Analysis of Current Research Related to the Impact of Low-Volume Procedures/Surgery and Care on Outcomes of Care

David R. Urbach MD MSc
Thérèse A. Stukel PhD
Ruth Croxford MSc
Nancy L. MacCallum MLIS

Institute for Clinical Evaluative Sciences
2075 Bayview Avenue, G Wing
Toronto, Ontario M4N 3M5
Tel: (416) 480-4055
Fax: (416) 480-6048

ICES Institute for Clinical Evaluative Sciences

Enhancing the effectiveness of health care for Ontarians through research

Table of Contents

List of Tables

List of Figures

Review of Research Relating Volume to Outcome in Health Care

Executive Summary

A large body of research relating volume to outcome in health care has developed over the past few decades. Numerous studies, comprising a variety of procedures, diseases, health outcomes and health care settings, have found that the clinical outcomes of health services are better when patients are treated by physicians or at hospitals that provide a large number of similar services. While the underlying mechanisms of the association between volume and outcome is still subject to debate, there is increasing interest in health policy measures intending to reduce the potential harm caused by health care provided in low volume units, such as regionalization of certain services to high volume centres, and public reporting of surgeon and hospital volumes.

In this report, we reviewed 161 articles that contained 313 analyses of possible volume-outcome associations. Studies evaluated a heterogeneous group of surgical procedures, medical interventions, health care programs, health care settings (e.g., intensive care), medical diagnoses and health conditions, and diagnostic testing (e.g., mammography). The most common categories were gastrointestinal operations (24.1%), vascular procedures (18.7%), diagnoses and other health conditions (13.9%), and heart operations (11.4%). The single most commonly studied operation was carotid endarterectomy (22 studies), followed by coronary artery bypass graft surgery (20 studies) and percutaneous coronary intervention (17 studies).

Most studies were published after 1995 and typically analyzed data from a number of years. Most (79.5%) were done in the United States, and 8.7% were done in Canada. Most studies were limited to geographic regions such as states or provinces. Many studies done in the US had exclusion criteria that may have affected their interpretation of volume-outcome associations and their generalizability. These include the frequent use of US Medicare data, typically limited to persons aged 65 years or older, or data from Veterans Administration hospitals that predominantly treat male patients.

Almost every study used a cross-sectional design, which does not provide information about causation. Innovative designs, such as quasi-experimental longitudinal designs (which measure the effect of recent volume on outcome as volume changes over time), were infrequently used but are probably the strongest non-experimental method available for providing evidence of a causal association between volume and outcome.

Most studies (72.7%) used administrative data. Of the 44 studies that used clinical data derived from observations on individual patients, only 17 (38.6%) used a prospective design. Information on data quality was not provided in most studies (67.7%), and in studies that described data quality, there were often important limitations such as missing data on the stage of disease in studies of cancer treatment, and the inability to link large numbers of subjects to physicians in studies of physician volume.

A synthesis of 313 analyses done in the 161 studies showed that 68.4% found statistically significant associations between volume and outcome. Many of the studies that did not find a statistically significant association had design flaws, such as low statistical power, that do not preclude the possibility that they simply failed to detect a true association. Hospital volume was found to be significantly associated with outcome in 67.3% of analyses, and physician volume was significantly associated with volume in 63.3% of analyses. In fact, only 4 analyses (1.3%) demonstrated a significant association between higher volume and *poorer* outcome, which is likely a chance finding.

Across all analyses, statistically significant volume-outcome associations were demonstrated for virtually every procedure or condition examined. Procedures that appeared not to have volume-outcome associations in some studies almost always had evidence supporting volume-outcome associations in other studies. In our review, we were not so much struck by the observation that volume-outcome associations were so prevalent, but by the remarkable finding that it was impossible to identify a health service that had been evaluated in more than one study that did *not* have a volume-outcome association.

Our interviews with researchers and opinion leaders identified several research priorities. All of the respondents felt that the descriptive literature regarding the association of volume with heath outcome is mature, and further descriptive research on volume-outcome associations is unlikely to move the research field forward in a way that will help inform health policy. In general, respondents felt that future research on volume and outcome in health care should be carefully aligned with health policy goals, and include: developing better models (that might include, in addition to volume, measures of process of care and outcomes) for the prediction of future outcomes of health services, examining the feasibility and consequences of regionalization strategies, investigating the process-of-care mechanisms underlying volume-outcome associations, extending research into defining appropriate volume thresholds for specific health services, and using experimental or quasi-experimental methods to test the assumption of causation for volume-outcome associations.

Introduction

Since an influential publication by Harold Luft appeared in the *New England Journal of Medicine* in 1979[1] showing that higher procedure volume was associated with better short-term clinical outcome for several surgical procedures, many studies have assessed the effect of volume on outcome in health care. These studies have examined a large variety of surgical procedures, diagnoses, clinical environments, and clinical programs. Investigators have examined different study populations and measured several outcomes in various health care contexts. Nevertheless, there is a striking similarity in the results of these studies: higher clinical volume appears to be associated with improved outcome. Many observers have suggested that the large and consistent effect seen in this large body of literature is too compelling to ignore. There has been increasing interest in the regionalization of certain surgical procedures at high-volume centres to improve the safety of complex surgery and other types of clinical care.[2-4]

While the observed relationship between volume and outcome is undoubtedly complex, two theories have been proposed as possible explanations for these associations, the ‘selective referral’ hypothesis and the ‘practice makes perfect’ hypothesis.[5-7] Both hypotheses assume that improved outcome in high volume units is due to better quality of care. The selective referral hypothesis suggests that high quality centres become high volume centres because of community recognition and more patient referrals from referring physicians. The practice makes perfect hypothesis suggests that centres improve the quality of care they provide by improving their processes of care due to increased experience and coordination, and by improvements in institutional resources that occur with increasing experience.

The literature relating volume to outcome in health care has generated a considerable amount of controversy. Health policy measures predicated on volume-outcome associations, such as volume-based regionalization, will affect the case load of at least some low volume physicians or institutions. Market-based strategies, such as public reporting of volumes, are unpalatable to some low volume providers who might also be concerned about reduced access to patients. There are also legitimate concerns about access of patients to regionalized services, including a possible increase in travel time for some patients,[8] patient preferences for local care,[9] and the possibility that moving some health services out of smaller hospitals may compromise their ability to provide related care due to loss of expertise or resources.[4]

In discussing the literature relating to volume in health care, it is important to remember that volume by itself does not “cause” health outcomes. Rather, volume is a proxy for a complex network of factors that are associated with outcome. Its close relationship with outcome has caused many to regard volume as a surrogate for quality in health care. However, to the extent that volume is a correlate of quality, it is an imperfect one, and elucidating the key structures and processes of care that affect outcome will ultimately require better measures of quality than volume. The reason that volume has become such a popular measure in health care relates to the ease with which it can be measured at the level of providers (as compared with measuring outcomes, or measuring the actual quality of care), especially with the availability of large electronic data sources.

In this report, we present a comprehensive review of the literature relating volume to outcome in health care. The principal objectives were to produce an exhaustive list of procedures and medical conditions that have been studied; to describe characteristics of the studies, including study design, source of data, data quality and general findings; and to identify current and ongoing areas of investigation.

Methods

We sought to identify studies of any surgical procedure, intervention or other type of health care service or program in which the outcomes of care were studied in relation to a measure of the volume of care. Our strategy in designing an electronic search of the health literature databases was to be as comprehensive as possible and to avoid *a priori* constraints regarding the procedures and types of medical care studied. Search strategies in previous syntheses of the volume-outcome literature have relied heavily on pre-specified procedures and care types.[10,11] Previous search algorithms were developed to identify those studies *of specified health interventions* in which health care outcome was analyzed in relation to a measure of volume.

We searched the MEDLINE and EMBASE electronic health databases, limiting study types to human and English language, and restricted to the publication years 1980 to 2004. The search was done by a health librarian. Initially, we searched text word and subject heading terms related to the concepts of volume and health outcomes, such as "volume," "frequency," "outcome," "mortality," and "survival." Searches using combinations of these terms yielded tens of thousands of citations. To narrow the scope of our search, we searched terms in titles only. The search strategy is described in the Appendix.

We scanned titles and abstracts of 639 articles identified in our electronic search and excluded articles whose principal focus was not on the evaluation of health care outcome in relation to a measure of the volume of care (for example, many articles appearing in our electronic search examined the outcome of a cancer treatment, typically radiotherapy, according to the volume of tumor present before treatment). We supplemented this search with a manual search of bibliographies of previous reviews[10,11] of the volume-outcome literature, bibliographies of included citations and consultation with researchers in this area.

All articles were reviewed by a biostatistician and health researcher. We developed a data collection form, and abstracted data from the standardized forms onto an electronic database for further analysis.

Results

This section describes the findings of our literature review. Based on our electronic literature search, we retrieved 184 full manuscripts. Twenty-three papers[12-34] were excluded because they were editorials, position statements, articles that did not examine a measure of volume or other types of articles in which original data were not presented. This section describes our analysis of the remaining 161 individual articles.

Procedures and diseases

Most articles (129, 80.1%) reported on a single procedure or diagnosis. Ten articles (6.2%) reported on 2 procedures, and 5 articles (3.1%) reported on 3 procedures. The largest number of procedures examined in a single study was 17, which was reported in 2 articles (1.2%).

Studies evaluated a heterogeneous group of surgical procedures, medical interventions, health care programs, health care settings (e.g., intensive care), medical diagnoses and health conditions, and diagnostic testing (e.g., mammography). Percentages given in the table below relate to the total number of studies analyzed (161), and percentages in subsequent tables relate to the number of procedures in each category.

Table 1. Procedures and conditions evaluated

	Number	Percent
Diagnoses and other health conditions	44	13.9
Gastrointestinal endoscopy	3	0.9
Gastrointestinal operations (excluding major pancreatic)	76	24.1
Heart operations	36	11.4
Lung operations	13	4.1
Neurosurgical and spinal procedures	8	2.5
Obstetric and gynecologic procedures	4	1.3
Orthopedic procedures	19	6.0
Other surgical procedures	3	0.9
Pancreatic surgery	16	5.1
Percutaneous coronary artery procedures	22	7.0
Urologic surgical procedures	13	4.1
Vascular procedures	59	18.7
Total	**316**	**100.0**

Diagnoses and other health conditions

Forty-four studies (13.9%) evaluated the effect of volume on the outcome of a health care program, setting, medical diagnosis or health conditions, and diagnostic test. Of these conditions, the most frequently studied were trauma care (7 studies, 15.9%), acute myocardial infarction (AMI ["heart attack"], 6 studies, 13.6%) and care for the human immunodeficiency virus (HIV) and the acquired immune deficiency syndrome (AIDS, 5 studies, 11.4%).

Types of care in this category included medical diagnoses for acute (acute myocardial infarction, community-acquired pneumonia) and chronic (cirrhosis) diseases, childbirth, intensive care (adult, pediatric, neonatal), multidisciplinary management of HIV/AIDS and cancer (breast and colorectal), management of trauma and injury, and mammography.

Table 2. Diagnoses and other health conditions

	Number	Percent
Acute myocardial infarction (AMI)	6	13.6
Birth	2	4.5
Birth, singleton at risk for neonatal intensive care unit (NICU) admission	2	4.5
Birth, singleton attended by a family physician	1	2.3
Birth, white singleton	1	2.3
Breast cancer	1	2.3
Cirrhosis	1	2.3
Community acquired pneumonia	2	4.5
Gall bladder and ulcer, combined	1	2.3
Gall bladder diagnosis, non surgical	1	2.3
Hip fracture	1	2.3
Injury	1	2.3
Human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS)	5	11.4
Intensive care	2	4.5

	Number	Percent
Pediatric intensive care	1	2.3
Peptic ulcer	1	2.3
Pneumocystis carinii pneumonia (PCP)	1	2.3
Primary colorectal cancer	1	2.3
Reading mammograms	2	4.5
Respiratory distress syndrome	1	2.3
Respiratory distress syndrome in neonates	1	2.3
Subarachnoid hemorrhage	1	2.3
Trauma	7	15.9
Ulcer, non surgical	1	2.3
Total	**44**	**100.0**

Gastrointestinal endoscopy

Three studies evaluated the effect of the volume of gastrointestinal endoscopic procedures (colonoscopy, endoscopic stent placement for pancreatic cancer, and endoscopic retrograde cholangio-pancreatography [ERCP]) on outcome.

Gastrointestinal operations (excluding major pancreatic surgery)

Seventy-six studies (24.1%) examined the effect of volume on outcome after gastro-intestinal operations. We examined pancreatic surgery separately from other gastrointestinal operations, since there is a large literature specific to volume-outcome associations relating to major pancreatic surgery. Some articles studied the same procedure but for different diagnoses, contexts, or extent. For example, some articles evaluated colon surgery for cancer only, whereas other articles studies colon surgery for any reason. Some articles on gallbladder surgery specified operations that were done using a laparoscopic (minimally invasive) approach, and other articles evaluated gallbladder surgery regardless of the surgical approach.

Of articles studying a specific procedure, the most commonly studied procedure was colectomy (removal of all or a portion of the large intestine, 10 studies, 13.2%), esophagectomy (removal of all or a portion of the esophagus, 9 studies, 11.8%), gastrectomy (removal of all or a portion of the stomach, 7 studies, 9.2%), and cholecystectomy (removal of the gallbladder, 6 studies, 7.9%). Colectomy and cholecystectomy are common and relatively low-risk operations. Gastrectomy and esophagectomy and less commonly done and are associated with a higher risk of surgical complications and death.

Table 3. Gastrointestinal operations (excluding major pancreatic surgery)

	Number	Percent
Appendectomy	3	3.9
Biliary tract surgery	1	1.3
Biliary tract surgery, anastomosis	1	1.3
Cholecystectomy	6	7.9
Cholecystectomy and common bile duct exploration	1	1.3
Cholecystectomy, laparoscopic	1	1.3
Cholecystectomy, open	1	1.3
Colectomy	10	13.2
Colectomy without cancer	1	1.3

	Number	Percent
Colectomy with cancer	5	6.6
Colon or rectal resection	3	3.9
Esophagectomy	9	11.8
Esophagectomy, cancer	1	1.3
Gastrectomy	7	9.2
Gastrectomy, partial	1	1.3
Gastrectomy, total	1	1.3
Gastric bypass	1	1.3
Hepatic lobectomy	2	2.6
Hernia repair	4	5.3
Intestinal operations	2	2.6
Large bowel operations	1	1.3
Liver transplant	1	1.3
Palliative bypass for pancreatic cancer	1	1.3
Rectal resection, cancer	4	5.3
Rectal resection, sphincter-preserving	1	1.3
Stomach operations	2	2.6
Stomach operations, cancer	1	1.3
Stomach operations, non-cancer	2	2.6
Vagotomy	1	1.3
Vagotomy and/or pyloroplasty	1	1.3
Total	**76**	**100.0**

Heart operations

Thirty-six studies (11.4%) evaluated heart operations. The most common cardiac operation studied was coronary artery bypass graft (CABG, 20 studies, 55.6%). Other procedures examined included aortic and mitral valve replacement, repair of congenital heart defects and heart transplantation.

Table 4. Heart operations

	Number	Percent
Coronary artery bypass graft	20	55.6
Aortic valve replacement	3	8.3
Arterial switch operation	1	2.8
Cardiac transplant	1	2.8
Ventricular septal defect repair	1	2.8
Heart transplant	1	2.8
Mitral valve replacement	2	5.6
Open heart surgery	1	2.8
Open heart valvuloplasty	1	2.8
Repair of congenital heart defects	3	8.3
Repair of congenital heart defects, closed heart surgery	1	2.8
Repair of congenital heart defects, open heart surgery	1	2.8
Total	**36**	**100.0**

Lung operations

Thirteen studies (4.1%) evaluated lung operations. Procedures included lobectomy (removal of a portion of one lung) and pneumonectomy (removal of an entire lung). Some studies were limited to procedures done only for a cancer diagnosis.

Table 5. Lung operations

	Number	Percent
Lobectomy	3	23.1
Lobectomy or pneumonectomy	3	23.1
Pneumonectomy	4	30.8
Pneumonectomy, cancer	1	7.7
Resection, lung cancer	2	15.4
Total	**13**	**100.0**

Neurosurgical and spinal procedures

Eight studies (2.5%) examined the effect of volume on the outcome of neurosurgical and spinal operations. The majority of these studies examined intracranial procedures.

Table 6. Neurosurgical and spinal procedures

	Number	Percent
Craniotomy, aneurysm or subarachnoid hemorrhage	1	12.5
Craniotomy, ruptured aneurysm	1	12.5
Craniotomy, unruptured aneurysm	1	12.5
Laminectomy	1	12.5
Resection, intracranial tumor	1	12.5
Surgical treatment unruptured intracranial aneurysm	1	12.5
Spine procedures	1	12.5
Trans-sphenoidal surgery	1	12.5
Total	**8**	**100.0**

Obstetric and gynecologic operations

Four studies (1.3%) evaluated volume-outcome relations for obstetric and gynecologic operations. Three of these studies evaluated hysterectomy (removal of the uterus) and one evaluated caesarian section.

Orthopedic procedures

Nineteen studies (6.0%) assessed orthopedic surgical procedures. The most frequent procedure was total hip replacement (8 studies, 42.1%). Other studies assessed surgery for fractures of the hip, femur and pelvis.

Table 7. Orthopedic procedures

	Number	Percent
Femur fracture	1	5.3
Hip fracture surgery	2	10.5
Hip fracture surgery with other trauma	1	5.3
Hip fracture surgery, no other trauma	1	5.3

	Number	Percent
Major hip and knee surgery	2	10.5
Minor hip and knee surgery	1	5.3
Pelvic fracture	1	5.3
Revision hip and knee replacement	1	5.3
Total hip replacement	8	42.1
Total knee arthroplasty	1	5.3
Total	**19**	**100.0**

Pancreatic surgery

Sixteen studies (5.1%) evaluated major pancreatic surgery. Most of these studies examined the "Whipple" procedure, which involves removal of the head of the pancreas and duodenum. This is an uncommon operation with a high risk of serious postoperative complications and death.

Table 8. Pancreatic surgery

	Number	Percent
Pancreatectomy	4	25.0
Pancreatectomy, cancer	1	6.3
Whipple procedure	10	62.5
Whipple procedure or total pancreatectomy	1	6.3
Total	**16**	**100.0**

Percutaneous coronary artery procedures

Percutaneous procedures on the coronary arteries were evaluated in 22 studies (7.0%). The most common specific procedure studied was percutaneous transluminal coronary angioplasty (PTCA) a procedure done to reduce narrowing in the arteries supplying the heart.

Table 9. Percutaneous coronary procedures

	Number	Percent
Angioplasty and thrombolysis	2	9.1
Cardiac catheterization	3	13.6
Coronary angioplasty	1	4.5
Coronary stent placement	1	4.5
Percutaneous coronary intervention	4	18.2
Percutaneous transluminal coronary angioplasty	9	40.9
Percutaneous transluminal coronary angioplasty without stent	1	4.5
Percutaneous transluminal coronary angioplasty with stent	1	4.5
Total	**22**	**100.0**

Urologic surgical procedures

Thirteen articles (4.1%) studied urologic procedures. The most common procedures studied were cystectomy (removal of the urinary bladder, 4 studies, 30.8%) and transurethral resection of the prostate (TURP, 4 studies, 30.8%).

