Published in final edited form as:
Lancet. 2013 September 28; 382(9898): 1121–1129. doi:10.1016/S0140-6736(13)61215-5.

Understanding regional variation in the use of surgery

John D. Birkmeyer, MD,
Center for Healthcare Outcomes & Policy, University of Michigan, Ann Arbor, MI

Bradley N. Reames, MD,
Center for Healthcare Outcomes & Policy, University of Michigan, Ann Arbor, MI

Peter McCulloch, MD,
Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

Andrew J. Carr, MA, ChM, FRCS, FMedSci,
Nuffield Department of Orthopaedic, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

W. Bruce Campbell, MS, FRCP, FRCS, and
National Institute for Health and Clinical Excellence, London, UK

John E. Wennberg, MD, MPH
The Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, NH

Abstract

The use of common surgical procedures varies widely across geographical regions. Differences in illness burden, diagnostic practices, and patient attitudes about medical intervention explain regional variation in surgery rates to only a small degree. Instead, current evidence suggests that surgical variation primarily reflects differences in physician beliefs about the indications for surgery and the extent to which patient preferences are incorporated into treatment decisions. These two components of clinical decision making help explain the "surgical signatures" of specific procedures, as well as why some consistently vary more than others. Variation in clinical decision making is in turn influenced by broader environmental factors, including technology diffusion, specialist supply and local training paradigms, financial incentives, and regulatory factors, which vary across countries. Better scientific evidence about the comparative effectiveness of surgical and non-surgical interventions may help mitigate regional variation, but

Correspondence: Dr. John Birkmeyer, Center for Healthcare Outcomes & Policy, Building 16 / 136E, 2800 Plymouth Road, Ann Arbor, MI 48109, 734-998-7470; jbirkmey@umich.edu.

Contributors
Dr. Birkmeyer conceived the narrative review and was the lead writer. Dr. Reames had primary responsibility for the literature review, with input from Dr. Wennberg. All authors assisted with interpretation of data, critical revision, and gave final approval of the manuscript.

Conflicts of interest
Dr. Birkmeyer is Chief Scientific Officer and has an equity interest in ArborMetrix, a software company that provides cost and quality decision support in specialty care. Dr Wennberg is a paid consultant to the Informed Medical Decision Foundation and receives royalties from Health Dialog. The Foundation and Health Dialog provide patient decision aids used to help patients make informed decisions when choosing treatments for preference-sensitive conditions. None of these groups were involved in the creation of this manuscript. The remaining authors have no conflicts of interest.

broader dissemination of shared decision making tools will be essential in reducing variation with preference-sensitive conditions.

Introduction

It is an uncomfortable fact that a patient's odds of undergoing surgery often depend more on where he lives than on his clinical circumstances. Almost eighty years ago, J. Alison Glover noted that tonsillectomy rates in Britain varied dramatically according to school district.[1] In 1936, for example, a child in Enfield was twenty times more likely to undergo the operation than a child in Hornsey, even though only seven miles separated the two districts. In a *Science* article published 40 years later, Wennberg and Gittelsohn noted that rates of tonsillectomy varied almost 12-fold across counties in rural Vermont, while other common procedures varied almost as much.[2] This study and subsequent analyses based on data from New England served to usher in the field of "small area analysis."[3–5]

Despite considerable advances in medical science, there is little evidence that regional variation in the use of surgery is shrinking over time. Although some studies have documented small improvements in medical practice variation,[6–8] the relative degree of variation in population-based rates of 10 common surgical procedures has been remarkably stable over the past 20 years, according to US Medicare data compiled longitudinally by the *Dartmouth Atlas of Healthcare.* In 2008–10, rates of hip replacement, coronary bypass surgery, prostatectomy and many other major procedures continued to vary at least four- to five-fold across hospital referral regions (Figure 1). Data from the United Kingdom documented similar degrees of variation in the use of surgery among 152 primary care trusts in 2009–10.[9, 10]

In this review, we examine the major determinants of regional variation in procedure rates. After considering the role of patient "demand," we review the evidence that surgical variation reflects differences in physicians' beliefs about the value of surgery in specific circumstances, and in the extent to which they incorporate patient preferences into surgical decisions. We consider how surgery rates—and regional variation—are influenced by external factors, such as new technology, surgeon supply, and financial incentives. Finally, while focusing on regional variation within countries, we consider the literature on small area analysis in the context of international differences in the use of surgery. Although others have suggested alternative models of medical practice variation,[11] this review focuses on the clinical decision making paradigm, which we believe best fits current evidence on the root causes of regional variation in the use of surgery.

