The Relationship between Volume and Health Outcomes

Report of Volume/Outcome Sub-Group

to

Advisory Group to
National Framework for
Service Change
NHS Scotland

Gordon D Murray
Professor of Medical Statistics, University of Edinburgh

Graham M Teasdale
President, Royal College of Physicians and Surgeons of Glasgow

February 2005

1. Introduction

In 1979 Luft and his colleagues focused attention on the possibility of a connection between an increased volume of clinical work and improved outcome. Since then, many further studies have reported this relationship. Nevertheless, certain issues identified by Luft remain to be clarified so that controversy, even conflict continues about the importance of the "complexity" of the condition and the risks of treatment, the existence of "thresholds" in the relationship, and the relative roles of individual practitioner versus unit or hospital volume. The need for risk stratification and the importance of using an appropriate index of outcome in investigating these issues has been recognised

Much of the relevant information is derived from routine statistics. These allow little opportunity for risk stratification and provide outcome only as mortality. For interventions with a low risk of death, large numbers must be studied in order not to overlook an effect that could be important if the intervention is very common. For many interventions, mortality is not an appropriate index and other indicators of effectiveness, of quality of care and of patient satisfaction are required.

In 1997, the NHS Centre for Reviews and Dissemination at the University of York published a systematic review of the evidence available on the volume/outcome relationship in health care, based upon data available up to 1996 (Sowden *et al* 1997). Although they identified many studies that they considered suitable, that showed a volume outcome association, in essence, the conclusion reached was that the bulk of research evidence was methodologically flawed and of little value in forming decisions about the planning of the delivery of health services. Since then, substantial additional literature has been published, including other comprehensive reviews

This review was undertaken in response to a request in the summer of 2004 from the Advisory Group to the National Framework for Service Change in the NHS in Scotland. The purpose was to re-review the research evidence, in particular taking account of papers published since 1997. In view of the time constraints placed by the timetable of the work of the Service Change Group, it was accepted that the approach taken would be that of a conventional narrative review, not a formal, comprehensive systematic review.

2. Background

The York Review (Sowden *et al*, 1997) reached three main conclusions on volume/outcome relationships:

(i) **Case-mix**: "*Most of the existing research, because it does not sufficiently take account of differences in case-mix, probably overestimates the impact of volume on the quality of care*." (Summary Report, Page 10).

(ii) **Causation**: "… *because none of the research indicates that increasing activity over time leads to improvements in clinical outcome, it is difficult to infer from results of cross-sectional studies which show better outcomes in higher-volume units that similar differences in outcomes can be expected by the expansion of an existing unit*." (Summary Report, Page 10).

(iii) **Thresholds**: "*The most that the research evidence can support is a conclusion that if there are significant quality gains from increased volume, these gains appear to be exhausted at relatively low threshold levels. Volumes of activity above these thresholds should be achievable without significant structural changes, but may require a more sharply defined internal division of labour across consultant staff (which may be consistent with increased sub-specialisation within disciplines*)." (Summary Report, Page 11).

3. Methods

A general literature search was undertaken on volume/outcome relationships together with more focused literature searches on methodological aspects of volume/outcome relationships, on studies relating to volume/outcome relationships which evaluated the impact of an intervention to increase regionalisation, and methodological aspects of assessing clinical 'learning curves'. These searches concentrated on publications from 1997 onwards, although some key references predate this.

These literature searches showed that more relevant papers have been published between 1997 and 2004 than were published up to and including 1996, and so it would indeed be a major undertaking to perform a formal systematic review of the more recent literature. This review was therefore selective, based on reading abstracts and obtaining full papers when appropriate and when they were relatively easily accessed. As most papers in this area are observational and cross-sectional by design, and are often very specific to a local area (e.g. a single US state), they tend to be published in low impact journals which are not held by the local University libraries. Thus most of the full papers reviewed are either from high impact journals or from journals whose contents can be accessed electronically.

4.1 Results

Over 500 abstracts were scanned and 50 full papers (Appendix I) reviewed in detail. The vast majority of the papers relate to surgical procedures with outcomes assessed in terms of short-term (in hospital or 30 day) mortality. There is also an increasing number of papers relating to surgical oncology with outcomes assessed in terms of long-term (5 years plus) survival.

Many common procedures have very low surgical mortality rates, so that even if a volume/outcome relationship does exist for mortality, it would be difficult to demonstrate and might be of limited clinical relevance. So studies focussed on morbidity outcomes are also becoming more common. For example, avoidance of a stoma in bowel surgery has a major impact on quality of life, making it a very relevant patient-centred outcome measure (Hodgson *et al*, 2003) and similar issues apply in prostate cancer (Begg *at al*, 2002). Sometimes, there may be a long delay between intervention and its consequence (eg in obstetrics) and / or the effect may be exerted indirectly (e.g. screening for cancer) so that 'intermediary' indices of patient outcome are needed.

