stay in the ruptured cases. Female sex, transfer from another acute hospital and age were also associated with increased odds of death	Sensitivity analyses were carried out to evaluate moving the 5 year volume cut off points for each category up or down from the original cut off points, the results remained the same After exclusions there were 5492 unruptured cases and 1203 ruptured cases Case-mix adjustment score = 2

Table 11 Abdominal aortic aneurysms (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hertzer et al (1984)[56] USA To examine outcomes for 3 types of vascular surgery according to surgeon experience	Patients undergoing AAA= 1053 Hospitals in Cleveland-Akron region = 28 Surgeons = 36	Retrospective analysis of The Cleveland Vascular Society computer registry (includes preoperative risks, operative circumstances) 1978 - 1981	Mortality	Surgeon volume per year: < 10 10 - 25 > 25	Patients were divided into either elective or emergency operations Fisher's exact	Overall mortality for elective operations was 6.5% Mortality rates were 4.7% for low volume surgeons, 15.9% for medium volume and 2.9% for high volume. Only the difference between medium volume surgeons and the 2 other groups were significant (p<0.001) Overall mortality for emergency operations was 32.9%, surgeon volume had no effect on mortality	Case-mix adjustment score = 1

Table 12: Vascular and cerebro vascular surgery

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients undergoing amputation of the lower limb (no current trauma) = 10267 Acute care hospitals in the United States = 973	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year (average n = 10.55) average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 14.4% The SMR for low v high volume hospitals was 1.11 v 0.94 p<0.05 (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For all risk patients in low volume hospitals the SMRs were above 1 and were lowest for high risk patients and below 1 for all patients in high volume hospitals and were lowest for medium risk patients (results presented graphically)	Case-mix adjustment score = 3

Table 12: Vascular and cerebro vascular surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing vascular surgery = 44786 Hospitals = 1309	Retrospective analysis of discharge abstract data from the Commission on Professional and hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: < 200 = 1291 hospitals, 39285 patients > 200 = 18 hospitals, 5501 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 12: Vascular and cerebro vascular surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients with subarachnoid haemorrhage (stroke) = 5049 Hospitals = 749	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	Volume was measured as a continuous variable	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was not significantly associated with outcomes	Case-mix adjustment score = 2

Table 12: Vascular and cerebro vascular surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with cerebro-vascular accident = 4395 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 201.82 (s.d.115.37) Physician volume = 7.60 (s.d.7.63)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between hospitals, but mortaility rate was inversely correlated with physician volume (coefficient = -0.149, p<0.05) For LOS hospital volume was significant (coefficient = 0.050, p<0.01) indicating that larger volume hospitals keep patients hospitalised for longer, but physician volume was non significant	Case-mix adjustment score = 1

Table 12: Vascular and cerebro vascular surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hertzer et al (1984)[56] USA To examine outcomes for 3 types of vascular surgery according to surgeon experience	Patients undergoing lower extremity revascularization = 1987 Hospitals in Cleveland-Akron region = 28 Surgeons = 36	Retrospective analysis of The Cleveland Vascular Society computer registry (includes preoperative risks, operative circumstances) 1978 - 1981	Mortality and morbidity (major amputation)	Surgeon volume per year: < 10 10 - 25 > 25	Patients were divided into either aortofemoral reconstruction or femoropopliteal and distal bypass Fisher's exact	Overall mortality for aortofemoral reconstruction was 3.5% and major amputation was 1.5% Mortality for low volume surgeons was 3.6% and for high volume was 4.4% (n.s) Overall mortality for femoropopliteal and distal bypass was 2.8% and major amputation was 6.0% Mortality did not differ significantly between surgeons, but the amputation rate of 2.8% for high volume surgeons was significantly better than that of lower volume surgeons; 6.3% and 9.3% ($p<0.001$)	Case-mix adjustment score = 1

Table 13 Respiratory

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with pneumonia = 5910 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 210.68 (s.d. 85.55) Physician volume = 8.23 (s.d.7.17)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	High volume hospitals had significantly greater mortality than low volume hospitals (coefficient = 0.372, p<0.01), but there were no significant differences between high and low volume physicians High volume hospitals and physicians had significantly longer LOS than low volume hospitals and physicians (coefficient = 0.059, p<0.001 and 0.037, p<0.01 respectively)	Case-mix adjustment score = 2

Table 13 Respiratory (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with chronic pulmonary disease = 1803 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 90.62 (s.d. 59.83) Physician volume = 6.56 (s.d. 6.76)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	High volume hospitals had significantly greater mortality than low volume hospitals (coefficient = 0.995, $p<0.05$), but there were no significant differences between high and low volume physicians There were no significant differences in LOS between high and low volume hospitals, but high volume physicians had significantly longer LOS than low volume physicians (coefficient = 0.058, $p<0.01$)	Case-mix adjustment score = 1

Table 13 Respiratory (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with respiratory failure = 1331 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 79.25 (s.d. 69.07) Physician volume = 15.03 (s.d. 30.09)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, or physicians High volume hospitals had a significantly longer LOS than low volume hospitals (coefficient = 0.119, p<0.05) as did high volume physicians (coefficient = 0.070, p<0.01)	Case-mix adjustment score = 1

Table 13 Respiratory (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with respiratory infection = 1381 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 67.66 (s.d. 50.35) Physician volume = 4.79 (s.d. 5.94)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, but high volume physicians had a significantly lower mortality rate than low volume physicians (coefficient = -0.332, p<0.01) There were no significant differences in LOS between high and low volume hospitals or physicians	Case-mix adjustment score = 1

Table 14: Gastric surgery

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients undergoing operations for ulcers = 26688 Acute care hospitals in the United States = 1100	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year (average = 24.26 average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 4.3% A non-significant difference in the SMR for low v high volume hospitals (1.05 v 0.97) (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For low risk patients the SMR was 1 in low volume hospitals, rising for medium risk patients and falling for high risk patients and in high volume hospitals the SMR was below 1 for all patients but greatest for medium risk patients (results presented graphically)	Case-mix adjustment score = 3

Table 14: Gastric surgery (cont)

Author, year, country and objectives	Procedure, diagnosis(n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hannan et al (1989)[32] USA To test the combined relationship of hospital and physician volume with in-hospital mortality rates	Patients undergoing partial gastrectomie = 1342 hospitals = 216 physicians = 828	Retrospective analysis of New York State discharge abstracts from the Statewide Planning and Research Cooperation System 1986	In-hospital mortality	A continuous volume measure was used which logistic regression defined as: hospital volume: >5, n = 124, cases = 307 6 - 8, n = 36, cases = 245 9 - 14, n = 38, cases = 414 15 - 36, n = 18, cases = 376 physician volume > 1, n = 517 < 1, n = 311 (low and high physician volume was also defined within each of the hospital volume categories)	Age, sex, race, admission status, upto 4 secondary diagnoses and procedures, and severity of illness. Logistic regression	The overall unadjusted mortality rate was 12.3% Age ($p<0.001$), number of secondary diagnoses ($p<0.001$) and disease stage ($p<0.01$) were significantly associated with mortality in the regression analysis as was physician volume ($p<0.01$) Low volume physicians risk adjusted mortality was 13.60 vs 9.60% for high volume physicians, the ratio of these percentages was 1.42 ($p = 0.05$) Patients treated by low volume physicians practising in hospitals with 5 or fewer cases had a risk adjusted mortality rate of 17.5% vs 12.3% for low volume physicians in high volume hospitals, the ratio of these percentages was 1.43 ($p = 0.05$)	There were originally 16 procedures included in this study, but only those 5 which had statistically significant volume-mortality relationships were presented Case-mix adjustment score = 3

Table 14: Gastric surgery (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Kelly et al (1986)[65] USA To examine the effects on post-surgical mortality rates of the volume of procedures performed by individual physicians and hospitals	Patients undergoing a stomach operation for ulcer diagnosis = 1742 Short term general hospitals = 98, physicians = 382	Retrospective analysis of data from the Hospital Cost and Utilization Project, based on discharge abstract records 1977	In-hospital mortality	Both hospital and physician volume were expressed as continuous variables Mean hospital volume = 38 (s.d. 23) Mean physician volume = 8 (s.d. 0.37)	Age, sex, number of diagnoses and stage of illness were controlled for in a logistic regression analysis where hospital and physician volume were included as independent variables	An inverse correlation between hospital volume and adjusted mortality Each additional 17 operations performed decreased the probablity of death by 1% (p=0.01) The volume of procedures performed by individual physicians did not have a statistically significant effect on mortality The number of diagnoses had a significant effect on mortality (p = 0.05); a patient with 3 diagnoses had between a 4 and 7% higher probability of experiencing a negative outcome than a patient with only 1 diagnosis Increasing age was associated with a higher probability of mortality (p = 0.05). Disease staging variables were significant (p = 0.05)	Case-mix adjustment score = 3

Table 14: Gastric surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing stomach operations = 9442 Hospitals = 656 Surgeons = 3735	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 14.11, s.d. 12.7) Surgeon volume: median number of procedures per year = 2, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-0.0773, n.s) and a lower proportion of patients operated on by low volume surgeons (0.1708, n.s.) were positively related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score = 2

Table 14: Gastric surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing vagotomy and/or pyloroplasty for duodenal ulcer = 4316 Hospitals = 938	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: <5 = 609 hospitals, 1287 patients > 5 = 285 hospitals, 2293 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume, but flattened out at relatively low volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 14: Gastric surgery (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing vagotomy = 8704 Hospitals = 1108	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: < 1 = 161 hospitals, 161 patients > 1 = 947 hospitals, 8543 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	No relation between volume and mortality was found (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 15: Cholecystectomy (and other gallbladder operations)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients undergoing gallbladder operations = 130749 Acute care hospitals in the United States = 1196	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year (average = 109) average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 1.1% The SMR for low v high volume hospitals was 1.19 v 0.93, $p<0.005$ (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For low, medium and high risk patients SMRs were above 1 in low volume hospitals, and medium risk patients had the greatest SMR and below 1 in high volume hospitals, where SMRs rose for medium and high risk patients (results presented graphically)	Case-mix adjustment score = 3

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hannan et al (1989)[32] USA To test the combined relationship of hospital and physician volume with in-hospital mortality rates for several procedures	Patients undergoing total cholecystectomies = 25091 Hospitals = 253 Physicians = 2322	Retrospective analysis of New York State discharge abstracts from the Statewide Planning and Research Cooperation System 1986	In-hospital mortality	Hospital volume: < 53, n = 50, cases = 2754 54 - 102, n = 112, cases = 5717 103 - 130, n = 29, cases = 3348 131 - 168, n = 21, cases = 3475 169 - 220, n = 19, cases = 3707 221 - 400, n = 22, cases = 6090	Age, sex, race, admission status, up to 4 secondary diagnoses and procedures, and severity of illness. Logistic regression	The overall unadjusted mortality rate was 1.4% Age (p<0.001), sex (p<0.001), admission status (p<0.01), number of seconday diagnoses (p<0.001) and disease stage (p<0.001) were significantly associated with mortality in the regression analysis as was hospital volume (p<0.01) Physician volume was not statistically significant in the regression analysis For hospitals performing 168 procedures or less per year the adjusted mortality rate was 1.52% vs 1.21% for hospitals performing more than 168 procedures , the ratio of these percentages was 1.26 (p = 0.05)	There were originally 16 procedures included in this study, but only those 5 which had statistically significant volume-mortality relationships were presented in the study. Case-mix adjustment score = 3

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Boles (1994)[45] USA To examine linkages between hospital volume and outcomes for several procedures	Patients undergoing cholecy-stectomy(n not given) Acute care general hospitals = 1751	Retrospective analysis of Medicare Hospital Information from the Health care Financing Administration 1988 and 1990	In-hospital mortality	Mean hospital volume: 3.607 (s.d. 0.671) Cross sectional and longitudinal assessments were made	Age, sex, comorbidity, type and source of admission, previous hospitalisations Linear regression (and regression lines)	In the cross sectional analysis at low levels of volume observed mortality was greater than predicted and at high volume levels were lower than expected ($p = 0.6$) Volume increased over time ($t = 3.707$, $p<0.0002$) but the risk adjusted mortality rate did not change significantly over time	Case-mix adjustment score = 2

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing cholecy-stectomy = 80587 Hospitals = 742 Surgeons = 7062	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 107.88, s.d. 87.8) Surgeon volume: median number of procedures per year = 7, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-0.1466) and a lower proportion of patients operated on by low volume surgeons (0.5236) were positively related to better outcomes, although not significantly so	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score = 2

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing cholecy-stectomy = 162572 Hospitals = 1481	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: < 1 = 3 hospitals, 3 patients > 1 = 1478 hospitals, 162569 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	No relation between volume and mortality was found (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing cholecy-stectomy and incision of common bile duct = 3580 Hospitals = 894	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: <5 = 609 hospitals, 1287 patients > 5 = 285 hospitals, 2293 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume, but flattened out at relatively low volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing biliary tract surgery = 8957 Hospitals = 1278	Retrospective analysis of discharge abstract data from the Commission on Professional and Hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: ≤ 10 = 1007 hospitals, 4500 patients > 10 = 271 hospitals, 4457 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume, but flattened out at relatively low volume (results presented graphically)	Differences between hospitals were not tested statistically Case mix adjustment score = 2

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Roos et al (1986)[71] Canada To examine patient, surgeon and hospital characteristics associated with post discharge complications	Patients undergoing cholecy-stectomy = 9318 Hospitals (n not given)	Retrospective analysis of Manitoba's health insurance database (hospital and medical claims information) 1974 - 1976 2-year before operation and 2 - year after operation histories were obtained	Post-operative comp-lications, readmission Two specialists were given histories based on claims data who indepen-dently judged whether or not the re-admissions were due to complica-tions	Hospital volume per year: > 100 < 100 Surgeon volume: > 20 < 20	Age, comorbidity, prior history, patient residence, hospital location Multiple regression	Hospital volume was not found to be significantly associated with the probability of complications but surgeon volume was. Patients operated on by physicians averaging 20 or fewer procedures per year were almost twice as likely to have complications (odds ratio 1.78, $p = 0.0001$) In the multiple regression analysis physician volume was found to be a significant predictor of readmissions (adjusted odds ratio 1.85, $p < 0.01$)	14% of patients were excluded and 14% were lost to follow-up leaving 6922 cases for analysis Case-mix adjustment score = 2

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged Medicare beneficiaries	Aged Medicare patients undergoing cholecy-stectomy = 34693 Short stay hospital = n not given	Retrospective analysis of the Medicare Statistical System of the health care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 5, 10513 patients 5 - 7, 8294 patients 8 -11, 7909 patients > 11, 7977 patients	Age and sex. Multiple regression	Age was significantly associated with mortality (p<0.001) There were no statistically significant associations between volume and 60 day mortality or in-hospital mortality	This was a 20% probability sample Case-mix adjustment score = 1

Table 15 Cholecystectomy (and other gallbladder operations) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1981)[70] USA To examine the relation between rates of postoperative wound infection and volume of surgery	Patients undergoing cholecy-stectomy = 4156 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: < 100 beds 100 - 300 > 300 and volume was specified within each size: 429 1397 2330 (respectively)	No patient risk factors were specified Logistic regression was used to examine the association of infection with frequency of an operation	Volume was found to be a significant predictor of postoperative wound infection (chi-square = 11.48 p = 0.0007)	Case-mix adjustment score = 0

Table 16: Appendicectomy

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing appendicectomy = 39545 Hospitals = 646 Surgeons = 6434	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 53.01, s.d. 41.5) Surgeon volume: median number of procedures per year = 4, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-0.3552 $p<0.01$) and a lower proportion of patients operated on by low volume surgeons (0.4515, $p<0.05$) were positively related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score =2

Table 16: Appendicectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986) [28] USA To model actual and expected deaths as a function of volume	Patients undergoing appendecec-tomy = 80211 Hospitals = 916	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 158 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.082$, $p<0.01$)	Case-mix adjustment score = 2

Table 16: Appendicectomy (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Shortell & LoGerfo (1981)[43] USA To examine the relationship among hospital structural characteristics , individual physician characteristics, medical staff organisation characteristics and quality of care in 2 conditions	Patients with a primary appendicec-tomy = 8695 East North Central region hospitals = 92	Retrospective analysis of case abstract data from the Commission on Professional and Hospital Activities (CPHA) 1973	Percentage of normal tissue removed	Physician volume was expressed as the number of patients in each diagnosis divided by the number of appropriate physicians on staff (mean number of patients per physician = 2.83, s.d. = 2.10)	Patients were stratified based on their age and sex (4 groupings) using multiple regression (age and sex - because correlated with the possibility of the appendix being normal or perforated)	Age did not have a statistically significant effect for males but females between 6 and 64 years of age were more than twice as likely to have normal tissue removed as females aged 5 years or younger or 65 and older The number of appendicectomies per surgeon was not significantly associated with the adjusted per cent normal tissue removed ratio (coefficient = 0.088)	This was a 35% sample of all hospitals in the East North Central region stratified by bed size, ownership and teaching status 512 eligible cases were excluded due to missing or miscoded data Case-mix adjustment score = 1

Table 16: Appendicectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1981)[70] USA To examine the relation of between rates of postoperative wound infection and volume of surgery	Patients undergoing appendicec-tomy = 3671 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: < 100 beds 100 - 300 > 300 and volume was specified within each size: 469 1134 2068 (respectively)	No patient risk factors were specified Logistic regression was used to examine the association of infection with frequency of an operation	Volume was found to be a significant predictor of postoperative wound infection (chi-square = 4.95 p = 0.02)	Case-mix adjustment score = 0

Table 17: Intestinal surgery (excluding cancer)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hannan et al (1989)[32] USA To test the combined relationship of hospital and physician volume with in-hospital mortality rates for several procedures	Patients undergoing partial colectomies = 10297 Hospitals = 250 Physicians = 1997	Retrospective analysis of New York State discharge abstracts from the Statewide Planning and Research Cooperation System 1986	In-hospital mortality	A continuous volume measure was used which logistic regression defined as: hospital volume: < 18, n = 93, cases = 951 19 - 40, n= 68, cases = 1948 41 - 100, n = 70, cases = 4594 101 - 170, n = 15, cases = 1950 171 - 278, n = 4, cases = 854 physician volume < 8, n = 1622 > 8, n = 375 (low and high physician volume was also defined within each of the hospital volume categories)	Age, sex, race, admission status, up to 4 secondary diagnoses and procedures, and severity of illness. Logistic regression	Overall unadjusted mortality rate was 6.0% Age (p<0.001), admission status (p<0.001), number of secondary diagnoses (p<0.001) and disease stage (p<0.001) were significantly associated with mortality in the regression analysis as was physician volume (p<0.01) For each hospital volume range, high volume physicians had lower risk adjusted mortality rates than low volume physicians, the ratio of these percentages was 1.26 (p = 0.05) Hospitals with an annual volume of 40 cases or less had a risk adjusted mortality rate of 8.3% vs 5.9% for hospitals with volumes of 40 and above, the ratio of these percentages was 1.41 (p = 0.05)	There were originally 16 procedures included in this study, but only those 5 which had statistically significant volume-mortality relationships were presented Case-mix adjustment score = 3

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients undergoing large bowel resection = 3297 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 135.95 (s.d. 72.12) Physician volume = 9.45 (s.d. 10.15)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, but high volume physicians had lower mortality rates than low volume physicians (coefficient = -0.225, p<0.05) There were no significant differences in LOS between high and low volume hospitals or physicians	Case-mix adjustment score = 2

Table 17: Intestinal surgery (excluding cancer) (continued)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with gastro-intestinal bleeding = 3258 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 123.88 (s.d. 62.95) Physician volume = 5.32 (s.d. 5.60)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	High volume hospitals had significantly greater mortality than low volume hospitals (coefficient = 1.151, $p<0.001$), but mortality did not differ significantly between physicians High volume hospitals had significantly longer LOS than than low volume hospitals(coefficient = 0.135, $p<0.001$) but LOS did not differ significantly between physicians	Case-mix adjustment score = 2

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Gordon et al (1995)[72] USA To study the cost and outcome of regionalisation in Maryland for pancreatico duodenctomies	Patients undergoing pancreatico-duodenc-tomies = 501 Acute care hospitals = 39	Retrospective analysis of Maryland Health Services Cost Review Commission 1988 - first half of 1993	In-hospital mortality, length of stay and intensive care unit length of stay	Hospital volume: < 20: n = 38, patients n = 230 >20 (regional centre RC): n = 1, patients n = 271	Age, sex, race, source of admission, source of payment and up to 4 comorbidities. Multiple linear regression models	Significant differences were found between patients treated at low volume hospitals compared with the RC. Patients at the RC were more likely to have been transferred from another hospital (p = 0.003) and to have hypertension (p <0.001). Patients at low volume hospitals were more likely to be black (p p<0.001), to be on Medicare (p = 0.01) and to have pulmonary disease (p = 0.01) Mortality: unadjusted mortality was 2.2% in the RC vs 13.5% in the low volume hospitals, a crude difference of 11.3% (p<0.001). The adjusted difference in mortality was 11.4% (p <0.001) Unadjusted mortality decreased with increasing volume: 1 - 5 cases per year = 19%, relative risk 8.7 6 - 10 cases: 14%, RR 6.5 11 - 15 cases 13%, RR 5.9% 16 - 20 cases 9%, RR 4.0 >20 cases 2% RR 1.0 For live discharges length of stay in hospital and in intensive care unit differed significantly between low volume hospitals and the RC: hospital: 22.5% to 27.9%, a crude difference of 5.4% and an adjusted difference of 5.7% (p<0.001) ICU: 1.8% to to 3.8%, a crude difference of of 2% and an adjusted difference of 1.9% (p<0.001)	Case-mix adjustment score = 2

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing intestinal operations (includes cancer) = 28486 Hospitals = 708 Surgeons = 5436	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 39.78, s.d. 38.5) Surgeon volume: median number of procedures per year = 3, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-0.5625 $p<0.01$) and a lower proportion of patients operated on by low volume surgeons (0.3358, n.s.) were positively related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score = 2

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing colectomy = 36083 Hospitals = 1390	Retrospective analysis of discharge abstract data from the Commission on Professional and hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: <50 = 1191 hospitals, 21523 patients > 50 = 199 hospitals, 14560 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume, but flattened out at relatively low volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing intestinal operations = 36860 Hospitals = 898	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 131 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.006$, $p<0.01$)	Case-mix adjustment score = 2

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged medicare beneficiaries	Aged Medicare patients undergoing resection of the intestine = 22560 Short stay hospital = n not given	Retrospective analysis of the Medicare Statistical System of the health care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 4, 6739 patients 4 - 6, 5899 patients 7 -10, 5152 patients > 10, 4653 patients	Age and sex. Multiple regression analysis	Age was significantly associated with mortality (p<0.001) High surgical volume was significantly associated with lower 60 day mortality (coefficient = -0.147, p<0.001) and lower in-hospital mortality (coefficient -0.176, p<0.001)	This was a 20% probability sample Case-mix adjustment score = 1