Table 10. Urologic surgical procedures

	Number	Percent
Cystectomy	4	30.8
Nephrectomy	3	23.1
Prostatectomy	2	15.4
Transurethral resection of the prostate	4	30.8
Total	**13**	**100.0**

Vascular procedures

Fifty-nine articles (18.7%) studied vascular procedures. The most commonly studied procedure was repair of abdominal aortic aneurysm (AAA, a ballooning of the aorta, the major blood vessel carrying arterial blood from the heart to the body), which may result in sudden rupture and death). Of 26 articles on AAA repair, 8 included both elective and ruptured aneurysms, 11 were restricted to elective procedures, and 7 only evaluated ruptured aneurysms.

Another common vascular procedure studied was carotid endarterectomy (CEA, a procedure done to remove fatty deposits from major arteries supplying the brain, to reduce the risk of stroke). Twenty-two studies (37.1%) evaluated the outcomes of carotid endarterectomy. Other procedures examined included procedures done to improve arterial blood flow to the legs, such as lower extremity arterial bypass (3 studies, 5.1%).

Table 11. Vascular procedures

	Number	Percent
Abdominal aortic aneurysm repair	8	13.6
Abdominal aortic aneurysm repair, elective	11	18.6
Abdominal aortic aneurysm repair, ruptured	7	11.9
Aorto-bifemoral bypass	1	1.7
Amputation, lower limb	3	5.1
Carotid endarterectomy	22	37.3
Endovascular therapy	1	1.7
Infrainguinal vascular reconstruction	1	1.7
Intra-abdominal artery operations	1	1.7
Lower extremity arterial bypass	3	5.1
Vascular surgery	1	1.7
Total	**59**	**100.0**

Other surgical procedures

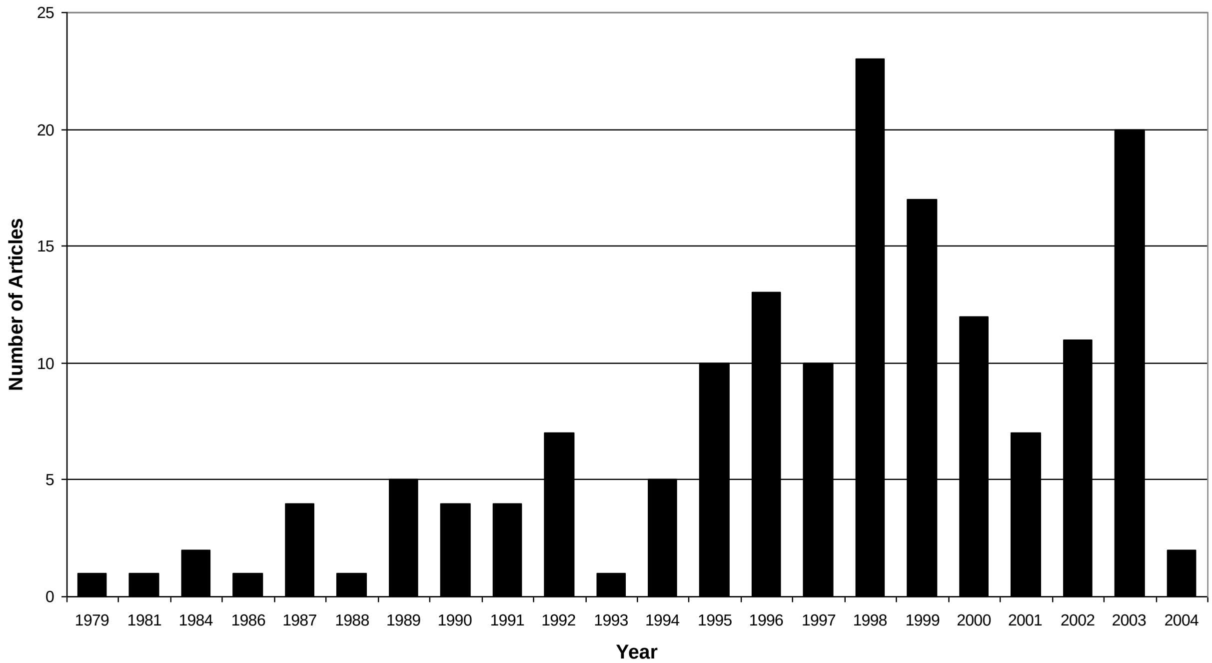
Three studies evaluated procedures that were not easily classified according to the procedure categories listed above. One study combined the outcomes of 9 surgical procedures,[6] one examined the outcomes of pelvic exenteration (removal of the rectum, urinary bladder and the prostate or uterus) and one examined breast cancer surgery.

Characteristics of the articles

Year of publication

The years of publication of the articles studied ranged from 1979 to 2004. Relatively few studies were published prior to 1995, and many more articles have appeared in recent years. The largest number of articles was published in 1998 (23 articles, 14.3%).

Figure 1. Number of articles published by year.



Years covered in the analysis

Most studies examined the outcomes of procedures done over a number of continuous years. One article compared 1996 with 1993, and one article did not mention the years covered. The average number of months included in the articles was 48.7 (just over 4 years), with a standard deviation of 37.1 months. The median number of months was 38, with a range from 3 months to 15 years. The 25th and 75th percentiles were 18 months and 6 years.

Thirty-seven studies (23.1%) covered one year or less. Approximately half (48.8%) covered between 1 and 5 years, and 28.1% covered more than 5 years.

Geography

Most studies were done in the United States (128 studies, 79.5%). The next most common country of origin for the articles reviewed was Canada (14 studies, 8.7%). One “international” study examined outcomes for patients of the US Department of Defense treated in hospitals in the US and Europe.[35] One “worldwide” study involved an international group of pediatric hospitals.[36] Three studies included explicit comparisons between countries, two between the US and Canada and one between the US and UK.

Table 12. Country of study

Country	Frequency	Percent
Canada	14	8.7
England	2	1.2
Europe	1	0.6
Finland	2	1.2
Germany	2	1.2
Japan	1	0.6
Netherlands	2	1.2
Norway	1	0.6
UK	3	1.9
UK/US	1	0.6
US	128	79.5
US/Canada	2	1.2
International (US based)	1	0.6
Worldwide (mostly US)	1	0.6

Among Canadian studies, 2 were done in single cities (Edmonton and Vancouver). Half of the Canadian studies examined data from Ontario, 2 were from Quebec and 2 were from Alberta (one of which looked at Northern Alberta). One study covered multiple provinces.

Of the 128 studies done in the US, 49 (38.3%) contained data from more than 2 states. Fifty-eight studies (45.3%) were done in a single state. The states most commonly studied in articles that examined one or a small number of states were California (22 studies), New York State (20 studies) and Maryland (10 studies).

Table 13. Geographic regions studied in volume-outcome studies done in the United States

Geographic region	Number	Percent
One state	58	45.3
Two states	3	2.3
Three or more states	49	38.3
Single city, county, or region	13	10.2
Multiple cities	1	0.7
Cooperative region	4	3.1
Total	**128**	**100.0**

Number of study subjects

There was substantial variation in the number of study subjects included, as well as the numbers of physicians and institutions. The median number of persons included in a volume-outcome analysis (calculated from the 288 of the 313 analyses where it was presented) was 8,224, the median number of physicians in the 74 analyses that reported it was 226, and the median number of institutions in the 262 analyses that reported it was 151. The table below presents these data according to the 313 analyses included in the161 articles.

Table 14. Number of subjects, physicians and institutions in the studies

Unit	Number of Analyses	Mean (SD)	Median	Range	Inter-quartile Range
Subject	288	44,947 (115,191)	8,224	60 – 974,803	2,341 – 35,607
Physician	74	1,161 (2,057)	226	7 – 8,027	75 – 825
Institution	262	508.5 (764.3)	151	1 – 4,587	41 – 742

Age of study subjects

Most studies (104, 64.6%) were not restricted with respect to the age of study subjects. Twenty-four (14.9%) studies were limited to adults (including 3 Veterans' Administration studies), 2 (1.2%) were limited to adults aged less than 65 years, 27 (16.8%) were limited to adults aged 65 years or older (26 of which used US Medicare data) and 4 studies (2.5%) were limited to children.

The age of included subjects is an important factor in volume-outcome studies, since older patients are at higher risk of adverse events. Younger patients may be less sensitive to variations in the quality of care than older patients. Therefore, studies limited to older patients are likely to find stronger volume-outcome associations than studies that do not limit the age of subjects.

Other restrictions of study subjects

Nineteen of 26 studies using US Medicare data used data from the entire US, while 7 were restricted to a single state (n=4), region (New England, n=1), a comparison of a high-volume state with a low-volume state (n=1), or a subset of 5 metropolitan areas and 5 states (n=1).

Insurance status was not an exclusion criterion in any of the studies done outside the US. Of the US studies, 26 were restricted to subjects enrolled in the US Medicare program. Although the US Medicare program provides health coverage for disabled persons, persons with end-stage renal disease and persons aged 65 years or older,[37] all studies using Medicare data excluded disabled subjects and those with end-stage renal disease. Two other studies were limited to subjects with specific health insurance coverage, one based in a Health Maintenance Organization (HMO) and one including subjects insured by the US Department of Defense.

Research design

No study used a true experimental design. Almost all analyses were cross-sectional studies of outcome according to a measure of volume. Only one study used a design intended to test the hypothesis that high volume is causally associated with improved outcome.[38] This analysis of surgery for hip fracture in Quebec used a longitudinal design, and tested whether hospital volume in the previous year (as a logarithm-transformed continuous variable) predicted in-hospital death or length of hospital stay. Since hospital volume changed over time, this study was intended to test whether changes in hospital volume were causally associated with outcome (this study found that while there was an overall association between volume and outcome, changes in hospital volume over time were not associated with change in the risk of adverse outcome).

Sources of data

Of the 161 articles, 117 (72.7%) used administrative data. Forty-one of these (35.0%) used US data with a nationwide target sample and 43 (36.8%) used data from a single US state, most commonly California, New York or Maryland. State databases typically excluded federal hospitals.

Forty-four studies (27.3%) used data derived from observations on individual patients (‘clinical’ data). Of these, 17 (38.3%) were conducted prospectively, with data acquisition occurring after the specification of a study hypothesis and study protocol. The other articles using clinical data were either retrospective or did not provide enough information to determine whether data were collected prospectively to test an *a priori* hypothesis. Five of the studies using ‘clinical’ data examined carotid endarterectomy (all retrospective), five examined trauma care (one of which was a prospective design), and two examined care for HIV/AIDS.

Table 15. Sources of data in the articles

Source of data	Number (Percent)
ADMINISTRATIVE	117 (72.7%)
US, national scope	41 (35.0%)
Medicare; Health Care Financing Administration (HCFA); Centers for Medicare & Medicaid Services	17
Healthcare Cost and Utilization Project (HCUP); Nationwide Inpatient Sample (NIS)	13
Commission on Professional and Hospital Activities (CPHA)	6
Surveillance, Epidemiology and End Results (SEER)/Medicare linked database	5
US, single state database	43 (36.8%)
California	13
New York	10
Maryland	10
Other states	10
Canada, Canadian Institute for Health Information (CIHI) or Quebec	8 (6.8%)
State/provincial cancer registry	5 (4.3%)
US/Canada multiple state/province	5 (4.3%)
Other	13 (11.1%)
Not stated	2 (1.7%)
CLINICAL	44 (27.3%)
Retrospective or not clearly specified	27 (61.4%)
Prospective	17 (38.6%)

Data quality

Assessment of the quality of the data included in the studies varied. Most of the articles (109, 67.7%) did not mention data quality assurance. Fourteen studies (8.7%) mentioned data quality, 16 (9.9%) indicated that an examination of coding accuracy was done, and 22 studies (13.7%) indicated a more rigorous assessment of data quality. Some articles in the latter category, however, are included because of data quality checks routinely done by a data collection agency, and not an active assessment of data validity by the study investigators.

Missing data occurred in many studies. In articles examining the outcomes of cancer surgery, data on tumor grade or location were missing in 10%, 47% and 61% of observations in three studies that mentioned it. In studies that examined physician volume, many subjects could not be linked to a surgeon, and many surgeons had missing data on surgeon characteristics. Another common problem with data linkage was linking Medicare data with data in the Surveillance, Epidemiology and End Results (SEER) program. The proportion of observations that could not be linked ranged from a low of 6% (SEER-Medicare) to highs of 42% and 50% in 2 studies for the failure to link a physician code to individual patients. Among studies reporting the proportion of observations for which linkage was impossible, the median and mean percentages were 20%. Three studies using surveys to augment data in administrative databases reported non-response rates of 14%, 15% and 60%.

Twenty-six studies mentioned missing information for variables not described above. The proportion of missing data varied from "negligible" in one study to a high of 41%. Occasionally, data were missing on key outcomes (such as vital status at 30 days), or important confounding variables such as age and gender. The mean percentage of missing data (among studies that reported it) was 7.1%, standard deviation 8.7%, median 4.5%, and inter-quartile range 2.9% to 8.0%.

Synthesis of study findings

Detailed results of the individual analyses of the studies are presented in the Appendix. We examined the results of 313 analyses done in the 161 studies. The table lists the analyses grouped by article, according to the first author of each article. The table also lists the date of study publication, procedure or diagnosis, country, years of study, the number of subjects, physicians and hospitals included, and a summary of the principal findings. To quantify the strength of associations between volume and outcome, a P value from the test of statistical significance from the principal analysis is included. When P values were not reported, they were inferred where possible from other reported statistics such as standard errors. We considered a P value of 0.05 or less to indicate a statistically significant result.

We did not perform a quantitative synthesis of the reported results of the studies, since there was substantial heterogeneity in the way results were reported, especially in the specification of volume categories and outcome measures. Others[39] have summarized results across studies as the absolute difference in outcome between the highest and lowest volume categories. However, this measure is sensitive to the manner in which volume categories were defined. For example, if a study demonstrates a volume-outcome association with a linear trend across volume categories, the absolute difference in the risk of outcome between the extreme volume categories will be much larger if five volume categories were used than if only two or three volume categories were used.

The majority of the analyses (214, 68.4%) showed significant associations between higher volume and better outcome. Of the remaining 99 analyses, 11 (11.1%) were indeterminate and 88 (88.9%) did not demonstrate a significant adverse effect of low-volume care, four of which showed a statistically significant association of higher volume with worse outcome. Even among the 88 analyses that did not find a significant relationship between high volume and better outcome, the procedures or diagnoses studied were usually found to have volume-outcome associations in other studies. We also examined analyses according to whether they measured the effect of hospital volume or physician volume. Among 275 analyses that examined hospital volume, 185 (67.3%) showed a statistically significant association with outcome. Among the 98 analyses examining physician volume, 62 (63.3%) showed a statistically significant association.

Among studies evaluating diagnoses and other health conditions and health care environments, 16 of 44 analyses (36.4%) did not find significant volume-outcome associations. The most common analysis in this category was trauma care, with six of seven studies (85.7%) being 'null'. Among gastrointestinal operations, 24 of 76 analyses (31.6%) did not demonstrate significant volume-outcome associations, including 4 of 9 analyses (44.4%) of cholecystectomy, 9 of 24 analyses (37.5%) of colon or rectal surgery, and 8 of 17 (47.1%) analyses of any type of gastric surgery. Ten of 36 analyses (27.8%) of heart operations were 'null', including 6 of 20 analyses (30%) of coronary artery bypass grafting (CABG). Seven of 13 analyses (53.8%) of major lung operations did not identify significant volume-outcome associations. Relatively few analyses of neurosurgical and spinal procedures (1 of 8, 12.5%), orthopedic procedures (5 of 19, 26.3%), pancreatic operations (3 of 16, 18.8%), percutaneous coronary procedures (3 of 22, 13.6%), and urologic procedures (3 of 13, 23.1%) were 'null' comparisons. Among analyses of vascular operations, 14 of 59 (23.7%) did not demonstrate significant associations, including 5 of 26 analyses (19.2%) of abdominal aortic aneurysm repair and 6 of 22 analyses (27.3%) of carotid endarterectomy.

Twenty-six analyses that examined the effect of hospital and provider volume together found that outcome was significantly associated with either hospital volume or provider volume, but not both. Half of these studies found outcome was significantly better for subjects treated at high-volume hospitals but not by high-volume providers, including two analyses of coronary artery bypass graft, two of abdominal aortic aneurysm repair, two of major pancreatic surgery, and analyses of percutaneous transluminal coronary angioplasty, carotid endarterectomy, palliative surgery for pancreatic cancer, hysterectomy, intestinal surgery, total hip arthroplasty and cardiac catheterization.

Thirteen of these 26 analyses demonstrated significant associations between improved outcome and higher provider volume, but not hospital volume. These included three analyses of acute myocardial infarction, three of abdominal aortic aneurysm repair, two of carotid endarterectomy, and analyses of coronary artery bypass graft, aortic valve replacement, esophagectomy, gastric surgery for morbid obesity and rectal cancer surgery.

Another 26 studies that examined the effect of hospital and provider volume together found that improved outcome was associated with both provider and hospital volume. Therefore, of 52 studies in which provider and hospital volume were studied together, 26 (50%) found that improved outcome was significantly associated with both high provider and institution volume, 13 (25%) found improved outcome was significantly associated with high provider volume but not hospital volume, and 13 (25%) found improved outcome was significantly associated with high institution volume but not provider volume.

Four analyses showed a statistically significant association between higher volume and poorer outcome. Two of these analyses were included in the same article. One article assessed 130 Whipple operations for cancer done between 1989 and 1994 at US Department of Defense hospitals around the world (but mostly in the US).[35] Thirty-day mortality was higher among the third of patients treated in high-volume hospitals as compared with low-volume hospitals (9% as compared with 6%, $P<0.05$). An article measuring in-hospital mortality for non-surgical care of 88,839 patients with gallbladder disease treated without surgery in 1,210 hospitals in 1972 found a higher mortality in the high-volume hospitals, with volume categorized as a binary variable.[5]

One article that examined 17 procedures and diagnoses contained two analyses that showed significant associations between higher volume and worse outcome.[40] Among patients treated in the US in 1972, higher volume was associated with a higher risk of in-hospital death for 5,049 subjects with subarachnoid hemorrhage treated in 749 hospitals, and for 80,211 subjects with acute appendicitis treated in 916 hospitals.[40] This small number of analyses with such discrepant results might easily be explained on the basis of random error, since the long-run probability that a typical analysis will produce a statistically significant result by chance alone (when there truly is no difference) is approximately 5%.

Our review provides considerable evidence for the ubiquity of volume-outcome associations for virtually every health service that has been investigated. Assuming that the lack of negative studies is not due to a publication bias against research with negative results, one might speculate that almost any health service subjected to sufficient scrutiny will be found to have a volume-outcome relationship.

Current initiatives

To review current research initiatives, we contacted researchers and opinion leaders in volume-outcome research and health policy making relating to the regionalization of health care. We interviewed John Birkmeyer, Professor of Surgery at the University of Michigan, an internationally recognized researcher on the effect of volume on outcome in surgical care at the hospital and surgeon level, and an influential policy advisor in the United States. We also interviewed Dr. Jack Tu, Team Leader, Canadian Cardiovascular Outcomes Research Team; Dr. Terry Sullivan, Vice President of Cancer Control and Research at Cancer Care Ontario; Dr. Bernard Langer, Senior Consultant in Surgery, Cancer Care Ontario; and representatives from the Department of Health Care and Epidemiology, University of British Columbia.

In March 1999, Cancer Care Ontario released a report from a task force it convened to review whether pancreatic cancer surgery should be regionalized in Ontario. The task force had broad representation and concluded that pancreatic cancer surgery should be restricted to surgeons with specific knowledge, skill, experience and commitment criteria, with defined surgical training, and an annual volume of at least 10 major pancreatic cancer operations per year (and an optimum volume of 50 operations per year). Appropriate hospital structure characteristics were also defined. The impact of the release of this report on the distribution pancreatic cancer surgery is uncertain, and is currently being evaluated.