4.2 Methodological Quality

The York Review (Sowden *et al*, 1997) was rightly critical of the methodological quality of volume/outcome studies published prior to 1997. There is clear evidence that case-mix adjustment using clinical data on individual patients leads to the most reliable results and that case-mix adjustment using only administrative data tends to overestimate the magnitude of volume/outcome relationships. Studies which use no case-mix adjustment at all are likely to overestimate such effects even more so. This is now widely recognised and methodological standards are higher in recent studies. For example, in the Gandjour review (2003), 16 of the 'best' 33 hospital volume/outcome studies published between 1990 and 2000 were published in 1999 or 2000.

Increasingly (and appropriately!) sophisticated statistical approaches are now being used to try to disentangle the complex issue of whether it is surgeon volume or hospital volume which drives the observed volume/outcome relationships (Birkmeyer *et al* 2003, Panageas *et al* 2003).

4.3 Findings

Methodologically flawed studies are still published, but there is now a strong core of methodologically sound papers which use high quality data and appropriate statistical methods to explore

volume/outcome relationships. These papers are based on either series of patients with data extracted from administrative systems (giving very large sample sizes but incomplete case-mix adjustment) or on series of patients with data extracted from clinical databases (giving smaller sample sizes but good case-mix adjustment). Even when one restricts attention to these higher quality studies there is still very strong evidence of an association between volume and outcome in the direction that high volume surgeons and high volume hospitals tend to have superior outcomes compared to low volume surgeons and hospitals. The magnitude of this effect, and how it depends on the clinical area, is discussed in Section 4.4.

Two particularly useful systematic reviews were identified. Halm *et al* (2002) reviewed studies published between January 1980 and December 2000 and Gandjour *et al* (2003) reviewed studies published between January 1990 and December 2000. Halm *et al* was a conventional systematic review covering 27 procedures and diagnoses. In the 135 studies that meet their criteria a statistically significant relationship between higher volume and better outcomes were found for 71% of studies of hospital volume (Appendix II) and for 69% for clinician volume (Appendix III). The review of Gandjour *et al* covered 34 diagnoses and interventions and included another 26 reports not analysed by Halm et al. In a total of 76 studies, higher hospital volume was statistically significantly better in 51, non-significantly better in 21, non-significantly worse in 3 and significantly worse in only one. These authors took the unusual additional approach of identifying the single most reliable study (based on criteria such as the quality of case-mix adjustment) for each of a number of procedures. In 20 such 'best' studies, high volume was significantly better in 10, non-significantly better in 6, non-significantly worse in 3 and significantly worse in one.

4.4 Thresholds

Ramsay *et al* (2001) undertook a systematic review of methods used to analyse learning curves in health care, and more recently, Cook *et al* (2004) proposed methods for adjusting for learning curve effects in randomised trials of surgical interventions.

There was no clear consensus on appropriate ways to analyse learning effects, with one of the major problems being that as experience is gained in a new technique it tends to be deployed for higher risk patients, meaning that outcomes can deteriorate as experience is gained. This requires careful case-mix adjustment to interpret correctly, but almost by definition there are insufficient data for such analyses early in the learning experience.

The Leapfrog Group, a large US-based consortium of health care purchasers (Birkmeyer *et al* 2004) places great importance on volume thresholds. However, the impression from the literature is that definitions of 'low volume' and 'high volume' relate more to *potential* volumes than to any objective evidence on the *actual* level of activity which is required to achieve and/or maintain competence. For example, a unit performing 100 carotid endarterectomies per year could be classified 'high volume' whereas a unit performing 400 coronary artery bypass graft procedures per year could be classified as 'low volume' (Gandjour *et al* 2003, Birkmeyer *et al* 2004).

Studies which present outcome data for a range of activity volumes, as opposed to a simple low volume/high volume dichotomy, do often report poor outcomes at low activity levels and then a levelling off with outcomes in moderate volume units being comparable to outcomes in high volume units. This is partly the result of a statistical artefact, with greater variation being observed in the small samples which derive from low volume units. However when this excess variation is taken into account there is still evidence of poorer outcomes being observed in very low activity units (see, for example, the review by Shahian and Normand, 2003).

4.5 Causation

In spite of there being very strong evidence of an association between increased volume and better outcomes, there are remarkably few studies which try to assess if this association is causal. It was stated in the York review that there was no evidence that increasing the volume of activity in a given unit would lead to an improvement in outcomes. This reflected a lack of evidence rather than

evidence of a lack of effect. Evidence in this area is still extremely limited, but a number of studies are beginning to appear which evaluate the impact of interventions designed to concentrate activity

Trauma systems are an area with a long history of regionalisation, and where different approaches adopted by different countries constitute a 'natural experiment' on the organisation of trauma care. Nathens *et al* (2004) review the history of trauma management in the US and in France, and demonstrate how outcomes of trauma victims improved in the US following the introduction of regionalisation. However, there was a substantial lag period between the introduction of regionalisation and an observed improvement in outcome.

The UK Neonatal Staffing Study Group (2002) reviews the evidence for regionalisation of neonatal intensive care units. The situation is complex, but evidence of volume/outcome relationships from older studies is not seen currently, and this is ascribed to lower volume units adopting developments in treatment which were initially used only in high volume units. Training and staffing levels appear to be more important than volume *per se*. The study also raises the caveats that high volume units with a large number of consultant staff had higher levels of nosocomial bacteraemia, and that units running close to capacity have worse outcomes then when there is 'slack' in the system. These findings on the importance of staffing levels are consistent with the analysis of English hospital death rates published by Jarman *et al* (1999).