Table 17: Intestinal surgery (excluding cancer) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1981)[70] USA To examine the relation between rates of postoperative wound infection and volume of surgery	Patients undergoing colon resection = 1331 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: < 100 beds 100 - 300 > 300 and volume was specified within each size: 58 294 979 (respectively)	No patient risk factors were specified Logistic regression was used to examine the association of infection with frequency of an operation	Volume was found to be a significant predictor of postoperative wound infection (chi-square = 16.56, p<0.0001)	Case-mix adjustment score = 0

Table 18 Hernia repair

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farley et al (1992)[23] USA To examine how patient outcomes for several procedures respond to changes in volume over time	Patients undergoing inguinal hernia repair = 37041 Short term general hospitals = 330	Retrospective analysis of data collected as part of the Hospital Cost and Utilisation Project (discharge abstract data) 1980 - 1987	Mortality	Volume was measured as a continuous variable Both within and between hospital differences were measured	Age, disease staging Multivariate regression	A significant relationship between increased volume over time and adjusted mortality was found (coefficient = -0.206 $p<0.05$) No significant effect of volume on adjusted mortality was found in the cross sectional analysis	Case-mix adjustment score = 2

Table 18 Hernia repair (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing hernia repair = 78377 Hospitals = 742 Surgeons = 7476	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 104.92, s.d. 86.9) Surgeon volume: median number of procedures per year = 6, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher hospital volume (-0.4750, p<0.01) and a lower proportion of patients operated on by low volume surgeons (0.9427, p<0.01) were positievly related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score = 2

Table 18: Hernia repair (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing inguinal hernia repair = 134497 Hospitals = 920	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 380 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.144$, $p<0.01$)	Case-mix adjustment score = 2

Table 18: Hernia repair (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged medicare beneficiaries	Aged Medicare patients undergoing inguinal hernia repair = 32721 Short stay hospital = n not given	Retrospective analysis of the Medicare Statistical System of the health care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 5, 10934 patients 5 - 7, 8221 patients 8 -11, 6355 patients > 11, 7211 patients	Age and sex Multiple regression	Age was significantly associated with mortality ($p<0.001$) There were no statistically significant associations between volume and 60 day mortality or in-hospital mortality	This was a 20% probability sample Case-mix adjustment score = 1

Table 18 Hernia repair (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1970)[70] USA To examine the relation between rates of postoperative wound infection and volume of surgery	Patients undergoing herniorr-haphy = 5432 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: < 100 beds 100 - 300 > 300 and volume was specified within each size: 498 1875 3059 (respectively)	No patient risk factors were specified Logistic regression was used to examine the association of infection with frequency of an operation	Volume was found to be a significant predictor of postoperative wound infection (chi-square = 13.95 p = 0.0002)	Case-mix adjustment score = 0

Table 19: Non-surgical gallbladder

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients with a non-surgical gallbladder diagnosis = 88839 Acute care hospitals in the United States = 1210	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year (average n = 73.42) average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 2.8% The SMR for low v high volume hospitals was 0.90 v 1.04, $p<0.05$ (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For all risk patients in low volume hospitals the SMRs were below1 and were lowest for high risk patients and in high volume hospitals SMRs were above 1 and were fairly flat (results presented graphically)	Case-mix adjustment score = 3

Table 20: Ulcer

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients with a non-surgical ulcer diagnosis = 138268 Acute care hospitals in the United States = 1214	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year (average n = 113.89) Average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 2.4% The SMR for low v high volume hospitals was 1.02 v 0.98, n.s (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For all risk patients in low volume hospitals the SMRs were slightly above 1 and were fairly falt and in high volume hospitals SMRs were around 1 except for medium risk patients who had a lower SMR (results presented graphically)	Case-mix adjustment score = 3

Table 21: Hip or knee replacement/arthroplasty

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Benjamin (1995)[73] USA To examine the relationship between hospital volume and length of stay associated with knee replacement surgery	Aged Medicare patients undergoing knee replacement =324	Retrospective analysis of data collected as part of the Knee Replacement Patient Outcomes Research Team (PORT), identified via Medicare files 1985 - 1989	Length of stay (LOS), post-operative complications	Volume was specified as a continuous variable (mean =35)	Age, sex, socio-economic status, pre-operative health (comorbidity index), type of knee replacement Regression analysis	Higher volume hospitals had a lower probability of a complication (-0.4141) and after adjusting for this a shorter length of stay (-0.0633, p<0.01)	Case-mix adjustment score = 3

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Boles (1994)[45] USA To examine linkages between hospital volume and outcomes for several procedures	Patients undergoing hip replacement (n not given) Acute care general hospitals = 1751 (not procedure specific)	Retrospective analysis of Medicare Hospital Information from the Health care Financing Administration 1988 and 1990	In-hospital mortality	Mean hospital volume = 3.556 (s.d. 0.849) Cross sectional and longitudinal assessments were made	Age, sex, comorbidity, type and source of admission, previous hospitalisations Linear regression (and regression lines)	No significant differences in mortality between hospitals. Volume increased over time ($t = 12.10$, $p<0.0001$) but the risk adjusted mortality rate did not change significantly over time	Case-mix adjustment score = 2

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farley et al (1992)[23] USA To examine how patient outcomes for several procedures respond to changes in volume over time	Patients undergoing hip replacement / arthroplasty = 130494 Short term general hospitals = 337	Retrospective analysis of data collected as part of the Hospital Cost and Utilisation Project (discharge abstract data) 1980 - 1987	Mortality	Volume was measured as a continuous variable Both within and between hospital differences were measured	Age, disease staging Multivariate regression	No significant relationship between increased volume over time and adjusted mortality was found No significant effect of volume on adjusted mortality was found in the cross sectional analysis	Case-mix adjustment score = 2

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing total hip replacement = 13767 Hospitals = 501 Surgeons = 2301	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 27.15, s.d. 34.6) Surgeon volume: median number of procedures per year = 3, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-0.4782, p<0.01) and a lower proportion of patients operated on by low volume surgeons (0.15457, n.s.) were positively related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score = 2

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing total hip replacement = 16339 Hospitals = 804	Retrospective analysis of discharge abstract data from the Commission on Professional and hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: < 50 = 725 hospitals, 10297 patients > 50 = 79 hospitals, 6042 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume, but flattened out at relatively low volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score = 2

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing hip replacement = 20429 Hospitals = 730	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 152 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.521$, $p<0.01$)	Case-mix adjustment score = 2

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients undergoing a major joint procedure = 4880 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 228.65 (s.d.121.43) Physician volume = 53.66 (s.d. 83.05)	Age and sex Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, but high volume physicians had significantly lower mortality than low volume physicins (coefficient = -0.467, $p<0.01$) High volume hospitals had significantly longer LOS than than low volume hospitals(coefficient = 0.032, $p<0.001$) but LOS did not differ significantly between physicians	Case-mix adjustment score = 1

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Lavernia & Guzman (1995)[75] USA To assess the effects of volume on the short-term outcome of primary and revision hip and knee arthroplasty as a function of surgeon and hospital volume	Patients undergoing primary or revision arthroplas-ties of the hip or knee primary = 19925 revision = 2536 Hospitals in Florida (n not given)	Retrospective analysis of discharge information from Florida's Agency of Health Care Administration 1992	In-hospital mortality, length of stay (LOS) and compl-ications	Hospitals and surgeons: < 10 = 101 patient(primary) 446 (revision) 10 - 100 = 4861 patients (primary), 1315 (revision) >100 = 14963 patients (primary), 775 (revision)	Primary and revision separated. Multivariate regression, Pearson correlations and t-test	Primary procedures: there was a significant difference in the mean complication rate between hospitals with low and high volume (p =0.0001) There were no statistically significant differences in average LOS or mortality between hospitals In surgeons there was a statistical trend showing a decreased complication rate among those that performed more than 100 procedures per year (p = 0.06). Surgeons that had fewer than 10 cases per year showed a significant increase in mortality (p = 0.003) and average LOS (p = 0.0001) compared to high volume surgeons Revision procedures: surgeons performing less than 10 procedures per year had significantly higher mortality rates (p = 0.009) than high volume hospitals and average LOS was longer (p = 0.06) There were no significant differences in mortality or LOS between hospitals	Case-mix adjustment score = 1

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged Medicare beneficiaries	Aged Medicare patients undergoing total hip arthroplasty = 9862 Short stay hospital = n not given	Retrospective analysis of the Medicare Statistical System of the health care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 6, 2643 patients 6 - 9, 2391 patients 10 -15, 2407 patients > 15, 2421 patients	Age and sex Multiple regression	Age was significantly associated with mortality (p<0.001) There were no statistically significant associations between volume and 60 day mortality or in-hospital mortality	This was a 20% probability sample Case-mix adjustment score = 1

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged Medicare beneficiaries	Aged Medicare patients undergoing hip arthroplasty (other) = 17628 Short stay hospital = n not given	Retrospective analysis of the Medicare Statistical System of the health care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 4, 4233 patients 4 - 7, 5248 patients 8 - 12, 4054 patients > 12, 4093 patients	Age and sex Multiple regression analysis	Age was significantly associated with mortality (p<0.001) There was a significant association between high volume and lower in-hospital mortality (coefficient = -0.16, p=0.003) and less mortality within 60 days (coefficient = -0.078, p=0.04)	This was a 20% probability sample Case-mix adjustment score = 1

Table 21: Hip or knee replacement/arthroplasty (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Fowles et al (1987)[74] USA To examine the relationship between hospital and surgeon volumes and mortality or major complications in hip surgery	Patients undergoing total hip replacement surgery = 1324 Acute care hospitals in northern California = 261 Surgeons = 399	Retrospective analysis of billing data for Medicare patients 1980	Death within 90 days and complica-tions (e.g. dislocation of the prosthesis and minor revision surgery)	144 surgeons performed 1 Medicare hip replacement 148 hospitals performed 1 or more procedures	No adjustments for case-mix were made Statistical technique not reported	8% of patients died during follow-up and 5% experienced complications A significant inverse relationship between surgeon volume and post-operative mortality and major complications were also inversely related to surgeons volume, no association for minor complications Same inverse relationship for hospital volume (no statistics reported)	Case-mix adjustment score = 0

Table 22: Hip fracture/fracture of the femur

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients with hip fracture = 52368 Acute care hospitals in the United States = 1169	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities study (PAS) 1972	In-hospital mortality	Two volume measures were developed: average number of patients treated per year (average n = 44.80) average number of patients treated per year, categorised by level of risk Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume	Age, sex, white blood cell count, blood pressure, urine sugar, urine albumin, temperature, haemoglobin, poverty indicator, height/weight index, diabetes drugs, thyroid drugs, stage of disease, secondary diagnoses and procedures Patients were also divided into 3 risk categories To adjust observed outcomes for differences in patient mix indirect standardised was used, using logistic regression	Overall mortality was 9.1% The SMR for low v high volume hospitals was 1.04 v 0.98, n.s (where above 1 indicates more deaths than expected and under 1 fewer deaths than expected) For low and medium risk patients in low volume hospitals the SMRs were above 1 and below 1 for high risk patients and in high volume hospitals the SMR was below 1 for all patients and remained fairly flat for all risk groups (results presented graphically)	Case-mix adjustment score = 3

Table 22: Hip fracture/fracture of the femur (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Boles (1994)[45] USA To examine linkages between hospital volume and outcomes for several procedures	Patients with hip fracture (n not given) Acute care general hospitals = 1751 (not procedure specific)	Retrospective analysis of Medicare Hospital Information from the Health care Financing Administration 1988 and 1990	In-hospital mortality	Mean hospital volume = 3.150 (s.d. 0.762) Cross sectional and longitudinal assessments were made	Age, sex, comorbidity, type and source of admission, previous hospitalisations Linear regression (and regression lines)	In the cross sectional analysis at low levels of volume observed mortality is greater than predicted and at high volume levels were lower than expected (0.001) Volume increased over time ($t = 3.887$, $p<0.0001$) but the risk adjusted mortality rate did not change significantly over time	Case-mix adjustment score = 2

Table 22: Hip fracture/fracture of the femur (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with a hip fracture = 2492 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 98.96 (s.d. 51.38) Physician volume = 10.93 (s.d. 14.74)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, but high volume physicians had significantly lower mortality than low volume physicians (coefficient = -0.425, $p<0.01$) High volume hospitals had significantly longer LOS than low volume hospitals (coefficient = 0.114, $p<0.01$) but high volume physicians had significantly shorter LOS than low volume physicians (coefficient = -0.047, $p<0.001$)	Case-mix adjustment score = 2

Table 22: Hip fracture/fracture of the femur (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1988)[76] USA To assess the relationship between hospital volume and patient outcomes for patients with hip fracture	Patients with a diagnosis of hip fracture = 44905 Short term hospitals = 704	Retrospective analysis of case abstract data from the Professional Activities study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS)	Volume categories were not specified	Age, sex, type of surgery and secondary diagnoses Two sets of simultaneous equations to analyse the causal direction of the volume - outcome relationship, where actual and expected outcomes and number of patients were used to calculate Z scores	Higher (log) volume led to reduced mortality (coefficient = -0.264 [0.092], p<0.01) and reduced LOS (coefficient = -0.349 [0.141], p<0.01) supporting a practice makes perfect hypothesis Volume was increased in hospitals that had better than expected mortality rates (coefficient = -0.227 [0.049], p<0.01) and lower than expected LOS (coefficient = -0.072 [0.011], p<0.01) supporting a selective referral hypothesis Patient outcomes also significantly affected outcome: diabetes (p<0.05) and heart disease (p<0.01)	No information given on how many cases were excluded due to missing or miscoded data Case-mix adjustment score = 2

Table 22: Hip fracture/fracture of the femur (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients with fracture of the femur = 46468 Hospitals = 910	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 115 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.034$, $p<0.05$)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Field et al (1990)[77] UK To examine how differing approaches to neonatal care affected survival after premature delivery in one region within the UK	Infants who required admission to baby care units of less than or equal to 28 weeks gestation = 4252 Consultant obstetric units = 17	Prospective cohort design where data was collected on every admission by 1 of 2 observers (2 doctors and 2 nurses) February 1987 - January 1988	Survival to discharge	Intensive care units where each unit carried out at least 500 days of ventilation annually: n = 5 Special care units where less than 500 days of ventilation per year were carried out = 12 Infants were divided into 4 groups: all care in IC units (n =1443), all care in special care units (n = 2500), antenatal transfers (n = 117), postnatal transfers (n = 192)	Birthweight, gestational age, respiratory distress at birth, cephalic or breech presentation, Apgar scores and multiple pregnancy X^2 test, Fisher's exact and Mann Whitney U test were used for analysing categorical data Confidence intervals for relative odds were estimated	Infants of ≤ 28 weeks gestation in IC units showed significantly better survival rates than infants treated in special care units (52% v 22% survivors) relative odds (95% CI) of dying in SC v IC units were 3.98 (1.55 - 10.18) Differences in survival between more mature infants were not significant Only the incidence of respiratory distress was significantly different between the 2 types of unit, indicating a worse potential for infants at IC units. IC units had 62 of 65 infants with respiratory distress and special care units had 28 of 37 ($p<0.01$) Infants of < 28 weeks gestation who were transferred either antenatally or postnatally to IC units showed improved survival compared with infants who remained in special care units (antenatal transfers: X2 = 9.48, 1df, $p<0.002$), relative odds of dying (95% CI) 5.44 (1.89 - 15.62) and (postnatal transfers: X2 = 11.31, 1df, $p<0.0008$), relative odds of dying (95% CI) 5.34 (1.97 - 14.45)	Case-mix adjustment score = 3

Table 23 Neonatal/perinatal care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farley et al (1992)[23] USA To examine how patient outcomes for several procedures respond to changes in volume over time	Neonates with acute respiratory distress syndrome = 56014 Short term general hospitals = 222	Retrospective analysis of data collected as part of the Hospital Cost and Utilisation Project (discharge abstract data) 1980 - 1987	Mortality	Volume was measured as a continuous variable Both within and between hospital differences were measured	Multiple birth, respiratory problems other than respiratory distress, hemolytic disease, other perinatal jaundice, intrauterine hypoxia, birth asphyxia and birthweight (if coded on the discharge abstract) Multivariate regression	A significant relationship between increased volume over time and adjusted mortality was found (coefficient = -0.415 $p<0.05$) No significant effect of volume on adjusted mortality was found in the cross sectional analysis	Case-mix adjustment score =2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
LeFevre (1992)[79] USA To determine the relationship between physician volume and perinatal outcome as measured by neonatal and perinatal mortality	All births to Missouri residents = 303,104 Level III obstetric/neonatal centres with new-born intensive care = 4 Physicians = 715	Retrospective analysis of fetal death certificates and birth certificates linked to infant death certificates 1984 - 1987	Neonatal/perinatal death (neonatal death = live birth with subsequent death <29 days) (perinatal death = fetal death or neonatal death)	Physician volume: < 100 (n = 10563 deliveries) 101 - 200 (n = 18019) 201 - 300 (n = 20603) 301 - 400 (n = 25894) > 400 (n = 135468)	Maternal race, age, education, smoking habits and birthweight (for neonatal death) Logistic regression	Neonatal mortality rate was 4.3 per 1000 live births and perinatal mortality was 10.2 per 1000 live births plus fetal deaths Volume was not a statistical predictor of death for either perinatal (odds ratio 1; 95% CI: 0.99 - 1.01) or neonatal death (odds ratio 1; 95% CI: 0.98 - 1.02) Including birthweight as an independent variable did not alter the pattern of results (odds ratio 1.0; 95% CI: 0.98 - 1.02) Using volume as a categorical variable the odds ratio did not differ significantly from 1 for any of the volume groups and the confidence intervals were narrow	4 level III (with IC) centres were excluded (n = 34,126) to help eliminate bias associated with transport of women in labour with premature infants (total of 30% excluded leaving 210547 births) Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
LeFevre et al (1992)[78] USA To examine the relationship between neonatal mortality and the level of perinatal services present in the hospital of birth	Singleton, live births = 354441 Missouri hospitals = 119	Retrospective analysis of birth certificates matched with death certificates January 1980 - December 1984	Mortality	Hospitals were classified according to available services and number of deliveries: Level IA: < 500 births per year & no intensive care facilities (n = 79 & 84171 births) Level IB:< 500 births (n = 11 & 39818 births) Level IIA: < 500 births (n = 17 & 101920 births) Level IIB: < 500 births (n = 8 & 95693 births) (Levels IB < IIB had increasing degrees of facilities) Level III: > 500 births and neonatal intensive care (n = 4 & 32839 births)	Race and birthweight Logistic regression	Both black (e.g. 18 deaths per 1000, 95% CI: 11.3 - 24.7 v 8.9, 95% CI: 6.7 - 11.1 in level IA v level III) and white (e.g 27, 95% CI: 23.9 - 30.1 v 17.3, 95% CI: 14.2 - 20.4 respectively) infants born at level IA hospitals who weighed less than 5lb had worse outcomes than those at level III hospitals There were no other statistically significant differences between remaining hospital levels at any birthweight	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Lumley [80] (1988) Australia To assess the perinatal outcome by size of hospital in one Australian state	Stillborn and liveborn infants of 500g or more in the first 28 days of life = 179628 Hospitals in Victoria (n not given)	Retrospective analysis of perinatal death certificates, autopsy reports and perinatal morbidity forms completed for all births in Victoria 1982 - 1984	Mortality and morbidity (as measured by the 5 minute Apgar score)	Hospitals were grouped by the number of births: < 100: country hospitals with level I perinatal services (9717 deliveries) 100 - 999: country hospitals with level II services and metropolitan and country hospitals with level I services (67077 deliveries) 1000 - 1999: metropolitan level II hospitals (44315 deliveries) > 2000: obstetric teaching hospitals with level III services (58521 deliveries)	Birthweight Confidence intervals for relative risk and birthweight standardised perinatal mortality ratio were calculated	Crude mortality rate ranged from 4.3 to 20.7 per 1000 with all except the largest hospitals having rates below the state figure of 13.3 per 1000 For infants of less than 1500g the birthweight specific mortality rate decreased significantly with increasing size Infants of 2500 - 2999g there was no overall trend but mortality was significantly lower in hospitals with fewer than 100 births a year (2.4, 11.1, 10.5, 10.5 respectively) All infants under 2500g had a better outcome in large hospitals when late transfers were taken into account (e.g. 1500-1999g, mortality was 192.3, 105.4, 55.4, 55.3 for <100, 100-999, 1000-1999, >2000 respectively For infants of 3000g and above mortality increased significantly with increasing hospital size (2.7, 2.7, 3.1, 4.3 respectively) Hospitals with 100 - 999 births had a significantly higher number of perinatal deaths than expected (ratio 114, 95% CI: 103 - 187, p<0.001) Hospitals with 2000+ births had a significantly lower than expected death rate (ratio 92, 95% CI: 186 - 99, p<0.02)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Babies with respiratory distress syndrome = 16373 Hospitals = 770	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 71 cases per year	Birthweight Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.080$, $p<0.01$)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Mayfield et al (1990)[81] USA To examine the relation of hospital obstetrical volume and neonatal nursery technology to perinatal outcome	White, singleton births over 1000 gm = 226164 Washington hospitals offering obstetrical care (included level I, level II and level III) = 90	Retrospective analysis of birth-death linked records and fetal death records 1980 -1983	Perinatal mortality (all fetal deaths over 1000 gm and all live births over 1000 gm which died within 1 week of birth)	Obstetrical volume per year was calculated as the average of all deliveries performed over the 4 year period: <200 (31 hospitals) 201 - 500 (12 hospitals) 501 - 1000 (14 hospitals) 1001 - 2000 (11 hospitals) >2000 (11 hospitals) Volume was specified within each level of care	Diabetes, hypertension, pyelonephritis, pregnancy induced hypertension, acute urinary tract infection, surgery, tumour, epilepsy, prior fetal death, pre-eclampsia, eclampsia, Rh-sensitization, abruptio-placenta, placenta previa or other medical problems (not defined) & birthweight Descriptive analyses and a log-linear model	The lower the birthweight the more likely the infant was to be born in a level III setting and larger volume facilities Infants under 2000 gm were twice as likely to die if born in a level I or II facility compared to a level III facility (e.g. 459 v 175 deaths per 1000 births) and had twice the mortality rate in facilities with less than 2000 deliveries per year when compared to those born in facilities with more than 2000 deliveries per year (numbers not given). There was little difference for infants over 2000 gm The regression model showed that birthweight was the most important variable, followed by nursery level but not volume on the perinatal mortality rate	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Williams (1979)[82] USA To examine the effectiveness of perinatal care in California via the development of an outcome based index	Births in California = 3370338 Hospitals in California = 504	Retrospective analysis of linked birth and death certificates from the California Department of Health 1960, 1965 - 1973	Perinatal mortality rate	Size of delivery service was measured by the average annual number of deliveries over the sample period by the number of years that a hospital had a viable maternity service	Birthweight, gender, race and plurality were used to construct an expected perinatal mortality rate Multiple regression	The risk adjusted perinatal mortality rate decreased with the size of delivery service until 2850 births then increased thereafter ($F = 27.28$, $p = 0.0001$) Mortality ratios were significantly higher in hospitals with large percentages of Hispanic mothers ($F = 51.33$, $p = 0.0001$)	2% of the sample were excluded for a variety of reasons (e.g. stillbirths less than 500g, non hospital births) Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hein (1978)[83] USA To review the quality of perinatal services in a rural midwestern state in America	All births in Iowa from 1972 - 1976: n = 204480 Iowa hospitals = 135	Retrospective analysis of perinatal birth and mortality statistics 1972 - 1976	Mortality	Hospitals were classified by deliveries per year: 1 - 99 100 - 249 250 - 499 500 - 999 1000 +	None, although numbers were given for each category of hospital by birthweight No statistics reported	Statewide mortality for 1976 was 9 per 1000 births In 1976 the neonatal death rate for low volume hospitals was 5.85% and in high volume hospitals was 12.26% However in 1976 the highest volume hospital had 45% of its births in infants weighing under 1500g compared with 3.25% of such births in the lowest volume hospitals	Case-mix adjustment score = 0