Perhaps the highest-profile application of volume-outcome research to health policy has been the “evidence-based hospital referral” initiative of The Leapfrog Group in the United States. The Leapfrog Group is a consortium of 150 large public and private health care purchasers representing over 40 million beneficiaries of health care in the US. The group was founded by The Business Roundtable, and is supported by the Robert Wood Johnson Foundation. To advance their mission of improving patient safety and raising the standards of health care, the group proposed this initiative to refer patients requiring five procedures to hospitals with specific volume thresholds.

Table 16. The Leapfrog Group "Evidence-Based Hospital Referral" procedures and recommended hospital volumes (2000 criteria)

Procedure	Volume threshold
Coronary artery bypass graft	≥450 /year
Percutaneous coronary intervention	≥400 /year
Abdominal aortic aneurysm repair	≥50 /year
Pancreatic resection	≥11 /year
Esophagectomy	≥13 /year
High-risk delivery:	
Expected birth weight < 1500 grams Gestational age < 32 weeks Pre-natal diagnosis of major congenital abnormality	Neonatal ICU with average daily census ≥15

In 2003, the evidence-based hospital referral guidelines were revised to include process of care measures, such as the use of beta-blockers, internal mammary artery graft and lipid-lowering therapy for coronary artery bypass surgery and aspirin therapy on discharge after percutaneous coronary intervention, and direct outcome measures such as being in the lowest quartile of operative mortality in the United States for coronary artery bypass surgery, and participation in the Society of Thoracic Surgeons national database. In 2005, a revision of these guidelines will include thresholds for surgeon volume in addition to hospital volume. It has been estimated that adherence to these guidelines would save the lives of 2,581[41] (2000 guidelines) and 7,818[42] (2003 guidelines) persons per year having surgery in the Unites States. The Leapfrog Group guidelines have been widely disseminated and are available at www.leapfroggroup.org.

In the United States, there are several ongoing large national projects involved in the collection of high quality data and analysis of the outcomes of surgical procedures, including the Society of Thoracic Surgeons National Cardiac Database (available at www.sts.org), and the US Veterans Affairs National Surgical Quality Improvement Program (NSQIP).[43]

All of the respondents felt that the descriptive literature regarding the association of volume with heath outcome is mature, and further descriptive research on volume-outcome associations is unlikely to move the research field forward in a way that will help inform health policy. In general, respondents felt that future research on volume and outcome in health care should be carefully aligned with health policy goals. Important research priorities included: developing better models (that might include, in addition to volume, measures of process of care and outcomes) for the prediction of future outcomes of health services, examining the feasibility and consequences of regionalization strategies, investigating the process-of-care mechanisms underlying volume-outcome associations, extending research into defining appropriate volume thresholds for specific health services, and using experimental or quasi-experimental methods to test the assumption of causation for volume-outcome associations.

Methodologic Review of Research Relating Volume to Outcome in Health Care

Executive Summary

There were major differences among the studies with respect to methods. When measuring volume, most studies examined the frequency of occurrence of a specific procedure or diagnosis, and not the frequency of all related procedures or diagnoses. One hundred and thirty-nine articles (86.3%) examined hospital volume as a possible predictor of outcome, and 67 articles (41.6%) examined provider volume. Approximately equal numbers of studies modeled volume as a continuous variable only (27.3%), collapsed volume into *post hoc* categories (30.4%) and collapsed volume into categories according to pre-specified cut points (29.2%). The remaining studies used a combination of approaches. Of the articles that used pre-specified cut points to define volume categories, 68.9% selected cut-points based on the volume distribution and 8.2% used categories based on an external standard such as the guidelines of a specialty society. The most common number of categories used when volume was categorized into equal strata according to the distribution of subjects was 4 (quartiles).

Death was the most common outcome measured in the studies (88.2%) and was most frequently defined as in hospital (58.5%), followed by 30-day death (16.9%), and in hospital or 30-day death (7.0%). Nineteen studies used survival analysis to estimate time to death. Other commonly measured outcomes were length of hospital stay, unplanned readmission to hospital after discharge, re-intervention, complications, costs or charges, and disease-specific measures. Composite outcomes were occasionally used, and multiple outcomes were measured frequently.

Many patient characteristics, such as age, sex, severity of illness, comorbid medical conditions and socioeconomic status, affect the risk of adverse outcomes of health care and should be controlled in a study examining health outcomes. Comorbid conditions were measured in 63.4% of studies, and when included in risk-adjustment models were entered as individual comorbidities (44.1%), a comorbidity index such as the Charlson score (32.4%) or other index, or a count of the number of comorbid diseases (13.7%).

Statistical analyses used in the studies we reviewed were often not conducted appropriately. The data encountered in volume-outcome analyses have idiosyncrasies (such as clustering of subjects, and the fact that inferences about quality are made at the level of a provider although outcomes are assessed at the level of individual patients) which complicate the statistical analysis. While most studies performed analyses that adjusted for potentially confounding factors using multivariable statistical models, only 13.7% of studies used state-of-the-art methods.

The most common way in which estimates of volume effects were presented were adjusted measures of effect, such as adjusted odds ratios, hazard ratios, or rate ratios. Among studies that presented adjusted estimates, the other measures used were adjusted outcomes, such as the adjusted risk of death, and observed-to-expected differences or ratios. 21.7% of studies presented unadjusted effects.

The substantial degree of variation we found among the methods used in the studies suggest that there is little uniformity among investigators regarding the appropriate analysis of volume-outcome analyses. Many studies had major methodologic problems that made it difficult to interpret their findings. The quality of future studies would be improved by the involvement of a methodologist in the study design and analysis, and by the use of modern analytic methods.

Introduction

Studies investigating the outcomes of health interventions in relation to the volume of similar interventions pose several unique challenges, including conceptual, methodological and statistical problems. Examples of conceptual problems include selecting appropriate procedures or health conditions for analysis, understanding the relation of volume to quality of care, operationalizing the concept of "volume" (average versus cumulative experience), and defining a level for the assignment of volume (for example, the physician, hospital, program, or geographic region). Methodological problems include accounting for differences in pre-existing risk between patients, problems with the quality and richness of the data available, and the choice of outcome measure. Statistical problems include issues such as the specification of volume in a statistical analysis and accounting for clustering and heterogeneity in the data typical of volume-outcome analyses.

In this report, we present a critical methodological review of the health literature relating volume to outcome in health care. We also present recommendations regarding appropriate methods for the design and analysis of these types of studies.

Methods

To identify studies looking at the relation between volume of care and health care outcomes, we conducted a systematic review of the literature. We searched the MEDLINE and EMBASE electronic health databases, limiting study types to human and English language, and restricted to the publication years 1980 to 2004. The search was done by a health librarian. Initially, we searched text word and subject heading terms related to the concepts of volume and health outcomes, such as "volume," "frequency," "outcome," "mortality," and "survival." Searches using combinations of these terms yielded tens of thousands of citations. To narrow the scope of our search, we searched terms in titles only. The search strategy is described in the Appendix.

We scanned titles and abstracts of 639 articles identified in our electronic search, and excluded articles whose principal focus was not on the evaluation of health care outcome in relation to a measure of the volume of care (for example, many articles appearing in our electronic search examined the outcome of a cancer treatment, typically radiotherapy, according to the volume of tumor present before treatment). We supplemented this search with a manual search of bibliographies of previous reviews[10,11] of the volume-outcome literature, bibliographies of included citations, and consultation with researchers in this area.

All articles were reviewed by a biostatistician and health researcher. We developed a data collection form and abstracted data from the standardized forms onto an electronic database for further analysis. For this analysis, we focused on methods used in the reports of the articles.

Results

This section describes the results of our methodologic review of the literature. The analyses presented here are based on the 161 articles retrieved in our literature search.

Definition of volume

Precisely how volume was defined was often difficult to ascertain in the studies. In general, volume appeared to be measured as "average" intensity of activity over a specified period, and not "lifetime" volume. Use of average volume instead of lifetime volume would not account for any effect of cumulative experience on outcome. When measuring volume, most studies examined the frequency of occurrence of a specific procedure or diagnosis, and not the frequency of all related procedures or diagnoses. For example, studies of coronary artery bypass graft (CABG) typically defined volume according to the volume of CABG, and not the volume of CABG plus other related operations such as heart valve replacement. To the extent that skills from experience with related procedures are transferable, use of specific-procedure volume instead of the more general related-procedure volume may be a source of error in studies if there is significant variation in the distribution of procedures done by providers.

Multi-year studies often referred to “annual volume” without specifying whether this indicated the average annual volume (the total volume over the study period divided by the number of years), or whether annual volume was calculated separately for each year. Some studies measured volume during the year prior to the hospitalization of the person whose outcome was being measured. The potential advantage of this latter method is that the volume measure is more relevant to the study subject, especially if a provider or institution volume changed over time. The definitions of volume are presented below according to the time span of the studies. In one study, the time period was not stated.

Studies of one year or less

Forty-three studies covered a period of time of one year or less. Thirty-one of these studies (72.1%) measured total volume, and 7 (16.3%) used the volume of Medicare patients. Six studies (13.9%) used a different definition of volume: the ‘familiarity index’ (the number of discharges of patients with AIDS per 10,000 hospital discharges), surgeon volume measured as the volume in the 12 months prior to an admission (but hospital volume measured over the whole period), patient-to-bed ratio, risk-adjusted volume, and high-risk volume. The definition of volume was not stated in two studies.

In addition to these volume categorizations, three studies measured the volume of patients in the same diagnostic group (for example, a study of coronary artery bypass graft also measured volume using the total number of patients with heart disease), two studies stratified by volume at different levels of risk, and one other study on AIDS care also used the familiarity index in addition to volume of patients with AIDS.

Studies covering one to two years

Twenty studies fell into this category, 17 of which (85%) measured average annual volume. One study measured volume in the year of a patient’s admission, and one measured the average volume of Medicare patients. Finally, a study on the outcomes of AIDS care used measures of volume, including the familiarity index in the second year of the study, volume in the second year of the study and cumulative (lifetime) volume.

Studies covering two to five years

Of the 56 studies covering this period of time, 41 (73.2%) measured average annual volume, 4 (7.1%) used average annual Medicare volume, 3 (5.4%) measured volume during the 12 months prior to hospital admission and 2 (3.6%) used volume in the same year as hospital admission. Two studies compared eight different volume definitions, one study categorized volume by type of hospital (small rural, academic, etc.) and three studies did not provide definitions. Some studies also used other measures of volume as independent variables for the prediction of outcome, including the volume of all related surgical procedures, and cumulative experience.

Studies covering longer than five years

Of 41 studies in this category, 27 (65.9%) measured average volume, 7 (17.1%) measured volume during the year of each subject's hospitalization and 3 (7.3%) measured average Medicare volume. One study measured the cumulative number of subjects treated up to the time of the hospitalization of the subject whose outcome was being assessed, one study measured volume during the year after the time period covered in the study and one study classified surgeons according to their higher volume in either the current year or previous year. One study did not provide a definition of volume. In addition to these volume measures, two studies also measured cumulative lifetime provider experience.

Level at which volume is measured

Clinical volume, as a measure of the frequency with which heath care services are provided, can be measured at several different levels. The most obvious levels are perhaps those of the health care provider (such as a surgeon), or an institution (such as a hospital). However, volume of services can be measured at other levels, such as a provider group, clinical unit, health program, hospital network or region. Which level to use for measuring volume as a predictor of outcome depends on the purpose of the analysis and the proposed conceptual framework of the role of physicians, programs, institutions, networks and their interactions in contributing to health outcomes. Since many analyses of volume and outcome in health care have focused on the role of regionalization of services or selective referral, studies have tended to focus on the level of the provider and institution.

Overall, 139 articles (86.3%) examined hospital volume as a possible predictor of outcome, and 67 articles (41.6%) examined provider volume. Over half of the 161 articles (94, 58.4%) looked only at hospital volume as a predictor of outcome. Forty-five articles (28.0%) examined both hospital and provider volume, and 22 articles (13.7%) examined only provider volume. Articles that evaluated only provider volume tended to be smaller studies with only one or a few hospitals.

Table 17. Level at which volume was measured in the studies

	Provider volume analyzed		
Hospital volume analyzed	**No**	**Yes**	**Total**
No	0	22	22 (13.7%)
Yes	94	45	139 (86.3%)
Total	94 (58.4%)	67 (41.6%)	161 (100%)

Of the 45 articles that examined institution and provider volume, 34 (75.6%) assessed the joint effects of both institution and provider volume simultaneously (for example, by including both institution and provider volume as independent variables in a multi-variable statistical model). Ten of the articles assessed institution and provider volume separately, and in one article it was not clear whether institution and provider volume were analyzed separately or jointly.

Some articles that examined hospital volume were, in effect, measuring the volume of a clinical unit (for example, intensive care unit) or program (for example, trauma care). No study specifically measured the volume at other levels of analysis such as a geographic region or regional hospital network.

Categorization of high and low volume

There are different ways of analyzing volume as potential predictor of outcome. One approach would be to treat volume as a continuous variable. In this case, the hypothesis that volume is associated with outcome could be tested by assessing whether outcome changes in a linear fashion according to different values of volume. Volume can also be treated as a categorical variable, where providers or hospitals are grouped into categories defined by a range of volumes. Sometimes, volume categories are defined *post hoc*, which can complicate the interpretation of an analysis since it allows an analyst to select cut-points for volume categories that optimize any volume-outcome association. A more valid method is to define volume categories *a priori*, before any data have been observed. Categorization of volume has the disadvantage of losing some information and reducing the statistical power of a study to detect an association. On the other hand, volume categories are easier to interpret by consumers of volume-outcome research, and analyses using volume categories are more amenable to health policy strategies such as regionalization or selective referral.

Often, volume categories are defined according to the distribution of subjects with respect to the volume of the provider or institution that treats them. Typically, subjects are assigned to categories to ensure approximately equal numbers of subjects in each category, according to the percentiles of the population (quantiles). When volume is categorized in this fashion, the highest volume category usually contains far fewer providers or hospitals than the lower volume categories.

Occasionally volume categories are defined on a substantive theoretical basis, for example by external volume standards from the literature or organizational guidelines (such as for trauma care or number of coronary bypass operations), by visual inspection of volume-outcome curves to identify inflection points, or by using statistical techniques to determine values of volume where the form of the relation between volume and outcome changes substantially.

Approximately equal numbers of studies modeled volume as a continuous variable only (44 studies, 27.3%), collapsed volume into *post hoc* categories (49 studies, 30.4%), and collapsed volume into categories according to pre-specified cut points (47 studies, 29.2%). The remaining studies used a combination of approaches. For example, some studies first modeled volume as a continuous variable to determine whether or not there was a volume-outcome relation, and then used volume categories to estimate risk ratios, or for a more interpretable presentation of the results. Six studies modeled volume as a continuous variable and then used cut-points to define volume categories on a *post hoc* basis; 11 studies used pre-specified cut points to define volume categories *a priori*, after first determining whether there was an association between volume and outcome by treating volume as a continuous variable. Three studies used a combination of *a priori* and *post hoc* volume categories, for example by categorizing hospital volume by quantiles and surgeon volume using arbitrary cut-points. In three studies it was not possible to determine how volume was categorized.

Table 18. Specification of volume in the studies

Volume categorization	Was volume modeled as a continuous variable: Yes	No	Total
	Number of studies		
None	44	0	44
Any			
Pre-specified categories only	9	47	56
Post hoc categories only	6	49	55
Both	0	3	3
Total	**59**	**99**	**158**

In the table, volume categorizations done on the basis of percentiles were classified as pre-specified, unless explicitly stated otherwise in the article (for example, one article categorized subjects according to tertiles but only after first observing the volume-outcome plots). Categorization done using cut-points other than percentiles or based on a volume distribution was classified as *post hoc*. We did not consider volume to be categorized for the purpose of the analysis, if volume was categorized solely for the purpose of graphical presentation.

Of the 59 studies using volume as a continuous variable, 20 (33.9%) used a logarithmic transformation, 9 (15.3%) used volume as a linear term and 2 (3.4%) used both (in one study, surgeon volume was log-transformed and hospital volume was treated as a linear variable). Two studies (3.4%) used an exponential transformation of volume (quadratic or cubic), one study (1.7%) used a statistical smoothing procedure (robust locally weighted scatterplot smoother, LOWESS) and one article (1.7%) modeled the reciprocal of volume. Of the remaining 24 studies (40.7%) where volume was a continuous variable but the form was not further specified, it is likely that volume was included as a linear term.

Of the 61 articles that used pre-specified cut points to define volume categories, 42 (68.9%) selected cut-points based on the volume distribution (quantiles or the mean volume). Five studies (8.2%) used categories based on an external standard such as the guidelines of a specialty society, and 2 studies (3.3%) used categories based on a previous study. Five studies (8.2%) examined more than one type of categorization based on multiple criteria (for example, a mix of categories based on an external standard, a previous study, and quantiles). The purpose of these studies was typically to compare the results of studies using different volume categories. Two studies (3.3%) categorized hospitals into a "high-volume or regional" category for comparison with a low-volume category, 3 (4.9%) used a different method of categorization than described above, and in 2 (3.3%) the rationale for categorization was not specified.

When quantiles were used, the number of volume categories ranged from 2 to 15 with a median number of 3 volume categories. Twenty-eight studies used 2 categories, 32 used 3 (variably described as 'tertiles' and 'terciles'), 37 used 4 (quartiles) and 11 studies used 5 (quintiles).

Very little research has evaluated optimal cut-points for defining volume categories that are useful for informing health policy. The bulk of the research we reviewed focused on determining whether volume-outcome associations existed, rather than selecting meaningful categories of volume based on a substantive theoretical framework, and determining how outcome varies according to these meaningful volume categories. There is empirical evidence that volume thresholds used in popular guidelines do not have a rational basis.[44,45] Rational selection of volume categories requires a conceptual framework of the relationship of volume-outcome associations with health policy measures. For example, health policy measures aimed at restricting certain types of care to only a few very high-volume centres will require different volume thresholds than policy measures aimed only at reducing the amount of activity at extremely low-volume units. Inspection of volume-outcome plots, and statistical procedures that yield nonparametric estimates of the shape of volume-outcome curves, may provide valuable information on volume thresholds associated with qualitative changes in the shape of the volume-outcome relationship, which should be the primary focus of policy-relevant research.

Outcomes measured

Death

Overall, 142 studies (88.2%) analyzed death, as a categorical variable (yes/no), as at least one of the outcome measures. Most frequently, death was defined as "in hospital" (83 articles, 58.5% of the 142 in which death was an outcome). Another 24 studies (16.9%) defined 30-day death and 10 (7.0%) used in-hospital or 30-day death. Five studies (3.5%) used "operative death" without further specification. Four studies (2.8%) used death in 30 days as a primary outcome and death within another time period (for example, 1 year) as a second outcome. Nine studies (6.3%) used death within some other period (14 days, 28 days, 3 months, 3 months and 1 year, stillborn, or within 1 week for newborns).

In-hospital death is commonly used in analyses of hospitalization databases that do not contain information on events after the hospital separation, but indicate whether death occurred during the hospitalization. In-hospital death also has the desirable characteristic of including subjects who died at the end of a prolonged hospitalization, since death occurring in hospital without a discharge following a health intervention can arguably be considered an adverse outcome. In contrast, 30-day mortality has the advantage of enumerating (for the purpose of measuring adverse outcomes) deaths that occurred within a short time of a health intervention, but did not occur during a hospitalization because the patient was discharged or transferred to another hospital. However, measurement of 30-day mortality requires a data source that contains information on out-of-hospital death, or the ability to track an individual between databases (such as a hospitalization database and a vital statistics database) using a unique identifier or a linkage strategy.