Nobilio *et al* (2004) report on the impact of regionalisation of cardiac surgery in an Italian region. They looked at patient outcomes, accessibility for patients and the efficiency of referral systems following the adoption of a "hub & spoke" model. The study does provide evidence of benefit, and the authors conclude that their findings suggest "*that policies aimed at increasing cooperation rather than competition among health service providers have a positive impact on quality of care*."

This latter finding is consistent with data from the Lothian Surgical Audit which was presented at the recent Consensus Conference of the Association of Surgeons of Great Britain and Ireland on "Modernising Medical Careers and General Surgery". (Robson et al, 2005). In the Lothian's experience, restructuring of emergency surgical care, focussed on sub-specialisation appropriate to upper and lower abdominal conditions, has led to improved quality of care and outcome.

4.6 Examples of the Magnitude of Volume/Outcome Associations

Halm *et al* (2002) summarised volume/outcome effects in terms of absolute differences in mortality between high and low volume hospitals (See Appendices II & III) and Gandjour *et al* (2003) presented

mortality rates for high volume relative to low volume hospitals. Absolute differences in mortality rates of the order of 10% are reported when high volume units are compared to low volume units in a number of complex high risk surgical procedures including paediatric cardiac surgery, surgery to repair ruptured abdominal aortic aneurysms, pancreatic cancer surgery and oesophageal cancer surgery. Relative differences in mortality rates of at least 10% are reported in a range of common lower risk procedures including percutaneous transluminal coronary angioplasty, carotid endarterectomy, knee replacement and surgery for hip fracture. Associations between 5 levels of relative volume and outcome, drawn from 2.5 million procedures for 14 interventions, in the study of Birkmeyer et al 2002, are illustrated in Appendix IV.

5. Commentary: Implications of Findings

The interpretation of the results of volume/outcome studies is complex. At the time of the York Review, methodological deficiencies in the evidence base meant that the studies had little if any relevance to health service planning. Recent improvements in the methodological rigour of at least a proportion of published volume/outcome studies mean that there is now a large body of credible evidence. This shows both substantial effects in a limited range of complex high risk surgical procedures and modest but clinically relevant effects in a wide range of more general procedures. Questions remain about the nature of the effects and their implications for service planning. Clarification of these depends upon somewhat different avenues of thinking.

The effects are likely to be most clear in circumstances where the condition is complex and its treatment associated with high risk, and where data from substantial numbers of patients are available, covering a wide spectrum of levels of volume. This is reflected in the abundance and consistency of evidence about complex, high risk surgical procedures. This is already accepted into service delivery. Indeed, in specialities such as neurosurgery, cardiac surgery and transplantation, the debate is not if they should be regionalised but if greater, even national concentration is appropriate. Furthermore, the relationship between increased volume and improved outcome in these circumstances is likely to be continuous, with improvement even at relatively high levels of experience. One exception may be if the volume becomes excessive, so that penalties of "overwork" lead to deterioration in outcome. For more common, less complex procedures, the improvement in outcome with increasing volume is likely to diminish beyond a certain threshold. In practice the majority of services will lie between these two extremes, which are illustrated graphically in Appendix V.

For some disorders, even though evidence is less abundant and the extent of the effect not dramatic (and hence less easy to show and more controversial), the consequences may still be important. For example, reduction by a few percent in mortality for myocardial infarction could save many lives in Scotland.

Another issue is the fact that the volume/outcome literature looks at average effects. Although high volume is associated with good outcome in general, there are low volume hospitals whose outcomes are superior to typical high volume hospitals and there are high volume surgeons with poor results who work within high volume hospitals. However this does not deflect the criticism that, in the NHS, too many operations are still taking place in hospitals with low volumes and that this may change if patient choice is effective (Taylor 2004).

Is it the volume of activity for an individual surgeon or physician which is important or the volume of the relevant unit or hospital? More studies have looked at hospital volume than have looked at surgeon volume. There is evidence that each can be influential, with perhaps hospital volume stronger but there is no consensus. A related point is if the surgeon/hospital needs to be 'high volume' for the procedure in question, or is high volume in general associated with good outcomes for all procedures? Urbach and Baxter (2004), for example, argue that volume in general is more important than volume for the specific procedure.

There is an increasing focus on indicators of outcome other than mortality. The occurrence of infection or other post-operative complications are generally applicable indices. Procedure - specific "clinical" indicators can include if intestinal surgery is followed by a stoma or if this is avoided by anastomosis, recurrence after procedures for hernia, varicose vein, and the persistence of reduced mobility and pain after orthopaedic surgery. Indices appropriate to 'medical' care are well established but rarely if ever available from routine sources of information. Quality control of diagnostic work is well established and recent evidence links volume with accuracy of breast screening. (Théberge et al 2005). In a service increasingly taking account of patient choice, satisfaction with the whole process of care will need to be taken into account, with the potential of benefit from high volume, highly experienced specialised team care to be set against local familiarity, and convenience.