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
The International Neonatal Network (1993)[84] UK To compare neonatal unit performance in 13 UK neonatal units	Infants of birthweight 1500g or less or gestational age less than 31 weeks = 1659 Neonatal units = 13	Retrospective cohort where data was abstracted form hospital records by 4 research assistants 1988 - 1990	Mortality	Tertiary neonatal intensive care units: n = 9 & 1439 infants Non-tertiary neonatal intensive care units: n = 4 & 220 infants	Birthweight and in a separate analysis CRIB (clinical risk index for babies): birthweight, gestation, congenital malformations, maximum and minimum fraction of inspired oxygen and maximum base excess during the first 12 hours The odds of hospital death, before and after adjustment for initial risk were compared	When no adjustments were made for infant risk there were no statistically significant differences in the odds of hospital death between tertiary and non-tertiary units After adjustment for birthweight alone the odds of hospital deaths in non-tertiary units compared with tertiary units were 1.45 (95% CI: 1.01 - 2.11, p = 0.04) After adjustment for CRIB the odds of hospital deaths in non-tertiary units compared with tertiary units were 2.12 (1.39 - 3.24, p = 0.0005)	There were 1548 cases with complete data sets (93%) Case-mix adjustment score = 3

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Becerra et al (1989)[85] South America To evaluate the regionalisation of perinatal health services operating in Puerto Rico	Single delivery births (337415) with 4361 neonatal deaths (birthweights 227 - 500g) Hospitals (n not given)	Retrospective analysis of linked birth-death certificates 1980 - 1984	Mortality	Hospitals were classified as: > 400 deliveries per year < 400 deliveries per year and by level of care: medical centres (level A) (all had > 400 deliveries) regional centres (level B) (all had > 400 deliveries) primary hospitals (level C)	Maternal age, marital status, pregnancy complications, birthweight, Apgar score at 5 minutes Regression was used to identify significant predictors of death, which were then used to indirectly standardise mortality by level of care. Z scores were calculated for each hospital as the ratio between the observed number of deaths to the square root of the expected n of deaths	Birthweight, Apgar score and pregnancy complications were statistically significant independent predictors of death ($p<0.05$) Level A hospitals had the highest proportion of low birth weight infants (10.6%) followed by level B (9.5%) and then level C (6.4%). The same held true for very low birth weight infants (1.4, 0.9%, 0.6%) Low birthweight infants were at no advantage in either level A or B hospitals. Infants of normal birthweight had a higher risk of death at level B, and even higher at level A More deaths occurred than expected in levels A and B after adjusting for birthweight, Apgar score and pregnancy complications (no figures given, results presented as graphs)	Case-mix adjustment score =2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Clarke et al (1993)[86] UK To evaluate mortality rates as a method of auditing obstetric and neonatal care (compares perinatal mortality rates between different maternity units)	Singleton perinatal deaths and their selected live born controls =1179 out of 114362 singleton births Hospitals = 39	Case control study where data was collected from case notes and from interviews with the women 1978 - 1987	Mortality	Hospitals were classified into: main consultant units (hospital 1 & 2): other consultant units (hospital 3, n = 28) GP units (n = 9)	Age, race, parity, occupation, height, antenatal care, diabetes, smoking status, history of infertility, social class, gestational diabetes, urinary tract infection, pre-eclampsia toxaemia, antepartum haemorrhage, length of inpatient care and birthweight Standardised perinatal mortality rates for delivery units were calculated using logistic regression to estimate relative risks adjusted for confounding factors and then standardising the rate	Women transferred to consultant units during pregnancy had a high perinatal mortality rate (16.8/1000 deliveries) The crude perinatal mortality rate in hospital 1 was a third higher (12.6/1000) than that in hospital 3 (9.4/1000) Perinatal mortality by place of booking showed little difference between GP units (9.3/1000) and consultant units (8.8/1000) Perinatal mortality rates by place of delivery showed differences between GP units (3.3/1000) and consultant units (9.4/1000) due to selective referral of high risk women from GP units Adjustment for risk factors made little difference to the rates except when deaths due to immaturity were adjusted for birthweight (e.g. crude rates (95% CI): 3.1 (2.6 - 3.6), 2.1 (1.5 - 2.7), 2.3 (1.5 - 3.1), 0.6 (0.2 - 1.0) v adjusted 1.9 (1.6 - 2.2), 2.5 (1.8 - 3.2), 5.7 (3.7 - 7.7), 0.8 (0.3 - 1.3) in hospital 1, 2, 3 and GP units respectively	Case-mix adjustment score =2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Gortmaker et al (1985)[87] USA To examine differentials in survival among low birth weight infants according to hospital of birth in 4 American states	Single and plural births < 2501g in birthweight = 53948 Hospitals in Louisiana, Ohio, Tennessee & Washington (n not given)	Retrospective analysis of linked birth and death certificate data 1978 - 1979	Survival in hospital and at birth to one year	Hospitals were classified as : level III all level I and II divided into urban or rural	Birthweight and date of mother's last menstrual period Mortality rates and survival curves were estimated for race and birthweight specific groups (1500 to 2000g and 2000 to 2500g) as well as for types of hospital Multivariate logistic regression and proportional hazards model	The survival of black and white of 1000 to 1500g born in level III centres was significantly greater than that of infants born in rural hospitals (p<0.0001). White infants born in level III centres experienced significantly greater 4 day survival than white infants born in other urban hospitals (p = 0.0001) although this was not true for black infants Infants < 1000g born in level III centres experienced significantly better survival than infants born in either urban or rural hospitals (e.g. 486/1000 v 608/1000 for level III v rural, p<0.05) Mortality throughout the first year of life also differed significantly between hospitals for white infants (e.g. 544/1000 v 615/1000, p<0.05) for level III v rural hospitals in infants < 1000g) but for black infants only differences between level III and rural were significant in infants 1000 to 15000g (143/1000 v 254/1000, p<0.001) When the data were analysed including gestation and plural or single birth the results remained very similar to the birthweight/race adjusted rates	Case-mix adjustment score =2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Grzybowski et al (1991)[88] Canada To determine whether a small isolated hospital can provide acceptably safe obstetric and perinatal care	Women beyond 20 weeks gestation = 286 A 21 bed general hospital	Population based prospective cohort study where data were collected from the chart's of all women and from the British Columbia Labour and Delivery and Newborn records January 1984 - December 1988	Mortality and adverse perinatal outcome (death, birth weight of less than 2500g, neonatal transfer or Apgar score of less than 7 at 5 minutes)	The outcomes of the 21 bed general hospital were compared with the outcomes from 2 other hospitals (reported in the literature) and those women who were transferred after admission to hospitals with fewer than 50 deliveries per year	Native and non-native populations were compared on a number of characteristics: age, parity, gestational age, birthweight, sex X^2 and Fisher's exact test	There were 6 perinatal deaths for a rate of 20.8 (95% CI: 4.4 - 37.2) There no statistically significant differences between native and non-native populations The hospital based rate of adverse perinatal outcome was 6.2% (12 out of 193 newborns, [not all infants were born at the general hospital] 95% CI: 2.8% - 9.6%) A community hospital in New York with 6856 deliveries had an adverse outcome rate of 4.4% and a rural practice had a rate of 7.4% among 635 pregnancies 12% were transferred after admission	Case-mix adjustment score =2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Horbar et al (1988)[89] USA, Canada & England To determine whether survival varied among infants born at 11 neonatal intensive care centres	Infants weighing 701 - 1500g = 1776 Neonatal intensive care centres = 11	Retrospective analysis of clinical databases and chart review 1983 - 1984	Survival and survival without supplemental oxygen on day 28 after birth	No information is given other than the number of intensive and intermediate care beds ranged from 16 - 60	Birthweight, gender and race Univariate associations between outcomes and birthweight, gender, race and centre were determined using X^2 analysis Logistic regression	85% of infants were alive on day 28 after birth and survival at individual centres ranged from 80% to 92% (n.s.) 60% of infants were alive without supplemental oxygen on day 28 and the rate ranged from 51% to 70% at individual centres; differences were significant (X^2 = 31.8, p<0.001) Differences in survival according to birthweight, gender and race were all statistically significant, with increased chance of survival as birthweight increased, being female and being black After adjusting for the effects of birthweight, gender and race there was a significant unexplained residual variation among the 11 centres in both survival (X^2 = 23.9, p<0.01) and survival without supplemental oxygen (X^2 = 44.2, p<0.0001) Predicted survival rates varied by as much as 15% and survival without supplemental oxygen by nearly 17% between centres	Data was solicited from individuals on a mailing list of an international meeting of neonatologists. 32 centres requested data forms and 12 self-selected centres provided data. Data from 1 centre was excluded as it was located in a high altitude area No information is given about the centres Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Lallo et al (1991)[90] Italy To evaluate the efficacy of availability of neonatal intensive care available in place of birth	Live births less than 1500g = 319 Hospitals (n not stated)	Cohort study of very low birthweight babies in one region Register for babies birth and death certificates 1987	1 year mortality	Obstetrician volume: less than 500 per year Centres were classified as: Level III: neonatal intensive care (accepted all babies) Level II: neonatal care (accepted babies but transferred to level III) Level I: no special facilities (did not accept low birth weight babies) Babies transferred from level I to II, not considered as transfers but as if they stayed at level I	Birthweight, gestational age, sex, type of delivery (single or plural) Logistic regression. Kruskall Wallis and Z statistic used for comparisons of averages	In level I most wards were low volume Level I transferred 76% of very low birth weight births , 59% in level II 83% at level I who survived at 4 hours were transferred to level III, 65% transferred in level II Mortality was not statistically significant for 3 groups, lowest level is in level II: level I was 55%, level II , 49 % and 57.9 % for level III Adjusted odds ratios: risk of mortality in level I and II is similar to level III Odds ratio after 4 hours: risk of mortality less for babies in levels 1 and 2 compared with those in level 3	5 babies excluded due to heart malformation Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Paneth et al (1982)[91] USA To examine mortality across 3 levels of care for low birthweight infants	Low birthweight infants who weighed 501 - 2250g at birth = 16365 Hospitals in New York city = 66	Retrospective analysis of vital records of New York City 1976 - 1978	Mortality	Hospitals were grouped according to their level of care: Level I: no special facilities (n = 32 & 3105 births) Level II: capabilities for care of most premature infants (n =20 & 5857 births) Level III: newborn intensive care units (n = 14 & 4598 births)	Race, sex, birthweight and gestational age A further 8 confounders were included in a second analysis (age, parity, education, marital status, private or hospital physician, Medicaid, medical complications and prenatal care) Logistic regression	Singletons: (n = 13560) The adjusted neonatal mortality rate for level III hospitals was 128.5 per thousand live births, (OR 1.0) significantly lower than the rates for both level II (168.1, OR 1.37, 95% CI: 1.19 - 1.58) and for level I (163.0, OR 1.32, 95% CI: 1.12 - 1.55) $p<0.001$ Adding in the additional 8 possible confounders did not change the pattern of results Multiple births: (n = 2805) There was a similar pattern of results to singletons: mortality was significantly higher at level II than level III hospitals ($p<0.05$) but at level I hospitals it was slightly lower but not significantly so (no figures presented)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Phibbs et al (1996)[92] USA To examine the effects of neonatal intensive care patient volume and the level of care available at the hospital of birth on neonatal mortality	Singleton births with conditions likely to result in NICU admission = 53229 Non-Federal hospitals in California with maternity services (n not given)	Retrospective analysis of birth certificate data, linked to infant death certificates and to infant discharge abstracts 1990	Mortality within the first 28 days of life or within the first year if continuously hospitalised	Hospitals were classified by the level of NICU care available (level I to level III) and by the average patient census in the NICU	Birthweight and mother's characteristics Multiple logistic regression	Patient volume and level of care both had significant effects on risk-adjusted neonatal mortality: compared to level I hospitals level III hospitals with an average census of at least 15 patients had lower risk adjusted mortality (OR 0.70, p = 0.002) Risk adjusted mortality for infants born in smaller level IIIs and in level IIs regardless of size was not significantly different from hospitals without a NICU	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Rosenblatt et al (1985)[93] New Zealand To assess whether the low volume of deliveries in small hospitals affects perinatal mortality in the context of a regionalised system of care	Births in New Zealand public maternity hospitals = 206054 births Hospitals = 111	Retrospective analysis of the National Health Statistics Centre of New Zealand (register of all births and perinatal deaths) 1978 - 1981	Perinatal mortality rate	Level 3 (tertiary care units) = 5 hospitals Level 2 (sub-regional referral) = 19 hospitals Level 1 (small, rural) = 87 hospitals Extent of regionalisation was assessed by determining proportion of mothers served by level 1 facilities delivered in level 2 or 3 hospitals Volume specified as: < 100 101 - 200 201 - 500 501 - 1000 1001 - 2000 > 2000	Birthweight (Statistical technique not reported)	Perinatal mortality rate was 12.0/1000 total births and diminished from 13.0 in 1978 to 10.5 in 1981 Crude perinatal mortality increased with hospital volume Level 1 hospitals had lower birthweight specific perinatal mortality rates than level 2 or 3 hospitals in all but the lowest birthweight categories. (Volume tended to cluster around hospital level with low volume in level 1 hospitals) Mortality was lower in lower volume hospitals (by birthweight) (e.g. in low volume hospitals for < 2500g babies the rate was 2.6 and in high volume hospitals 5.8) in all but lowest birthweight babies (428.6 v 421.1)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Verloove-Vanhorick et al (1988)[94] Netherlands To evaluate the influence of place of birth on mortality	Newborn infants born alive in the Netherlands at a gestational age of less than 32 weeks and/or with a birthweight of less than 1500g =1338 Hospitals = 133	Retrospective analysis of the data recorded by attending paediatricians as part of the Project on Preterm and Small for Gestational Age Infants January - December 1983	In-hospital mortality	Hospitals were classified into 1 of 3 levels: Level 1: perinatal intensive care units (n = 8 and 482 infants) Level II: regional teaching hospitals (n = 19 and 359 infants) Level III: hospitals with no facilities for special care (n = 106 and 497 infants) All infants were assigned to their hospital of birth regardless of transfer	Gestational age, birthweight, sex of infant and multiple pregnancy were included in the first logistic regression and in a second logistic regression 22 potential confounding variables were included Odds ratios were calculated with level III mortality as the baseline	Crude mortality rates between the three hospital levels did not differ significantly (around 26%) In the first logistic regression (controlling for 4 possible confounders) the odds for in-hospital mortality was significantly greater in level I (OR 1.62, 95% CI: 1.12 - 2.34, $p<0.01$) and in level II (OR 1.56, 95% CI: 1.04 - 2.33, $p<0.05$) when compared with level III In the second logistic regression (controlling for 22 possible confounders) the odds for in-hospital mortality was significantly greater in level I (OR 1.80, 95% CI: 1.1 - 3.0, $p<0.05$) and in level II (OR 1.90, 95% CI: 1.1 - 3.2, $p<0.05$) when compared with level III	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Viisainen et al (1994)[95] Finland To evaluate whether hospitals of different levels are equally safe places to give birth in a regionalised system of care	Women who gave birth in Finland = 123,065 Maternity hospitals = 53	Retrospective analysis of the Finnish Medical Birth Registry and National Education Registry 1987 - 1988	Mortality	Hospitals were classified into 1 of 4 levels of care: Level 1A: community (n = 5 and 113 average annual deliveries) Level 1B: primary (n = 25 and 614 deliveries) Level 2: secondary (n = 18 and 1505 deliveries) Level 3: tertiary care units (n = 5 and 3478 deliveries)	Age, marital status, education, previous pregnancies and previous still-births, birthweight Logistic regression, using level III hospitals as the reference	Low birthweight and preterm infants were concentrated in level III hospitals Crude perinatal mortality rate was significantly higher in level III hospitals than in level II, level IB and level IA hospitals (13.5%, 6.5%, 5.0%, 4.4%, respectively. p<0.001) Birthweight specific mortality rates showed that very low birthweight newborns had a better rate of survival in level III hospitals than in other hospitals (376.5%, 508.5%, 857.1%, 1000% respectively, p<0.001), but for low birth weight infants level II hospitals had a lower mortality rate than level III (61.4%, 43.7%, 68.3%, 83.3% respectively, p<0.001) For normal birthweight infants mortality was lowest in level II hospitals and highest in level III hospitals (3.1%, 2.1%, 2.3%, 2.7% respectively, only level II was significant p<0.01)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Bowes (1981)[96] USA To assess the impact of regionalisation of perinatal health care on the neonatal mortality rates in Colarado	Live births: n = 319040 Hospitals (n not given)	Retrospective analysis of vital records (birth and linked death certificates) in two 4 year periods (1971 - 1974 and 1975 - 1978) during which regionalisation of perinatal health care occurred on a gradual, nonplanned, voluntary basis	Mortality	Hospital outcomes were compared in 1971 - 1974 (n = 154208 births) and 1975 - 1978 (n = 164832 births) and hospitals were divided into: level III (n = 3) level II (n = 7) level I (n not given)	Very low birth weight (VLBW) infants were analysed separately Neonatal death rates were expressed as the number of deaths per 1000 live births (Statistical technique not reported)	There was a decline in the neonatal mortality rate from 13.4 to 6.9 over the 8 year period There was an increase in VLBW infants in level III hospitals from 2.8% to 4.8% of total live births and a decrease from 1.6% to 1.1% in level I hospitals during the respective study periods A decrease in neonatal mortality from 9.7 to 6.7 between 1974 and1978 which was greatest in level I hospitals (8.3 to 4.6) while there was no decrease in level III hospitals (14.5 & 14.5) In VLBW infants there was a decrease in neonatal mortality from 318 to 220 between 1974 and 1978, which was greatest in level II hospitals (346 to 196) There was a significant difference in the percentage of infants dying within the first 2 hours of life in level I and level III hospitals in 1978 (70% v 30%, $p<0.05$)	Case-mix adjustment score = 1

Table 23 Neonatal/perinatal maternity care (level of care) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Black & Fyfe (1984)[97] Canada To determine the relationship between the level of services provided in small communities and the obstetric and neonatal outcomes in Ontario	Newborn and obstetric patients (northern and southern Ontario) = 123081 Ontario hospitals (n not given)	Retrospective analysis of Hospital Medical records Institute files and data from Registrar General (for comparison) April 1980 - March 1982	Perinatal deaths (stillbirths through to 28 days after birth)	Hospitals in northern Ontario were classified as: level 0 to 1D: < 1000 deliveries per year level II : > 1000 deliveries per year Mortality rates were also compared with those published by the Ontario registrar general for 1980 (southern Ontario)	None, although Indian reserves and communities isolated from hospitals were calculated separately X^2 used to compare community types	Results are presented for communities grouped according to the level of obstetric service available locally The perinatal loss rates in each type of community were not significantly different (e.g. in hospitals with < 1000 deliveries = 13.89 (95% CI: 2.87 - 40.60) and for hospitals with > 1000 deliveries = 12.13 (10.37 - 13.89) In northern Ontario the perinatal death rate per 1000 births was 11.38 v 14.05 in southern Ontario The Indian population did not affect the mortality rate found for the entire population	Information was retrieved by place of residence not by hospital of delivery Case-mix adjustment score = 0

Table 23 Neonatal/perinatal maternity care (regionalisation) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Commentary
Hein (1980)[98] USA To evaluate the outcomes of a voluntary system of regionalised perinatal health care in Iowa	Live births (= 85988) and neonatal deaths (= 937) Iowa hospitals (n not given)	Retrospective analysis of matched birth and death certificates 1972 and 1978	Neonatal mortality rates	Data from 1972 (prior to the perinatal programme) were compared with data from 1978 (the most recent available data post the programme) Hospitals were classified as: small community hospitals with < 500 deliveries per year large community hospitals with > 500 deliveries level II regional centres tertiary centre	Mortality rates were presented for infants weighing <1500g and for infants weighing >1500g The ratio of neonatal deaths to the number of < 1500g birthweight, live infants was compared in the various hospital categories No statistics were reported	There was a decline in mortality of 40% from 1972 to 1978 In 1972 the crude neonatal death rate for small, large, level II hospitals and the tertiary centre were: 12.8, 14.4, 14.6, 14.4 and in 1978 were 5.2, 8.6, 10.0, 17.8 indicating a lower mortality rate in 1972 for small hospitals but similar rates for the other categories. In 1978 the non-centre hospitals had lower mortality rates The survival rates in 1972 for < 1500g infants in were: 27.3, 34.0, 31.0, 51.7 and for infants weighing < 1500g were 92.9, 94.5, 92.6, 94.0 in small, large, level II hospitals and the tertiary centre respectively In 1978 survival rates were 59.2%, 45.7, 52.5, 56% for < 1500g and were 96.2, 97.2, 96.9, 96.5 for infants weighing > 1500g in small, large, level II hospitals and the tertiary centre The mortality risk ratios for neonatal deaths in 1972 were: 1.54, 1.22, 1.13, 1.07 and in 1978 were: 1.21, 1.11, 0.85, 0.60, for small, large, level II hospitals and the tertiary centre respectively	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (regionalisation) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hein & Burmeister (1986)[99] USA To evaluate the effects of the Iowa regional perinatal care system on perinatal outcomes	Live births, fetal and neonatal deaths with specified weights = 87213 Iowa hospitals (n not stated)	Before/after study design using birth and mortality data 1978 (prior to initiation of the programme) 1982 (year for which data available after initiation of the programme)	Mortality	Hospitals were stratified by level: level I: (< 500 births per year and > 500 births per year level II level III	Birthweight Cochran's method was used to compare differences in 1972 and 1982 neonatal mortality rates which were assessed according to effect of changes in birthweight specific mortality rates	Neonatal and fetal mortality decreased overall between 1972 and 1982: neonatal mortality in 1972 v 1982 was: chi-square = 120.48 ($p<0.001$) and the association was 100.38 ($p<0.001$). Fetal mortality in 1972 v 1982 was (X^2 38.07, $p<0.001$) and the association was (X^2 9.16, $p<0.005$) In 1972 neonatal mortality rates were consistent among the various hospital categories: the total X^2 comparing rates in 1972 with 1982 across hospital categories was 193.96 ($p<0.001$). The X^2 assessing the average change was 144.48 ($p<0.001$). (Rates were 12.8, 14.4, 14.6, 14.4 in 1972 for level I (< 500, > 500), II, III respectively For fetal deaths the X^2 was significant (45.95, $p<0.001$). (Rates were 9.2, 11.2, 10.8, 14.6 in 1972 and in 1982: 5.6, 5.1, 6.9, 22.1 for level I (< 500, > 500), level II and level III respectively In 1982, 78.2% of all very low birthweight births were in level II and III hospitals. For very low birth weight babies % survival in 1972 was 25.3, 34.4, 31.1, 51.7 for level I (< 500, > 500), level II, level III and in 1982 was 46.9, 51.2, 61.2, 67.8 showing improved survival for very low birth weight infants	Case-mix adjustment score =2