Nineteen studies used survival analysis to estimate time to death. Of these, 10 also analyzed short-term mortality as a binary variable. When more than one definition of "death" was used as an outcome measure, we considered it as a single outcome. Among the 19 studies, 9 measured time to death from any cause, and did not also look at short-term death as an outcome, 7 measured time to death from any cause, in addition to short-term death, 2 measured procedure- or disease-specific survival in addition to short-term death, and one measured overall and disease-specific death in addition to short-term death. In summary, death, measured either as time to death or as a dichotomous outcome, was an outcome measure in 151 studies (93.8%), mostly as a short-term outcome.

Length of stay, complications, and resource use

Length of hospital stay was examined in 43 studies (26.7%), and unplanned hospital readmission during a specified period of time after discharge was examined in 5 studies (3.1%). Re-intervention was examined as an outcome in 22 studies (13.7%). Examples of re-interventions were coronary artery bypass graft surgery following PCI or PTCA 16 studies, revision following hip or knee replacement (3 studies), re-operation (1 study) and other (2 studies).

Thirty-four studies (19.3%) looked at complications. These were usually stroke and/or myocardial infarction (MI, 13 studies). Three studies used “adverse outcome”, defined as a stroke or death, and one study used “stroke or untoward outcome”. Nine studies used some other specific definition for complications, and eight studies examined “any” operative complication (that is, they did not specify what constituted a complication or did not specify complications which were specific to the procedure).

Twenty studies (12.4%) examined cost or charges. Nine explicitly examined charges and the rest either examined cost, or justified the use of charges as a proxy for cost on the basis that state regulation of hospital charges (as in some studies done in Maryland) keeps constant the charge-to-cost ratio.

A variety of other outcomes were measured in the articles we examined. Three studies of percutaneous coronary interventions evaluated the degree of residual coronary artery stenosis. Four studies examined discharge other than to home as an adverse outcome, and one study of the outcomes of obstetric care used 5-minute Apgar score of less than 7 and admission to a neonatal intensive care unit. Two studies of surgery for rectal cancer measured the number of procedures done without a colostomy (“sphincter sparing procedures”) as a proportion of the total number of rectal cancer operations, which otherwise involve complete removal of the rectum and anus and require a permanent colostomy. Two studies measured the accuracy of mammogram reading, one of which also measured the proportion of breast operations done using breast-conserving surgery. One study measured the rates of Caesarian delivery and vaginal delivery after a prior Caesarian section, and one study used consultation with an obstetrician as a measure of outcome. No study examined quality of life as an outcome.

Composite outcomes (permutations of events that constitute an adverse outcome) were measured in some studies. In two studies of percutaneous coronary interventions, either of coronary artery bypass surgery or death was considered an adverse outcome (“CABG or death”). In two studies that measured death as the principal outcome, death or a length of hospital stay greater than the 90^{th} percentile was measured as a secondary outcome. Two studies used the outcome of “any adverse event, including death, transfer, or complication”. One study of percutaneous coronary interventions measured the rate of “no death, myocardial infarction, or coronary artery bypass graft surgery, and residual stenosis less than 50%”.

Seventy-five studies (46.6%) examined more than one outcome. Most commonly the two outcomes of death and complications were examined (19 studies). Sixteen studies looked at death and length of hospital stay, and eight studies measured death, length of stay and cost/charges. Four studies with multiple outcomes did not involve death as one of the outcomes.

Measurement of pre-existing risk

Many patient characteristics, such as age, sex, severity of illness, coexisting medical conditions and socioeconomic status, affect the risk of adverse outcomes of health care. In studies that attempt to explain variation in patient outcome in terms of factors such as provider volume, it is important to ensure that apparent associations between volume and outcome are not unduly influenced by these patient-level characteristics. Importantly, patient characteristics affecting the risk of adverse outcomes can confound the results of an analysis only to the extent that they are distributed differently among providers or institutions. The degree to which patient severity of illness varies between providers and institutions is a matter of debate. However, studies that measure and control for these factors as completely as possible have more credibility than those studies in which confounding by differences in patient risk factors is likely. Statistical procedures used to control for pre-existing differences in subjects risk factors between providers are usually termed “risk-adjustment” or “case-mix adjustment”.[46]

Seven articles we examined (4.3%) did not examine pre-existing patient risk factors that might have affected outcome. A further three studies used a rudimentary method of accounting for pre-existing risk. One study stratified subjects according to whether or not they had had an acute myocardial infarction, one adjusted for the reason for the procedure, and one adjusted for case complexity as a binary variable (simple/complex).

Nine studies adjusted for age and for comorbidities but did not adjust for sex, and one study adjusted for sex and comorbidities but omitted age. The remaining studies adjusted for at least age and sex. In some cases age and sex were not explicitly included in a statistical risk-adjustment model because an overall risk score was calculated in which age and sex went into the calculations of risk. One study adjusted only for age and sex but no other variables.

We assessed the proportion of studies that measured clinical factors, defined as factors measured in individual persons by health care providers, such as such as blood pressure, left ventricular ejection fraction, and the presence of cardiogenic shock. Twenty-five studies (15.5%) adjusted for clinical factors.

A large majority of the studies (138, 85.7%) adjusted for acuity, comorbidities, or both. In our analysis, we defined ‘acuity’ as a measure of the severity of the problem for which the person was being treated, and ‘comorbidity’ as a measure of other diseases or conditions that might affect a subject’s risk of adverse outcomes. Twenty studies (12.4%) did not adjust for either acuity or comorbidities. Thirty-nine studies (24.2%) adjusted for acuity, 62 (38.5%) adjusted for comorbidities and 40 (24.8%) adjusted for both. Overall, 79 studies (49.1%) adjusted for acuity, 102 (63.4%) adjusted for comorbidity and 141 (87.6%) adjusted for something beyond age, sex and race.

A common difficulty in the analysis of administrative data is determining whether conditions other than the primary problem represent pre-existing comorbid diseases, or diseases that occurred as a consequence of the care provided. The reason for this is that comorbidities are typically measured in administrative databases by enumerating relevant codes in diagnosis fields that indicate diagnoses other than the one associated with the presenting problem. In some databases (for example, Medicare) pre-existing conditions are not distinguished from conditions whose onset occurred after hospitalization. For example, a diagnosis code indicating ‘acute myocardial infarction’ may not distinguish between pre-existing heart disease, or an acute myocardial infarction that occurred as a complication of care. Since risk-adjustment is typically done to “hold constant” the effect of comorbid conditions in an analysis, adjusting for comorbidities that are actually complications will diminish estimated associations between a measure of quality of care and outcome. Most studies using Medicare data addressed this problem by considering only hospitalizations occurring prior to the event of interest in the measurement of comorbidity.

Of the 102 studies that adjusted for comorbidities, 45 (44.1%) adjusted for individual comorbidities (that is, individual comorbidities were entered into the model). The most common comorbidity index used was the Charlson comorbidity index,[47] a measure that has been adapted for use in administrative data analysis.[48,49] Thirty-three articles (32.4%) used the Charlson index, 9 (8.8%) used some other index, and 14 (13.7%) used a count of the number of comorbidities.

One study (1.0%) did not indicate the form that the comorbidity adjustment took. In two instances, the “count” of the number of comorbidities was a simple none/one or more, or one/multiple dichotomy, and in several studies the “individual comorbidities” which were adjusted for consisted of a single condition (for example, diabetes mellitus).

Table 19. Adjustment for comorbid medical conditions in the studies in which comorbidity was measured

Method	Number	Percent
Individual comorbidities entered individually	45	44.1
Comorbidity index		
Charlson score	33	32.4
Other index	9	8.8
Count of comorbid conditions	14	13.7
Not stated	1	1.0
Total	**102**	**100.0**

Statistical analysis

Statistical model

The most common statistical method for estimating volume-outcome associations is by use of a regression model. Fitting a regression model to a data set allows an analyst to estimate the relationship between an exposure variable of interest (for example, volume) and an outcome, while taking into account a number of other factors that might influence the relationship between the exposure and the outcome.

Ten studies (6.2%) did not use a statistical model, and in two studies (1.2%), use of a statistical model was not felt to be appropriate. In studies that used a statistical model, the form of the model typically reflected the outcome being estimated. For example, analyses of the risk of short-term death or other dichotomous variables usually used logistic regression, and analyses of time to death used Cox proportional-hazards models or other survival analysis models.

There are several features of volume-outcome studies that pose important challenges for appropriate statistical analysis. In a typical analysis, a health outcome of interest is measured at the level of an individual patient. Since patients may differ considerably in “compositional” characteristics that also influence the outcome of interest (such as their age, sex or severity of illness), variation in these factors must also be accounted for in an analysis. Usually a simple regression modeling approach is used, with provider and institution characteristics attributed to the individual patient, and inferences of provider or institution effects are estimated in a multivariable regression model taking other potential confounders into account.

There are several problems with this approach. First, clustering of patients within providers and clustering of providers within institutions often results in correlation of individuals nested within higher-level units. The resulting lack of independence of observations in a conventional regression analysis of these data is a violation of an important assumption of regression models, and may result in underestimation of the standard errors of regression coefficients and therefore an overestimation of the statistical significance of apparent effects. Second, conventional statistical models assign provider characteristics to the patient-level, implicitly treating characteristics of the provider as though they were characteristics of the patient. This results in artificially inflating the amount of real information that is available about the impact of provider characteristics on patient outcomes. Finally, heterogeneity of effects is not explicitly modeled using conventional analysis, and may bias estimates of the regression coefficients if a single level analysis is done.[50-55]

Over half the studies (91, 56.5%) used a simple model that did not adjust for the potential of clustering of patients within hospitals and/or physicians. In this category, we also included some studies with rudimentary analyses such as chi-square tests, in addition to studies that analyzed data using regression models.

Studies used various methods to account for the fact that although outcomes are assessed at the level of individual patients, inferences about quality are made at the level of a provider or a hospital. Twenty-one studies (13.0%) used the hospital or the surgeon (in one instance) as the unit of analysis. Twenty-two studies (13.7%) used a modeling strategy that explicitly accounted for the clustering of patients within provider or hospital groups, either by use of hierarchical or mixed-effects models (identified if they explicitly describe appropriate statistical procedures such as SAS 'PROC MIXED' or software packages such as MLWin), or by use of generalized linear models and general estimating equations. Fifteen studies (9.3%) recognized the problem of clustering and used some other adjustment method of adjustment (for example, overdispersed binomial model, robust error estimation or some other indication that clustering had been thought of).

Evaluation of the statistical model

Since regression models vary in their ability to explain variation and to predict outcome, it is helpful to summarize the performance of a statistical model. Two important properties of regression models in the analysis of health care outcomes are discrimination and calibration. 'Discrimination' refers to the ability of a model to distinguish between subjects who experience an adverse outcome and those who do not, and may be measured by summary statistics such as a C-statistic. 'Calibration' refers to how closely the outcomes predicted by a model match the actual outcomes of subjects in the analysis, and is often summarized using a Hosmer-Lemeshow goodness of fit test (which compares predicted and observed outcomes according to deciles of predicted risk estimated by a logistic regression model).

Overall, 47 studies (29.2%) reported using some sort of goodness of fit or diagnostic evaluation of the statistical model used. In the 161 studies, 28 (17.4%) reported a C-statistic and/or an R-square value. Thirteen (8.1%) performed a Hosmer-Lemeshow goodness-of-fit test, and three (1.9%) checked the proportional hazards assumption of the Cox survival analysis. Twelve studies (7.5%) performed sensitivity analyses, analyzed subsets or used a validation sample to check the model for robustness in some way. Six studies (3.7%) reported some other form of model diagnostic evaluation (for example, assessed the residuals or examined for outliers).

Presentation of results

We used a hierarchical classification for the presentation of the results of the volume-outcome analysis in the studies we reviewed. If an article reported adjusted effects of any sort, we classified them as reporting adjusted effects. We defined adjusted measures as an 'adjusted measure of effect' (a measure of the degree to which the risk of an outcome changes according to different levels of volume, such as odds ratios, relative risks, rate ratios or hazard ratios), or an 'adjusted outcome' (a measure of the adjusted risk of the outcome according to volume categories). Studies that reported both adjusted effects and adjusted outcomes were classified as reporting adjusted measures of effect.

Almost half of the articles (77, 47.8%) reported adjusted effects. The reported measure of effect depended on the outcome and the model form (for example, odds ratios were estimated from logistic regression models and hazard ratios were estimated from Cox proportional-hazards models. Thirty-three studies (20.5%) reported adjusted outcomes. Thirteen (8.1%) reported observed/expected differences or observed/expected ratios. These studies generally reported indirectly standardized rates, or used the Z-score or two-stage approach, in which the difference between observed and expected deaths is determined for each hospital, based on patient risk factors. Thirty-five studies (21.7%) did not report adjusted results, but reported raw rates, or showed the results graphically. Three studies (1.9%) used some other form of reporting.

Table 20. Measure used to present the results of the volume-outcome analysis in the studies

Measure of estimated effect	Number	Percent
Adjusted measure of effect	77	47.8
Adjusted outcome	33	20.5
Observed-expected difference or ratio	13	8.1
Unadjusted outcome	35	21.7
Other	3	1.9
Total	**161**	**100.0**

Recommendations

Defining ‘best practices’ for volume-outcome research depends, to a large extent, on the goal of the research. In general, studies may seek to determine whether a volume-outcome association exists or may seek to predict the effect on health outcomes of altering the distribution of low volume providers or institutions. In either case, large datasets are usually required to provide stable estimates of relatively infrequent outcome events. Practically speaking, this implies the use of administrative data, although there are now several sources of augmented data sets, clinical databases and large voluntarily-reported procedure registries that are potential sources of high-quality data.

If the goal of an analysis is simply to test whether an association exists, volume should be specified initially as a continuous variable. If volume is categorized for an analysis, the categories should be specified *a priori* and ideally should reflect a meaningful theoretical framework rather than simply be based on the distribution of subjects among units with different clinical volumes. If volume is measured as total volume or average volume over a study period, there should be evidence that volume did not vary substantially over the study period. If volume does vary substantially, a more current measure of volume, such as provider volume during the 12 months prior to the treatment of each study subject, should be used. Outcomes should be selected carefully, and the choice of outcome should reflect an understanding of the importance of volume and related factors on the risk of adverse outcome. Statistical analysis should account for the possible confounding effects of all important measurable risk factors such as age, sex, comorbid conditions, severity of presenting illness, acuity and potentially also measures of socioeconomic status and race. Comorbid diseases should appear in risk adjustment models only if they reflect pre-existing disease and not complications or adverse effects of medical care. Incorporation of clinical data into the risk adjustment model may improve the model discrimination and fit, but there is no evidence that use of more detailed clinical data substantially affects the results of volume-outcome analyses. Analyses should account for the effect of clustering of subjects within providers and institutions. If provider volume is assessed in addition to hospital volume, a statistical model should account for both provider and hospital volume jointly, including their possible interaction effects, and preferably using a hierarchical or mixed model to model possible heterogeneity in outcome between different providers.

In studies linked closely to health policy measures such as regionalization, selection of the procedures and health conditions of interest and the choice of outcome are extremely important. Volume cut points defining low-volume units must be considered carefully and should balance the goals of regionalization with feasibility and the geographic distribution of hospitals and the population.

Studies seeking to determine whether there is a causal association between volume and outcome should use a design other than the cross-sectional design used in virtually all volume-outcome analyses. An excellent design to test hypotheses about causation is a quasi-experimental design, where the outcome of each study subject is assessed according to the current volume of the provider or hospital. Especially in cases where volumes change substantially over time, a quasi-experimental longitudinal design provides an excellent means of controlling for confounding factors and testing whether secular changes in volume within a hospital result in changes in outcome.

Appendices

Appendix 1. Search Strategy

We searched MEDLINE and EMBASE, limited to English-language studies done on human subjects, publication years 1980 to April 2004.

Search 1:
(volume [ti] OR frequent [ti] OR frequency [ti] OR statistics [ti]) AND (outcome [ti] OR outcomes [ti])

Search 2:
volume [ti] AND mortality [ti]

Search 3:
volume [ti] AND survival [ti]

Appendix 2. Summary of the Results of the Studies

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
Luft[1]	1979	Open heart surgery	US	1974-75	27,471		587	Outcome improves with higher volume.	< 0.001
		Vascular surgery	US	1974-75	44,786		1,309	Outcome improves with higher volume.	< 0.001
		TURP	US	1974-75	86,714		1,217	Outcome improves with higher volume.	< 0.001
		CABG	US	1974-75	17,165		182	Outcome improves with higher volume.	< 0.001
		Colectomy	US	1974-75	36,083		1,390	Outcome improves with higher volume.	< 0.001
		Biliary tract surgery	US	1974-75	8,957		1,278	No relationship between volume and outcome.	NS
		Total hip replacement	US	1974-75	16,339		804	Outcome improves with higher volume.	< 0.001
		AAA	US	1974-75	4,624		692	Outcome improves with higher volume.	< 0.001
		Vagotomy and/or pyloroplasty	US	1974-75	4,316		938	No relationship between volume and outcome.	NS
		Cholecystectomy and incision of common bile duct	US	1974-75	3,580		894	Outcome improves with higher volume.	< 0.05
		Vagotomy	US	1974-75	8,704		1,108	No relationship between volume and outcome	NS
		Cholecystectomy	US	1974-75	162,572		1,481	No relationship between outcome and volume	NS
Luft[40]	1987	AAA	US	1972	6,065		736	Supports a selective referral effect.	
		AMI	US	1972	98,066		906	Supports practice makes perfect.	
		Cirrhosis	US	1972	24,228		913	No relationship between volume and outcome	
		Fractured femur	US	1972	46,468		910	Supports a selective referral effect.	
		Peptic ulcer	US	1972	142,870		913	Supports a selective referral effect.	
		Respiratory distress syndrome	US	1972	16,373		770	Supports practice makes perfect.	
		Subarachnoid hemorrhage	US	1972	5,049		749	Inverse referral effect (poorer outcomes with increased volume).	
		Angiography and cardiac catheterization	US	1972	26,678		360	? (missing from outcome table).	
		Appendectomy	US	1972	80,211		916	Inverse referral effect (poorer outcomes with increased volume).	
		CABG	US	1972	5,172		114	Supports a selective referral effect.	