It becomes a value judgement to explain the observed volume/outcome relationships. The two widely cited explanations are:- 'practice makes perfect' and 'selective referral' (i.e. patients are selectively referred to clinicians or hospitals that have historically achieved good outcomes). The former would suggest that volume/outcome associations are causal but the latter would imply that the observed associations are artefactual rather than causal. There are also issues such as aspects of process

and/or structure which are associated with high volume (e.g. a large well staffed intensive care unit) and which might lead indirectly to better outcomes.

A major current problem in applying these findings is the shortage of evidence supporting the hypothesis that the volume/outcome association is a causal association, whereby manipulating volume will have a beneficial impact on outcome. It should be noted that the problem is a lack of evidence rather than clear evidence of a lack of a causal effect. There is some limited evidence accumulating to support the association as being causal, but a great deal more research is needed in this area. The extent to which benefits can be achieved through diffusion of 'best practice' from 'centres of excellence' needs to be defined. Rigorous evidence of the effectiveness of the Clinical Networks established since 1998 would be important. Ultimately, the quality and safety of any service or procedure can be assured only by focussed audit.

Service planners may complain of the lack of clear cut, quantitative relationships, particularly concerning thresholds. In practice, the responsibility perhaps now lies with planners to specify the extent of effect that will be crucial in their decision making. Evidence may need to be stronger if it is the only or main factor for change in an existing arrangement. Where reconfiguration is needed in response to other factors, a general presumption of volume/outcome relationship is a reasonable starting point, and the issue may be more what level of effect is relevant in the circumstances under consideration. The more sensitive the indicator of outcome used, and the more common the condition, the longer the list of interventions to which the volume/outcome effect will be relevant.

6. Conclusions from Review

Returning to the three conclusions extracted from the York Review in Section 2, it is clear that the concerns over case-mix adjustment no longer hold. There is now a core of studies of adequate methodological quality to establish striking volume/outcome associations in certain complex high risk surgical procedures and more modest but clinically relevant effects in a wide range of common procedures. The size of the effect is influenced by the index of outcome and the range of volume considered.

There is still only limited evidence to suggest that the observed associations are causal, and that interventions to manipulate volume can lead to better outcomes. It is, however, very important to note that the issue here is that evidence is sparse, rather than there being strong evidence of a lack of a causal association. The relevance of the observed volume/outcome relationships to health

service planning depends crucially on how one interprets the underlying mechanisms which generate the associations.

The recent literature appears, in general, to provide limited support for the final conclusion of the York Review. Benefits arising from manipulation of volume are likely to be most clearly apparent at a relatively low level of volume. The rate of improvement may then diminish but further improvement can still occur.

References

Begg CB, Riedel ER, Bach PB, et al. Variations in morbidity after radical prostatectomy. *N Eng. J Med* 2002; **346**: 1138-1144.

Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. *N Eng J Med* 2002; **346**: 1128-1137.

Birkmeyer JD and Dimick JB. Potential benefits of the new Leapfrog standars: effect of process and outcome measures. *Surgery* 2004; **135**: 569-575.

Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. *N Eng J Med* 2003; **349**: 2117-2127.

Cook JA, Ramsay CR and Fayers P. Statistical evaluation of learning curve effects in surgical trials. *Clinical Trials* 2004; **1**: 421-427.

Gandjour A, Bannenberg A and Lauterbach KW. Threshold volumes associated with higher survival in health care. A systematic review. *Med Care* 2003; **41**: 1129-1141.

Halm EA, Lee C and Chassin MR. Is volume related to outcome in health care? A systematic review and methodological critique of the literature. *Ann Intern Med* 2002; **137**: 511-520.

Hodgson DC, Zhang W, Zaslavsky AM, et al. Relation of hospital volume to colostomy rates and survival for patients with rectal cancer. *J Natl Cancer Inst* 2003; **95**: 708-716.

Jarman B, Gault S, Alves B, et al. Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999; **318**: 1515-1520.

Luft HS, Bunker, JP, Einthoven AC. Should operations be regionalised? The relationship between surgical volume and mortality. *N Eng J Med* 1979; **301** 1364-1369.

Nathens AB, Brunet FP and Maier RV. Development of trauma systems and effect on outcomes after injury. *Lancet* 2004; **363**: 1794-1801.

Nobilia L, Fortuna D, Vizioli M, et al. Impact of regionalisation of cardiac surgery in Emilia-Romagna, Italy. *J Epidemiol Community Health* 2004; **58**: 97-102.

Panageas KS, Schrag D, Riedel E et al. The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. *Ann Intern Med* 2003; **139**: 658-665.

Ramsay CR, Grant AM, Wallace SA, et al. Statistical assessment of the learning curves of health technologies. *Health Technology Assessment* 2001; **5** (12): 1-98.

Robson AJ, Richards JMJ, Nixon SJ and Paterson-Brown, S. Specialisation of emergency surgery in the management of perforated peptic ulcer disease. *Br J Surg* 2005; **92**, Suppl 1: 147.

Shahian DM and ST Normand. The volume-outcome relationship: from Luft to Leapfrog. *Ann Thorac Surg* 2003; **75**: 1048-1058.