Table 23 Neonatal/perinatal maternity care (regionalisation) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Nicholls et al (1993)[100] UK To compare the outcomes of neonates with gastroschisis delivered in a regional obstetric and neonatal centre without the facility of on site surgery with the outcomes of neonates in peripheral hospitals who then underwent postnatal transfer	Neonates with gastroschisis who underwent postnatal transfer to the Birmingham Children's Hospital = 43	Retrospective analysis of case notes at the Birmingham Children's hospital 10 year period (dates not given)	Primary closure rates, time to full enteral feeding and mortality	A regional obstetric and neonatal centre (without on site surgery) was compared with peripheral hospitals (both without groups underwent postnatal transfer)	The two groups of transfers were compared with respect to gestational age, birthweight, caesarean section rate and time to operation using unpaired t-tests, Mann-Whitney U test and X^2 test	There were no statistically significant differences in gestational age, birth weight, caesarean section rate or time to operation between the regional centre and the peripheral hospitals Primary closure rates were 89% for the regional centre and 94% for peripheral hospitals (n.s.) Mean time to full enteral feeding was 24 days for the regional centre and 23 days for those delivered peripherally (n.s.) There were 3 deaths in the peripheral hospitals and no deaths in the regional centre (n.s.)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (regionalisation) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Siegel et al (1985)[101] USA To evaluate the impact of a rural regional perinatal care (RPC) programme	Resident fetal deaths and live births: pre-implementation period (= 4093 pilot and 3301 control) Hospitals (n not given)	Quasi-experimental with a matched control region using an interrupted time series design (interruption = introduction of the RPC programme) Vital statistics data and systematic chart recordings by physicians and nurses were used 1968 - 1974 pre-regionalisation 1975 - 1979 post-regionalisation 1978 - 1979 morbidity data collected	Yearly fetal, neonatal and perinatal mortality trends and short term obstetric and newborn morbidity data (presence of 1 or more high risk conditions)	Two regions were compared: 1 with regional perinatal care and 1 without	Regions were compared on a number of variables and results presented for white and non-whites Pre- implementation of the RPC perinatal mortality was 31% v 28.8% for pilot v control regions Segmented regression was used to detect changes in trends and test for differences between pilot and control changes Birthweight specific mortality rate % reductions for fetal deaths and live births were compared between the 2 regions for pre and post-programme periods	No statistically significant differences were found in mortality trends between pilot and control regions; both declined over the study period Percentage declines in each birthweight-race category were similar in the pilot and control regions (e.g. non-white, low birthweight = 14.2% pilot region v 14.1% control region) The incidence of prenatal morbidity was almost identical in pilot and control regions (31% v 30.8% respectively) Intrapartum morbidity was significantly lower in the pilot region (34.5% v 48.1%) than in the control region ($p < 0.001$) Newborn morbidity was also lower in the study region (23.6% v 32.7%, $p < 0.001$)	Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (regionalisation) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Siegel et al (1985)[102] (extension of Siegel 1985) USA To evaluate the impact of a regional perinatal care programme on 1 year developmental, neurologic, maternal-infant attachment outcomes and physical health measures	All very low birthweight infants and a stratified random sample for infants weighing >1500g = 447	Assessments were made in infant's homes by 1 of 3 developmental specialists and included a medical and social interview, Receptive-Expressive Emergent Language scale, Bayley Scales of Infant Development, neurologic examination, Maternal Attachment Scale, selected physical health measures (weight, head circumference and hematocrit) November 1978 - October 1979	Neurologic, maternal-infant attachment and physical health measures	As previous study	As previous study General linear model techniques were used to explore the presence of main observer effects and interaction effects between the developmental specialist and characteristics of the mother	Receptive language was significantly better in the study region (mean = 13.5 (SD 1.9) v 12.9 (1.6), $p<0.05$) than in the control region. In infants weighing <1500g mean receptive language scores were not significantly different between study and control regions Maternal attachment was significantly greater in the study region, for both sets of infants (mean = 2.9 (SD 0.8) v 2.2 (0.9), $p<0.05$) in study v control respectively) and in infants weighing < 1500g (mean = 3.0 (0.7) v 2.3 (0.8), $p< 0.05$) No statistically significant differences between study and control regions were observed for Bayley Mental and Motor scores, abnormal neurologic signs and the physical health measures	All survivors and the samples within the study or control regions were compared on maternal education, age, parity, infant race and sex and no statistically significant differences were detected Case-mix adjustment score = 2

Table 23 Neonatal/perinatal maternity care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Linderkamp et al (1991)[103] Germany To examine the effect of regionalisation on perinatal mortality	Live births = 6013 Hospitals = n not given	No information given 1983 - 1989	Mortality	Town with perinatal centre compared with other hospitals in control region	No adjustments No statistical technique reported	Introduction of perinatal centre (in 1986) decreased mortality from 28% to 12% whereas in control region in same time span it decreased from 29% to 23% (1983 - 1989)	Case-mix adjustment score = 0

Table 24: Paediatric intensive care

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Pollack et al (1994)[104] USA To study the association of pediatric intensive care units (ICU) and general hospital characteristics with risk adjusted mortality	Pediatric intensive care patients = 5415 PICU = 16	Prospective cohort - medical records of consecutive cases from a random sample following a national survey of all hospitals with PICUs in 1989 December 1989 - January 1992	Mortality adjusted for physiologic status, diagnosis and other mortality risk factors	ICU volume ranged from 13 patients per month to 63 patients per month:	Age, PRISM score, operative status, emergency or elective status, pre-ICU care area, transport method to hospital, pre-ICU cardiac massage, system of disease and cause of disease Logistic linear regression	Mortality ranged from 2% to 16% ($p<0.0001$) for the 16 different ICUs Variables significantly associated with mortality risk were: PRISM score ($p = 0.0001$), operative status ($p= 0.003$), pre-hospital cardiac massage ($p = 0.0001$), admission to ICU from inpatient unit ($p = 0.009$), presence of oncological disease ($p = 0.005$), chromosomal anomalies ($p = 0.0004$), acute endocrine disease ($p = 0.004$), acute multisystem disease ($p = 0.01$) A non-statistically significant association was found between mortality and volume per month	The reliability of data collection was checked in a random sample of 30 cases per unit Case-mix adjustment score = 3

Table 24: Paediatric intensive care (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Pollack et al (1991)[105] USA To compare outcomes from pediatric intensive care in tertiary and non-tertiary hospitals	Critically ill children under 18 years of age = 463 Hospitals = 74	Prospecitve cohort - medical records were used to compare outcomes in tertiary vs non-tertiary hospitals June 1986 - November 1986	Mortality rates	Tertiary hospitals: n = 3, patients = 128 non-tertiary: n = 71, patients = 335	PRISM Logistic regression	Patients in tertiary hospitals were significantly younger ($p<0.0001$), more likely to be seen in a referring hospitals ED before transfer ($p<0.0001$), to require operations on day of admission ($p<0.05$) and to have received intensive care for a longer duration ($p<0.0001$) There were 24% deaths in tertiary hospitals compared with 6% in non-tertiary hospitals ($p<0.0001$) The distribution of severity of illness differed between tertiary and non-tertiary patients ($p<0.0001$): 55% of non-tertiary patients had mortality risks <1% vs 29% of tertiary patients, 22% of tertiary patients had mortality risks >50% vs 3% of non-tertiary patients When comparing observed mortality with predicted there were 30 deaths and 98 survivors observed in tertiary hospitals compared with 29 deaths and 98 survivors predicted (n..s). There were 20 deaths and 315 survivors observed in non-tertiary and 14 deaths and 321 survivors expected ($p<0.05$) Mortality was more frequent in non-tertiary hospitals than in tertiary hospitals when adjusted for severity of illness (odds ratio 1.1 (low risk), 2.3 (moderate risk) and 8 (high risk) compared with crude odds of 0.7, 2.25 and 2.15 respectively	Case-mix adjustment score= 3

Table 25: Adult intensive care

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Jones & Rowan (1995)[106] UK To examine the association between the volume of patients admitted to intensive care units and mortality	Patients over 15 years of age = 11612 Intensive care units = 26	Data were derived from the Intensive Care Society's UK APACHE II Study	Mortality	The mean monthly volume ranged from 8.3 to 37.7	Socio-demographic, physiological, and diagnostic data were used for the calculation of APACHE II scores (severity of illness score) Patients were also divided into surgical and non-surgical Severity standardised mortality ratios were calculated for each unit	A non-statistically significant relation between volume and severity adjusted mortality rates	19% (8796) were excluded Case-mix adjustment score = 3

Table 26: Prostate

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Wennberg et al (1987)[108] USA To examine differences in death rates between hospitals following prostatectomy	Patients undergoing prostate-ctomy = 4570	Retrospective analysis of Medicare and the Manitoba Health Services Commission claims data (and discharge abstracts) 1975 - 1977	Death within 90 days	Hospital volume: < 40 per year 40 - 90 > 90	Age, medical history, cardiovascular diagnoses, nursing home resident, comorbidity, type of operation (open or transurethral) Logistic regression	There was a non-significant increase in risk adjusted mortality in low volume hospitals (hospitals with more than 90 operations v hospitals with less than 40 = odds ratio 1.26, and hospitals with 40-90 v hospitals with less than 40 = odds ratio 1.66)	335 patients were excluded Case-mix adjustment score = 3

Table 26: Prostate (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing transurethral prostatectomy (TURP) = 41821 Hospitals = 631 Surgeons = 2892	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 65.76, s.d. 357.7) Surgeon volume: median number of procedures per year = 7, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher hospital volume (-0.5867 $p<0.01$) and a lower proportion of patients operated on by low volume surgeons (0.9829, $p<0.05$) were positively related to better outcomes	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score =2

Table 26: Prostate (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Luft et al (1979)[7] USA To examine the relationship between volume and mortality for a number of procedures	Patients undergoing transurethral resection = 86714 Hospitals = 1217	Retrospective analysis of discharge abstract data from the Commission on Professional and hospital Activities (CPHA) 1974 and 1975	In-hospital mortality	Volume was defined as: > 200 = 1162 hospitals, 71964 patients > 200 = 55 hospitals, 14750 patients	Age, sex, race and number of secondary diagnoses Expected death rates were calculated to correct for differences in case-mix and actual and expected death rates were plotted for patients in each volume category	Mortality decreased with increasing volume (results presented graphically)	Differences between hospitals were not tested statistically Case-mix adjustment score =2

Table 26: Prostate (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing prostate-ctomy = 58083 Hospitals = 756	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 113 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.076$, $p<0.01$)	Case-mix adjustment score = 2

Table 26: Prostate (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Roos et al (1986)[71] Canada To examine patient, surgeon and hospital characteristics associated with postdischarge complications	Patients undergoing prostatec-tomy = 4232 Hospitals (n not given)	Retrospective analysis of Manitoba's health insurance database (hospital and medical claims information) 1974 - 1976 2-year before operation and 2 - year after operation histories were obtained	Post-operative comp-lications, readmission Two specialists were given histories based on claims data who indepen-dently judged whether or not the re-admissions were due to complica-tions	Hospital volume per year: > 100 < 100 Surgeon volume: > 50 < 50	Age, comorbidity, prior history, patient residence, hospital location Multiple regression	Hospital or surgeon volume were not found to be significantly associated with the probability of complications	31% of patients were excluded and 8% were lost to follow-up leaving 2721 cases for analysis Case-mix adjustment score = 2

Table 26: Prostate (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Roos & Ramsey (1987)[107] Canada To provide quantitative estimates of the short-term and long-term risks associated with prostate surgery	Patients undergoing prostatec-tomy for non-malignant conditions = 2699 Hospitals in Manitoba (n not given) Surgeons (n not given)	Retrospective analysis of Manitoba insurance claims data 1974 - 1976	90-day mortality, repeat prostatectomy (within 8 years post-operative), stricture dilations post-operatively and cystoscopy post-operatively	Surgeon volume defined as: < 85 = 1067 patients (n performed in 1974 only) > 85 = 1555 patients (n performed in 1974 only)	Age, pre-existing cardiovascular disease and type of case Logistic regression was used, where patient factors and surgeon experience were entered as independent variables	Patients treated by surgeons who performed 85 procedures or more per year were at greater risk for stricture dilations (OR1.6, $p<0.001$) and for cystoscopy post-operatively (OR 1.6, $p<0.001$) than were patients treated by lower volume surgeons Age was a strong predictor of mortality: men aged 75 and over were more likely to die within 90-days post-operatively than younger men (OR 5.2, $p<0.001$) Prior hospitalisation for cardiovascular disease was a significant predictor of mortality (OR 3.0, $p<0.001$) as was surgery within 4 days (OR 0.4, $p<0.01$), which was also a predictor of stricture dilations post-operatively (OR 1.7, $p<0.001$)	1482 eligible cases were excluded from the study due to being not covered by the insurance plan, disagreement over the type of procedure carried out, diagnosis of malignancy and those undergoing radical and perineal procedures Case-mix adjustment score =2 Same data source as Roos (1986) but excluded any malignant conditions

Table 26: Prostate (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Riley & Lubitz (1985)[33] USA To examine the relation between surgical volume and mortality for eight procedures in aged Medicare beneficiaries	Aged Medicare patients undergoing transurethral resection of the prostate (TURP) = 55742 Short stay hospital = n not given	Retrospective analysis of the Medicare Statistical System of the health care Financing Administration 1979 - 1980	In-hospital mortality and 60 day mortality	< 9, 15687 patients 9 - 14, 13951 patients 15 - 22 12213 patients > 22, 13891 patients	Age and sex Multiple regression analysis	Age was significantly associated with mortality (p<0.001) High surgical volume was significantly associated with lower in-hospital mortality (coefficient = -0.22, p<0.001) and 60 day mortality (coefficient = -0.08, p<0.01)	This was a 20% probability sample Patients with cancer included. Case-mix adjustment score = 1

Table 27: Kidney/urinary tract infection and urology

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with kidney/urinary tract infection = 2530 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 99.48 (s.d. 44.43) Physician volume = 4.26 (s.d. 4.56)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals or physicians There were no statistically significant differences in LOS between high and low volume hospitals or physicians	Case-mix adjustment score = 2

Table 27 Kidney/urinary tract infection and urology (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Munoz et al (1990)[110] USA To examine the relationship between the volume of urologic procedures by an individual urologist and outcome	Patients undergoing urology operations = 2980 1 Medical centre in New York	Retrospective analysis of patient records in one medical centre January 1985 - December 1987	Mortality	Surgeon volume (during the 3 year study period): 5 patients or less 8 patients or more	Age, emergency or non-emergency, DRG weight index, and number of procedures and diagnoses were compared by t-tests and ANOVA	Non-emergency patients of high volume physicians had a significantly lower mortality rate than patients of low volume physicians (0.5% v 4.5%, p<0.01) however patients of low volume physicians also had a significantly higher number of diagnosis codes (2.34 v 1.92, p<0.01) For emergency patients there were no statistically significant differences in mortality or in the number of diagnosis codes	Case-mix adjustment score = 2

Table 28: Hysterectomy

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hughes et al (1987)[35] USA To analyse the relationship between the proportion of patients operated on by low volume surgeons, hospital volume and patient outcomes for several procedures	Patients undergoing hysterec-tomy = 105550 Hospitals = 736 Surgeons = 8027	Retrospective analysis of discharge abstracts of the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1982	In-hospital mortality and length of stay (LOS) = poor outcome rate	Hospital volume: (mean = 142.44, s.d. 157.4) Surgeon volume: median number of procedures per year = 8, used to differentiate less experienced surgeons	A risk factor matrix was developed based on age, sex, type of procedure or diagnosis and secondary diagnoses Actual and expected outcomes were calculated and converted to a Z score which was the dependent variable in the regression analysis	Higher (log) hospital volume (-0.7998, $p<0.01$) and a lower proportion of patients operated on by low volume surgeons (0.4663) were positively related to better outcomes (n.s.)	Patient records were excluded if data were missing on age, sex, discharge status or LOS (number excluded not given) Case-mix adjustment score = 2

Table 28: Hysterectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients undergoing hysterec-tomy = 180464 Hospitals = 915	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 784 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.132$, $p<0.01$)	Case-mix adjustment score = 2

Table 28: Hysterectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Roos et al (1986)[71] Canada To examine patient, surgeon and hospital characteristics associated with postdischarge complications	Patients undergoing hyster-ectomy = 6609 Hospitals = (not stated)	Retrospective analysis of Manitoba's health insurance database (hospital and medical claims information) 1974 - 1976 2-year before operation and 2 - year after operation histories were obtained	Post-operative complica-tions, readmission Two specialists were given histories based on claims data who indepen-dently judged whether or not the re-admissions were due to complica-tions	Hospital volume per year: > 100 < 100 Surgeon volume: > 20 < 20	Age, comorbidity, prior history, patient residence, hospital location Multiple regression	Surgical volume both among hospitals and surgeons was a significant predictor of the probability of complications. Lower volume hospitals (odds ratio 1.44, $p = 0.03$) and surgeons (1.37, $p = 0.03$) had readmission rates for complications about 40% higher than their counterparts performing more operations In the multiple regression analysis hospital and physician volume were found not to be significant predictors of readmissions	23% of patients were excluded and 15% were lost to follow-up leaving 4353 cases for analysis Case-mix adjustment score = 2

Table 28: Hysterectomy (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1981)[70] USA To examine the relation between rates of postoperative wound infection and volume of surgery	Women undergoing abdominal hysterectomy = 5117 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: < 100 beds 100 - 300 > 300 specified within each size: 369 1532 3216 (respectively)	No patient risk factors were specified Logistic regression	Volume was found to be a significant predictor of postoperative wound infection (chi-square = 203.45, p<0.0001)	Case-mix adjustment score = 0

Table 29: Caesarean section

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1981)[70] USA To examine the relation between rates of postoperative wound infection and volume of surgery	Women undergoing caesarean section = 3478 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: <100 beds 100-300 >300 and volume was specified within each size: 218 1411 1849 (respectively)	No patient risk factors were specified Logistic regression	Volume was not found to be a significant predictor of postoperative wound infection (chi-square = 0.52 p= 0.4)	Case-mix adjustment score = 0

Table 30: Trauma care

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Nicholl et al (1995)[123] UK To measure the effectiveness of the experimental trauma system set up in the North West Midlands	Patients with major trauma = 2646 Trauma system Non-trauma system hospitals	Prospective comparative study using inquest registers of coroners, accident and emergency registers, basic injury data, inpatient notes, ambulance service patient report forms and hospital PAS system 1990 - 1993	Overall mortality, avoidable deaths (in a random sub-sample of 402 patients) and residual disability and residual disability in a random sample of survivors Follow-up was for 6 months post incident	Experimental trauma system (n = 1143 patients) compared with 2 comparator regions (n = 1503 patients) (without trauma systems)	Patients were divided into 2 age groups (< 55 and > 55) and divided by blunt or penetrating injuries, injury severity (ISS), and revised Trauma Score (RTS) were calculated Expected deaths were compared with observed deaths using age, ISS and RTS scores	In the trauma system overall mortality was 43% and in the comparison regions was 45%, crude mortality was 37.2/100 and 36.0/100 respectively Mortality rates standardised for severity variables did not differ significantly between the 2 regions. The estimated change in probability of dying in trauma system compared to the comparison regions was 1% per year (95% CI: 4%, 2%) There were no significant differences in morbidity outcomes between trauma system and comparison regions, although small changes were seen on the OPCS disability scale in favour of the trauma system No statistically significant difference in mortality from major trauma between high and low volume A+E departments with volumes ranging from <10 per year to >90 per year in 3 regions (ref).	Case-mix adjustment score = 3

Table 30: Trauma care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Waddell et al (1991)[131] Canada To compare the results of one Metropolitan Canadian trauma hospital with the broad American experience	Trauma victims (of which 89% were blunt trauma) = 752 1 Toronto tertiary care unit (900 beds)	Retrospective analysis of the unit's prospectively managed database of the outcomes of trauma victims which was then compared with outcomes from the Multiple Trauma Outcome study 1986 - 1989	Probability of survival, calculated using TRISS methodology	Toronto tertiary care unit (over the 4 year study period the number of patients averaged 15 per month) compared with American trauma care	Severity of injury (assessed using Injury Severity Score [ISS], Trauma Score [TS] and Revised Trauma Score [RTS]) The Z statistic was used to compare survival outcomes in the Toronto centre with those of the Multiple Trauma Outcome Study (MTOS)	The overall mortality rate was 15.8% There were no statistically significant differences between the actual and predicted death rate for the 4 year period or for any individual year (range: -1.5 - 1.26) Comparing outcomes in 1986 with outcomes in 1989 (in patients with similar TS and ISS scores) there were no statistically significant differences in outcomes despite a doubling in volume levels	Case-mix adjustment score = 3

Table 30: Trauma care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Konvolinka et al (1995)[119] USA To measure the relationships between patient volume and outcome in Pennsylvania's trauma centres	Trauma patients who were severely and urgently injured = 12037 Trauma centres = 24 (11 Level I, 11 Level II and 2 paediatric)	Retrospective analysis of Pennsylvania's Trauma Outcome Study (PTOS) data (data from all Pennsylvania's accredited trauma centres) 1988 - 1989	Severity controlled measure of survival/ death (W)	Annual trauma centre volume for severely injured patients ranged from 60 to 844 The number of trauma surgeons ranged from 3 to 10	Injury severity was calculated using the Abbreviated Injury Scale (blunt and penetrating injuries were separated) Stepwise linear regression was used to select independent variables for the predictive model Correlation coefficient (R^2) (% of variation in outcome explained by the regression) measured the model's fit	The annual number of seriously injured patients per surgeon was the single most important variable for predicting survival - the regression explained 36% of the variance (W = -0.3312 + 0.0200). W increased by 1 for each increment of 50 patients per year For adult blunt injured patients the n of seriously blunt injured patients per surgeon was a significant predictor of survival (W = -0.3638 + 0.0248) the equation explained 6% of the variance in W. W increased by 1 for each increase of 40 patients with serious blunt injury per year Volume was not a significant contributor to prediction of survival for adult patients with penetrating injuries or for paediatric patients Institutional volume was not significant in any of the analyses	Case-mix adjustment score = 2