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
		Cholecystectomy	US	1972	102,917		914	Supports practice makes perfect.	
		Hernia repair	US	1972	134,497		920	No relationship between outcome and volume.	
		Hysterectomy	US	1972	180,464		915	Supports both practice makes perfect and selective referral.	
		Intestinal operations	US	1972	36,860		898	Supports practice makes perfect.	
		Stomach operations	US	1972	24,072		864	Supports practice makes perfect.	
		Total hip replacement	US	1972	20,429		730	Supports both practice makes perfect and selective referral.	
		TURP	US	1972	657		756	Supports a selective referral effect.	
Flood[5]	1984	Intra-abdominal artery operations	US	1972	9,532		645	Outcome improves with higher volume.	< 0.001
		Total hip replacement	US	1972	13,424		702	Results not given.	
		Gallbladder operations	US	1972	130,749		1,196	Outcome improves with higher volume.	< 0.005
		Stomach operations for ulcer	US	1972	26,688		1,100	No effect of volume on outcome.	> 0.05
		Large bowel operations	US	1972	16,110		984	Results not given.	
		Hip fracture surgery with other trauma	US	1972	6,925		886	Results not given.	
		Hip fracture surgery, no other trauma	US	1972	52,368		1,169	No effect of volume on outcome.	> 0.05
		Amputation of lower limb	US	1972	10,267		973	Outcome improves with higher volume.	< 0.05
		Amputation of lower limb with current trauma	US	1972	881		217	Results not given.	
		Stomach operations, cancer diagnosis	US	1972	1,500		377	Results not given.	
		Stomach operations, non-cancer	US	1972	7,148		875	Results not given.	
		Large bowel operations, cancer	US	1972	17,872		1,040	Outcome improves with higher volume.	< 0.05
		Large bowel operations, other diagnoses	US	1972	6,575		858	Results not given.	
		Pelvic fracture	US	1972	18,033		1,113	Results not given.	
		Femur shaft fracture	US	1972	13,677		976	Results not given.	
		Gall bladder diagnosis, non-surgical	US	1972	88,839		1,210	Outcome worsens with higher volume.	< 0.05

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
		Ulcer diagnosis, non surgical	US	1972	138,268		1,214	No relationship between volume and outcome.	> 0.05
Flood[6]	1984	9 categories of surgical patients, combined	US	1972	266,944		1,196	Outcome improves with higher volume.	0.0001
		Gall bladder and ulcer, combined	US	1972	227,107		1,215	Outcome improves with higher volume.	0.0001
Thiemann[56]	1999	AMI	US	1994-95	98,898		n/a	Outcome improves with higher volume.	< 0.001
Sainsbury[57]	1995	Invasive breast cancer	England	1979-88	12,861	180	n/a	Outcome improves with higher volume, surgeon.	~0.0004
Simunovic[58]	2000	Resection, rectal cancer	Canada	1990	1,072		124	No relationship between volume and outcome.	
Brook[59]	1990	Carotid endarterectomy	US	1981	1,171	n/a	n/a	No relationship between volume and outcome.	0.9
Karp[60]	1998	Carotid endarterectomy	US	1993	1,945		n/a	Outcome improves with higher volume.	0.02 for the effect of volume using a chi-square trends test (i.e., not adjusted)
Roohan[61]	1998	Surgery for breast cancer	US	1984-89	47,890		266	Outcome improves with higher volume.	< 0.0001
Simons[62]	1997	Sphincter-sparing procedures, resection	US	1988-92	1,731			Outcome improves with higher volume.	< 0.001
Schrag[63]	2000	Resection, colon cancer	US	1991-96	27,986		611	Outcome improves with higher volume.	< 0.001
Malenka[64]	1999	PCI	US	1994-96	15,080	47	5	No relationship between volume and outcome.	non-emergency CADG, p = 0.051 MI: p = 0.065
Mcgrath[65]	2000	PCI, adjusting for use of coronary stent	US	1997	167,208	6,534	1,003	Low physician volume increases risk of CABG, no effect on mortality. Low hospital volume increases risk of mortality, no effect on CABG, independent of use of stent. Hospital volume more important than	< 0.001

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
								physician volume, when looking at the combined mortality/CABG outcome.	
Hannan[66]	1991	CABG	US	1989	12,448	126	30	Outcome improves with higher volume, surgeon and hospital.	Better outcomes with both higher surgeon volume and higher hospital volume
Ho[67]	2000	PTCA	US	1984-96	353,488		129	Outcome improves with higher volume.	
Hsia[68]	1998	Carotid endarterectomy	US	1985-96	900,000		2,700	Outcome improves with higher volume.	< 0.001
Hamilton[69]	1998	Hip fracture surgery	Canada	1990-93	7,483		68	No relationship between volume and outcome.	0.22 for LOS, 0.33 for death
Farley[70]	1992	AMI	US	1980-87	974,803		426	Outcome improves with higher volume.	< 0.0001
		Hernia repair	US	1980-87	37,041		330	Increased volume <u>within</u> a hospital leads to reduced mortality	0.032
		CABG	US	1980-87	146,890		62	Increased volume <u>within</u> a hospital leads to reduced mortality; correlation between volume and the residuals indicates selective referral.	< 0.0001
		Total hip replacement	US	1980-87	130,494		337	No effect of volume on outcome.	0.16
		Respiratory distress syndrome in neonates	US	1980-87	56,014		222	Increased volume <u>within</u> a hospital leads to reduced mortality.	0.039
Bennett[71]	1989	PCP infection	US	1986-87	257	15		Outcome improves with higher volume.	< 0.01
Bennett[72]	1992	PCP infection	US	1987	3,126		73	Outcome improves with higher volume.	< 0.001
Stone[73]	1992	AIDS	US	1987-88	300		41	Outcome improves with higher volume.	< 0.05
Turner[74]	1992	AIDS	US	1986-87	10,538		258	Outcome improves with higher volume.	0.009
Hogg[75]	1998	AIDS	Canada	1987-93	38,075		513	Outcome improves with higher volume.	0.014
Cunningham[76]	1999	AIDS	US	1994	7,901		333		< 0.0001
Ritchie[77]	1993	PTCA	US	1989	24,883		110	No relationship between volume and outcome.	For CABG, < 0.001; for mortality 0.42 for

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
									AMI and 0.98 for non-AMI groups
Ritchie[78]	1999	PTCA	US	1993-94	163,527		214	Outcome improves with higher volume.	< 0.01
Jollis[79]	1994	PTCA	US	1987-90	217,836		1,194	Outcome improves with higher volume.	< 0.001
Jollis[80]	1997	Coronary angioplasty	US	1992	97,478	6,115	984	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	< 0.001
Phillips[81]	1995	PTCA	US	1989	24,856		110	Outcome improves with higher volume.	< 0.001 for all outcomes?
Hannan[82]	1997	PTCA	US	1991-94	62,670		31	Outcome improves with higher volume.	
Maynard[83]	1999	PTCA	US	1993, 1996	39,195			No relationship between volume and outcome.	< 0.001
Showstack[84]	1987	CABG	US	1983	18,996		77	Outcome improves with higher volume.	Poor outcome: < 0.001
Hannan[85]	1989	CAGB	US	1986	9,774	353	27	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	
		AAA	US	1986	1,635	508	170	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	
		Partial gastrectomy	US	1986	1,342	828	216	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	
		Colectomy	US	1986	10,297	1,997	250	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
		Cholecystectomy	US	1986	25,091	2,322	253	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	
Grumbach[86]	1995	CABG	US/ Canada	1987-89	43,837		157	In NY and CA, adjusted mortality decreased as volume increased. In Canada (no very low-volume hospitals) there was no volume-outcome effect.	NY, CA < 0.001. Canada 0.59
Shroyer[87]	1996	CABG	US	1987-92	22,021		44	No relationship between volume and outcome.	Patient level analysis: 0.10
Sollano[88]	1999	CABG	US	1990-95	97,137		31	No relationship between volume and outcome.	
		AAA elective	US	1990-95	9,847		195	Outcome improves with higher volume, hospital.	< 0.001
		Repair of congenital heart defects	US	1990-1995	7,199		16	Outcome improves with higher volume, hospital.	< 0.001
Hannan[89]	1998	Repair of congenital heart defects	US	1992-95	7,169		16	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	
Jenkins[90]	1995	Repair of congenital heart defects	US	1989	2,833		37	Outcome improves with higher volume, hospital.	Mortality 0.01
Edwards[91]	1991	Carotid endarterectomy	US	1979-88	11,199	190	n/a	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	Physician volume: 0.02 for mortality, 0.008 for stroke, < 0.001 for LOS and cost. Hospital volume: < 0.001 for LOS and

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
Cebul[92]	1998	Carotid endarterectomy	US	1983-84	678	478	115	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	0.004 for cost. 0.006
Fisher[93]	1989	Carotid endarterectomy	US	1983-84	2,089		139	Outcome improves with higher volume, hospital.	0.03 when volume modeled as continuous variable.
Hannan[94]	1998	Carotid endarterectomy	US	1990-95	8,207	462	153	Outcome improves with higher volume, hospital. Outcome improves with higher volume, surgeon.	< 0.05
Kantonen[95]	1998	Carotid endarterectomy	Finland	1991-94	1,600	104	23	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	0.005 for deviance between Lowess curve and line showing no volume effect.
Kempczinski[96]	1986	Carotid endarterectomy	US	1983-84	656	61	16	No relationship between volume and outcome.	
Khuri[97]	1999	AAA	US	1991-93	3,767		107	No relationship between volume and outcome.	
		Infrainguinal vascular reconstruction	US	1991-93	12,535		107	No relationship between volume and outcome.	
		Carotid endarterectomy	US	1991-93	10,173		93	No relationship between volume and outcome.	
		Lobectomy/ pneumonectomy	US	1991-93	4,890		107	No relationship between volume and outcome.	
		Open cholecystectomy	US	1991-93	7,113		124	No relationship between volume and outcome.	

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
		Lap cholecystectomy	US	1991-93	8,602		123	No relationship between volume and outcome.	
		Colectomy	US	1991-93	13,310		125	No relationship between volume and outcome.	
		Total hip replacement	US	1991-93	8,241		101	No relationship between volume and outcome.	
Kirshner[98]	1989	Carotid endarterectomy	US	1984-85	1,035	22	6	No relationship between volume and outcome.	
Manheim[99]	1998	Lower extremity arterial bypass	US	1982-94	100,963		27	Outcome improves with higher volume, hospital.	< 0.001
		Carotid endarterectomy	US	1982-94	106,493		27	Outcome improves with higher volume, hospital.	< 0.001
		AAA, unruptured	US	1982-94	35,130		27	Outcome improves with higher volume, hospital.	0.02
		AAA, ruptured	US	1982-94	7,327		27	Outcome improves with higher volume, hospital.	< 0.001
Perler[100]	1998	Carotid endarterectomy, elective	US	1990-95	9,918		48	Outcome improves with higher volume, hospital.	< 0.0001
Richardson[101]	1989	Carotid endarterectomy	US	1983-94	705	98	41	No relationship between volume and outcome.	0.28
Wennberg[102]	1998	Carotid endarterectomy	US	1992-93	113,300		2,699	Outcome improves with higher volume, hospital.	< 0.001
Katz[103]	1994	AAA, unruptured	US	1980-90	8,185		n/a	Outcome improves with higher volume, hospital.	< 0.001
		AAA, ruptured	US	1980-90	1,829		n/a	Outcome improves with higher volume, hospital.	< 0.001
Hannan[104]	1992	AAA, unruptured	US	1985-87	3,570	391	152	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	0.005
		AAA, ruptured	US	1985-87	954	236	134	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	0.0002
Amundsen[105]	1990	AAA, both	Norway		444		26	Outcome improves with higher volume, hospital.	Ranges from NS to 0.03.

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
Wen[106]	1996	AAA, ruptured	Canada	1988-92	1,203			No relationship between volume and outcome.	Mortality 0.02, LOS < 0.0001
		AAA, unruptured	Canada	1988-92	5,492			Outcome improves with higher volume.	Mortality 0.02, LOS < 0.0001
Kazmers[107]	1996	AAA, both	US	1991-93	3,687		116	Outcome improves with higher volume, hospital.	~0.003
Rutledge[108]	1996	AAA, ruptured	US	1988-93	1,480	n/a	157	Outcome improves with higher volume, hospital.	0.025
Dardik[109]	1998	AAA, ruptured	US	1900-95	527	226	47	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	Hospital volume ranges from 0.1 to 0.8; surgeon volume ranges from 0.05 to 0.46
Begg[110]	1998	Pancreatectomy, cancer	US	1984-93	742		258	Outcome improves with higher volume.	0.004
		Esophagectomy, cancer	US	1984-93	503		190	Outcome improves with higher volume.	< 0.001
		Pneumonectomy, cancer	US	1984-93	1,375		313	No relationship between volume and outcome.	0.32
		Resection, hepatic, cancer	US	1984-93	801		286	Outcome improves with higher volume.	0.04
		Pelvic exenteration, cancer	US	1984-93	1,592		347	Outcome improves with higher volume.	0.04
Gordon[111]	1995	Whipple	US	1988-93	501		39	Outcome improves with higher volume, hospital.	< 0.001
Birkmeyer[112]	1999	Whipple	US	1992-95	7,229		1,772	Better survival in higher volume hospitals, volume particularly important after treatment for benign pancreatic disease.	< 0.001
Birkmeyer[113]	1999	Whipple	US	1992-95	7,229		1,243	Better outcomes with higher volume.	< 0.001
Imperato[114]	1996	Whipple	US	1991-94	579		118	Mortality lowest in regional centres; among low-volume hospitals, mortality decreases with increasing volume. LOS shorter in the regional hospitals.	Mortality 0.006; LOS 0.0007.

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
Lieberman[115]	1995	Whipple or total pancreatectomy	US	1984-91	1,972	748	184	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	Hospital volume 0.05.
Sosa[116]	1998	Whipple	US	1990-95	449	n/a	48	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	< 0.001
		Palliative bypass (pancreatic cancer)	US	1990-95	589	n/a	48	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	0.03
		Endoscopic or percutaneous stent (pancreatic cancer)	US	1990-95	198	n/a	48	No relationship between volume and outcome.	
Simunovic[117]	1999	Whipple	Canada	1988-94	842		68	Outcome improves with higher volume.	< 0.001
Wade[35]	1996	Whipple	World-wide (mostly US)	1989-94	130		n/a	Outcome worsens with higher volume.	0.05
Gordon[118]	1999	Excision of esophagus	US	1989-97	518		51	Outcome improves with higher volume.	< 0.001
		Gastrectomy	US	1989-97	705		51	No relationship between volume and outcome.	
		Colectomy	US	1989-97	1,015		51	No relationship between volume and outcome.	
		Hepatic lobectomy	US	1989-97	293		51	Outcome improves with higher volume.	< 0.01
		Biliary tract surgery, anastomosis	US	1989-97	938		51	Outcome improves with higher volume.	< 0.01
		Whipple	US	1989-97	1,092		51	Outcome improves with higher volume.	< 0.001
Patti[119]	1998	Esophagectomy	US	1990-94	1,561		273	Mortality lowest in highest volume group; % discharged home and total charges increase with increasing volume; LOS and complications don't vary with volume, but mortality following complications varies significantly with volume (lowest in highest volume group).	< 0.001

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
Harmon[120]	1999	Resection, colorectal	US	1992-96	9,739	812	50	Outcome improves with higher volume, hospital or surgeon.	< 0.01
Parry[121]	1999	Primary colorectal cancer	UK	1993	927	112	39	No relationship between volume and outcome, hospital or surgeon.	
Hannan[122]	2002	Colectomy	US	1994-97	22,128	2,052	229	Outcome improves with higher volume, hospital or surgeon.	< 0.001
		Lobectomy	US	1994-97	6,954	373	178	Outcome improves with higher volume, hospital or surgeon.	0.006
		Gastrectomy	US	1994-97	3,711	1,114	207	Outcome improves with higher volume, hospital or surgeon.	< 0.001
Romano[123]	1992	Resection, pulmonary	US	1983-86	12,439		389	Generally better outcome with increased hospital volume.	< 0.001 to 0.05, depending on procedure
Taylor[124]	1997	Major hip and knee surgery	US	1992-94	632,533		3,842	Outcome improves with higher volume.	< 0.01
		Minor hip and knee surgery	US	1992-94	270,296		3,866	Outcome improves with higher volume.	
		Spine procedures	US	1992-94	105,740		2,456	Outcome improves with higher volume.	
Lavernia[125]	1995	Primary hip and knee replacement	US	1992	19,925	n/a	n/a	Outcomes improve with higher volume.	
		Revision hip and knee replacement	US	1992	2,536	n/a	n/a	Outcomes improve with higher volume.	
Kreder[126]	1998	Total hip replacement	Canada	1992	3,645	329	90	No relationship between volume and outcome, hospital or surgeon.	> 0.05
Kreder[127]	1997	Total hip replacement	US	1988-1991	7,936	494	67	Low-volume surgeons fared worst, little effect of hospital volume, no surgeon*hospital volume interaction.	
Ellison[128]	2000	Radical prostatectomy	US	1989-95	66,693			Mortality was low in high-volume hospital , low- and medium-volume hospitals didn't differ; and the effect of volume was consistent for all patient age groups. LOS fell with increasing volume. Costs were low at high-volume hospitals, low- and medium-volume hospitals didn't differ from one another.	< 0.001

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
Yao[129]	1999	Prostatectomy	US	1991-94	101,604		2,849	Better outcomes with increased hospital volume.	< 0.001
Thorpe[130]	1994	TURP	UK	1991	1,396		12	Better outcomes with increased hospital volume.	< 0.001
Taylor[131]	1997	Craniotomy, aneurysmal subarachnoid hemorrhage	US	1984-91	47,408		n/a	For surgically treated patients, better outcome with high volume and academic affiliation; for medically treated patients, probably no relationship with volume.	0.01 for surgical
Solomon[132]	1996	Craniotomy, SAH	US	1987-93	15,376		110	Better outcomes with higher volume.	< 0.01
Pronovost[133]	1999	AAA	US	1994-96	2,606	n/a	39	Surgeon volume was not significant low hospital volume associated with higher mortality but shorter time in ICU, and not significantly related to total LOS.	Relating hospital volume to mortality, 0.0006; relating hospital volume to ICU stay, 0.05; relating hospital volume to total LOS, 0.19.
Kitahata[134]	1996	AIDS	US	1984-94	403	125	1	Survival improved with physician experience.	0.02
Casale[135]	1998	AMI	US	1993	30,205	n/a	n/a	Lower mortality with high-volume physicians, no effect of hospital volume.	0.03
McGrath[136]	1998	PCI	US	1990-93	12,118	31	5	Angiographic and clinical success rates improved with higher volume; post-operative MI had an upward trend with increased volume but CABG had downward trend; no effect of volume on mortality.	Angio-graphic success 0.006; clinical success 0.03; new MI 0.06;

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
									CABG < 0.001; mortality 0.8
Ruby[137]	1996	Carotid endarterectomy	US	1985-91	3,880	226	32	Major complication and length of stay fell with increasing volume. However, these findings may be spurious, due to failure to adjust for surgeon specialty.	0.002 for major complica-tion vs volume; 0.001 for long LOS vs volume.
Porter[138]	1998	Resection, low anterior or abdominoperineal	Canada	1983-90	683	52	5	Better survival (overall disease specific mortality and time to local recurrence) better with high volume, and with specialist training; no interaction between volume and training. Perioperative mortality not affected by either volume or training.	Time to disease-specific death: 0.03 for training and 0.006 for volume. Time to local recurrence: 0.001 for training and < 0.001 for volume.
Hannan[139]	1995	CABG	US	1978-92	57,187	131	30	Worse results for lower volume surgeons (this was not the main purpose of the study). Among low-volume surgeons there are subgroups (formerly high volume, experienced but new to NY, and those who have been low volume for an extended period of time). Of these, the 'always low-volume' did worst.	Not stated.
O'Neill[140]	2000	Carotid endarterectomy	US	1984-95	12,725	482	153	No effect of surgeon volume on mortality, overall 'bad outcome' higher for volume of	NS for volume

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
								1-2, but then flat.	NS; 0.01 for total bad out-come.
Tu[141]	1998	Carotid endarterectomy	US/ Canada	1992-93				Outcomes better in historically high volume, high quality than in low/low.	CA (< 0.05), ON (< 0.05), not NY (0.30)
Hamilton[38]	1997	Hip fracture surgery	Canada	1990-93	7,383		66	Apparent volume-outcome relationship appears to be due to quality.	
Hughes[142]	1987	Cardiac catheterization	US	1983	76,584	2,987	150	Outcome improves with higher volume, hospital and surgeon.	< 0.01
		Appendectomy	US	1983	39,545	6,434	646	Outcome improves with higher volume, hospital and surgeon.	< 0.05
		CABG	US	1983	29,503	800	120	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	< 0.01
		Cholecystectomy	US	1983	80,587	7,062	742	No relationship between volume and outcome, hospital or surgeon.	
		Hernia repair	US	1983	78,377	7,476	742	Outcome improves with higher volume, hospital and surgeon.	< 0.01
		Hysterectomy	US	1983	105,550	8,027	736	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	< 0.01
		Intestinal operations	US	1983	28,486	5,436	708	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	< 0.01
		Stomach operations	US	1983	9,442	3,735	656	No relationship between volume and outcome, hospital or surgeon.	
		Total hip replacement	US	1983	13,767	2,301	501	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	< 0.01

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
		TURP	US	1983	41,821	2,892	631	Outcome improves with higher volume, hospital and surgeon.	< 0.01
Kelly[143]	1987	CABG	US	1977	3,883	99	26	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	0.0008
		Cardiac catheterization	US	1977	4,835	145	39	Outcome improves with higher volume, hospital. No relationship between volume and outcome, surgeon.	0.054
		AMI	US	1977	11,033	926	146	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	0.003
Zelen[144]	1991	CABG	US	1986	3,408	49	8	No low-volume hospital had high-volume surgeons. Trend to poorer outcomes with low-volume hospitals, but not significant.	
Hosenpud[145]	1994	Cardiac transplant	US	1987-91	7,878		150		< 0.001
Goodney[146]	2003	Carotid endarterectomy	US	1994-99	n/a		n/a	Outcome improves with higher volume.	
		Lower extremity arterial bypass	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		AAA, elective	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		CABG	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Aortic valve replacement	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Mitral valve replacement	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Colectomy	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Gastrectomy	US	1994-99	n/a		n/a	Effect of volume depends on procedure.	
		Esophagectomy	US	1994-99	n/a		n/a	Outcome improves with higher volume.	
		Resection, pancreatic	US	1994-99	n/a		n/a	No relationship between volume and outcome.	