Sowden AJ, Grilli R and Rice N. The relationship between hospital volume and quality of health outcomes. CRD report 8, part 1. York: Centre for Reviews and Dissemination, 1997.

Taylor R. "Clinical Quality, where volume makes a difference" *Health Service Journal* 9 Dec 2004 p23.

Théberge I, Hébert-Goteau N, Langlois Et al. Volume of screening mammography and performance in the Quebec population Breast Cancer Screening Programme. *CMAJ*, 2005, **172**: 195-9.

The UK Neonatal Staffing Study Group. Patient volume, staffing, and workload in relation to risk-adjusted outcomes in a random stratified sample of UK neonatal intensive care units: a prospective evaluation. *Lancet* 2002; **359**: 99-107.

Urbach DR and Baxter NN. Does it matter what a hospital is "high volume" for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data. *BMJ* 2004; **328**: 73

Appendix I

List of 50 Articles Reviewed

for this Report

1. Bachmann MO, Alderson D, Edwards D, et al. Cohort study in South and West England of the influence of specialization on the management and outcome of patients with oesophageal and gastric cancers. *Br J Surg* 2002; **89**: 914-922.

2. Begg CB, Riedel ER, Bach PB, *et al*. Variations in morbidity after radical prostatectomy. *N Eng. J Med* 2002; **346**: 1138-1144.

3. Begg CB, Scardino PT. Taking stock of volume-outcome studies. *J.Clin.Oncol.* 2003; **21**: 393-4.

4. Berman MF, Solomon RA, Mayer SA, *et al*. Impact of hospital-related factors on outcome after treatment of cerebral aneurysms. *Stroke* 2003; **34**: 2200-7.

5. Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcome measures. *Surgery* 2004; **135**: 569-575.

6. Birkmeyer JD, Finlayson EV, Birkmeyer CM. Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative. *Surgery* 2001; **130**: 415-422.

7. Birkmeyer JD, Siewers AE, Finlayson EV, *et al*. Hospital volume and surgical mortality in the United States. *N Eng J Med* 2002; **346**: 1128-1137.

8. Birkmeyer JD, Stukel TA, Siewers AE, *et al*. Surgeon volume and operative mortality in the United States. *N Eng J Med* 2003; **349**: 2117-2127.

9. Birkmeyer JD, Warshaw AL, Finlayson SR, *et al*. Relationship between hospital volume and late survival after pancreaticoduodenectomy. *Surgery* 1999; **126**: 178-183.

10. 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**: 1071-1076.

11. Chang R-KR, Klitzner TS. Can regionalization decrease the number of deaths for children who undergo cardiac surgery? A theoretical analysis. *Paediatrics* 2002; **109**: 173-181.

12. Chen EW, Canto JG, Parsons LS, *et al*. Relation between hospital intra-aortic balloon counterpulsation volume and mortality in acute myocardial infarction complicated by cardiogenic shock. *Circulation* 2003; **108**: 951-957.

13. Cook JA, Ramsay CR and Fayers P. Statistical evaluation of learning curve effects in surgical trials. *Clinical Trials* 2004; **1**: 421-427.

14. Cowan JA Jr, Dimick JB, Leveque JC, *et al*. The impact of provider volume on mortality after intracranial tumor surgery. *Neurosurgery* 2003; **52**: 48-53.

15. Daley J. Invited commentary: quality of care and the volume-outcome relationship--what's next for surgery? *Surgery* 2002;**131**:16-8.

16. Dimick JB, Cowan JA, Jr., Ailawadi G, *et al*. National variation in operative mortality rates for esophageal resection and the need for quality improvement. *Arch.Surg.* 2003; **138**: 1305-1309.

17. Dimick JB, Cowan JA, Jr., Upchurch GR, Jr., Colletti LM. Hospital volume and surgical outcomes for elderly patients with colorectal cancer in the United States. *J.Surg.Res.* 2003; **114**: 50-66.

18. Dimick JB, Stanley JC, Axelrod DA, *et al*. Variation in death rate after abdominal aortic aneurysmectomy in the United States: impact of hospital volume, gender, and age. *Ann Surg* 2002; **235**: 579-585.

19. Feasby TE, Quan H, Ghali WA. Hospital and surgeon determinants of carotid endarterectomy outcomes. *Arch of Neurology* 2002; **59**: 1877-1881.

20. Finlayson EV,.Birkmeyer JD. Effects of hospital volume on life expectancy after selected cancer operations in older adults: a decision analysis. *J.Am.Coll.Surg.* 2003; **196**: 410-417.

21. Gandjour A, Bannenberg A, Lauterbach KW. Threshold volumes associated with higher survival in health care. A systematic review. *Med Care* 2003; **41**: 1129-1141.

22. 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**: 1155-1162.

23. Goodney PP, O'Connor GT, Wennberg DE, Birkmeyer JD. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? *Ann.Thorac.Surg.* 2003; **76**: 1131-1136.

24. Gutow AP. Surgeon volume and operative mortality. *N.Engl.J.Med.* 2004; **350**:1256-8.

25. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodological critique of the literature. *Ann Intern Med* 2002; **137**: 511-520.