Table 30: Trauma care (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Smith et al (1990)[128] USA To analyse the impact of volume on mortality for seriously injured trauma patients in a large urban system	Seriously injured trauma patients = 1643 Level I trauma centres in Chicago = 7	Retrospective analysis of patient data collected by the Chicago Department of Health 1987 - 1988	In-hospital mortality	Trauma centres were classified into: High volume: > 200 (seriously injured patients over the 22 month study period), 1348 patients, 4 centres Low volume: < 140 (seriously injured patients over the 22 month study period), 295 patients, 3 centres	The code '99' [immediately life-threatening] or the TS<13 [Champion Trauma Score of 12 or less] were used to classify patients according to severity Pearson correlation coefficients, categorical analysis and Tobit analysis was used (to construct a model to predict mortality)	A significant inverse relationship was found between volume and unadjusted mortality ($r = -0.90$, $p = 0.005$), when mortality was adjusted for patient severity this relationship was no longer statistically significant ($r = -0.64$, $p = 0.11$) Mortality rates for '99' and TS <13 patients were significantly different: 13.8% v 23.2% respectively and low volume centres had a high % of TS < 13 patients Mortality odds ratios for patients taken to low volume trauma centres v high volume trauma centres were: unadjusted OR 1.71 (95% CI: 1.29 - 2.27, $p = 0.001$) and adjusted OR 1.3 (95% CI: 1.0. - 1.66, $p = 0.04$) indicating that risk of death was 30% greater at low volume centres Tobit analysis showed the relationship between volume and mortality to be significant, accounting for around 30 - 40% of the observed variation in mortality rates	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified)(cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Guirguis et al (1990)[115] Canada To compare trauma care at 2 Canadian hospitals with the national standards reported in the Major Trauma Outcome Study (MTOS)	Consecutive trauma patients with multiple system injuries = 274 1 designated level I trauma centre and 1 non-designated official trauma centre National standards from MTOS	Retrospective analysis of patients' records April - July 1987 (trauma centre) April 1987 - October 1988 (non-trauma centre)	Survival	1 designated trauma centre (106 cases) and 1 non-designated trauma centre (168 cases) were compared v national standards from the MTOS	Trauma score, revised trauma score and injury severity score (ISS) were calculated TRISS methodology was used to calculate the probability of survival The Z statistic was used to compare survival rates and the M statistic to measure the match of ISS between the 2 hospitals and MTOS (t-test was also used to compare ISS)	There were 10 deaths in the trauma centre, mean trauma score was 14.2 and the ISS was 19.7 There were 35 deaths in the non-trauma centre, mean Revised trauma score was 6.1 and ISS was 26.4 The M statistic for trauma centre patients was 0.92 indicating a good match of ISS with MTOS, and for non-trauma centre patients the M statistic was 0.56 indicating a poor match with MTOS population. Non-trauma centre patients had more severe injuries than the MTOS population ($p<0.001$) Trauma centre patients' Z score was -0.05 v 1.20 for non-trauma centre patients: there were no statistically significant differences in survival between either hospital and MTOS outcomes	Case-mix adjustment score = 3

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hunt et al (1995)[117] Australia To assess the effect of regionalisation on the outcome of patients with severe head injuries within the central Sydney area	Patients with head injuries who had an admission Glasgow Coma Score (GCS) <9, systolic blood pressure (SBP) > 90 and an Injury Severity Score (ISS) > 15 = 88 1 trauma centre in Sydney	Before/after study where data were collected prospectively both before and after regionalisation of trauma services as part of the Hospital Trauma Data Registry Period 1: 18 months prior to January 1992 Period 2: 18 months from January 1992	Mortality, time from injury to decrompres-sion of intracranial mass lesions and transfer numbers	Pre-regionalisation = 50 Post-regionalisation = 38	Patients were stratified according to age, sex, ISS, GCS mode of injury and intracranial pathology Chi-squared test and Fisher's exact test were used to calculate the statistical significance of differences in mortality rates and Mann-Whitney U test was used to assess the significance of differences between times from injury to decompression of mass lesions	Stratification variables were comparable in both groups Overall mortality fell from 21/50 (42% to 10/38 (26%) over the 2 phases of the study (n.s.) In patients having decompression of an intracranial mass lesion mortality was 7/15 (47%) in phase 1 and 3/9 (33%) in phase 2 (n.s.) The median time from injury to decrompression was 2h 47m in phase 1 and 2h 21m in phase 2 (n.s.) There were 11 transfers over the study period and a significant difference ($p<0.01$) in the time taken for primary retrieval patients to undergo craniotomy compared with transferred patients	278 (76%) eligible cases were excluded due to none fulfilment of the entry criteria Case-mix adjustment score = 3

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Tator et al (1995)[130] Canada To examine the effectiveness of an acute spinal cord injury unit (ASCIU)	Patients with acute spinal cord injury (ASCI) = 552 1 ASCIU and 2 hospitals (prior to the establishment of the ASCIU)	Before/after study (where data was collected prospectively for the ASCIU group and retrospectively from medical records for the pre-ASCIU group) 1947 - 1981	Mortality, length of stay and neurological recovery (extended follow-up)	ASCIU = 201 patients (1974 - 1981) Pre-ASCIU = 351 patients admitted to 2 hospitals (1947 - 1973)	Incidence of pre-existing spinal abnormalities, type of accident, distance and time interval from accident to hospital, level and severity of cord injury, level and type of vertebral column injury, incidence of non-spinal injuries (an injury severity score was calculated and total trauma burden was calculated using an anatomical injury severity score [ISS]) Outcomes were compared using $X2^2$ test and t-test, linear and multiple regression	The significant differences reported between the 2 groups were: type of accidents (more motor vehicle and sports accidents in ASCIU group) ($p = 0.001$), the ASCIU group were admitted to hospital sooner ($p < 0.0001$), the total trauma burden was 26.7 in the pre-ASCIU group and 24.8 in the ASCIU group ($p = 0.01$) Mortality: in the pre-ASCIU group unadjusted mortality was 14% v 7.5% in the ASCIU group ($p = 0.02$). Mortality classified by severity score was: Grade 1 = 22% in the pre-ASCIU group v 25% in the ASCIU group and for Grades 2 - 10 in the pre-ASCIU group it was 10% v 7.8% in the ASCUI group LOS: in the pre-ASCIU group mean LOS was 86.9 days v 48.2 days in the ASCIU group ($p<0.001$) Neurological recovery: mean recovery was 13% for the pre-ASCIU group v 28.8% for the ASCIU group ($p<0.001$) Multiple regression showed increasing age and ISS score increased mortality ($p = 0.0001$). There was a shorter length of stay in the ASCIU group ($p = 0.0001$), for the non-ASCIU group LOS was dependent upon ISS, age and sex and LOS was increased for more severe injuries, younger age and male patients. Greater neurological recovery was associated with less severe ASCI and ISS and with being treated in the ASCIU There was no significant correlation between LOS when plotted against year. Total mortality and neurological recovery declined throughout the period, but when each group considered separately the effect disappeared	28 (5%) patients were lost to follow-up: 5.6% in pre-ASCIU and 4% in ASCIU (only applicable for neurological recovery) Case-mix adjustment score = 3

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Cales (1984)[112] USA To investigate the mortality of trauma before and after implementation of a regional trauma system	Patients who had suffered a motor vehicle accident and subsequently died = 118 Prior to regionalisation patients were transported to the closest hospital After regionalistion there were 5 designated trauma facilities	Before/after study design, where patients' pre-hospital record chart and coroners' records were reviewed 1977 - 1978 and 1980 - 1981 The regional trauma system was introduced in 1980	Potential 'salvage-ability' of the patient, as judged by 1 of 6 physician reviewers	58 patients were included prior to regionalisation 60 patients included post-regionalisation	Hospital Trauma Index was used to calculate Injury Severity Scores (ISS) (based on the degree of anatomic and physiological injury) Linear regression was used to analyse actual death rates for the period preceding regionalisation and death rates were projected if there had been no change. The actual and projected death rates post-regionalisation were compared using chi-square, Fisher's exact and Mann Whitney U test	There were no statistically significant differences in the aetiology of deaths before and after regionalisation 29 of the 118 deaths were judged potentially salvageable. Following regionalisation the proportion dropped from 34% (20/58) to 15% (9/60) ($p<0.02$) 7 of the 9 potentially salvageable deaths occurred in 13 patients treated in non-trauma facilities and 2 potentially salvageable deaths occurred in 47 patients treated in trauma facilities ($p<0.0002$) The median ISS rose from 42.5 pre-regionalisation to 52 post-regionalisation ($p<0.03$). Potentially salvageable patients scored lower than non salvageable patients in the first period (median 41.5 v 44.5, $p<0.04$) and in the second period (34 v 54, $p<0.003$) The 1982 death rate for vehicular trauma dropped to 12.37 compared to a projected rate of 15.80 ($p<0.02$)	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Champion et al (1992)[113] USA To examine trauma care in one level I trauma centre over a 6 year period of increasing commitment to trauma care	Consecutive patients with blunt injuries = 2608 1 level I trauma centre	Retrospective analysis of the hospital's trauma registry (included information from patients' charts and autopsy reports) over a 6 year period where a regional trauma system was instituted in 1979 1977 - 1982	Survival	Prior to 1977 there was no trauma service or trauma system, in 1977 a trauma service was implemented and in 1979 a regional trauma care system was instituted which had designated trauma centres n = 1016 pre 1979 n = 1592 from 1979 - 1982	Injury severity (ISS) and age Patients were divided into set 1 (all survivors with hospital length of stay of 1 day or more and all deaths = 1870) and set 2 (all survivors with ISSs greater than 15 and all deaths = 738) Z and W statistics were used to compare patient outcomes by year of admission with the predictions of norms based on the entire study period	Survival, when controlled for severity mix showed a trend of improvement during the 6 year period Data from years 1 and 2 when compared with data from years 5 and 6 showed a significant difference in survival, with an average of 13.44 more survivors per 100 patients treated per year with an ISS greater than 15	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (continued)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Cooper et al (1993)[114] USA To compare trauma care in a state that lacked a well organised trauma system with national norms (from the National Paediatric Trauma Registry [NPTR])	Trauma patients aged 15 years and under = 31332 New York State Hospitals (n not given) Paediatric trauma centres (participating in NPTR] (n not given)	Retrospective analysis (cross sectional) of hospital discharge abstracts (New York State hospitals) and NPT registry 1989 (New York State hospitals) 1985 onwards (NPTR)	Mortality	New York State hospital's outcomes (14234 patients) (no organised trauma system) compared with paediatric trauma centre's outcomes (17098 patients)	Injury Severity Score (ISS) was calculated for each patient Relative risk of mortality between the 2 groups was calculated	Trauma centres treated more children with brain and internal injuries which were more complex and severe than New York State hospitals The crude mortality rates differed between New York State hospitals and trauma centres (e.g. for brain injuries mortality was 8.6% in trauma centres and 1% in New York State hospitals) Survival stratified by ISS was 10 times greater overall in trauma centres for patients with either brain or internal injuries of moderately great severity and skeletal injuries of moderately great severity than in New York State hospitals (e.g. for brain injuries mortality 0.8% in trauma centres and 4.7 in New York hospitals) Mortality rates were similar between trauma centres and New York State hospitals for most diagnoses	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Hill et al (1995)[116] Australia To evaluate the outcome of patients admitted with haemorrhagic shock to a trauma centre	Patients with major injury (defined by a severity score [ISS] of greater than 15) = 308 1 trauma centre in Sydney	Before/after study where data were collected prospectively, both before and after regionalisation of trauma services as part of the Hospital Trauma Data Registry Period 1: July 1990 - December 1991 Period 2: January 1992 - June 1993	Overall mortality and mortality from blood loss (in patients who were in a state of haemorrhagic shock)	Pre-regionalisation: n = 149 Post-regionalisation: n = 159	Patients were stratified according to age, sex, mechanism of injury and ISS Patients were further stratified according to haemorrhagic shock on arrival Chi-square and Fisher's exact test were used to compare differences in mortality rates between the 2 study periods	The median ISS was significantly lower in patients admitted during the second study period as compared with the first (25 v 21, $p = 0.02$) Overall mortality was reduced from 31% to 11% ($p<0.001$) over the 2 phases of the study Mortality from blood loss in the 40 shocked patients fell from 10/25in the first period to 2/15 ($p = 0.07$) in the second period	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Linn et al (1977)[120] USA To examine whether burn patients in hospitals with special care facilities do better than patients in hospitals without such facilities	Burn patients =1297 Florida state hospitals = 73	Retrospective analysis of medical records and interviews with hospital administrators as part of the Florida Burn Survey 1 year study (date not given)	Mortality, complica-tions, length of stay (LOS)	Hospitals with special burn facilities (n = 3) were compared with hospitals without special burn facilities (n = 70)	Age, sex, race, percentage of total burn and percentage of full thickness burn Severity variables were held constant in an analysis using covariant techniques and secondly randomly selected pairs of patients from each group of hospitals were matched by the total % of BSA burned and their outcomes compared	Unadjusted mortality was 13.4% in hospitals with burn care v 4.1% in hospitals without (p<0.001). Complication rates (e.g. septicaemia) were 8.2% v 2.1% (p<0.001) respectively and LOS was 22.9 v 14.6 days (p<0.001) in special v no burn care hospitals After adjustment mortality was 9.8% v 4.9% (p<0.001), complication rates were 7.3% v 2.4% (p<0.001), LOS was 21.9 v 14.8 days (p<0.001) respectively in hospitals with special burn facilities v no special burn facilities	359 out of 1656 patients (22%) were excluded because they were admitted with burns occurring more than 2 days prior to admission, were transferred or had an uncertain date of burn. Comparisons were made between original and reduced sample and no statistically significant differences were found Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Mullins et al (1994)[122] USA To determine if risk of death for hospitalised injured patients changes when an urban trauma system is implemented	Hospitalised patients with at least one discharge diagnosis indicating injury = 70350 Acute care hospitals in Oregon = 18	Hospitals were compared before, during and after implementation of a trauma system using the Oregon Hospital Discharge Index 1984 - 1985 (before) 1986 - 1987 (during) 1990 - 1991 (after)	In-hospital mortality	Hospitals were classified into 3 groups: level I trauma centres (the 2 level I trauma centres that functioned as such since 1986) group 2 (applicant trauma centres - the 3 hospitals that were trauma centres during 1986 to 1988 group 3 (nontrauma centres - the 13 hospitals that did not receive trauma patients)	Age (children, adults, elderly), sex, comorbidity (chronic obstructive pulmonary disease, congenital coagulopathy, chronic liver disease and cirrhosis and ischaemic heart disease) and a computer programme was used for injury severity scoring, classifying into minor, serious and multiple injuries Multiple regression	Over the time period there was an increase in the number of patients with severe injuries and the number who died Adjusted risk of death for patients admitted to level I trauma centres declined after the trauma centre was established but did not in other hospitals (OR = 0.65, 95% CI: 0.51 - 0.81) For all patients classified as seriously injured adjusted risk of death was lower in 1990 and 1991 compared with the 2 earlier periods (OR = 0.83, 95% CI: 0.70 - 0.99) For children the adjusted risk of death was lower in 1990 and 1991 compared to earlier periods (OR = 0.47, 95% CI: 0.26 - 0.84)	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Sampalis et al (1992)[124] Canada To compare the mortality in a sample of severely injured patients from a large metropolitan area without regionalised trauma care to that predicted according to the Major Trauma Outcome study (MTOS)	Patients with severe trauma = 355 (this is a random sample of the total sample of 8007 patients in the study period) Acute care hospitals in Montreal = 26	Retrospective analysis of patient records April 1987 - March 1988	Mortality	The outcomes of patients in Montreal with no regionalised trauma service were compared with outcomes in the MTOS study Montreal hospitals were also classified as level I compatible (n = 3), level II compatible (n = 8) and level III compatible (n = 15)	Injury severity scores were calculated Z statistic and standardised mortality rates (SMRs) as indicators of excess mortality were calculated Multivariate logistic regression was used to provide adjusted estimates of odds ratios for death in excess of that expected by standardisation to the MTOS	A total of 70 deaths occurred in the sample of 355. According to ISS, age and prevalence of penetrating trauma, a total of 38.6 deaths were predicted from the probabilities of the MTOS. A significant overall excess mortality was observed ($Z = 6.77$, $SMR = 1.81$, $p<0.05$), which gives an excess mortality of 81% Being treated at a level I or II hospital was associated with a decreased adjusted standardised odds of dying when compared with patients treated at level III compatible hospitals (OR 29.9, 95% CI: 2.7 - 33.3), although it was not statistically significant	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Sampalis et al (1995)[125] Canada To evaluate the impact of trauma centre designation on mortality	Patients with major trauma = 446 3 Montreal area hospitals 2 Montreal area level I trauma centres	Before/after study where outcomes were compared before and after designation of trauma centres using chart review performed by a medical archivist and entered into a database for the Quebec Provincial Trauma Registry 1987 and 1993	Mortality	Before designation: 3 hospitals, 158 patients After designation: 2 trauma centres, 288 patients	The two groups were compared in terms of age, mechanism of injury, body regions injured and Injury Severity (ISS) using bivariate analyses. Logistic regression was used to produce odds ratios for mortality that were adjusted for the above variables	There were significant differences between the 2 groups of patients in: age, which was lower in 1987 than in 1993 (30.9 v 40 years, p<0.001), more patients in 1987 suffered stabbing injuries (18% v10%, p = 0.02) and in 1993 more patients suffered falls injuries (25% v 13%, p = 0.003). There were significantly lower rates of head or neck injury (40% v 55%, p = 0.003) and extremity (47% v 76%, p = 0.001) in the 1987 group and lower rates of abdominal (38% v 15%, p = 0.0001) and external (54% v 15%, p = 0.0001) injuries in the 1993 group. There were no statistically significant differences in ISS between groups There was a significant difference in mortality (20% v 10%, p = 0.0006). the crude OR was 2.10 [95% CI: 1.22 - 3.62] indicating a two-fold increase in risk of dying in 1987 Adjusted OR for mild injuries was not statistically different. For patients with ISS between 16 - 25, OR 2.79 [0.96 - 8.11, p = 0.05], indicating a three-fold increase in the risk of dying in 1987 and in patients with an ISS of 26 - 49 OR 2.77 [1.08 - 3.13, p = 0.03] indicating a three-fold increase in risk of death Logistic regression produced estimates of adjusted mortality (associated with 1987 group) of 3.25 [1.62 - 6.52, p = 0.009]	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Schiowitz (1990)[126] USA To examine the requisite annual caseload for individual trauma surgeons	Trauma admissions = 85 1 level II institution in a voluntary regional trauma system, in Pennsylvania	Consecutive trauma admissions over a 3 year period were identified and outcomes were compared with reports from other trauma facilities Years of study not given	Mortality and number of preventable deaths	1 level II trauma facility (average of 28 trauma patients annually) was compared with other trauma facilities in 2 comparison groups	Age, trauma score, and severity of injury (ISS) TRISS was used to determine the probability of survival and a blind review of deaths occurring was carried out to determine the number preventable	Over the 3 year period the mortality rate was 7% (6 patients), while mortality for the comparisons were 10.5% (Yale study) and 10% (Major trauma outcome study) ISS was 18.6% overall compared with 22.5% in Yale comparison 5 of the 6 patients who died had a probability of survival of < 50% and might be statistically predicted to die. In the blinded review no death was judged preventable	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Shackford et al (1986)[127] USA To compare care before and after implementation of a trauma system	Major trauma victims = 1707 Acute care hospitals in San Diego = 30 Trauma centres = 5	Before/after study comparing outcomes pre-regionalisation and post regionalisation of trauma services Data collected via review of medical records, autopsy reports and trauma Registry form March - June 1982 August - December 1984	Mortality classified as preventable or non-preventable	Before regionalisation = 30 hospitals, 341 patients After regionalisation = 5 trauma centres (unclear whether the 5 centres were in the original sample), 1366 patients	Trauma score and Injury Severity Score (ISS) were compared in before and after patients Data were compared using the Chi-square analysis or t-test	Overall mortality reduced from 26.4% to 8.2% after implementation of the trauma system and preventable deaths were reduced from 13.6% to 2.7% In 1982 21.6% of deaths were felt to be preventable v 9.8% in 1984 ($p<0.01$) and 78.5% v 90.2% deaths were felt to be non-preventable The mean trauma score in 1982 was 11.2% and in 1984 was 14.1%, therefore a sub-group of the 1984 patients were selected which contained all patients with an ISS of > 13 ($n = 576$) and compared with the 1982 group. In 1982, 26.4% died v 19.4% in 1984(sub-group)	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Wright et al (1984)[134] Canada To evaluate the efficacy of the Ontario Emergency Service System in trauma management	Trauma victims who died 155 Regional trauma unit Referring hospitals = 36	Retrospective analysis coroner's office records 1977 - 1980	Mortality	Regional trauma unit (52 cases) compared with 36 referring hospitals (103 cases)	Age, sex and nature of injury Post-mortem review was used where an injury severity score (ISS) was assigned and where ISS refers to a given ISS above which 50% of a given age group would be expected to succumb to their injury	The average ISS for referring hospitals was 33 and for the trauma unit was 53 (more severe injuries in the trauma unit) Of trauma unit injuries 5.8% had an ISS below their LD_{50} with 53% below their LD_{50} in referring hospitals	Case-mix adjustment score = 2

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Anderson et al (1988)[111] UK To identify deaths in hospital that might have been preventable had a fully staffed and equipped American style trauma centre been available	Consecutive deaths after injury in 11 coroner's districts in England and Wales = 1000 Hospitals = 28	Retrospective analysis of pathologists records 1986 (11 - 39 months onwards in order to accrue the necessary number of deaths)	Preventable deaths (as assessed by 4 independent assessors) if patient treated at a fully staffed and equipped American style trauma centre)	Non-trauma facilities (UK) compared with American trauma centres	An injury severity score was calculated for each case Central nervous system deaths were analysed separately to other deaths No statistical analysis reported	514 patients were admitted to hospital and subsequently died (486 died at the scene or on arrival). In 335 (65%) central nervous system injury was the cause of death Of 514 deaths 102 (20%) were judged by all assessors to be preventable and 68 (13%) by 3 out of 4 assessors (total = 170) 69 of the preventable deaths were in teaching hospitals and 101 in no-teaching hospitals	Case-mix adjustment score = 1

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Jos et al (1989)[118] Netherlands To evaluate trauma care in the Netherlands	Severely injured trauma patients sustaining blunt or penetrating injuries = 148 Trauma hospitals in the Netherlands: 12	Retrospective analysis of hospital records and autopsy records October 1984 - October 1985	Potentially preventable, preventable or non-preventable deaths and management errors	Hospitals were compared where bed size was: 376 beds, n = 5 and 13 patients 709 beds, n = 3 and 50 patients 860 beds and university, n = 4 and 85 patients	Age, sex, pre-existent disease, mechanism of injury, time of accident, time of admission, diagnoses, airway and ventilation management, volume replacement, diagnostic procedures and therapeutic procedures Mann Whitney U test and Chi-square test were used to compare outcomes between hospital categories	Significantly higher percentage of management errors in small general hospitals (72%) than in large general (29%) and in university hospitals (34%) A (possibly or definitely) preventable cause of death was identified in 25% of all fatalities and there was a significantly higher preventable death rate in small hospitals (48%) than in large (14%) and university (19%) hospitals	Case-mix adjustment score = 1