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
		Nephrectomy	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Lobectomy, pulmonary	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Pneumonectomy	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
		Cystectomy	US	1994-99	n/a		n/a	No relationship between volume and outcome.	
Birkmeyer[147]	2002	Carotid endarterectomy	US	1998-99	136,049	8,818	n/a	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	
		CABG	US	1998-99	220,592	2,772	n/a	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	
		Aortic valve replacement	US	1998-99	42,541	2,440	n/a	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	
		AAA, elective	US	1998-99	39,794	6,276	n/a	Outcome improves with higher volume, hospital and surgeon.	
		Resection, lung cancer	US	1998-99	24,092	4,178	n/a	Outcome improves with higher volume, hospital and surgeon.	
		Cystectomy	US	1998-99	6,340	2,918	n/a	Outcome improves with higher volume, hospital and surgeon.	
		Esophagectomy	US	1998-99	1,640	997	n/a	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	
		Resection, pancreatic	US	1998-99	3,060	1,624	n/a	Outcome improves with higher volume, hospital and surgeon.	
Birkmeyer[148]	2002	CABG	US	1994-99	901,667		1068	Outcome improves with higher hospital volume.	< 0.0001
		Aortic valve replacement	US	1994-99	151,610		1,069	Outcome improves with higher hospital volume.	< 0.0001

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
		Mitral valve replacement	US	1994-99	64,935		1,050	Outcome improves with higher hospital volume.	< 0.0001
		Carotid endarterectomy	US	1994-99	479,289		2,990	Outcome improves with higher hospital volume.	0.004
		Lower extremity arterial bypass	US	1994-99	266,570		3,184	Outcome improves with higher hospital volume.	< 0.0001
		AAA, elective	US	1994-99	139,850		2,819	Outcome improves with higher hospital volume.	< 0.0001
		Colectomy	US	1994-99	299,960		4,587	Outcome improves with higher hospital volume.	< 0.0001
		Gastrectomy	US	1994-99	31,435		3,423	Outcome improves with higher hospital volume.	< 0.0001
		Esophagectomy	US	1994-99	6,337		1,575	Outcome improves with higher hospital volume.	< 0.0001
		Resection, pancreatic	US	1994-99	12,105		1,868	Outcome improves with higher hospital volume.	< 0.0001
		Nephrectomy	US	1994-99	59,495		3,292	Outcome improves with higher hospital volume.	0.01
		Cystectomy	US	1994-99	22,349		2,422	Outcome improves with higher hospital volume.	< 0.0001
		Resection, pulmonary	US	1994-99	74,080		2,753	Outcome improves with higher hospital volume.	< 0.0001
		Pneumonectomy	US	1994-99	10,410		1,877	Outcome improves with higher hospital volume.	< 0.0001
de Gara[149]	2003	Gastrectomy	Canada	1991-97	~550	84	4	No effect of volume (using either cut point) on (unadjusted) 30-day or 5-year survival.	
Hodgson[150]	2003	Resection, rectal	US	1994-97	7,257		367	Outcomes better with high volume; high volume favors all patients equally (not just some subsets).	< 0.0001
Harcourt[151]	2003	Breast cancer	US	1980-94	2,409		9	There was no correlation between volume and mortality.	
Bach[152]	2001	Resection, lung cancer	US	1985-96	2,118		76	Higher volume associated with better outcomes.	0.02 - 0.002
Margulies[153]	2001	Trauma	US	1998-99	1,754	86	5	Volume not associated with outcome. Caseload not significant in even the subgroup of patients for whom surgeon experience should have mattered the most.	0.44

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
Canto[154]	2000	Angioplasty	US	1994-98			422	For angioplasty, outcome is better in high volume hospitals. For thrombolytic therapy, outcome does not depend on volume.	< 0.0001 for angio-plasty, 0.67 for throm-bolytics.
Konvolinka[155]	1995	Trauma	US	1988-89	36,346		24	Survival improves with increasing surgeon experience with serious trauma patients overall, and for the blunt injured subset, but not for penetrating or pediatric patients.	
Glance[156]	2003	CABG	US	1996	20,078		32	For the 95% of patients with risk < 9%, the benefit from high volume is greatest for those at least risk; for those with risk > 25%, benefit grows as risk grows; for those in the middle, no benefit due to volume.	< 0.001 for volume* risk interaction.
Barker[157]	2003	Transphenoidal surgery	US	1996-2000	5,497	825	538	Mortality, complication rate, discharge home, and LOS better with higher volume (hospital or physician). Hospital charges don't vary by volume age*volume interaction: older patients benefit most from high volume.	0.02 to < 0.001
Hoh[158]	2003	Endovascular therapy	US	1996-2000	421	75	81	Hospital mortality: significantly related to only one of the 8 definitions of hospital (not surgeon) volume. Discharge home: better outcomes with increased volume, no matter how defined, both hospital and physician. LOS shorter and charges less at high volume hospitals. Improvement with volume was continuous - no cut point was found.	Mortality: 0.04; Discharge home, hospital volume, < 0.001; LOS 0.001.
Finlayson[159]	2003	Colectomy	US	1995-97	120,270		1,082	No relationship between volume and outcome.	
		Gastrectomy	US	1995-97	16,081		911	No relationship between volume and outcome.	
		Esophagectomy	US	1995-97	5,282		603	Outcome improves with higher volume.	< 0.0001
		Resection, pancreatic	US	1995-97	3,414		483	Outcome improves with higher volume.	< 0.0001

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
		Nephrectomy	US	1995-97	23,278		820	No relationship between volume and outcome.	
		Cystectomy	US	1995-97	4,937		590	No relationship between volume and outcome.	
		Lobectomy, pulmonary	US	1995-97	11,000		674	No relationship between volume and outcome.	
		Pneumonectomy	US	1995-97	11,000		674	No relationship between volume and outcome.	
Marrie[160]	2003	Community acquired pneumonia	Canada	1994-99	43,642		121	No effect of volume on 3-day death. Lower in-hospital mortality with higher physician volume. One-year mortality higher with increased physician volume (their claim).	< 0.0001 for in-hospital mortality.
Urbach[161]	2003	Esophagectomy	Canada	1994-98	613		47	Outcome improves with higher volume.	0.04
		Resection, colorectal	Canada	1994-98	18,898		134	No relationship between volume and outcome.	0.54
		Whipple	Canada	1994-98	586		49	No relationship between volume and outcome.	0.08
		Pneumonectomy	Canada	1994-98	5,156		54	No relationship between volume and outcome.	0.07
		AAA, unruptured	Canada	1994-98	6,279		57	Outcome improves with higher volume.	< 0.01
Sava[162]	2003	Trauma, severe	US	1990-2001	20,695	21	1	Low-volume physicians had higher mortality for 4 of the 6 injury patterns, but was never significant. Z-score for high-volume physicians was 1.88 vs 0.5 for low-volume (high value means fewer deaths than predicted, so both groups had "too few" deaths), but p-value for z = 1.88 is 0.06, so not significant.	0.06
Dimick[163]	2003	Aorto-bifemoral bypass	US	1997	3,073		483	Hospital mortality improved in high-volume hospital, no effect of volume on probability of long LOS.	0.04
Barker[164]	2003	Surgical treatment of Unruptured intracranial aneurysm	US	1996-2000	3,498	585	463	Hospital mortality: no relation to volume (surgeon or hospital). Discharge home: better outcome with higher volume (when tested jointly, only hospital volume significant).	Discharge home: 0.02; for hospital volume

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of			Results/Comments	P-value
					Subj	Phys	Hosp		
								Most adverse outcomes not related to volume: when they were, higher volume hospitals were better. LOS did not vary with volume. Charges were higher in high-volume hospital. Optimal cut-point for hospital volume was 2 clippings/year; no single cut-point for surgeon volume.	charges: 0.005
Ho[165]	2003	Whipple	US	1988-98	6,652		500	Mortality decreases with increased volume years of experience independently predicts mortality and experience mitigates some of the disadvantages of a lower volume hospital.	< 0.001
London[166]	2003	Injury	US	1998-99	98,245		38	Little effect of volume on mortality except for most severely injured (ISS > 15). More seriously injured patients do worse at high volume centres, and also have longer LOS	Data to support this are not given.
Cowan[167]	2003	Resection, intracranial tumor	US	1996-97	7,547	657	379	High volume (physician or hospital) associated with better outcome. High volume associated with decreased LOS (no other information was provided). Correlation between hospital volume and physician volume = 0.5	Hospital volume 0.04; surgeon volume 0.01
Damhuis[168]	2002	Resection, gastric cancer	Nether-lands	1987-97	1,978		22	Only 2 hospitals differed from the others (p ~ 0.05 without controlling for multiple testing) and the worst hospital had medium volume. Since hospitals were included in the model using dummy variables, volume was not tested. The graphics show no relationship between volume and outcome.	
Spiegelhalter[169]	2001	Repair of congenital heart defects, open heart surgery	England	1991-94	n/a		12	Outcome improves with higher volume.	< 0.001
		Repair of congenital heart defects, closed heart surgery	England	1991-94	n/a		12	Ambiguous.	
Tu[170]	2001	AMI	Canada	1992-97	98,194	5374	195	Hospital volume not a predictor after adjusting for physician volume.	< 0.001

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
								Outcomes improve with increasing volume; no apparent threshold after which volume doesn't matter; volume effect depends on physician specialty.	
van Lanschot[171]	2001	Esophagectomy	Nether-lands	1993-98	1,900		n/a	Better outcomes with increased hospital volume, despite the fact that high volume centers get more advanced tumors.	< 0.0001
Magid[172]	2000	Primary angioplasty	US	1994-99	62,299		446	Volume*procedure interaction: in low volume hospital, no association between treatment choice and outcome; in medium and high volume hospital, angioplasty patients did better. No volume effect on stroke, CABG decreased as volume increased graphically: mortality after angioplasty decreases with volume, probably no volume-outcome relation for thrombolytics.	< 0.001 for improve-ment in outcome due to angioplasty in medium- and high-volume hospital
Swisher[173]	2000	Esophagectomy	US	1994-96	340		25	Mortality and cost depend specifically the volume of esophagectomies (not overall surgical volume).	0.03 (mortality); 0.003 (charges)
Cooper[174]	2000	Trauma, inpatients	US	1994-95	26,973		43	Using ACS criterion, high and low volume centres do not differ. Using tertiles of high severity patients, the middle tertile had higher adjusted mortality than expected. Using tertiles of total volume, the low volume tertile had fewer deaths than expected and the highest volume had high mortality.	Using volume of severely injured, 0.03 for the middle-volume tertile; using total volume, 0.01 for the low-volume group and 0.06 for the high-

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
									volume group
Edwards[175]	1999	Liver transplant	US	1992-94	7,363		99	Mortality decreases with increasing volume up to 20 transplants/year and then stabilizes. Low-volume centers affiliated with high-volume centers do just as well as high-volume centers, therefore, centers were dichotomized into low-volume unaffiliated vs all others.	< 0.001
Kantonen[176]	1997	AAA, elective	Finland	1991-94	929	105	25	No relationship between volume and outcome, hospital. Outcome improves with higher volume, surgeon.	0.01
		AAA, ruptured	Finland	1991-94	454	105	25	No relationship between volume and outcome, hospital or surgeon.	
Phibbs[177]	1996	Birth, singleton at risk for NICU admission	US	1990	53,229		n/a	Compared to high-volume level III NICU, all others (including low-volume level III) have higher mortality; costs and LOS do not vary with level or size.	< 0.05
Bates[178]	1996	Open heart valvuloplasty	US	1987-89	2,633		n/a	Outcome improves with higher volume	< 0.01
		CABG	US	1987-89	5,635		n/a	No relationship between volume and outcome	
		Colectomy with cancer	US	1987-89	5,324		n/a	No relationship between volume and outcome	
		Colectomy w/o cancer	US	1987-89	2,299		n/a	No relationship between volume and outcome	
		Amputation above the knee	US	1987-89	5,359		n/a	No relationship between volume and outcome	
Kagan[179]	1994	Trauma, head injury	US	1986-88	4,667		10	Better outcomes with increased hospital volume.	< 0.01
Mayfield[180]	1990	Birth, white singleton	US	1990-96	226,164		90	They were not able to interpret their model. They claim that infants do best in high-volume facilities, but their outcome (deaths saved) doesn't support their claim.	
Iapichino[181]	2004	ICU	Europe	1994-95	12,615		89	Better outcomes with higher volume.	< 0.0001
Schrag[182]	2003	Resection, colon cancer	US	1991-96	24,166			Both hospital and surgeon volume significant for all outcomes.	Hospital volume <

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
								Effect of surgeon volume attenuated when hospital volume added.	0.001; surgeon volume ~ 0.02.
Brown[183]	2003	PTCA w/o stent	US	1997	18,940		140	Outcome improves with higher volume.	< 0.001
		PTCA with stent	US	1997	25,336		130	Outcome improves with higher volume.	0.01
Darling[184]	2000	Carotid endarterectomy	US	1990-96	2,152	7	1	Though volume rose, death due to stroke fell between the two time periods, particular in the asymptomatic group.	0.0005
Courcoulas[185]	2003	Gastric bypass	US	1999-2001	4,674	129	73	Better outcomes with higher volume surgeons, hospital volume is NS.	Surgeon volume 0.02; hospital volume 0.16.
Dimick[186]	2002	AAA	US	1994-96	2,987		52	Better outcomes with increased hospital volume.	0.02
Ellis[187]	1997	PTCA	US	1993-94	12,985	38	5	Better outcomes with increased hospital volume (both death and rate of adverse events), but not correlated to years of experience.	0.01 (death); < 0.001 (all adverse outcomes)
Esserman[188]	2002	Reading mammograms	UK/US	1996	60	60		Low- and medium-volume US radiologists did worse than UK radiologists (who were all high volume).	< 0.001 (Bonne-feronni correction for multiple testing)
Fujita[189]	2002	Total gastrectomy	Japan	1995-98	136	21	1	Better outcomes obtained by the high volume surgeons.	Discrimi-nant analysis < 0.01
Dardik[190]	1999	AAA, elective	US	1990-95	2,335	219	46	Better outcomes with high hospital volume and high surgeon volume. No difference in LOS by hospital volume, but costs are higher in low volume hospitals	Mortality 0.04 for hospital volume and

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
								(unadjusted); both costs and LOS are higher for low-volume surgeons.	0.01 for surgeon volume LOS and costs, < 0.0001 for surgeon volume.
Williams[191]	1991	CABG	US	1985-87	4,613	14	5	Adjusting for hospital in the model, mortality was not related to surgeon volume.	0.6
Garcia[192]	2001	Birth	US	1996	63,143		52	Non-significant trends for caesarian and episiotomy vs. volume. Fewer lacerations with low volume. Complication rate was highest in the medium-volume group, and NS different from very high volume in the low-volume group.	~ 0.03 for laceration; ~ 0.12 for complica-tions (low relative to very high volume) and < 0.001 for complica-tions (medium relative to very high volume)
Laffel[193]	1992	Heart transplant	US	1984-86	1,123	n/a	56	Better results once a centre has some experience, high-risk patients benefit from increased prior experience with low-risk patients. Training of cardiologist matters, and so does training of the transplant coordinator +BX139.	Death rate in first 5 patients compared to death rate later, 0.002. Death rate in high risk patient, in

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
									center with low experience vs high experience, 0.04.
Jones[194]	1995	ICU	UK	2-year period	8,796		26	Correlation between volume and standardized mortality NS (low power, only 26 observations); correlation between volume and outcome much stronger for surgical cases than for non-surgical.	Overall 0.37; for surgical cases 0.08.
Rathore[195]	2004	CABG	US	1998-2000	228,738		393	Better outcomes with increased hospital volume.	< 0.0001
Kastrati[196]	1998	Coronary stent placement	Germany	1992-97, 1998	3,409	10	1	Better outcomes with increased surgeon volume, form of the volume/outcome relationship depends on definition of volume; CART was used to identify cut points of how much volume is needed to improve outcome.	
Klein[197]	1997	Coronary interventional procedures	US	1993-95	1,389	9	1	Trend to better (unadjusted) volumes with increasing volume (r = 0.56, but p = 0.12 due to low power); 5 surgeons fall in the "borderline" volume category according to the standards, but their results are significantly better than the results shown in other studies.	
Klein[198]	2002	Singleton births attended by a family physician	Canada	1997-98	4,444	152	1	Volume NS for maternal complications, 5 min Apgar score; high volume family physicians consulted /transferred delivery less often to specialist.	< 0.001
Ko[199]	2002	Resection, colon cancer	US	1996	22,408			Better outcomes with increased hospital volume and with increased surgeon volume.	Hospital volume ~ 0.05; surgeon volume 0.01
Kreder[200]	2003	Total knee arthroplasty	Canada	1992-96	14,352	267	88	No association between volume and 3-	Revision at

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Phys	Hosp	Results/Comments	P-value
								month mortality, readmission for infection; revision at 1 year and 3 years, related to hospital volume, not surgeon volume; complications related to hospital volume, and are better for low-volume hospitals; LOS related to surgeon and hospital volume.	1 yr, 0.05; LOS < 0.001.
McKee[201]	2002	Mammogram detectable breast cancer	US	1995-96	955	261		No relationship between positive rate and radiologist volume, surgeon volume, or center; case volume affected use of breast conserving surgery, with higher rates for higher volume surgeons (radiologist and centre volumes were NS).	For volume * use of breast - conserving surgery,. 0.04.
Richardson[202]	1998	Trauma	US	1995-96	2,330	13	1	No effect of volume or years of experience on mortality or morbidity.	None
Schrag[203]	2002	Resection, rectal	US	1992-96	2,815		420	Thirty-day mortality not associated with hospital or surgeon volume. Two-year mortality (yes/no) and overall survival associated with surgeon volume, not hospital volume. Abdominoperineal resection not related to volume.	Two-year mortality (yes/no) .0.004 for surgeon volume; overall survival 0.02 for surgeon volume
Scott[36]	2001	Ventricular septal defect repair	Inter-national (US based)	1992-96	1,679		24	No relationship between volume and outcome.	
		Arterial switch operation	Inter-national (US based)	1992-96	494		24	Outcome improves with higher volume.	0.01
Shook[204]	1996	PTCA	US	1991-94	2,350		1	Outcome improves with higher volume.	0.0001
Smith[205]	1990	Trauma, severe	US	198788	1,643		7	Adjusting for patient mix, correlation between mortality rate and volume is -0.65,	NS

Author	Pub Year	Procedure or Diagnosis	Country	Study Years	Number of Subj	Number of Phys	Number of Hosp	Results/Comments	P-value
								p = 0.12 (n = 7). Comparing low vs high using a 2x2 table, p = 0.07, OR = 1.3.	
Tilford[206]	2000	Pediatric ICU	US	1993	11,106	38	16	Outcome improves with higher volume.	0.05
Hermanek[207]	1996	Resection, colorectal	Germany	1984-86	2,347	14	7	No relationship between volume and outcome, hospital or surgeon.	
Farber[208]	1981	Hysterectomy	US	1977-79	5,117		22	Outcome improves with higher volume.	< 0.0001
		Resection, colon	US	1977-79	1,331		22	Outcome improves with higher volume.	< 0.0001
		Hernia repair	US	1977-79	5,432		22	Outcome improves with higher volume.	0.0002
		Cholecystectomy	US	1977-79	4,156		22	Outcome improves with higher volume.	0.0007
		Appendectomy	US	1977-79	3,671		22	Outcome improves with higher volume.	0.03
		Laminectomy	US	1977-79	2,756		22	No relationship between volume and outcome.	0.06
		Cesarean section	US	1977-79	3,478		22	No relationship between volume and outcome.	0.47
Hughes[209]	1988	Hip fracture	US	1982	44,905		704	Results support both selective referral and practice makes perfect.	< 0.01

References

1. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med* 1979;301:1364-9.