26. Hannan EL, Wu C, Ryan TJ, *et al*. Do hospitals and surgeons with higher coronary artery bypass graft surgery volumes still have lower risk-adjusted mortality rates? *Circulation* 2003; **108**: 795-801.

27. Ho V. Certificate of need, volume, and percutaneous transluminal coronary angiography outcomes. *Am Heart J* 2004; **147**: 442-448.

28. Hodgson DC, Fuchs CS, Ayanian JZ. Impact of patient and provider characteristics on the treatment and outcomes of colorectal cancer. *J.Natl.Cancer Inst.* 2001; **93**: 501-515.

29. Hodgson DC, Zhang W, Zaslavsky AM, *et al*. Relation of hospital volume to colostomy rates and survival for patients with rectal cancer. *J Natl Cancer Inst* 2003; **95**: 708-716.

30. Jarman B, Gault S, Alves B, *et al*. Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999; **318**: 1515-1520.

31. Kizer KW. The volume-outcome conundrum. *N.Engl.J.Med.* 2003;**349**:2159-61.

32. Liu JH, Etzioni DA, O'Connell JB, *et al*. Using volume criteria: Do Californian hospitals measure up? *J Surg Research* 2002; **113**: 96-101.

33. Matz PG. Editorial comment--Spontaneous subarachnoid hemorrhage: volume, experience, and outcome. *Stroke* 2003; **34**:2206-7.

34. Meyerhardt JA, Catalano PJ, Schrag D, *et al*. Association of hospital procedure volume and outcomes in patients with colon cancer at high risk of recurrence. *Ann Intern Med* 2003; **139**: 649-657.

35. Nathens AB, Brunet FP and Maier RV. Development of trauma systems and effect on outcomes after injury. *Lancet* 2004; **363**: 1794-1801.

36. Nobilio L, Fortuna D, Vizioli M, *et al*. Impact of regionalisation of cardiac surgery in Emilia-Romagna, Italy. *J Epidemiol Community Health* 2004; **58**: 97-102.

37. Nobilio L,.Ugolini C. Different regional organisational models and the quality of health care: the case of coronary artery bypass graft surgery. *J.Health Serv.Res.Policy* 2003; **8**: 25-32.

38. Panageas KS, Schrag D, Riedel E *et al*. The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. *Ann Intern Med* 2003; **139**: 658-665.

39. Parks RW, Bettschart V, Frame S, *et al*. Benefits of specialisation in the management of pancreatic cancer: results of a Scottish population-based study. *Br J Cancer* 2004; **91**: 459-465.

40. Peterson ED, Coombs LP, DeLong ER, *et al*. Procedural volume as a marker of quality for CABG surgery. *JAMA* 2004; **291**: 195-201.

41. Ramsay CR, Grant AM, Wallace SA, *et al*. Statistical assessment of the learning curves of health technologies. *Health Technology Assessment* 2001; **5** (12): 1-98.

42. Shahian DM and ST Normand. The volume-outcome relationship: from Luft to Leapfrog. *Ann Thorac Surg* 2003; **75**: 1048-1058.

43. Smink DS, Finkelstein JA, Kleinman K, Fishman SJ. The effect of hospital volume of pediatric appendectomies on the misdiagnosis of appendicitis in children. *Pediatrics* 2004; **113**: 18-23.

44. Soljak M. Volume of procedures and outcome of treatment. *BMJ* 2002; **325**: 787-788.

45. The UK Neonatal Staffing Study Group. Patient volume, staffing, and workload in relation to risk-adjusted outcomes in a random stratified sample of UK neonatal intensive care units: a prospective evaluation. *Lancet* 2002; **359**: 99-107.

46. Urbach DR, Baxter NN. Does it matter what a hospital is “high volume” for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data. *BMJ* 2004; **328**: 737-740.

47. Vaughan Sarrazin MS,.Rosenthal GE. Hospital volume and outcome after coronary angioplasty: Is there a role for certificate of need regulation? *Am.Heart J.* 2004; **147**: 383-385.

48. Wibe A, Eriksen MT, Syse A, *et al*. Effect of hospital caseload on long-term outcome after standardization of rectal cancer surgery at a national level. *Br J Surg* 2004; **92**: 217-224.

49. Wigmore SJ, Madhavan K, Currie EJ, *et al*. Does the subspeciality of the surgeon performing primary colonic resection influence the outcome of patients with hepatic metastases referred for resection? *Ann Surg* 1999; *230*: 759-765.

50. Yeast JD, Poskin M, Stockbauer JW, *et al*. Changing patterns in regionalization of perinatal care and impact on neonatal mortality. *Am J Obstet Gynecol* 1998; **178**: 131-135.