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Lowe et al (1983)[121] USA To evaluate care of motor vehicle accident victims	Patients with severe motor vehicle accident injuries = 659 Oregon hospitals = 22	Retrospective analysis of ambulance, hospital, emergency department records and autopsy reports 1979	Mortality and whether the outcome was appropriate	Hospitals were classified according to size: < 200 beds = 14 hospitals, 268 patients > 200 beds = 8 hospitals, 391 patients	Injury severity A trauma review panel evaluated whether the outcome seemed appropriate for the severity of injury (no statistical information given)	The overall mortality rate was 20.5%; in small hospitals it was 26% and in large hospitals 17% ($p<0.01$). The level of injury was not statistically different between hospitals 25% of fatalities were judged by the trauma review panel to be inappropriate and 16% of patients overall to have inappropriate outcomes	Case-mix adjustment score = 1

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Spain et al (1984)[129] USA To evaluate the effectiveness of the trauma care system in the Hudson Valley Emergency Medical Services region, where there is no designated regional trauma care centre	Trauma patient = 421 Acute care hospitals = 34	Retrospective analysis of autopsy reports 1979 - 1980	Mortality, defined as: preventable, possibly preventable and not preventable as assessed by 5 physician evaluators	Mortality was compared with mortality from San Francisco where there is a designated regional trauma care centre (425 cases)	Injury severity (No statistics reported)	10 deaths in the 132 patients who arrived alive at the hospital were judged by panel members as preventable (all had multiple system injury) The number of preventable deaths in San Francisco was 19	Case-mix adjustment score = 1

Table 30: Trauma care (volume not specified)(cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
West et al (1979)[132] USA To compare the results of trauma care in San Francisco County (SF)with the results of care in Orange County (OC)	Victims of motor vehicle accidents = 182 1 trauma centre (SF) 39 receiving hospitals (OC)	Cross sectional comparison of death certificates, coroner's reports and autopsy data (and hospital records for SF) 1974 (OC) 1974 - 1975 (SF)	Preventable, potentially preventable and non-preventable deaths	Trauma centre (92 patients) compared with 39 receiving hospitals (90 patients)	Injury severity (ISS) was calculated for each death (No statistical information given)	Deaths were classified into non-CNS and CNS Non-CNS: injuries in OC were less severe than in SF(37 v 45, p<0.03) There were 30 deaths in OC and 16 in SF 11 deaths in OC were preventable v 0 preventable deaths in SF, 11 deaths in OC were potentially preventable v 0 deaths in SF, 8 deaths were non-preventable in OC v 30 deaths in SF CNS: there were 60 deaths in OC and 76 deaths in SF Injuries in OC were less severe (38 v 46.5, p<0.01) (information on preventable, potentially preventable and non-preventable deaths not given)	Case-mix adjustment score = 1

Table 30: Trauma care (volume not specified) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
West et al (1983)[133] USA To evaluate the first years experience of a regional system of trauma care and compare trauma survival with and without a system of trauma care	Fatally injured patients = 29 Trauma centres = 5 (number of non-trauma hospitals not given)	Retrospective analysis of autopsy reports comparing trauma centre outcomes with non-trauma centre outcomes The regional system of trauma care (post 1980) was also compared with the trauma care provided pre-1980 June 1980 - June 1981 compared with 1974 and 1978 - 1979	Survival measured as preventable or non-preventable deaths	Trauma centres (n = 23 patients) compared with non-trauma centres (n = 4 patients) (2 patients were transferred to trauma centres)	Age and interval from injury to death was recorded for each patients Fisher's exact test was used to compare survival with and without a system of trauma care	9% of trauma centre deaths were judged to be potentially preventable, 67% of deaths in non-trauma centres were judged to be preventable and the 2 deaths in transferred patients were both judged to be preventable. Percentage of deaths in trauma and non-trauma were statistically significant ($p<0.01$) In 1974, 73% of deaths were preventable and in 1978 - 1979, 71% were preventable Comparing preventable deaths in 1980 with preventable deaths in 1974 and 1978 - 1979 the percent of preventable deaths occurring in trauma centres was reduced to 9% ($p<0.001$).	Case-mix adjustment score = 0

Table 31: AIDS

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Stone et al (1992)[137] USA To determine whether there is an association between mortality and hospital experience for all AIDS related diagnoses	All AIDS related diagnoses = 300 (with a total of 806 hospitalisa-tions) Massachusett shospitals = 40	Retrospective analysis of case records from the Massachusetts AIDS Surveillance Programme (MASP) (population based disease surveillance registry) January 1987 - December 1988	In-hospital mortality and 30-day mortality	Hospitals were ranked according to experience with AIDS patients The top 8 hospitals (20%) were ranked as high experience (HEH): 43 - 229 cases per year, with a total of 504 admissions The remaining 32 hospitals were ranked as low experience (LEH): 1 - 42 cases per year with a total of 302 admissions	The Severity Classification for AIDS Hospitalizations (SCAH) was used which stratifies patients into 3 risk stages based on diagnosis codes The Justice stage assessment (JSA) was also used on a sub-set of 22 hospitals and on 184 patients (521 hospitaliza-tions) which uses lab. and clinical data to claissfy patients into 1 of 3 stages of increasing disease severity Logistic regression	The SCAH was a significant predictor of inpatient and 30-day mortality, as was the JSA (<0.01) Unadjusted inpatient mortality at HEH was 10% and at LEH was 19%. Unadjusted relative risk (RR) of dying at low vs high experience hospitals was 2.16 (95% CI: 1.43 - 3.26) Comparing inpatient mortality for low vs high experience hospitals within strata (including SCAH), 17 out of 29 strata had significantly lower mortality at HEH Unadjusted RR of 30-day mortality from a LEH vs HEH was 1.93 (95% CI: 1.31 - 2.84) Comparing 30-day mortality for low vs high experience hospitals within strata (including SCAH), 14 out of 29 strata had significantly lower mortality at HEH Using logistic regression to predict inpatient mortality, being a LEH was a significant predictor (RR 2.92, 95% CI: 1.37 - 6.22) as was SCAH stage, transmission mode of IV drug use, blood product recipient and heterosexual contact ($p<0.05$). For 30-day mortality, being a LEH was a predictor (RR 2.51, 95% CI: 1.22 - 5.17) as was SCAH stage and age greater than 40 years at diagnosis	Case-mix adjustment score = 3

Table 31: AIDS (cont)

Author, year, country and objectives	Procedures, diagnoses (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Bennett et al (1989)[136] USA To evaluate the experience of 257 patients with AIDS and pneumocystis carinii pneumonia treated at 15 California hospitals	AIDS patients with pneumocystis carinii pneumonia = 257 California hospitals = 15	Retrospective analysis of the Uniform Hospital Discharge Data Set October 1986 - October 1987	In-hospital mortality	Hospitals were classified according to AIDS familiarity: high AIDS familiarity: >30 HIV related discharges per 10000 total hospital discharges (n = 7). Patients = 221 low AIDS familiarity: <30 HIV related discharges (n = 8). Patients = 36	Markers of severity of illness: age, sex, use of hospital intensive care unit, type of admission , presence of concomitant infection or tumour and prior hospitalizations Logistic regression	Significant differences were found in unadjusted mortality rates between high and low familiarity hospitals (12% vs 33%, p = 0.01) Significant differences were also found in concomitant infection/tumour rate (25% vs 42%, p = 0.04) and in the male/female ratio (97% vs 86% male, p = 0.01) in high and low familiarity hospitals respectively The following variables were significant when predicting in-hospital mortality: high familiarity (p <0.01), transfer admission (p = 0.02), intensive care use (p <0.01) and premarker admission (p <0.01) Logistic regression indicates that after adjustment the chances of dying in low familiarity hospitals are 3.6 times higher than at high familiarity hospitals (95% CI 1.9-6.3)	Case-mix adjustment score = 2

Table 32: Cataract surgery

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Schein et al (1994)[139] USA To examine associations between surgical technique, patient and surgeon characteristics and clinical outcomes of cataract surgery	Medicare beneficiaries (aged 50+) undergoing first eye cataract surgery = 772 Ophthalmologists = 75	Prospective cohort study, where ophthalmologists were randomly selected (stratified) from 3 sites, chosen to reflect low, medium and high cataract surgery rates Data on each patient were collected prospectively and included preoperative clinical status, intraoperative techniques, and outcomes July - December 1991	Adverse events and visual acuity (Snellen) at 4-months postoperative	Ophthalmologist volume: Moderate: 51 - 200: n = 30 and 255 patients High: 201 - 399: n = 27 and 290 patients Very high: 400 or more: n = 18 and 227 patients	Age, sex, baseline visual acuity, type of cataract and presence of other eye disease (and any adverse events were included for measurement of visual acuity) Multiple linear or logistic regression	High and very high volume surgeons had a greater rate of adverse events than did moderate volume surgeons (figures not given). The most common adverse event was posterior capsular opacification and the presence of cortical cataract was associated with increased risk of posterior capsular opacification. Higher volume surgeons also had an increased likelihood of posterior capsular opacification (high volume OR 2.54, 95% CI: 1.43 - 4.49 and very high volume OR 3.75, 95% CI: 2.11 - 6.69) Older age (coefficient -1.19 [SE 0.22], $p<0.05$), being female (0.39 [0.14], $p<0.05$), adverse events after surgery (-0.55 [0.16], $p<0.05$) and other eye disease (-0.73 [0.15], $p<0.05$) were all associated with poorer improvement in visual acuity There was no statistically significant association between volume and visual acuity (high volume 0.12 [0.17] and very high volume -0.24 [0.19])	There were no differences between the ophthalmologists selected and those not selected Out of 772 eligible patients data was available for only 717 (93%), therefore 7% are excluded from the study Case-mix adjustment score = 3

Table 33: Cancer (breast)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Sainsbury et al (1995)[145] UK To examine variation in survival between individual surgeons within one regional health authority	Women with invasive breast cancer = 12861 Consultants = 180	Retrospective analysis of a population based cancer registry 1976 - 1992	5 year survival	Surgeon volume per year: < 10 10 - 29 30 - 49 > 50	Age, stage, tumour grade, socioeconomic deprivation Linear regression	The largest differences in survival were due to case-mix (e.g. disease extent - metastases OR 4.46 (4.05 - 4.92) Patients treated by surgeons with caseloads >29 had significantly better survival than patients treated by surgeons with caseloads <10. The difference was reduced slightly when case-mix and type of treatment were included but was still significant (risk ratio 0.85, 0.77 - 0.93)	Case-mix adjustment score = 3
Sloan et al (1986)[17] USA To examine variations in hospital mortality rates	Women with breast cancer undergoing mastectomy not given Hospital (n not given)	Retrospective analysis of discharge abstracts from the Commission on Professional and Hospital Activities 1972 and 1981	In-hospital mortality rate	Mean annual volume per year: Low = 25 (1972) 27 (1981) Medium = 83 (1972) 50 (1981) High = no high volume (1972) 10 (1981)	Age, presence of other diagnosis Multiple regression	A curvilinear relationship between the number of procedures performed and mortality rates; hospitals with low and high mortality rates tended to have low mean volume while those with medium mortality rates had the highest volume	Case-mix adjustment score = 2

Table 33: Cancer (breast) (cont)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Boffetta et al (1993)[140] Italy To examine the association between breast cancer survival and patient characteristics and type of hospital	Women with breast cancer = 5012 Hospitals (n not given)	Retrospective analysis of population based registry 1979 - 1981	Survival at 1, 3 and 5 years post diagnosis	Hospital volume per 2 years: > 500 diagnoses 200 - 499 100 - 199 < 100	Age at diagnosis Proportionate hazard models were applied to obtain multivariate estimates of risk of dying during follow-up	The relative risk of dying in low v high volume hospitals did not differ significantly: RR of death for centres with < 100 patients (1.08, 95% CI: 0.90 - 1.28) 100 - 199 patients (1.15, 95% CI: 0.95 - 1.39) 200 - 499 patients (1.02, 95% CI: 0.84 - 1.24)	5% of eligible patients were excluded. Case-mix adjustment score = 1

Table 33 Cancer (colorectal)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Flood et al (1984)[63] USA To examine the relation between hospital volume and outcomes for selected diagnostic categories	Patients with colon cancer =17872 Acute care hospitals in the United States = 1040	Retrospective analysis of abstract data from the Commission of Professional Hospital Activities (CPHA) and from the Professional Activities Study (PAS) 1972	In-hospital mortality	Patients in hospitals with less than the mean hospital volume were compared with those in hospitals with greater than the mean hospital volume (mean = 17.18)	Age, sex, stage, white blood cell count, systolic and diastolic blood pressure, urine sugar and albumin, hemoglobin, history of operations, oral antidiabetics, insulin, and thyroid drugs given, secondary diagnoses and procedures Logistic regression	Standardised in-hospital mortality ratios were 1.14 and 0.94 in low-volume and high-volume hospitals, respectively ($p<0.05$).	Case-mix adjustment score = 3

Table 33: Cancer (colorectal) (cont)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Kelly & Hellinger (1986)[65] USA To examine the effects on post-surgical mortality rates of the volume of several procedures performed by individual physicians and hospitals	Patients with colorectal cancer = 2612 Hospitals = 116 Physicians = 434	Retrospective analysis of data from the Hospital Cost and Utilisation Project, based on discharge abstract records 1977	In-hospital mortality	Surgeon and hospital volume expressed as a continuous variable Mean hospital volume 50.360 (s.d. 36.4) Mean surgeon volume = 8.416 (s.d. 5.7)	Age, sex, stage and number of diagnoses Logistic regression	No association between mortality and hospital or surgeon volume was detected	Case-mix adjustment score = 3
McCardle & Hole (1991)[172] UK To assess the differences among surgeons in post-operative complications, mortality and survival	Patients undergoing surgery for colorectal cancer = 645 Consultant surgeons = 13	Prospective cohort study 1974 - 1979	Post-operative complications, post-operative mortality (within 30 days) and survival (up to 10 years)	Surgeon volume ranged from 21 to 98 procedures over the study period	Age, admission status, pre-existing cardiac or respiratory disease, presence of local spread of tumour, Dukes stage, degree of differentiation of tumour, metastatic spread Cox's proportional hazards model	Significant case-mix adjusted variation in overall post-operative mortality (8% to 30%), local recurrence (0% to 21%) anstomotic leak (0% to 25%) and survival at 10 years (e.g. (20% to 63%) was found between the 13 surgeons. These variations were not shown to be correlated with volume	Case-mix adjustment =3

Table 33: Cancer (colorectal) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Sagar et al (1996)[150] UK To compare surgeon outcome in patients undergoing colorectal resection	Patients undergoing laparotomy with colorectal resection = 438 (some patients had malignancy present) Surgeons = 5	Cohort study where data were collected prospectively by audit clerks and surgeons Year (not given)	Mortality and morbidity (eg infection, wound, chest, cardiac, renal thrombo-embolic or neurologic complica-tions or anastomotic leak)	Surgeon volume: 44 (surgeon D) 86 (surgeon C) 90 (surgeon B) 108 (surgeon A) 110 (surgeon E)	POSSUM system was used which includes: physiologic assessment (age, cardiac status, respiratory status, blood pressure, pulse, Glasgow coma score, blood urea, serum potassium, hemoglobin, white cell count, electrocardiogram) and operative severity score (operation, number of operations, volume of blood loss, peritoneal contamination, malignancy, timing of operation) Observed outcome was compared with expected outcome	Unadjusted morbidity varied from 13.6 (surgeon D) to 30.6% (surgeon A) and unadjusted mortality varied from 4.5 (surgeon D) to 6.9% (surgeon C). Risk adjusted morbidity or mortality did not differ significantly between surgeons No significant differences in case mix for surgeons, but surgeons A and D performed more emergency cases	Case-mix adjustment score = 3

Table 33: Cancer (colorectal) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients undergoing large bowel resection = 3297 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 136 (s.d = 72.1) Mean physician volume =9.5 (s.d =10.1)	Age, sex and comorbidity	No significant differences in mortality between high and low volume hospitals. The odds of mortality decreased as physician volume increased (-0.225, $p<0.05$) No significant differences in LOS between high and low volume hospitals or physicians	Case-mix adjustment score = 2
Hermanek et al (1995)[151] Germany To examine the importance of prognostic and surgeon factors in treatment of rectum carcinoma	Patients with carcinoma of the rectum = 594 Hospitals = 7 Surgeons = 43	Prospective cohort study (Study Group Colorectal Carcinoma) 1984-1986	5-year survival and locoregional recurrence	Surgeon volume over study period: <15 (n=29 surgeons) >15 (n=14 surgeons)	Tumour stage and site Multiple logistic regression	Locoregional recurrence correlated with 5-year survival (0.73, $p<0.005$). Surgeons with low recurrence rates had higher survival rates Figures for low volume surgeon group were roughly in the middle of the distribution presented for individual high volume surgeons	Case-mix adjustment score = 2 This was 67% sample of the original 887 patients to ensure homogeneity

Table 33: Cancer (colorectal) (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Lothian Borders Health Boards (1995)[152] UK To examine the results of surgery in patients with large bowel cancer	Patients undergoing large bowel cancer resections = 750 Consultant general surgeons =28	Prospective audit 1990 - 1992	Anastomosis that leaked	5 consultant surgeons (cared for 50% of patients) compared with 23 consultant surgeons who cared for the remaining 50% of patients	Patients were divided into emergency, delayed emergency and elective procedures Statistical technique not reported	The 5 consultants were no more likely to achieve an anastomosis than the other 23 consultants (73% v 65%), but when an anastomosis was performed by one of the 5 it was less likely to leak (4.2% v 14%, $p<0.05$)	Case-mix adjustment score = 1

Table 33: Cancer (pancreatic)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Yeo et al (1995)[148] USA To compare outcomes in patients undergoing pancreaticoduo-denectomy	Patients with tumour of pancreas, bile duct, ampulla, duodenum undergoing pancreatico-duodenectomy = 145 1 hospital in Baltimore Surgeons = 5	Patients recruited in a RCT of pancreaticogas-trostomy vs pancreaticoje-junostomy and data were collected prospectively May 1993 - January 1995	Incidence of pancreatic fistula	Surgeon volume over study period: 9 patients 14 17 29 76	Age, sex, race, preoperative history (including jaundice, weight loss, abdominal pain, smoking, prior abdominal surgery, hypertension, alcohol use, peptic ulcer, and diabetes), preoperative laboratory values Patients with different type of cancer or benign tumours (pancreas, bile duct, ampulla, duodenum) were included Logistic regression	Incidence of pancreatic fistula was related to surgeon volume: OR for patients treated by surgeons with 9 patients (11.62, 95% CI: 1.3-1.06) 14 patients (6, 95% CI: 0.9-41.3) 17 patients (12.96, 95% CI: 2.1-78.3) 29 patients (3.83, 95% CI: 0.7-20.8)	Case-mix adjustment score = 3 17% of eligible patients were excluded

Table 33: Cancer (pancreatic) (cont)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Lieberman et al (1995)[142] USA To examine the effect of hospital and surgeon volume on mortality after pancreatic resection for the treatment of pancreatic cancer	Patients with pancreatic cancer = 1972 Hospitals = 184 Surgeons = 748	Retrospective analysis of discharge abstracts from the Statewide Planning and research Cooperative System in New York 1984 - 1991	In-hospital mortality	Surgeon volume over the study period: < 9 patients (low) 9 - 41 (medium) > 41 (high) Hospital volume over the study period: < 10 (minimal) 10-50 (low) 51 - 80 (medium) > 81 (high)	Age, sex, race, admission status, transfer status, number of secondary diagnoses Logistic regression	Surgeon volume: standardised mortality rates were 13%, 9.7% and 6% for low, medium and high volume respectively For hospitals: standardised mortality rates were 18.9%, 11.8%, 12.9% and 5.5% for minimal, low, medium and high volume respectively. Differences across volume categories were statistically significant. However, only hospital volume appeared to be associated with mortality in the regression analysis	Case-mix adjustment score = 2

Table 33: Cancer (malignant teratoma)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Harding et al (1993)[141] UK To examine the contribution of prognostic factors, treatment centre and protocol treatment on survival.	Males with maglignant teratoma = 454 Cancer units = 5	Retrospective analysis of population based cancer registry 1975 - 1989	5 year mortality	The comparison was between the centre recruiting the majority of patients (53%) v other centres	Age, time from first symptom to diagnosis, site and volume of disease, tumour marker concentrations (to classify stage) Stepwise Cox regression	Mortality was lower in patients treated at the centre with higher recruitment; OR: 0.38 (95% CI: 0.23-0.61)	Case-mix adjustment score = 3 3% of eligible patients were excluded

Table 33: Cancer (oesphageal)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Matthews et al (1986)[143] UK To examine whether the relative experience of the surgeon has an effect on the operative mortality or longer term survival after oesophageal resection for carcicoma.	Patients with oesophageal cancer = 1143	Retrospective analysis of population based cancer registry 1957-1976	Operative (within 30 days) mortality and 5 years mortality rate	Surgeon volume: average number of resections per year: 1 or less (120 surgeons, 329 resections in study period) 2 or 3 (7 surgeons, 252 resections) 4 or 5 (0 surgeons) 6 or more (4 surgeons, 538 resections)	Patients were compared on the following variables: age, sex, site, histological type of tumour and node involvement, duration of symptoms, curative or palliative resection, but survival rates were only adjusted for age X^2 and t-test were used	No statistically significant differences were detected in patient risk factors Results were combined for 1, 2 or 3 resections per year There was a 39% operative death rate in patients treated by surgeons performing <3 operations v 22% in those treated by surgeons with higher volume of operations ($p<0.001$) Five year mortality was 89% in patients treated by surgeons performing <3 operations vs 85% in those treated by surgeons with higher volumes ($p<0.05$) After exclusion of operative deaths from the analysis, rates were 82% and 81%, respectively (ns)	Case-mix adjustment score = 3 2% of eligible patients were excluded from the study

Table 33: Cancer (stomach)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Kelly & Hellinger (1986)[65] USA To examine the effects on post-surgical mortality rates of the volume of several procedures performed by individual physicians and hospitals	Patients with stomach cancer = 341 Hospitals = 69 Physicians = 193	Retrospective analysis of data from the Hospital Cost and Utilisation Project, based on discharge abstract records 1977	In-hospital mortality	Surgeon and hospital volume expressed as a continuous variable Mean hospital volume = 9.645 (s.d 7.356) Mean surgeon volume = 2.733 (s.d 2.131)	Age, sex, stage and number of diagnoses Logistic regression	No statistically significant relationship between hospital or surgeon volume and mortality was detected	Case-mix adjustment score = 3

Table 33: Cancer (lung)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Romano & Mark (1992)[144] USA To examine patient and hospital characteristics related to mortality after lung cancer resection	Patients undergoing lung cancer resection = 12439 Acute care hospitals - 389	Retrospective analysis of hospital discharge abstracts from the Office of Statewide Health Planning and Development 1983-1986	In-hospital mortality	Hospital volume per year: < 9 resections 9-16 resections 17-24 resections >24 resections	Age, sex, race, location of the tumour, presence of chronic illnesses (renal failure, alcohol or drug dependence, liver disease, congenital or nutritional coagulophaties), extended procedures Logistic regression	Risk of death was significantly related to volume in patients undergoing lesser resections: 9-16 resections/yr: OR 0.7 (95% CI: 6-1.0) 17-24 resections/yr: OR 0.6 (95% CI: 0.5-0.8) >24 resections/yr: OR 0.6 (95% CI: 0.4-0.8 In patients undergoing penumonectomies, only in high volume hospital (ie, >24 res/yr) mortality was lower: OR 0.8 (95% CI: 0.5-1.2); 17-24 resections/yr: OR 0.8 (95% CI: 0.5-1.3); >24 resections/yr: OR 0.6 (95% CI: 0.4-1.0)	Case-mix adjustment score = 2