2. Birkmeyer JD. High-risk surgery--follow the crowd. *JAMA* 2000;283:1191-3.

3. Epstein AM. Volume and outcome--it is time to move ahead. *N Engl J Med* 2002;346: 1161-4.

4. Birkmeyer JD. Should we regionalize major surgery? Potential benefits and policy considerations. *J Am Coll Surg* 2000;190:341-9.

5. Flood AB, Scott WR, Ewy W. Does practice make perfect? Part I: The relation between hospital volume and outcomes for selected diagnostic categories. *Med Care* 1984;22:98-114.

6. Flood AB, Scott WR, Ewy W. Does practice make perfect? Part II: The relation between volume and outcomes and other hospital characteristics. *Med Care* 1984;22:115-25.

7. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health Serv Res* 1987;22:157-82.

8. Birkmeyer JD, Siewers AE, Marth NJ, Goodman DC. Regionalization of high-risk surgery and implications for patient travel times. *JAMA* 2003;290:2703-8.

9. Finlayson SR, Birkmeyer JD, Tosteson AN, Nease RF, Jr. Patient preferences for location of care: implications for regionalization. *Med Care* 1999;37:204-9.

10. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high volume hospitals: estimating potentially avoidable deaths. *JAMA* 2000;283:1159-66.

11. Hewitt M for the Committee on Quality of Health Care in America and the National Cancer Policy Board. *Interpreting the Volume-Outcome Relationship in the Context of Health Care Quality. Workshop Summary*. Washington, DC: Institute of Medicine, 2000.

12. Bunker JP, Luft HS, Enthoven A. Should surgery be regionalized? *Surg Clin North Am* 1982;62:657-68.

13. Flood AB, Ewy W, Scott WR, Forrest WH, Jr., Brown BW, Jr. The relationship between intensity and duration of medical services and outcomes for hospitalized patients. *Med Care* 1979;17:1088-102.

14. Hannan EL. The relation between volume and outcome in health care. *N Engl J Med* 1999;340:1677-9.

15. Sowden AJ, Deeks JJ, Sheldon TA. Volume and outcome in coronary artery bypass graft surgery: true association or artefact? *BMJ* 1995;311:151-5.

16. Hillner BE, Smith TJ, Desch CE. Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. *J Clin Oncol* 2000; 18:2327-40.

17. Shackley P, Slack R, Booth A, Michaels J. Is there a positive volume-outcome relationship in peripheral vascular surgery? Results of a systematic review. *Eur J Vasc Endovasc Surg* 2000;20:326-35.

18. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high-volume hospitals: estimating potentially avoidable deaths. *JAMA* 2000;283:1159-66.

19. Hunt DL, McKibbon KA. Locating and appraising systematic reviews. *Ann Intern Med* 1997;126:532-8.

20. Meade MO, Richardson WS. Selecting and appraising studies for a systematic review. *Ann Intern Med* 1997;127:531-7.

21. Brott T, Thalinger K. The practice of carotid endarterectomy in a large metropolitan area. *Stroke* 1984;15:950-5.

22. *Coronary Artery Bypass Surgery in New York State: 1995-1997*. New York: New York State Department of Health, 2000.

23. Hofer TP, Hayward RA. Identifying poor-quality hospitals. Can hospital mortality rates detect quality problems for medical diagnoses? *Med Care* 1996;34:737-53.

24. Zalkind DL, Eastaugh SR. Mortality rates as an indicator of hospital quality. *Hosp Health Serv Adm* 1997;42:3-15.

25. Dickersin K, Berlin JA. Meta-analysis: state-of-the-science. *Epidemiol Rev* 1992;14:154-76.

26. Nunley RM, Lachiewicz PF. Mortality after total hip and knee arthroplasty in a medium-volume university practice. *J Arthroplasty* 2003;18:278-85.

27. Berger DH, Ko CY, Spain DA. Society of University Surgeons position statement on the volume-outcome relationship for surgical procedures. *Surgery* 2003;134:34-40.

28. Carter D. The surgeon as a risk factor. *BMJ* 2003;326:832-3.

29. Luft HS, Romano PS. Chance, continuity, and change in hospital mortality rates. Coronary artery bypass graft patients in California hospitals, 1983 to 1989. *JAMA* 1993;270:331-7.

30. Steiner CA, Bass EB, Talamini MA, Pitt HA, Steinberg EP. Surgical rates and operative mortality for open and laparoscopic cholecystectomy in Maryland. *N Engl J Med* 1994; 330:403-8.

31. Hong D, Tandan VR, Goldsmith CH, Simunovic M. Users' guide to the surgical literature: how to use an article reporting population-based volume-outcome relationships in surgery. *Can J Surg* 2002;45:109-15.

32. Kalant N, Shrier I. Volume and outcome of coronary artery bypass graft surgery: are more and less the same? *Can J Cardiol* 2004;20:81-6.

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

34. Ghali WA, Quan H, Brant R. Coronary artery bypass grafting in Canada: hospital mortality rates, 1992-1995. *CMAJ* 1998;159:926-30.

35. Wade TP, Halaby IA, Stapleton DR, Virgo KS, Johnson FE. Population-based analysis of treatment of pancreatic cancer and Whipple resection: Department of Defense hospitals, 1989-1994. *Surgery* 1996;120:680-5.

36. Scott WA, Fixler DE. Effect of center volume on outcome of ventricular septal defect closure and arterial switch operation. *Am J Cardiol* 2001;88:1259-63.

37. Bodenheimer T, Grumbach K. Paying for health care. *JAMA* 1994;272:634-9.

38. Hamilton BH, Hamilton VH. Estimating surgical volume--outcome relationships applying survival models: accounting for frailty and hospital fixed effects. *Health Econ* 1997; 6:383-95.

39. Halm EAM, Lee CM, Chassin MRM. Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature. *Ann Intern Med* 2002; 137:511-20.

40. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health Serv Res* 1987;22:157-82.

41. Birkmeyer JD, Finlayson EVA, Birkmeyer CM. Volume standards for high-risk surgical procedures: Potential benefits of the Leapfrog initiative. *Surgery* 2001;130:415-22.

42. Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: Effect of process and outcomes measures. *Surgery* 2004;135:569-75.

43. Fink AS, Campbell DAJr, Mentzer RMJr, Henderson WG, Daley J, Bannister J. The National Surgical Quality Improvement Program in non-veterans administration hospitals: initial demonstration of feasibility. *Ann Surg* 2002;236:353.

44. Shahian DM, Normand S-LT. The volume-outcome relationship: from Luft to Leapfrog. *Ann Thorac Surg* 2003;75:1048-58.

45. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The Leapfrog volume criteria may fall short in identifying high-quality surgical centres. *Ann Surg* 2003;238: 447-55.

46. Iezzoni LI. *Risk Adjustment for Measuring Health Care Outcomes*. Ann Arbor, MI: Health Administration Press, 1994.

47. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chron Dis* 1987;40:373-83.

48. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: Differing perspectives. *J Clin Epidemiol* 1993;46:1075-9.

49. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613-9.

50. Duncan C, Jones K, Moon G. Context, composition and heterogeneity: Using multilevel models in health research. *Soc Sci Med* 1998;46:97-117.

51. Goldstein H. *Multilevel Statistical Models*. 2nd ed. New York: Edward Arnold, 1995.

52. Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society A* 1996;159:385-443.

53. Normand S-LT, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: Issues and applications. *J Am Stat Assoc* 1997;92:803-14.

54. Rice N, Leyland A. Multilevel models: applications to health data. *J Health Serv Res Pol* 1996;1:154-64.

55. Moerbeek M, van Breukelen GJP, Berger MPF. A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *J Clin Epidemiol* 2003;56:341-50.

56. Thiemann DR, Coresh J, Oetgen WJ, Powe NR. The association between hospital volume and survival after acute myocardial infarction in elderly patients. *N Engl J Med* 1999; 340:1640-8.

57. Sainsbury R, Haward B, Rider L, Johnston C, Round C. Influence of clinician workload and patterns of treatment on survival from breast cancer. *Lancet* 1995;345:1265-70.

58. Simunovic M, To T, Baxter N, et al. Hospital procedure volume and teaching status do not influence treatment and outcome measures of rectal cancer surgery in a large general population. *J Gastrointest Surg* 2000;4:324-30.

59. Brook RH, Park RE, Chassin MR, Kosecoff J, Keesey J, Solomon DH. Carotid endarterectomy for elderly patients: predicting complications. *Ann Intern Med* 1990;113: 747-53.

60. Karp HR, Flanders WD, Shipp CC, Taylor B, Martin D. Carotid endarterectomy among Medicare beneficiaries: a statewide evaluation of appropriateness and outcome. *Stroke* 1998;29:46-52.

61. Roohan PJ, Bickell NA, Baptiste MS, Therriault GD, Ferrara EP, Siu AL. Hospital volume differences and five-year survival from breast cancer. *Am J Public Health* 1998; 88:454-7.

62. Simons AJ, Ker R, Groshen S, et al. Variations in treatment of rectal cancer: the influence of hospital type and caseload. *Dis Colon Rectum* 1997;40:641-6.

63. Schrag D, Cramer LD, Bach PB, Cohen AM, Warren JL, Begg CB. Influence of hospital procedure volume on outcomes following surgery for colon cancer. *JAMA* 2000; 284:3028-35.

64. Malenka DJ, McGrath PD, Wennberg DE, et al. The relationship between operator volume and outcomes after percutaneous coronary interventions in high volume hospitals in 1994-1996: the northern New England experience. Northern New England Cardiovascular Disease Study Group. *J Am Coll Cardiol* 1999;34:1471-80.

65. McGrath PD, Wennberg DE, Dickens JD, Jr., et al. Relation between operator and hospital volume and outcomes following percutaneous coronary interventions in the era of the coronary stent. *JAMA* 2000;284:3139-44.

66. Hannan EL, Kilburn H, Jr., 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. *Med Care* 1991;29:1094-107.

67. Ho V. Evolution of the volume-outcome relation for hospitals performing coronary angioplasty. *Circulation* 2000;101:1806-11.

68. Hsia DC, Moscoe LM, Krushat WM. Epidemiology of carotid endarterectomy among Medicare beneficiaries: 1985-1996 update. *Stroke* 1998;29:346-50.

69. Hamilton BH, Ho V. Does practice make perfect? Examining the relationship between hospital surgical volume and outcomes for hip fracture patients in Quebec. *Med Care* 1998;36:892-903.

70. Farley DE, Ozminkowski RJ. Volume-outcome relationships and in-hospital mortality: the effect of changes in volume over time. *Med Care* 1992;30:77-94.

71. 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;261:2975-9.

72. Bennett CL, Adams J, Gertler P, et al. Relation between hospital experience and in-hospital mortality for patients with AIDS-related Pneumocystis carinii pneumonia: experience from 3,126 cases in New York City in 1987. *J Acquir Immune Defic Syndr* 1992;5:856-64.

73. Stone VE, Seage GR, III, Hertz T, Epstein AM. The relation between hospital experience and mortality for patients with AIDS. *JAMA* 1992;268:2655-61.

74. Turner BJ, Ball JK. Variations in inpatient mortality for AIDS in a national sample of hospitals. *J Acquir Immune Defic Syndr* 1992;5:978-87.

75. Hogg RS, Raboud J, Bigham M, Montaner JS, O'Shaughnessy M, Schechter MT. Relation between hospital HIV/AIDS caseload and mortality among persons with HIV/AIDS in Canada. *Clin Invest Med* 1998;21:27-32.

76. Cunningham WE, Tisnado DM, Lui HH, Nakazono TT, Carlisle DM. The effect of hospital experience on mortality among patients hospitalized with acquired immunodeficiency syndrome in California. *Am J Med* 1999;107:137-43.

77. Ritchie JL, Phillips KA, Luft HS. Coronary angioplasty. Statewide experience in California. *Circulation* 1993;88:2735-43.

78. Ritchie JL, Maynard C, Chapko MK, Every NR, Martin DC. Association between percutaneous transluminal coronary angioplasty volumes and outcomes in the Healthcare Cost and Utilization Project 1993-1994. *Am J Cardiol* 1999;83:493-7.

79. 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;331:1625-9.

80. Jollis JG, Peterson ED, Nelson CL, et al. Relationship between physician and hospital coronary angioplasty volume and outcome in elderly patients. *Circulation* 1997;95:2485-91.

81. 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. *Med Care* 1995;33:502-14.

82. Hannan EL, Racz M, Ryan TJ, et al. Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. *JAMA* 1997;277:892-8.

83. Maynard C, Every NR, Chapko MK, Ritchie JL. Institutional volumes and coronary angioplasty outcomes before and after the introduction of stenting. *Eff Clin Pract* 1999; 2:108-13.

84. Showstack JA, Rosenfeld KE, Garnick DW, Luft HS, Schaffarzick RW, Fowles J. Association of volume with outcome of coronary artery bypass graft surgery. Scheduled vs nonscheduled operations. *JAMA* 1987;257:785-9.

85. Hannan EL, O'Donnell JF, Kilburn H, Jr., Bernard HR, Yazici A. Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. *JAMA* 1989;262:503-10.

86. Grumbach K, Anderson GM, Luft HS, Roos LL, Brook R. Regionalization of cardiac surgery in the United States and Canada. Geographic access, choice, and outcomes. *JAMA* 1995;274:1282-8.

87. Shroyer AL, Marshall G, Warner BA, et al. No continuous relationship between Veterans Affairs hospital coronary artery bypass grafting surgical volume and operative mortality. *Ann Thorac Surg* 1996;61:17-20.

88. Sollano JA, Gelijns AC, Moskowitz AJ, et al. Volume-outcome relationships in cardio-vascular operations: New York State, 1990-1995. *J Thorac Cardiovasc Surg* 1999; 117:419-28.

89. Hannan EL, Racz M, Kavey RE, Quaegebeur JM, Williams R. Pediatric cardiac surgery: the effect of hospital and surgeon volume on in-hospital mortality. *Pediatrics* 1998;101: 963-9.

90. Jenkins KJ, Newburger JW, Lock JE, Davis RB, Coffman GA, Iezzoni LI. In-hospital mortality for surgical repair of congenital heart defects: preliminary observations of variation by hospital caseload. *Pediatrics* 1995;95:323-30.

91. Edwards WH, Morris JA, Jr., Jenkins JM, Bass SM, MacKenzie EJ. Evaluating quality, cost-effective health care. Vascular database predicated on hospital discharge abstracts. *Ann Surg* 1991;213:433-8.

92. Cebul RD, Snow RJ, Pine R, Hertzer NR, Norris DG. Indications, outcomes, and provider volumes for carotid endarterectomy. *JAMA* 1998;279:1282-7.

93. Fisher ES, Malenka DJ, Solomon NA, Bubolz TA, Whaley FS, Wennberg JE. Risk of carotid endarterectomy in the elderly. *Am J Public Health* 1989;79:1617-20.

94. Hannan EL, Popp AJ, Tranmer B, Fuestel P, Waldman J, Shah D. Relationship between provider volume and mortality for carotid endarterectomies in New York state. *Stroke* 1998;29:2292-7.

95. Kantonen I, Lepantalo M, Salenius JP, Matzke S, Luther M, Ylonen K. Influence of surgical experience on the results of carotid surgery. The Finnvasc Study Group. *Eur J Vasc Endovasc Surg* 1998;15:155-60.

96. Kempczinski RF, Brott TG, Labutta RJ. The influence of surgical specialty and caseload on the results of carotid endarterectomy. *J Vasc Surg* 1986;3:911-6.

97. Khuri SF, Daley J, Henderson W, et al. Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. *Ann Surg* 1999;230:414-29.

98. Kirshner DL, O'Brien MS, Ricotta JJ. Risk factors in a community experience with carotid endarterectomy. *J Vasc Surg* 1989;10:178-86.

99. Manheim LM, Sohn MW, Feinglass J, Ujiki M, Parker MA, Pearce WH. Hospital vascular surgery volume and procedure mortality rates in California, 1982-1994. *J Vasc Surg* 1998;28:45-56.

100. Perler BA, Dardik A, Burleyson GP, Gordon TA, Williams GM. Influence of age and hospital volume on the results of carotid endarterectomy: a statewide analysis of 9918 cases. *J Vasc Surg* 1998;27:25-31.

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

102. Wennberg DE, Lucas FL, Birkmeyer JD, Bredenberg CE, Fisher ES. Variation in carotid endarterectomy mortality in the Medicare population: trial hospitals, volume, and patient characteristics. *JAMA* 1998;279:1278-81.

103. 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;19:804-15.

104. Hannan EL, Kilburn H, Jr., O'Donnell JF, et al. A longitudinal analysis of the relationship between in-hospital mortality in New York State and the volume of abdominal aortic aneurysm surgeries performed. *Health Serv Res* 1992;27:517-42.

105. 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? Members of the Norwegian Abdominal Aortic Aneurysm Trial. *Acta Chir Scand* 1990;156:323-7.

106. Wen SW, Simunovic M, Williams JI, Johnston KW, Naylor CD. Hospital volume, calendar age, and short term outcomes in patients undergoing repair of abdominal aortic aneurysms: the Ontario experience, 1988-92. *J Epidemiol Community Health* 1996;50: 207-13.

107. Kazmers A, Jacobs L, Perkins A, Lindenauer SM, Bates E. Abdominal aortic aneurysm repair in Veterans Affairs medical centers. *J Vasc Surg* 1996;23:191-200.

108. Rutledge R, Oller DW, Meyer AA, Johnson GJ, Jr. A statewide, population-based time-series analysis of the outcome of ruptured abdominal aortic aneurysm. *Ann Surg* 1996; 223:492-502.

109. Dardik A, Burleyson GP, Bowman H, et al. Surgical repair of ruptured abdominal aortic aneurysms in the state of Maryland: factors influencing outcome among 527 recent cases. *J Vasc Surg* 1998;28:413-20.

110. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA* 1998;280:1747-51.

111. Gordon TA, Burleyson GP, Tielsch JM, Cameron JL. The effects of regionalization on cost and outcome for one general high-risk surgical procedure. *Ann Surg* 1995; 221:43-9.

112. Birkmeyer JD, Warshaw AL, Finlayson SR, Grove MR, Tosteson AN. Relationship between hospital volume and late survival after pancreaticoduodenectomy. *Surgery* 1999; 126:178-83.

113. Birkmeyer JD, Finlayson SR, Tosteson AN, Sharp SM, Warshaw AL, Fisher ES. Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. *Surgery* 1999;125:250-6.

114. Imperato PJ, Nenner RP, Starr HA, Will TO, Rosenberg CR, Dearie MB. The effects of regionalization on clinical outcomes for a high risk surgical procedure: a study of the Whipple procedure in New York State. *Am J Med Qual* 1996;11:193-7.

115. Lieberman MD, Kilburn H, Lindsey M, Brennan MF. Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. *Ann Surg* 1995;222:638-45.

116. Sosa JA, Bowman HM, Gordon TA, et al. Importance of hospital volume in the overall management of pancreatic cancer. *Ann Surg* 1998;228:429-38.

117. Simunovic M, To T, Theriault M, Langer B. Relation between hospital surgical volume and outcome for pancreatic resection for neoplasm in a publicly funded health care system. *CMAJ* 1999;160:643-8.

118. Gordon TA, Bowman HM, Bass EB, et al. Complex gastrointestinal surgery: impact of provider experience on clinical and economic outcomes. *J Am Coll Surg* 1999;189:46-56.

119. Patti MG, Corvera CU, Glasgow RE, Way LW. A hospital's annual rate of esophagectomy influences the operative mortality rate. *J Gastrointest Surg* 1998;2:186-92.

120. Harmon JW, Tang DG, Gordon TA, et al. Hospital volume can serve as a surrogate for surgeon volume for achieving excellent outcomes in colorectal resection. *Ann Surg* 1999; 230:404-11.

121. Parry JM, Collins S, Mathers J, Scott NA, Woodman CB. Influence of volume of work on the outcome of treatment for patients with colorectal cancer. *Br J Surg* 1999;86:475-81.

122. Hannan EL, Radzyner M, Rubin D, Dougherty J, Brennan MF. The influence of hospital and surgeon volume on in-hospital mortality for colectomy, gastrectomy, and lung lobectomy in patients with cancer. *Surgery* 2002;131:6-15.

123. Romano PS, Mark DH. Patient and hospital characteristics related to in-hospital mortality after lung cancer resection. *Chest* 1992;101:1332-7.

124. Taylor HD, Dennis DA, Crane HS. Relationship between mortality rates and hospital patient volume for Medicare patients undergoing major orthopaedic surgery of the hip, knee, spine, and femur. *J Arthroplasty* 1997;12:235-42.

125. Lavernia CJ, Guzman JF. Relationship of surgical volume to short-term mortality, morbidity, and hospital charges in arthroplasty. *J Arthroplasty* 1995;10:133-40.

126. Kreder HJ, Williams JI, Jaglal S, Hu R, Axcell T, Stephen D. Are complication rates for elective primary total hip arthroplasty in Ontario related to surgeon and hospital volumes? A preliminary investigation. *Can J Surg* 1998;41:431-7.

127. Kreder HJ, Deyo RA, Koepsell T, Swiontkowski MF, Kreuter W. Relationship between the volume of total hip replacements performed by providers and the rates of post-operative complications in the state of Washington. *J Bone Joint Surg Am* 1997;79: 485-94.

128. Ellison LM, Heaney JA, Birkmeyer JD. The effect of hospital volume on mortality and resource use after radical prostatectomy. *J Urol* 2000;163:867-9.

129. Yao SL, Lu-Yao G. Population-based study of relationships between hospital volume of prostatectomies, patient outcomes, and length of hospital stay. *J Natl Cancer Inst* 1999; 91:1950-6.

130. Thorpe AC, Cleary R, Coles J, Vernon S, Reynolds J, Neal DE. Deaths and complications following prostatectomy in 1400 men in the northern region of England. Northern Regional Prostate Audit Group. *Br J Urol* 1994;74:559-65.

131. Taylor CL, Yuan Z, Selman WR, Ratcheson RA, Rimm AA. Mortality rates, hospital length of stay, and the cost of treating subarachnoid hemorrhage in older patients: institutional and geographical differences. *J Neurosurg* 1997;86:583-8.

132. Solomon RA, Mayer SA, Tarmey JJ. Relationship between the volume of craniotomies for cerebral aneurysm performed at New York state hospitals and in-hospital mortality. *Stroke* 1996;27:13-7.

133. Pronovost PJ, Jenckes MW, Dorman T, et al. Organizational characteristics of intensive care units related to outcomes of abdominal aortic surgery. *JAMA* 1999;281:1310-7.

134. Kitahata MM, Koepsell TD, Deyo RA, Maxwell CL, Dodge WT, Wagner EH. Physicians' experience with the acquired immunodeficiency syndrome as a factor in patients' survival. *N Engl J Med* 1996;334:701-6.

135. Casale PN, Jones JL, Wolf FE, Pei Y, Eby LM. Patients treated by cardiologists have a lower in-hospital mortality for acute myocardial infarction. *J Am Coll Cardiol* 1998;32: 885-9.

136. McGrath PD, Wennberg DE, Malenka DJ, et al. Operator volume and outcomes in 12,998 percutaneous coronary interventions. Northern New England Cardiovascular Disease Study Group. *J Am Coll Cardiol* 1998;31:570-6.

137. Ruby ST, Robinson D, Lynch JT, Mark H. Outcome analysis of carotid endarterectomy in Connecticut: the impact of volume and specialty. *Ann Vasc Surg* 1996;10:22-6.

138. Porter GA, Soskolne CL, Yakimets WW, Newman SC. Surgeon-related factors and outcome in rectal cancer. *Ann Surg* 1998;227:157-67.

139. Hannan EL, Siu AL, Kumar D, Kilburn H, Jr., Chassin MR. The decline in coronary artery bypass graft surgery mortality in New York State. The role of surgeon volume. *JAMA* 1995;273:209-13.

140. O'Neill L, Lanska DJ, Hartz A. Surgeon characteristics associated with mortality and morbidity following carotid endarterectomy. *Neurology* 2000;55:773-81.

141. Tu JV, Hannan EL, Anderson GM, et al. The fall and rise of carotid endarterectomy in the United States and Canada. *N Engl J Med* 1998;339:1441-7.

142. Hughes RG, Hunt SS, Luft HS. Effects of surgeon volume and hospital volume on quality of care in hospitals. *Med Care* 1987;25:489-503.

143. Kelly JV, Hellinger FJ. Heart disease and hospital deaths: an empirical study. *Health Serv Res* 1987;22:369-95.

144. Zelen J, Bilfinger TV, Anagnostopoulos CE. Coronary artery bypass grafting. The relationship of surgical volume, hospital location, and outcome. *N Y State J Med* 1991; 91:290-2.

145. Hosenpud JD, Breen TJ, Edwards EB, Daily OP, Hunsicker LG. The effect of transplant center volume on cardiac transplant outcome. A report of the United Network for Organ Sharing Scientific Registry. *JAMA* 1994;271:1844-9.

146. Goodney PP, Stukel TA, Lucas FL, Finlayson EV, Birkmeyer JD. Hospital volume, length of stay, and readmission rates in high-risk surgery. *Ann Surg* 2003;238: 161-7.

147. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med* 2003;349:2117-27.

148. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med* 2002;346:1128-37.

149. de Gara CJ, Hanson J, Hamilton S. A population-based study of tumor-node relationship, resection margins, and surgeon volume on gastric cancer survival. *Am J Surg* 2003;186: 23-7.

150. Hodgson DC, Zhang W, Zaslavsky AM, Fuchs CS, Wright WE, Ayanian JZ. Relation of hospital volume to colostomy rates and survival for patients with rectal cancer. *J Natl Cancer Inst* 2003;95:708-16.

151. Harcourt KF, Hicks KL. Is there a relationship between case volume and survival in breast cancer? *Am J Surg* 2003;185:407-10.

152. Bach PB, Cramer LD, Schrag D, Downey RJ, Gelfand SE, Begg CB. The influence of hospital volume on survival after resection for lung cancer. *N Engl J Med* 2001;345:181-8.

153. Margulies DR, Cryer HG, McArthur DL, Lee SS, Bongard FS, Fleming AW. Patient volume per surgeon does not predict survival in adult level I trauma centers. *J Trauma* 2001;50:597-601.

154. Canto JG, Every NR, Magid DJ, et al. The volume of primary angioplasty procedures and survival after acute myocardial infarction. National Registry of Myocardial Infarction 2 Investigators. *N Engl J Med* 2000;342:1573-80.

155. Konvolinka CW, Copes WS, Sacco WJ. Institution and per-surgeon volume versus survival outcome in Pennsylvania's trauma centers. *Am J Surg* 1995;170:333-40.

156. Glance LG, Dick AW, Mukamel DB, Osler TM. Is the hospital volume-mortality relationship in coronary artery bypass surgery the same for low-risk versus high-risk patients? *Ann Thorac Surg* 2003;76:1155-62.

157. Barker FG, Klibanski A, Swearingen B. Transsphenoidal surgery for pituitary tumors in the United States, 1996-2000: mortality, morbidity, and the effects of hospital and surgeon volume. *J Clin Endocrinol Metab* 2003;88:4709-19.

158. Hoh BL, Rabinov JD, Pryor JC, Carter BS, Barker FG. In-hospital morbidity and mortality after endovascular treatment of unruptured intracranial aneurysms in the United States, 1996-2000: effect of hospital and physician volume. *AJNR Am J Neuroradiol* 2003;24:1409-20.

159. Finlayson EV, Goodney PP, Birkmeyer JD. Hospital volume and operative mortality in cancer surgery: a national study. *Arch Surg* 2003;138:721-5.

160. Marrie TJ, Carriere KC, Jin Y, Johnson DH. Mortality during hospitalisation for pneumonia in Alberta, Canada, is associated with physician volume. *Eur Respir J* 2003; 22:148-55.

161. Urbach DR, Bell CM, Austin PC. Differences in operative mortality between high- and low-volume hospitals in Ontario for 5 major surgical procedures: estimating the number of lives potentially saved through regionalization. *CMAJ* 2003;168:1409-14.

162. Sava J, Kennedy S, Jordan M, Wang D. Does volume matter? The effect of trauma surgeons' caseload on mortality. *J Trauma* 2003;54:829-33.

163. Dimick JB, Cowan JA, Jr., Henke PK, et al. Hospital volume-related differences in aorto-bifemoral bypass operative mortality in the United States. *J Vasc Surg* 2003;37:970-5.

164. Barker FG, Amin-Hanjani S, Butler WE, Ogilvy CS, Carter BS. In-hospital mortality and morbidity after surgical treatment of unruptured intracranial aneurysms in the United States, 1996-2000: the effect of hospital and surgeon volume. *Neurosurgery* 2003;52: 995-1007.

165. Ho V, Heslin MJ. Effect of hospital volume and experience on in-hospital mortality for pancreaticoduodenectomy. *Ann Sur*. 2003;237:509-14.

166. London JA, Battistella FD. Is there a relationship between trauma center volume and mortality? *J Trauma* 2003;54:16-24.

167. Cowan JA, Jr., Dimick JB, Leveque JC, Thompson BG, Upchurch GR, Jr., Hoff JT. The impact of provider volume on mortality after intracranial tumor resection. *Neurosurgery* 2003;52:48-53.

168. Damhuis RA, Meurs CJ, Dijkhuis CM, Stassen LP, Wiggers T. Hospital volume and post-operative mortality after resection for gastric cancer. *Eur J Surg Oncol* 2002;28: 401-5.

169. Spiegelhalter DJ. Mortality and volume of cases in paediatric cardiac surgery: retro-spective study based on routinely collected data. *BMJ* 2002;324:261-3.

170. Tu JV, Austin PC, Chan BT. Relationship between annual volume of patients treated by admitting physician and mortality after acute myocardial infarction. *JAMA* 2001; 285: 3116-22.

171. van Lanschot JJ, Hulscher JB, Buskens CJ, Tilanus HW, ten Kate FJ, Obertop H. Hospital volume and hospital mortality for esophagectomy. *Cancer* 2001;91:1574-8.

172. Magid DJ, Calonge BN, Rumsfeld JS, et al. Relation between hospital primary angioplasty volume and mortality for patients with acute MI treated with primary angioplasty vs thrombolytic therapy. *JAMA* 2000;284:3131-8.

173. Swisher SG, Deford L, Merriman KW, et al. Effect of operative volume on morbidity, mortality, and hospital use after esophagectomy for cancer. *J Thorac Cardiovasc Surg* 2000;119:1126-32.

174. Cooper A, Hannan EL, Bessey PQ, Farrell LS, Cayten CG, Mottley L. An examination of the volume-mortality relationship for New York State trauma centers. *J Trauma* 2000;48: 16-23.

175. Edwards EB, Roberts JP, McBride MA, Schulak JA, Hunsicker LG. The effect of the volume of procedures at transplantation centers on mortality after liver transplantation. *N Engl J Med* 1999;341:2049-53.

176. Kantonen I, Lepantalo M, Salenius JP, Matzke S, Luther M, Ylonen K. Mortality in abdominal aortic aneurysm surgery--the effect of hospital volume, patient mix and surgeon's case load. *Eur J Vasc Endovasc Surg* 1997;14:375-9.

177. 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. *JAMA* 1996;276:1054-9.

178. Bates EW, Berki SE, Homan RK, Lindenauer SM. The challenge of benchmarking: surgical volume and operative mortality in Veterans Administration Medical Centers. *Best Pract Benchmarking Healthc* 1996;1:34-42.

179. Kagan RJ, Baker RJ. The impact of the volume of neurotrauma experience on mortality after head injury. *Am Surg* 1994;60:394-400.

180. Mayfield JA, Rosenblatt RA, Baldwin LM, Chu J, Logerfo JP. The relation of obstetrical volume and nursery level to perinatal mortality. *Am J Public Health* 1990;80: 819-23.

181. Iapichino G, Gattinoni L, Radrizzani D, et al. Volume of activity and occupancy rate in intensive care units. Association with mortality. *Intensive Care Med* 2004;30:290-7.

182. Schrag D, Panageas KS, Riedel E, et al. Surgeon volume compared to hospital volume as a predictor of outcome following primary colon cancer resection. *J Surg Oncol* 2003;83: 68-78.

183. Brown DL. Analysis of the institutional volume-outcome relations for balloon angioplasty and stenting in the stent era in California. *Am Heart J* 2003;146:1071-6.

184. Darling RC, III, Kreienberg PB, Roddy SP, et al. Analysis of the effect of asymptomatic carotid atherosclerosis study on the outcome and volume of carotid endarterectomy. *Cardiovasc Surg* 2000;8:436-40.

185. Courcoulas A, Schuchert M, Gatti G, Luketich J. The relationship of surgeon and hospital volume to outcome after gastric bypass surgery in Pennsylvania: a 3-year summary. *Surgery* 2003;134:613-21.

186. Dimick JB, Pronovost PJ, Cowan JA, Ailawadi G, Upchurch GR, Jr. The volume-outcome effect for abdominal aortic surgery: differences in case-mix or complications? *Arch Surg* 2002;137:828-32.

187. Ellis SG, Weintraub W, Holmes D, Shaw R, Block PC, King SB, III. Relation of operator volume and experience to procedural outcome of percutaneous coronary revascularization at hospitals with high interventional volumes. *Circulation* 1997;95: 479-84.

188. Esserman L, Cowley H, Eberle C, et al. Improving the accuracy of mammography: volume and outcome relationships. *J Natl Cancer Inst* 2002;94:369-75.

189. Fujita T, Yamazaki Y. Influence of surgeon's volume on early outcome after total gastrectomy. *Eur J Surg* 2002;168:535-8.

190. Dardik A, Lin JW, Gordon TA, Williams GM, Perler BA. Results of elective abdominal aortic aneurysm repair in the 1990s: A population-based analysis of 2335 cases. *J Vasc Surg* 1999;30:985-95.

191. Williams SV, Nash DB, Goldfarb N. Differences in mortality from coronary artery bypass graft surgery at five teaching hospitals. *JAMA* 1991;266:810-5.

192. Garcia FA, Miller HB, Huggins GR, Gordon TA. Effect of academic affiliation and obstetric volume on clinical outcome and cost of childbirth. *Obstet Gynecol* 2001; 97: 67-76.

193. Laffel GL, Barnett AI, Finkelstein S, Kaye MP. The relation between experience and outcome in heart transplantation. *N Engl J Med* 1992;327:1220-5.

194. Jones J, Rowan K. Is there a relationship between the volume of work carried out in intensive care and its outcome? *Int J Technol Assess Health Care* 1995;11:762-9.

195. Rathore SS, Epstein AJ, Volpp KG, Krumholz HM. Hospital coronary artery bypass graft surgery volume and patient mortality, 1998-2000. *Ann Surg* 2004;239:110-7.

196. Kastrati A, Neumann FJ, Schomig A. Operator volume and outcome of patients undergoing coronary stent placement. *J Am Coll Cardiol* 1998;32:970-6.

197. Klein LW, Schaer GL, Calvin JE, et al. Does low individual operator coronary interventional procedural volume correlate with worse institutional procedural outcome? *J Am Coll Cardiol* 1997;30:870-7.

198. Klein MC, Spence A, Kaczorowski J, Kelly A, Grzybowski S. Does delivery volume of family physicians predict maternal and newborn outcome? *CMAJ* 2002;166:1257-63.

199. Ko CY, Chang JT, Chaudhry S, Kominski G. Are high-volume surgeons and hospitals the most important predictors of in-hospital outcome for colon cancer resection? *Surgery* 2002;132:268-73.

200. Kreder HJ, Grosso P, Williams JI, et al. Provider volume and other predictors of outcome after total knee arthroplasty: a population study in Ontario. *Can J Surg* 2003;46:15-22.

201. McKee MD, Cropp MD, Hyland A, Watroba N, McKinley B, Edge SB. Provider case volume and outcome in the evaluation and treatment of patients with mammogram-detected breast carcinoma. *Cancer* 2002;95:704-12.

202. Richardson JD, Schmieg R, Boaz P, et al. Impact of trauma attending surgeon case volume on outcome: is more better? *J Trauma* 1998;44:266-71.

203. Schrag D, Panageas KS, Riedel E, et al. Hospital and surgeon procedure volume as predictors of outcome following rectal cancer resection. *Ann Surg* 2002;236:583-92.

204. Shook TL, Sun GW, Burstein S, Eisenhauer AC, Matthews RV. Comparison of per-cutaneous transluminal coronary angioplasty outcome and hospital costs for low-volume and high-volume operators. *Am J Cardiol* 1996;77:331-6.

205. 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. *J Trauma* 1990;30:1066-75.

206. Tilford JM, Simpson PM, Green JW, Lensing S, Fiser DH. Volume-outcome relation-ships in pediatric intensive care units. *Pediatrics* 2000;106:289-94.

207. Hermanek P, Hohenberger W. The importance of volume in colorectal cancer surgery. *Eur J Surg Oncol*. 1996;22:213-5.

208. Farber BF, Kaiser DL, Wenzel RP. Relation between surgical volume and incidence of postoperative wound infection. *N Engl J Med*. 1981;305:200-4.

209. Hughes RG, Garnick DW, Luft HS, McPhee SJ, Hunt SS. Hospital volume and patient outcomes. The case of hip fracture patients. *Med Care*. 1988;26:1057-67.