Appendix II

Findings in Articles Relating Hospital Volume and Death

(Summary of Data in Halm et al 2002)

Procedures or Diagnosis	Number of Studies	Number with Significant Volume/Outcome Association	Median Average Mortality	Median Absolute Difference High-Low Volume
Ruptured Abdominal Aortic Aneurysm	8	2	50%	8%
AIDS	6	6	17%	9%
Myocardial Infarction	2	2	14.5%	2.3%
Oesophageal Cancer	3	3	14%	12%
Cerebral Aneurysm	3	3	14%	8%
Gastric Cancer	3	1	11%	6.5%
Pancreatic Cancer	10	9	10%	13%
Unruptured Abdominal Aortic Aneurysm	8	7	7.5%	3.3%
Paediatric Cardiac	3	3	7%	11%
Colorectal Cancer	10	4	6%	2%
Lung Cancer	4	2	5.5%	2%
Coronary Bypass	8	6	4%	1.6%
Limb Vascular	2	1	3.5%	1.2%
Prostrate TUR	2	2	1.9%	0.8%
Carotid Surgery	15	7	1.8%	0.4%
Coronary Angioplasty	9	5	1.4%	0.2%
Hip Replacement	8	3	0.8%	0.7%
Hip Fracture	2	2	NA	NA
Prostatectomy Open	2	2	0.3%	0.15%
Breast Cancer Surgery	1	1	-	60% higher 5 year mortality
Knee Replacement	1	1	0.2%	0.1%

Appendix III

Findings in Articles Relating Physician Volume and Death

(Summary of Data in Halm et al 2002)

Procedures or Diagnosis	Number of Studies	Number with Significant Volume/Outcome Association	Median Average Mortality	Median Absolute Difference High-Low Volume
Ruptured Abdominal Aortic Aneurysm	3	3	54%	14.5%
Pancreatic Cancer	2	1	10.5%	8.5%
Gastric Cancer	2	2	9.2%	4.8%
Unruptured Abdominal Aortic Aneurysm	1	0	7.6%	3.2%
Paediatric Cardiac	1	1	6.8%	2.9%
Coronary Bypass	3	3	3.7%	2.2%
Colorectal Cancer	5	4	3%	1.9%
Carotid Surgery	12	7	2%	1.4%
Lung Cancer	1	0	1.9%	1.1%
Coronary Angioplasty	5	1	1.0%	0.06%
Hip Replacement	3	2	0.4%	0.8%
AIDS	1	1	NA	NA
Breast Cancer	1	1	NA	NA
Myocardial Infarction	1	1	NA	NA

Appendix IV

Adjusted Mortality among Medicare Patients (1994-1999) according to Quintile of Hospital Volume

(Reproduced from Birkmeyer et al 2002)

Figure 1: Peripheral Vascular Procedures (Panel A) and Cardiac Procedures (Panel B)