Table 33: Cancer (childhood)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Stiller & Draper (1989)[146] UK To examine the survival rates for children treated at centres of different sizes	Children under the age of 15 years with acute lymphoblastic leukaemia = 4070	Retrospective analysis of the UK population based cancer registry 1971 - 1982	5 years mortality rate	Hospital volume: average of at least 6 patients per year average of at least 1 patient per year average of <1 patient per year	Age, sex, white cell count were compared but only age was controlled for X^2 test used	Patients treated at centres with higher volume experienced lower mortality rates (X^2 for trend with 1 degree of freedom = 38.4)	Case-mix adjustment score = 1
Stiller & Lennox (1983)[147] UK To examine the effects on survival of variations in the planned treatment of childhood medulloblastoma	Children undergoing surgery for medulloblastoma = 368 Hospitals (n not given)	Retrospective analysis of the UK population based cancer registry 1971 - 1977	Five year survival rates	Centres were divided into neurosurgical and radiotherapy and then by volume for study period: 1 - 5 6 - 10 11 - 20 over 20	None Statistical technique used was not reported	Overall, 5 year survival was 32% No differences in survival rates were detected between the 4 volume groups for either type of centre	Case-mix adjustment score = 0

Table 33: Cancer (oncologic procedures)

Author, Year, Country and Objectives	Procedure, Diagnosis (n), Hospitals (n)	Design and Data Source and Year(s) of Study	Outcomes Measured	Volume Measure and Cut-Point Used to Define High and Low Volume (n above and below)	Variables Controlled for and Statistical Technique Used	Results	Comments
Munoz et al (1990)[149] USA To examine the relationship between the volume of oncologic procedures by an individual surgeon and outcome	Adult patients undergoing elective oncologic procedures with a principal diagnosis of cancer = 2627 1 medical centre in New York	Retrospective analysis of patient records in one medical centre January 1985 - December 1987	Mortality and hospital length of stay	Surgeon volume (during the 3 year study period): 5 patients or less (n=907 cases and 57 surgeons) 8 patients or more (n=1408 cases and 17 surgeons)	Age, DRG, weight index. T-tests and ANOVA	Mortality for low volume surgeons was 4.3% compared with 1.3% for high volume surgeons (p<0.01) Mean adjusted hospital length of stay for low volume surgeons was 7.58 days and 5.76 for high volume (p<0.01) Low volume surgeons treated more younger patients with a greater number of diagnosis and procedure codes than higher volume surgeons (p<0.01)	Case-mix adjustment score = 2

Table 34: Miscellaneous

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Burns & Wholey (1991)[25] USA To compare the ability of hospital and physician characteristics to explain variations in mortality and length of stay	Patients with dehydration = 3167 General hospitals with 50 or more beds in one Western state (n not given)	Retrospective analysis of patient discharge data, and hospital data from the Annual Guide January - December 1988	Mortality and length of stay (LOS)	Mean hospital volume = 137.48 (s.d.77.36) Physician volume = 4.75 (s.d. 4.64)	Age, sex and comorbidity Multiple regression and ANCOVA were used to model mortality and LOS as a function of patient, hospital and physician characteristics	There were no significant differences in mortality between high and low volume hospitals, but high volume physicians had significantly higher mortality than low volume physicins (coefficient = 0.356, $p<0.01$) There were no statistically significant differences in LOS between high and low volume hospitals, but high volume physicians had significantly longer LOS than low volume physicians (coefficient = 0.110, $p<0.001$)	Case-mix adjustment score =2
Maerki et al (1986)[28] USA To model actual and expected deaths as a function of volume	Patients with cirrhosis = 24228 Hospitals = 913	Retrospective analysis of patient abstracts from the Professional Activity Study of the Commission on Professional and Hospital Activities (CPHA) 1972	Mortality	The low/high volume level was 41 cases per year	Age, sex, multiple diagnoses, admission blood pressure Actual and expected death rates were modelled as a function of hospital volume using regression	Volume was significant indicating improved outcomes at higher volume hospitals ($R^2 = 0.113$, $p<0.01$)	Case-mix adjustment score = 2

Table 34: Miscellaneous (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Munoz et al (1990)[153] USA To analyse the relationship of neurosurgical volume, hospital costs and outcome	Adult neurosurgical patients = 1002 1 academic medical centre in New York	Retrospective analysis of diagnosis group related data January 1985 - December 1987	Mortality and length of stay (LOS)	Low surgeon volume = less than 5 patients during the 3 year period high surgeon volume = more than 8 patients during the 3 year period	Age, DRG weight index, secondary diagnoses and procedures and patients were also divided into emergency v non-emergency admissions T-tests and/or ANOVA were used to compute differences between groups of patients	Non-emergency patients: Patients of low volume surgeons were significantly younger and had more comorbidity ($p<0.01$) than patients of high volume physicians, although there were no significant differences between DRG weight index (2.35 v 2.58 for low v high respectively) Mortality for patients of low volume surgeons was 1.4% v 1.3% for high volume surgeons (n.s.) and LOS was 6.88 days v 4.43 days respectively ($p<0.01$) Emergency patients: Patients of low volume surgeons had significantly more comorbidity ($p<0.01$) than patients of high volume surgeons, but no significant differences in DRG weight index (2.06% v 3.50% for low v high respectively) Mortality for patients of low volume surgeons was 4.2% v 5.8% for high volume surgeons (n.s.)and LOS was 5.72 days v 5.09 days respectively (n.s.)	Case-mix adjustment score = 2

Table 34: Miscellaneous (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Paganini et al (1993)[154] Argentina To study the relationship between structural hospital characteristics, specialisation and mortality and weight gain	Children under 1 year with acute diarrheral disease (ADD) = 3434 Not for profit hospitals in greater Buenos Aires = 14 Wards = 23	Analysis of routine data (no details of whether prospective or retrospective or dates of data)	Mortality rate and average weight gain for the ward	Number of beds/wards/% of patients with ADD (mean = 20.4;sd = 13.7) and number of beds/hospital (mean = 385.3; sd = 177.8) treated as continuous variables with no cut off	Age, nutritional status, severity of illness Multiple regression analysis of relative risk of death and average weight gain	Number of hours of nursing per patient was associated with a reduction in death (RR = 0.45; 95% CI: 0.21 - 0.94 and staff bed ratio an increase in risk (RR = 5.84; 95% CI: 2.34 - 14.54) of death Significant weight gain for wards with higher % of ADD cases per ward (B =0.36, p<0.05) Size of ward or hospital were not related to outcome	Case-mix adjustment score = 2

Table 34: Miscellaneous (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Farber et al (1981)[70] USA To examine the relation between rates of postoperative wound infection and volume of surgery	Patients undergoing laminec-tomy = 2756 Community hospitals = 22	Retrospective analysis of data from the Virginia Statewide Infection Control Programme (prospectively collected) January 1977 - May 1979	Rates of post-operative wound infection	Hospitals were grouped according to size: < 100 beds 100 - 300 > 300 and volume was specified within each size: 77 1021 1658 (respectively)	No patient risk factors were specified Logistic regression was used to examine the association of infection with frequency of an operation	Volume was found to be a significant predictor of postoperative wound infection (chi-square = 3.68 p = 0.05)	Case-mix adjustment score = 0

Table 34: Miscellaneous (cont)

Author, year, country and objectives	Procedure, diagnosis (n), hospitals (n)	Design and data source and year(s) of study	Outcomes measured	Volume measure and cut-point used to define high and low volume (n above and below)	Variables controlled for and statistical technique used	Results	Comments
Sanger & Clyne (1991)[155] UK To assess and compare the safety and efficiency of a surgical unit at a district general hospital and a community hospital for elective surgery	Patients undergoing elective general surgical operations = 730 District general hospital = 1 Community hospital = 1	Retrospective analysis of computerised hospital activity analysis data and monthly audit meeting records for all inpatients in the 2 hospitals April 1987 - March 1988	30-day mortality, morbidity (wound infection rate), duration of stay and patient satisfaction	District general hospital: n = 350 patients Community hospital: n = 380 patients	None reported Comparative analysis of the 2 hospitals	There were no post-operative deaths in the district general hospital and 1 death in the community hospital There were no differences in wound infection rate, or duration of hospital stay between hospitals 3 patients needed urgent transfer to the district general hospital following operations at the community hospital The community hospital was preferred by most patients for routine surgical operations	Case-mix adjustment score = 0

APPENDIX 1 SEARCH STRATEGY

Database: Medline <1992 to February 1996>

Set	Search	Results
1	(centrali#ation or central#ed).tw.	408
2	centralized hospital services/	39
3	(regional#ation or regional#ed).tw.	208
4	(regional adj3 (centre or center or centres or centers)).ti,ab,sh.	497
5	(regional adj3 (service or services or unit or units)).ti,ab,sh.	253
6	(regional adj3 (facility or facilities)).ti,ab,sh.	19
7	(quantity adj2 quality)ti.ab.sh.	497
8	speciali#ation.tw.	674
9	((physician or hospital or surgical or patient or surgeon) adj volume.t	59
10	((surgeon or physician) adj variability).ti,ab,sh.	4
11	(high adj volume).ti,ab,sh.	250
12	(volume adj outcome).ti,ab,sh.	8
13	(selective adj referral).ti,ab,sh.	8
14	(practice adj2 perfect).ti,ab,sh.	18
15	(decentrali#ation or decentral#ed).tw.	275
16	regional health planning/	196
17	regional medical programs/	184
18	exp treatment outcome/	23897
19	exp quality assurance, health care/	12275
20	hospital mortality/	1345
21	mortality/	1719
22	survival rate/	13879
23	exp quality of health care/	309997
24	exp "health status indicators"/	922
25	exp "health facilities"/	31027
26	laboratories, hospital/	341
27	animal/	416681
28	human/	977901
29	27 not (28 and 27)	313294
30	or/2,6,9-10,12-14	149
31	or/1,3-5,7-8,11,15-17	3281
32	or/18-25	333330
33	(specialist adj (centre or centers or center or centres or unit or units)).t	67
34	or/30,33	216
35	31 and 32	1587
36	or/34-35	1763
37	36 not (26 or 29)	1700
38	34 not (26 or 29)	201
39	37 not 38	1499
40	from 39 keep 3,5,7-9,11,14-16,19-20,22-26	16
41	exp laboratories/	1346
42	39 not 41	1485
43	(tissue adj volume).ti,ab,sh.	142
44	42 not 43	1484

Database: Medline <1992 to February 1996>

Set	Search	Results
1	centralized hospital services/	39
2	(regional adj3 (facility or facilities)).ti,ab,sh.	19
3	((physician or hospital or surgical or patient or surgeon) adj volume).t	59
4	((surgeon or physician) adj variability).ti,ab,sh.	4
5	(volume adj outcome).ti,ab,sh.	8
6	(selective adj referral).ti,ab,sh.	8
7	(practice adj2 perfect).ti,ab,sh.	18
8	or/1-7	149
9	((centrali#ation or centrali#ed) adj5 (care or service or services or faci	127
10	((regionali#ation or regionali#ed) adj 5 (care or service or services or f	79
11	(regional adj3 (centre or center or centres or centers)).ti	86
12	(regional adj3 (service or services or unit or units)).ti.	46
13	(speciali#ation adj5 (care or service or services or facilities or system	59
14	(high adj volume adj5 (care or service or services or facilities or syste	20
15	((decentrali#ation or decentrali#ed) adj5 (care or service or services o	92
16	regional health planning/	196
17	regional medical programs/	184
18	or/9-17	828
19	(quantity adj2 quality).ti,ab,sh.	497
20	exp treatment outcome/	23897
21	exp quality assurance, health care/	12275
22	hospital mortality/	1345
23	mortality/	1719
24	survival rate/	13879
25	exp quality of health care/	309997
26	exp "health status indicators"/	922
27	or/19-26	317536
28	laboratories, hospital/	341
29	(blood adj volume).sh.	0
30	exp laboratories/	1346
31	(tissue adj volume).ti,ab,sh.	142
32	or/28-31	1488
33	animal/	416681
34	human/	977901
35	33 not (33 and 34)	313294
36	(18 and 27) not (32 or 35 or 8)	375

File 155:MEDLINE(R) 1966-1996/May W4
File 151:Hlth.Plan&Admin 1975-1995/Nov
File 159:Cancerlit(R) 1963-1996/Feb
File 159: See HELP NEWS 159 for NCI message
File 73:EMBASE 1974-1996/Iss 12
File 35:Dissertation Abstracts Online 1861-1996/Apr

?S CENTRALIZED HOSPITAL SERVICES/DE
S1 516 CENTRALIZED HOSPITAL SERVICES/DE
?s regional (3w)(facility or facilities)

270645 REGIONAL
107987 FACILITY
222560 FACILITIES
S2 435 REGIONAL (3W)(FACILITY OR FACILITIES)
?s (physician or hospital or surgical or patient or surgeon) ()volume$

156705 PHYSICIAN
623020 HOSPITAL
624771 SURGICAL
1270326 PATIENT
38122 SURGEON
0 VOLUMES
S3 0 (PHYSICIAN OR HOSPITAL OR SURGICAL OR PATIENT OR SURGEON)
()VOLUME$
?s (physician or hospital or surgical or patient or surgeon) ()volume?

156705 PHYSICIAN
623020 HOSPITAL
624771 SURGICAL
1270326 PATIENT
38122 SURGEON
555816 VOLUME?
S4 571 (PHYSICIAN OR HOSPITAL OR SURGICAL OR PATIENT OR SURGEON) () VOLUME?
?s (surgeon or physician ()variability

38122 SURGEON
156705 PHYSICIAN
97736 VARIABILITY
S5 20 (SURGEON OR PHYSICIAN(() VARIABILITY
?s volume()outcome

514358 VOLUME
276677 OUTCOME
S6 33 VOLUME()OUTCOME
?s selective()referral

241667 SELECTIVE
55884 REFERRAL

APPENDIX 2 INDIVIDUALS CONTACTED

Dr David Field
Department of Child Health
Leicester Royal Infirmary

Professor John Nichol
Medical Care Research Unit
University of Sheffield Medical School
Sheffield

Dr Kathy Rowan
ICNARC
Tavistock House
Tavistock Square
London

Dr William Tarnow-Mordi
Department of Child Health
Ninewells Hospital
Dundee

Professor David Yates
North Western Injury Research Centre
Hope Hospital
Salford

APPENDIX 3 VOLUME AND OUTCOME DATA EXTRACTION SHEET

Authors

Year

Title

1 Objectives of the study

2 Hospital/physician information

a Type (n) of hospital/centres included

university
general
specialised
other (specify)

Type of physicians

surgeons
specialists (specify)
other (specify)

b Criteria for including hospitals

voluntary
random sample
geographical location
specific database
other (specify)

3 Year(s) during which study was carried out:

4 Patient information

a Type of procedure, diagnosis, condition, treatment, etc: n

5 **Methodological quality of the study**

a Type of study design:

RCT
controlled trial
before/after
prospective cohort
retrospective cohort
case control
other
not clear

b Source of patient information:

clinical records
clinical database (registry)
administrative database
insurance claims forms
other (specify)

c Number (%) of eligible patients included at the start of the study:

Number
%

if exclusions, reasons given

d Adjustment for confounding variables:

none
age/sex
comorbidity
specific prognostic variables (specify)
severity of comorbidity
other (specify)

6 **Volume**

a Volume measure:
continuous
categorical
other (specify)

b Cut-point used to define high and low volume:

hospital
physician

c Rational for cut-point
expressed by authors
based on formal analysis of the data
other (specify)

d Number of hospitals/physicians above and below cut-point

hospital
physicians
above
below

e Number of patients above and below cut-point:

above
below

7 Outcomes measured:

morbidity (specify)
mortality
in hospital
30-day
other (specify)

psycho-social

process (specify)

other (specify)

8 Statistical methods used

9 Results

10 Comments

APPENDIX 4 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 three-point classification scale.

A statistical model was used to assess whether there was any systematic changes in the estimates of the volume effect as the degree of adjustment for patient case-mix was improved. The dependent variable used in the analyses was the proportion of patients who died in low and high volume hospitals in each study. A logistic regression model was used to relate the proportion dead to a covariate indicating whether the hospitals were high or low volume and to a further covariate representing a volume by case-mix adjustment interaction term. This latter covariate was included to assess whether increased case-mix adjustment had any impact on the estimated effect of volume. The models also included a set of dummy variables for each study included in the analysis to adjusted for any observed 'study effects' not related to volume or case-mix.

The model was adapted by further including a volume by year of data collection interaction term substituting for the volume by case-mix term. In this way, any association between year of data collection and volume could be assessed. The null hypothesis being that case-mix adjustment would not improve over time and hence no association between year of data collection and volume effects would be observed.

The statistical models were fitted to data from the nine studies with the high-low volume cut-off point near 200 cases per year. All statistical analyses were performed using GLIM statistical software.[169] Over dispersion (residual heterogeneity) was accounted for by appropriate rescaling of standard errors (in GLIM this is achieved by rescaling standard errors by the square root of the residual deviance divided by the residual degrees of freedom[170]). The significance of the variables was assessed from z-scores calculated as the ratio of the effect sizes to the rescaled standard errors.[171]

REFERENCES

1. Bunker JP, Luft HS, Enthoven A. Should surgery be regionalised? *Surgical Clinics of North America* 1982;62:657-668.

2. Centre for Health Economics. *Review of the Royal College and Specialist Associations recommendations having volume and concentration implications.* York: University of York, 1996.

3. Department of Health. *A policy framework for commissioning cancer services.* London: Department of Health, 1994.

4. Banta D, Bos M . The relation between quantity and quality with coronary artery bypass graft (CABG) surgery. *Health Policy* 1991;18:1-10.

5. American College of Surgeons. Guidelines for minimal standards in cardiac surgery. *American College of Surgeons Bulletin* 1984;January:67-69.

6. American College of Cardiology/American Heart Association Task Force. Guidelines for percutaneous transluminal coronary angioplasty. *JACC* 1988;12:529-545.

7. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalised? The empirical relationship between surgical volume and mortality. *NEJM* 1979;301:1364-1369.

8. Office of Technology Assessment. *The Quality of Medical Care: Information for Consumers.* Washington DC: United States Government Printing Office, 1988.

9. Black N, Johnston A. Volume and outcome in hospital care: evidence, explanations and implications. *Health Services Management Research* 1990;3:108-114.

10. Houghton A. Variation in outcome of surgical procedures. *British Journal of Surgery* 1994;81:653-660.

11. Ministry of Health. *Tertiary services review: the final report.* Wellington, New Zealand: Ministry of Health, 1994

12. Banta DH, Engel GL, Schersten T. Volume and outcome of organ transplantation. *International Journal of Technology Assessment in Health Care* 1992;8:490-505.

13. Stiller CA. Centralised treatment, entry to trials and survival. *Br J Cancer* 1994;70:352-362.

14. Jarhult J. The importance of volume for outcome in cancer surgery - an overview. *European Journal of Surgical Oncology* 1996;22:205-215.

15. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: Practice makes perfect or selective referral patterns? *Health Services Research* 1987;22:157-182.

16. Luft HS. The relation between surgical volume and mortality: an exploration of causal factors and alternative models. *Medical Care* 1980;18:940-959.

17. Sloan FA, Perrin JM, Valvona J. In-hospital mortality of surgical patients: is there an empiric basis for standard setting? *Surgery* 1986;99:446-453.

18. Freeland MS, Hunt SS, Luft HS. Selective contracting for hospital care based on volume, quality and price: prospects, problems and unanswered questions. *Journal of Health Politics, Policy and Law* 1987;12:409-426.

19. 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.

20. Hannan EL, Kilburn H, Bernard H, O'Donnell JF, Lukacik G, Shields EP. Coronary artery bypass surgery: the relationship between inhospital mortality rate and surgical volume after controlling for clinical risk factors. *Medical Care* 1991;29:1094-1107.

21. Rosenfeld 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;77:498-500.

22. Kelly JV, Hellinger FJ. Heart disease and hospital deaths: an empirical study. *Health Services Research* 1987;22:369-395.

23. Farley DE, Ozminkowski RJ. Volume-outcome relationships and inhospital mortality: The effect of changes in volume over time. *Medical Care* 1992;30:77-94.

24. Williams SV, Nash DB, Goldfarb N. Differences in mortality from coronary artery bypass graft surgery at five teaching hospitals. *JAMA* 1991;266:810-815.

25. Burns LR, Wholey DR. The effects of patient, hospital, and physician characteristics on length of stay and mortality. *Medical Care* 1991;29:251-271.

26. Leape LL, Hilborne LH, Park RE, et al. The appropriateness of use of coronary artery bypass graft surgery in New York State. *JAMA* 1993;269:753-760.

27. Hannan EL, Siu AL, Kumar D, Kilburn H, Chassin MR. The decline in coronary artery bypass graft surgery mortality in New York State. *JAMA* 1995;273:209-213.

28. Maerki SC, Luft HS, Hunt SS. Selecting categories of patients for regionalization. *Medical Care* 1986;24:148-158.

29. Showstack JA, Rosenfeld KE, Garnick DW, Luft HS, Schaffazrick RW, Fowles J. Association of volume with outcome of coronary artery bypass graft surgery, Scheduled vs non-scheduled operations. *JAMA* 1987;257:785-789.

30. Johnson AN. The relationship between volume, quality and outcome in hospital care delivery [PhD]. Minnesota: Minnesota, 1988. 159 p.

31. Grumbach K, Anderson GM, Luft HS, Roos LL, Brook R. Regionalization of cardiac surgery in the United States and Canada. *JAMA* 1995;274:1282-1288.

32. 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.

33. Riley G, Lubitz J. Outcomes of surgery among the Medicare aged: surgical volume and mortality. *Health Care Financing Review* 1985;7:37-47.

34. 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.

35. Hughes RG, Hunt SS, Luft HS. Effect of surgeon volume and hospital volume on quality of care in hospitals. *Medical Care* 1987;25:339-358.

36. Clark RE. Outcome as a function of annual coronary artery bypass graft volume. *Ann Thorac Surg* 1996;61:21-26.

37. Shroyer ALW, Marshall G, Warner BA, et al. No continuous relationship between veterans affairs hospital coronary artery bypass grafting surgical volume and operative mortality. *Ann Thorac Surgeons* 1996;61:17-20.

38. Jenkins KJ, Newburger JW, Lock JE, Davis RB, Coffman GA, Iezzoni LI. In-hospital mortality for surgical repair of congenital defects: preliminary observations of variation by hospital caseload. *Pediatrics* 1995;95:323-330.

39. Hosenpud JD, Breen TJ, Edwards EB, Daily OP, Hunsicker LG. The effect of transplant centre volume on cardiac transplant outcome. *JAMA* 1994;271:1844-1849.

40. Laffel GL, Barnett A I, Finkelstein S, Kaye M. The relation between experience and outcome in heart transplantation. *NEJM* 1992;327:1220-1225.

41. Barbash GI, White HD, Modan M, Diaz R, Hampton JR, Heikkila J. Outcome of thrombolytic therapy in relation to hospital size and invasive cardiac services. *Arch Intern Med* 1994;154:2237-2242.

42. Reznik R, Ring I, Fletcher P, Berry G. Mortality from myocardial infarction in different types of hospitals. *BMJ* 1987;294:1121-1125.

43. Shortell SM, LoGerfo JP. Hospital medical staff organisation and quality of care: results for myocardial infarction and appendectomy. *Medical Care* 1981;19:1041-1053.

44. Doherty JG, Dawson F, Kerr F. Permanent pacemaker practice at a Scottish district general hospital between 1987 and 1993. *Br Heart J* 1995;73:475-478.

45. Boles MD. A causal model of hospital volume, structure and process indicators and surgical outcomes [PhD]. Richmond: Virginia Commonwealth, 1994. 291 p.

46. Adams DF, Fraser DB, Abrams HL. The complications of coronary arteriography. *Circulation* 1973;September:609-618.

47. Kelsey SF, Mullin SM, Detre KM, et al. Effect of investigator experience on percutaneous transluminal coronary angioplasty. *Am J Cardiol* 1984;53:56C-64C.

48. Kimmel SE, Berlin JA, Laskey WK. The relationship between coronary angioplasty procedure volume and major complications. *JAMA* 1995;274:1137-1142.

49. Jollis JG, Peterson ED, DeLong ER, et al. The relation between the volume of coronary angioplasty procedures at hospitals treating Medicare beneficiaries and short-term mortality. *N Engl J Med* 1994;331:1625-1629.

50. Phillips KA, Luft HS, Ritchie JL. The association of hospital volumes of percutaneous transluminal coronary angioplasty with adverse outcomes, length of stay and charges in California. *Medical Care* 1995;33:502-514.

51. Ritchie JL, Phillips KA, Luft HS. Coronary angioplasty. Statewide experience in California. *Circulation* 1993;88:2735-2743.

52. Talley JD, Mauldin PD, Leesar MA, Becker ER. A prospective randomised trial of 0.010 versus 0.014 balloon PTCA systems and interventional fellow versus attending physician as primary operator in elective PTCA: Economic, technical and clinical end points. *Journal of Interventional Cardiology* 1995;8:623-632.

53. Shook TL, Sun GW, Burstein S, Eisenhauer AC, Matthews RV. Comparison of percutaneous transluminal coronary angioplasty outcome and hospital costs for low volume and high volume operators. *Am J Cardiol* 1996;77:331-336.

54. Hartz AJ, Kuhn EM, Kayser KL, Pryor DP, Green R, Rimm AA. Assessing providers of coronary revascularization: a method for peer review organizations. *American Journal of Public Health* 1992;82:1631-1640.

55. Edwards WE, Morris JA, Jenkins JM, Bass SM, MacKenzie EJ. Evaluating quality, cost-effective health care. *Ann Surg* 1991;213:433-439.

56. Hertzer NR, Avellone JC, Farrell CJ, et al. The risk of vascular surgery in a metropolitan community. With observations on surgeon experience and hospital size. *Journal of Vascular Surgery* 1984;1:13-21.

57. Kempezinski RF, Brott TG, Labutta RJ. The influence of surgical speciality and caseload on the results of carotid endarterectomy. *Journal of Vascular Surgery* 1986;3:911-916.

58. Segal HE, Rummel L, Wu B. The utility of PRO data on surgical volume: the example of carotid endarterectomy. *QRB* 1993;May:152-157.

59. Ruby ST, Robinson D, Lynch JT, Mark H. Outcome analysis of carotid endarterectomy in Connecticut: the impact of volume and speciality. *Annals of Vascular Surgery* 1996;10:22-26.

60. Slavish LG, Nicholas GG, Gee W. Review of a community hospitals experience with carotid endarterectomy. *Stroke* 1984;15:956-959.

61. Richardson JD, Main KA. Carotid endarterectomy in the elderly population: a statewide experience. *J Vasc Surg* 1989;9:65-73.

62. Amundsen S, Skjaerven R, Trippestad A, Soreide O. Abdominal aortic aneurysms. Is there an association between surgical volume, surgical experience, hospital type and operative mortality? *Acta Chir Scand* 1990;156:323-328.

63. Flood AB, WR S, 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.

64. Katz DJ, Stanley JC, Zelenock GB. Operative mortality rates for intact and ruptured abdominal aortic aneurysms in Michigan: an eleven year statewide experience. *J Vasc Surg* 1994;19:804-817.

65. Kelly JV, Hellinger FJ. Physician and hospital factors associated with mortality of surgical patients. *Medical Care* 1986;24:785-800.

66. Pilcher DB, Davis JH, Ashikage T, et al. Treatment of abdominal aortic aneurysm in an entire state over 7 and a half years. *The American Journal of Surgery* 1980;139:487-494.

67. Veith FJ, Goldsmith J, Leather RP, Hannan EL. The need for quality assurance in vascular surgery. *J Vasc Surg* 1991;13:523-526.

68. Wu wen S, Simunovic M, Williams JI, Johnston KW, Naylor CD. Hospital volume, calendar age and short term outcomes in patients undergoing repair of abdominal aneurysms: the Ontario experience 1988-1992. *Journal of Epidemiology and Community Health* 1996;50:207-213.

69. Ouriel K, Geary K, Green RM, Fiore W, Geary JE, DeWeese JA. Factors determining survival after ruptured aortic aneurysm: the hospital, the surgeon and the patient. *J Vasc Surg* 1990;11:493-496.

70. Farber BF, Kaiser DL, Wenzel RP. Relation between surgical volume and incidence of postoperative wound infection. *NEJM* 1981;305:200-204.

71. Roos LL, Cageorge SM, Roos NP, Danzinger R. Centralization, certification, and monitoring. Readmissions and complications after surgery. *Medical Care* 1986;24:1044-1066.

72. Gordon TA, Burleyson GP, Tielsch JM, Cameron JL. The effects of regionalization on cost and outcome for one general high-risk surgical procedure. *Annals of Surgery* 1995;221:43-49.

73. Benjamin G. Three essays on volume, complications and hospital resource use: the case of knee replacement surgery [PhD]: Indiana University, 1995. 114 p.

74. Fowles J, Bunker JP, Schurman DJ. Hip surgery data yield quality indicators. *Business and Health* 1987;June:44-46.

75. Lavernia CJ, Guzman JF. Relationship of surgical volume to short-term mortality, morbidity and hospital charges in arthroplasty. *The Journal of Arthroplasty* 1995;10:133-140.

76. 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.

77. Field D, Hodges S, Mason E, Burton P. Survival and place of treatment after premature delivery. *Archives of Disease in Childhood* 1990;66:408-411.

78. LeFevre M, Sanner L, Anderson S, Tsutakawa R. The relationship between neonatal mortality and hospital level. *The Journal of Family Practice* 1992;35:259-264.

79. LeFevre ML. Physician volume and obstetric outcome. *Medical Care* 1992;30:866-871.

80. Lumley J. The safety of small maternity hospitals in Victoria 1982-84. *Community Health Studies* 1988;12:386-393.

81. Mayfield JA, Rosenblatt RA, Baldwin LM, Chu J, Logerfo JP. The relation of obstetrical volume and nursery level to perinatal mortality. *American Journal of Public Health* 1990;80:819-823.

82. Williams RL. Measuring the effectiveness of perinatal medical care. *Medical Care* 1979;17:95-110.

83. Hein HA. The quality of perinatal care in small rural hospitals. *JAMA* 1978;240:2070-2072.

84. 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.

85. Becerra JE, Saliceti NP, Smith JC. Evaluation of a regionalized perinatal care system through linked infant birth and death certificates: Puerto Rico 1980-84. *PRHSJ* 1989;8:305-311.

86. Clarke M, Mason ES, MacVicar J, Clayton DG. Evaluating perinatal mortality rates: effects of referral and case mix. *BMJ* 1993;306:824-827.

87. Gortmaker S, Sobol A, Clark C, Walker D, Geronimus A. The survival of very low birth weight infants by level of hospital of birth: a population study of perinatal systems in four states. *Am J Obstet Gynecol* 1985;152:517-124.

88. Grzybowski SCW, Cadesky AS, Hogg WE. Rural obstetrics: a 5year prospective study of the outcomes of all pregnancies in a remote northern community. *Can Med Assoc J* 1991;144:987-996.

89. Horbar JD, McAuliffe TL, Adler SM, et al. Variability in 28 day outcomes for very low birth weight infants: an analysis of 11 neonatal intensive care units. *Pediatrics* 1988;82:554-975.

90. Lallo DD, Bertollini R, Perucci CA, et al. Mortalita e livello delle cure in neonatidi peso molto basso. Uno studio di populazione. *Epidemiologia & Prevenzione* 1992;14:35-39.

91. Paneth N, Kiely JL, Wallenstein S, Marcus M, Pakter J, Susser M. Newborn intensive care and neonatal mortality in low birth weight infants. *NEJM* 1982;307:149-155.

92. Phibbs CS, Bronstein JM, Buxton E, Phibbs RH. The effects of patient volume and level of care at the hospital of birth on neonatal mortality: is competition killing babies? *JAMA* 1996;276:1054-1059.

93. Rosenblatt RA, Reinken J, Shoemack P. Is obstetrics safe in small hospitals? Evidence from New Zealand's regionalised perinatal system. *The Lancet* 1985:429-432.

94. Verloove-Vanhorick PS, Verwey RA, Ebeling MCA, Brand R, Ruys JH. Mortality in very preterm and very low birth weight infants according to place of birth and level of care: results of a national collaborative survey of preterm and very low birth weight infants in the Netherlands. *Pediatrics* 1988;81:404-411.

95. Viisainen K, Gissler M, Hemminki E. Birth outcomes by level of obstetric care in Finland: a catchment area based analysis. *Journal of Epidemiology and Community Health* 1993;48:400-405.

96. Bowes WA. A review of perinatal mortality in Colarado, 1971 to 1978 and its relationship to the regionalization of perinatal services. *Am J Obstet Gynecol* 1981;141:1045-1052.

97. Black DP, Fyfe IM. The safety of obstetric services in small communities in northern Ontario. *Can Med Assoc J* 1984;130:571-578.

98. Hein HA. Evaluation of a rural perinatal care system. *Pediatrics* 1980;66:540-546.

99. Hein HA, Burmeister LF. The effect of ten years of regionalized perinatal health care in Iowa, USA. *Eur J Obstet Gynecol Reprod Biol* 1986;21:33-48.

100. Nicholls G, Upadhyaya V, Gornall P, Buick RG, Corkey JJ. Is specialist centre delivery of gastroschisis beneficial? *Archives of Disease in Childhood* 1993;69:71-73.

101. Siegel E, Gillings D, Campbell S, Guild P. A controlled evaluation of rural regional perinatal care: impact on mortality and morbidity. *AJPH* 1985.

102. Siegel E, Gillings D, Campbell S, Guild P. Controlled evaluation of rural regional perinatal care: developmental and neurologic outcomes at 1 year. *Pediatrics* 1986;77:187-195.

103. Linderkamp O, Stolz W, Bastert G. The effect of regionalization of perinatal medicine on the mortality and long-term problems of very low birth weight infants. *Zent bl Gynakol* 1991;113:1351-1360.

104. Pollack MM, Cuerdon TT, Patel KM, Ruttimann UE, Getson PR, Levetown M. Impact of quality of care factors on pediatric intensive care unit mortality. *JAMA* 1994;272:941-946.

105. Pollack E, Ford-Jones L, Corey M, et al. Use of the pediatric risk of mortality score to predict nosocomial infection in a pediatric intensive care unit. *Critical Care Medicine* 1991;19:160-165.

106. 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* 1995;11:762-769.

107. Roos NP, Ramsey EW. A population based study of prostatectomy: outcomes associated with differing surgical approaches. *The Journal of Urology* 1987;137:1184-1188.

108. Wennberg JE, Roos N SL, Schori A, Jaffe R. Use of claims data systems to evaluate health care outcomes. Mortality and reoperation following prostatectomy. *JAMA* 1987;257:933-936.

109. Heaton J, Williams M, Long A, Dixon P, Brettle A. *Measuring the health outcomes of total hip replacement through the commissioning process.* Outcome Measurement Reviews. No 1. Leeds: UK Clearing House for Health Outcomes, May 1995.

110. Munoz E, Boiardo R, Mulloy K, Goldstein J, Brewster JG, Wise L. Economies of scale, physician volume for urology patients and DRG prospective hospital payment system. *Urology* 1990;36:471-476.

111. Anderson ID, Woodford M, Dombal FT, Irving M. Retrospective study of 1000 deaths from injury in England and Wales. *BMJ* 1988;296:1305-1308.

112. Cales RH. Trauma mortality in Orange County: the effect of implementation of a regional trauma system. *Annals of Emergency Medicine* 1984;13:1-10.

113. Champion HR, Sacco WJ, Copes WS. Improvement in outcome from trauma centre care. *Arch Surg* 1992;127:333-338.

114. Cooper A, Barlow B, DiScala C, String D, Ray K, Mottley L. Efficacy of pediatric trauma care: results of a population based study. *Journal of Pediatric Surgery* 1993;28:299-305.

115. Guirguis EM, Hong C, Liu D, Watters JM, Baillie F, McIntyre RW. Trauma outcome analysis of two Canadian centres using the TRISS method. *The Journal of Trauma* 1990;30:426-429.

116. Hill DA, West RH, Roncal S. Outcome of patients with haemorrhagic shock: an indicator of performance in a trauma centre. *JR Coll Surg Edinb* 1995;40:221-224.

117. Hunt J, Hill D, Besser M, West R, Ronical S. Outcome of patients with neurotrauma: the effect of a regionalized trauma system. *Aust NZ J Surg* 1995;65:83-86.

118. Jos MD, Draaisma TH, Haan AFJ, Goris JA. Preventable trauma deaths in the Netherlands - prospective multicentre study. *The Journal of Trauma* 1989;29:1552-1557.

119. Konvolinka CW, Copes WS, Sacco WJ. Institution and per-surgeon volume versus survival outcome in Pennsylvania's trauma centres. *The American Journal of Surgery* 1995;170:333-340.

120. Linn BS, Stephenson SE, Bergsttesser PR, Smith J. Are burn units the best places to treat burn patients. *Journal of Surgical Research* 1977;23:1-5.

121. Lowe DK, Gately HL, Goss R, Frey CL, Peterson CG. Patterns of death, complication and error in the management of motor vehicle accident victims: implications for a regional system of trauma care. *The Journal of Trauma* 1983;23:503-507.

122. Mullins RJ, Veum-Stone J, Helfand M, et al. Outcome of hospitalized injured patients after institution of a trauma system in an urban area. *JAMA* 1994;271:1919-1924.

123. Nicholl J, Turner J, Dixon S. *The cost-effectiveness of the regional trauma system in the North West Midlands.* Sheffield: University of Sheffield, 1995.

124. Sampalis JS, Lavoie A, Williams JI, Mulder DS, Kalina M. Standardized mortality ratio analysis on a sample of severely injured patients from a large Canadian city without regionalized trauma care. The Journal of Trauma 1992;33:205-212.

125. Sampalis JS, Lavoie A, Boukas S, et al. Trauma center designation: initial impact on trauma related mortality. *The Journal of Trauma: Injury, Infection and Critical Care* 1995;39:232-239.

126. Schiowitz MF. Patient throughput and mortality rate in a trauma service. *Br J Surg* 1990;77:497-498.

127. Shackford SR, Hollingworth-fridlund P, Cooper GF, Eastman AB. The effect of regionalization upon the quality of trauma care as assessed by concurrent audit before and after institution of a trauma system: a preliminary report. *The Journal of Trauma* 1986;26:812-820.

128. Smith RF, Frateschi L, Sloan EP, et al. The impact of volume on outcome in seriously injured trauma patients: two years experience of the Chicago trauma system. *The Journal of Trauma* 1990;30:1066-1077.

129. Spain DM, Fox RI, Marcus A. Evaluation of hospital care in one trauma care system. *American Journal of Public Health* 1984;74:1122-1125.

130. Tator CH, Duncan EG, Edmonds VE, Lapczak LI, Andrews DF. Neurolgical recovery, mortality and length of stay after acute spinal cord injury associated with changes in management. *Paraplegia* 1995;33:254-262.

131. Waddell TK, Kalman PG, Goodman SJL, Girotti MJ. Is outcome worse in a small volume Canadian trauma centre? *The Journal of Trauma* 1991;31:958-961.

132. West JG, Trunkey DD, Lim RC. Systems of trauma care. A study of two counties. *Arch Surg* 1979;114:455-460.

133. West JG, Cales RH, Gazzangia AB. Impact of regionalization. The Orange County experience. *Arch Surg* 1983;118:740-744.

134. Wright CS, McMirtry RY, Pickard J. A postmortem review of trauma mortalities - a comparative study. *The Journal of Trauma* 1984;24:67-68.

135. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. *The Journal of Trauma* 1987;27:370-378.

136. Bennett CL, Garfinkle JB, Greenfield S, et al. The relation between hospital experience and in-hospital mortality for patients with AIDS-related PCP. *JAMA* 1989;261:2975-2979.

137. Stone VE, Seage GR, Hertz T, Epstein AM. The relation between hospital experience and mortality for patients with AIDS. *JAMA* 1992;268:2655-2661.

138. Mathews CW, Ferdon E, Bennett C, Kanouse D. Evaluating institutional performance in AIDS-associated pneumocystis carinii pneumonia: a risk adjustment approach. *J Clin Epidemiol* 1989;42:421-425.

139. Schein OD, Steinberg EP, Javitt JC, et al. Variation in cataract surgery practice and clinical outcomes. *Opthamology* 1994;101:1142-1152.

140. Boffetta P, Merletti F, Winkelmann R, Magnani C, Cappa APM, Terracini B. Survival of breast cancer patients from Piedmont, Italy. *Cancer Causes Control* 1993;4:209-215.

141. Harding MJ, Paul J, Kaye SB. Management of malignant teratoma: does referral to a specialist unit matter? *The Lancet* 1993;341:999-1002.

142. Lieberman MD, Kilburn H, Lindsey M, Brennan MF. Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. *Annals of Surgery* 1995;222:638-645.

143. Matthews HR, Powell DJ, McConkey CC. Effect of surgical experience on the results of resection for oesophageal carcinoma. *Br J Surg* 1986;73:621-623.

144. Romano PS, Mark DH. Patient and hospital characteristics related to in-hospital mortality after lung cancer resection. *Chest* 1992;101:1332-1337.

145. Sainsbury R, Howard B, Rider L, Johnston C, Round C. Influence of clinician workload and patterns of treatment on survival from breast cancer. *The Lancet* 1995;345:1265-1270.

146. Stiller CA, Draper GJ. Treatment centre size, entry to trials, and survival in acute lymphoblastic leukaemia. *Archives of Disease in Childhood* 1989;64:657-661.

147. Stiller CA, Lennox EL. Childhood medulloblastoma in Britain 1971-77: analysis of treatment and survival. *British Journal of Cancer* 1983;48:835-841.

148. Yeo CJ, Cameron JL, Maher MM, et al. A prospective randomized trial of pancreaticogastrostomy versus pancreaticojejunostomy after pancreatico-duodenectomy. *Annals of Surgery* 1995;222:580-592.

149. Munoz E, Mulloy K, Goldstein J, Tenenbaum N. Costs, quality and the volume of surgical oncology procedures. *Arch Surg* 1990;125:360-363.

150. Sagar PM, Hartley MN, MacFie J, Taylor BA, Copeland GP. Comparison of individual surgeon's performance. *Dis Colon Rectum* 1996;39:654-658.

151. Hermanek P, Wiebelt H, Staimmer D, Riedi S. Prognostic factors of rectum carcinoma - experience of the German multicentre study SGCRC. *Tumori* 1995;81 Supplement:60-64.

152. The Consultant Surgeons and Pathologists of the Lothian and Borders Health Boards. Lothian and Borders large bowel cancer project: immediate outcome after surgery. *British Journal of Surgery* 1995;82:888-890.

153. Munoz E, Boiardo R, Mulloy K, Goldstein J, Tenenbaum N, Wise L. Economies of scale, physician volume for neurosurgery patients and the diagnosis related group prospective hospital payment system. *Neurosurgery* 1990;26:156-161.

154. Paganini JM. Quality and efficiency in hospitals. *Boletin de la Oficina Sanitaria Panameri Cana* 1993;115:482-510.

155. Sanger R, Clyne CA. The surgical value of community hospitals: a closer look. *Ann.R.Coll.Surg.Eng.* 1991;73:77-80.

156. Luft HS, Garnick DW, Mark DH, McPhee SJ. *Hospital volume, physician volume and patient outcomes. Assessing the evidence.* Michigan: Health Administration Press, 1990.

157. Staff of the Stanford Centre for Health Care Research. Comparison of hospitals with regard to outcomes of surgery. *Health Services Research* 1976;11:112-127.

158. Choudhry NK, Wright JG, Singer PA. Outcome rates for individual surgeons: concerns about accuracy, completeness and consequences of disclosure. *Surgery* 1994;115:406-408.

159. Rice N, Leyland A. Multilevel models: applications to health data. *J Health Serv Res Policy* 1996;1:154-164.

160. Effective Health Care. Total hip replacement. *Effective Health Care* 1996;2(7):1-12.

161. Fink A, Yano EM, Brook RH. The condition of the literature on differences in hospital mortality. *Medical Care* 1989;27:315-336.

162. Iezzoni LI, Ash A, Coffman GA, Moskowitz MA. Predicting in-hospital mortality: a comparison of severity measurement approaches. *Medical Care* 1992;30:347-359.

163. Fisher ES, Wennberg JE. Administrative data in effectiveness studies: the prostatectomy assessment. in Heithoff KA, Lohr KN, eds. Effectiveness and outcomes in health care. Washington DC: National Academy Press, 1990.

164. Concato DR, Horwitz RI, Feinstein AR, Elmore JG, Schiff SF. Problems of comorbidity in mortality after prostatectomy. *JAMA* 1992;267:1077-1082.

165. Seagroatt V. Mortality after prostatectomy: selection and surgical approach. *Lancet* 1995;346:1521-1524.

166. Davenport RJ, Dennis MS, Warlow CP. Effect of correcting outcome data for case-mix: an example from stroke medicine. *BMJ* 1996;312:1503-1505.

167. Green J, Wintfeld N. Report cards on cardiac surgeons. Assessing New York states' approach. *NEJM* 1995;332:1229-1232.

168. Localio AR, Hamory BH, Sharp TJ, Weaver SL, TenHave TR, Landis JR. Comparing hospital mortality in adult patients with pneumonia. A case study of statistical methods in a managed care programme. *Ann Intern Med* 1995;122:125-132.

169. Baker RJ. *Glim 3.77 reference manual.* Oxford: Numerical Algorithims Group, 1985.

170. Aitkin M, Anderson D, Francis B,.Hinda J. *Statistical modelling in GLIM.* Oxford: Oxford University Press, 1989;213-6.

171. McCullagh P, Nelder JA. *Generalised linear models.* London: Chapman and Hall, 1983.

172. McArdle CS, Hole D. Impact of variability among surgeons on postoperative morbidity and mortality and ultimate survival. *BMJ* 1991;302:1501-1505.