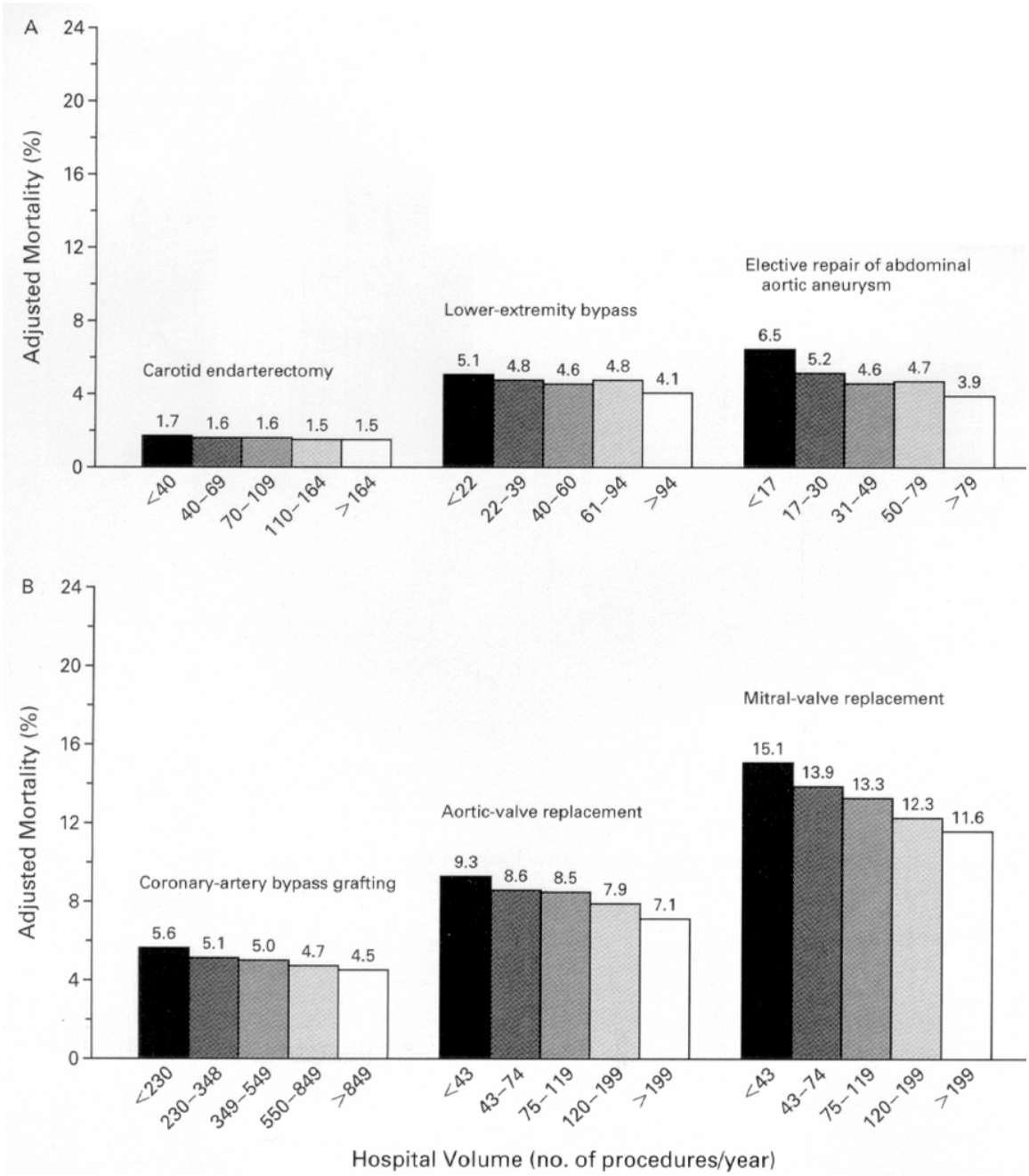
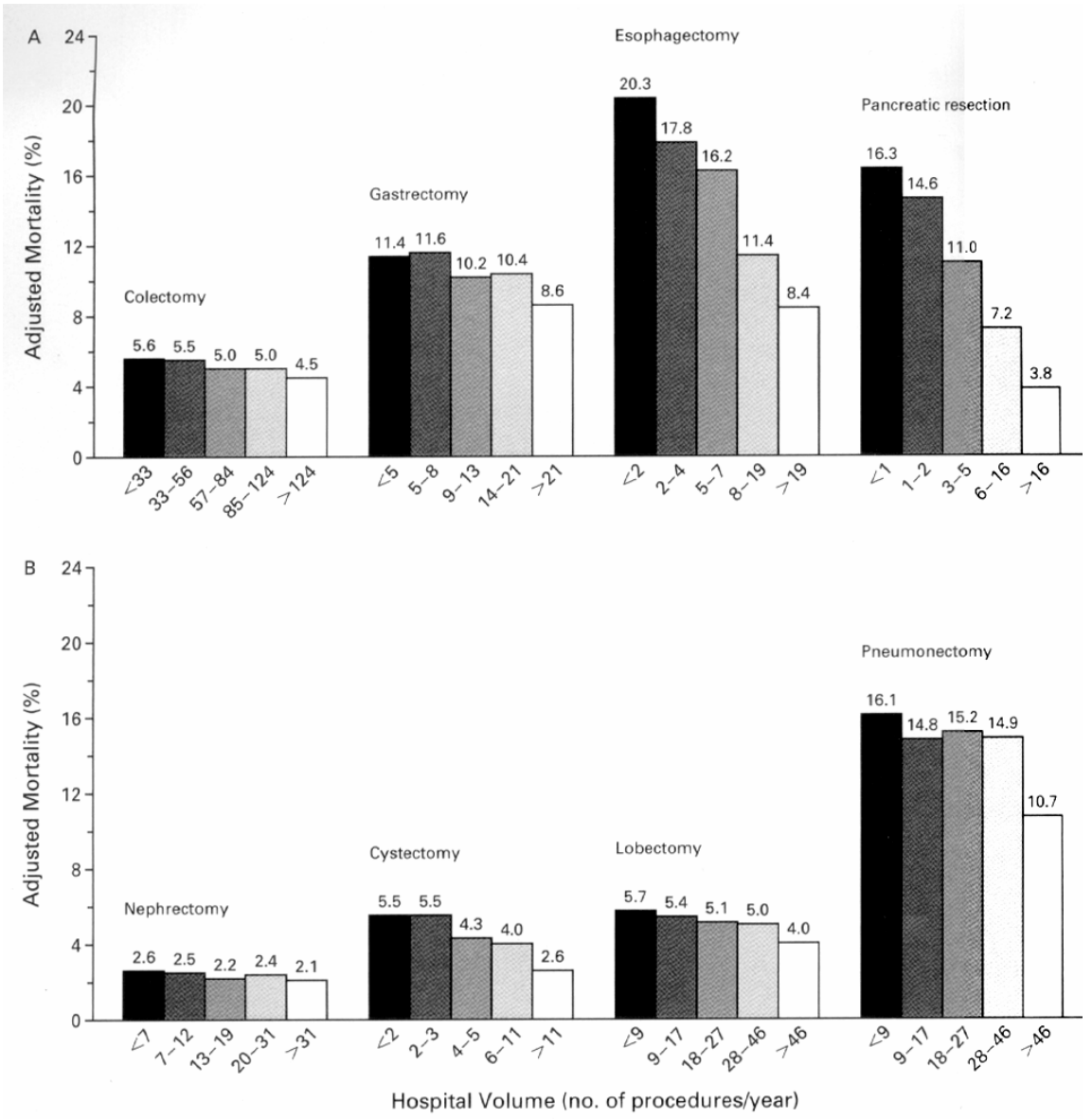


Figure 2: Resection of Gastrointestinal Cancers (Panel A) and Resection of Other Cancers (Panel B)



Appendix V

Illustrative Volume/Outcome Relationships for Procedures Ranging from Simple and Common To Complex and Uncommon