Variation in patient demand for surgery

Some regions may have higher rates of specific procedures because of greater demand for surgery there. We use the term "demand" here in the broad sense, to reflect patient-related factors influencing the real or perceived need for surgery—"upstream" of decisions made by surgeons. Such factors include the true incidence of surgically-treatable disease, the frequency in which sub-clinical disease is detected with medical testing, and the willingness

of patients to undergo surgical intervention. Each of these factors has been tied to regional rates of surgery in various clinical contexts.

The role of regional variation in disease incidence is most obvious in clinical conditions for which surgery is almost universally required or recommended, like hip fracture. In the US, regional rates of hip fracture repair in the elderly are almost linearly correlated with the regional incidence of hip fracture. Rates of hip fracture surgery in Hawaii are at least 60% lower than elsewhere in the US, not because orthopedists there are more conservative in their management, but because fewer patients have hip fractures.[12] The role of disease incidence is further suggested by correlations between surgical rates and disease risk factors. For example, the southeastern part of the US has markedly higher rates of all types of cardiovascular interventions than other regions of the country, but also much higher prevalence of cardiovascular risk factors, including cigarette smoking, obesity, and diabetes.

Even when the true prevalence of disease varies little by geography, the number of surgically treatable patients could vary according to regional differences in diagnostic testing of patients with asymptomatic or subclinical disease. For example, while there is little evidence that the true incidence of prostate cancer varies widely within countries, rates of prostate specific antigen (PSA) screening—the most common means by which this disease is detected—differ markedly. Moreover, these variations in screening are strongly correlated with variations in prostatic biopsy and resection rates.[13–15] Lu-Yao and colleagues found that PSA screening rates were five times higher in Seattle compared to Connecticut, which helped explain a five-fold spread in prostatectomy rates between the two cities.[14] Further evidence of the testing-treatment “cascade” is found in the use of revascularization interventions for coronary artery disease, which vary markedly. Regional rates of myocardial revascularization (percutaneous coronary interventions or coronary artery bypass surgery) are highly correlated ($R^2 = 0{\cdot}84$) with the regional frequency of cardiac catheterization, which provides both the diagnosis and the “road map” for revascularization.[16]

Finally, differences in patients’ willingness to undergo surgical intervention may play a part in regional variation in procedure rates. Hawker et al. conducted a population-based study of patients residing in two areas of Ontario, Canada with low and high rates of hip and knee arthroplasty. Based on functional status assessments and x-rays, a slightly higher percentage of residents aged over 55 years in the high rate area had a potential clinical need for surgery than those in the low rate area (3·6% vs. 2·9%, respectively). For those judged clinically appropriate for surgery, the investigators then presented patients with detailed information about the nature of and risks and benefits of joint replacement. Only 8·5% of patients in the low-rate area expressed that they were “definitely willing” to have surgery, while 14·9% of patients in the high-rate area expressed that preference.[17]

Although these and other studies have highlighted the potential influence of illness burden and patient demand, regional differences in such factors have tended to be small relative to differences in surgery rates and fail to fully explain regional variations. For example, the higher prevalence of cardiovascular disease may account for the generally higher intervention rates in the southeastern US, but it does not explain why the use of

cardiovascular surgery in adjoining counties within the southeast can differ by several-fold.[18] Moreover, a larger number of studies have failed to identify important differences in the clinical demand for surgery across regions with differing rates. For example, based on detailed household interviews conducted in adjacent areas of Vermont, Wennberg and Fowler found no evidence that disparities in illness burden explained wide regional variation in rates of different types of procedures.[19]

Variation in physician beliefs about the clinical indications for surgery

The most obvious, and most important, reason for regional variation in surgery rates is that physicians have differing attitudes and beliefs about the indications for surgery. In some instances, variation in procedure rates reflects differences in decisions by physicians about whether to refer patients to surgeons in the first place. In the seminal studies of the 1930s, Glover et al. concluded that variation in tonsillectomy across English school districts primarily reflected differences in the judgments of school health officers—the medical physicians responsible for diagnosis and referral of children for tonsillectomy. Supporting that assertion, tonsillectomy rates in specific school districts essentially changed overnight as one health officer was replaced with another.[1]

More direct evidence supporting Glover's hypothesis about the role of physician judgment appeared forty years later. This time, however, the research implicated the surgeons as well as referring physicians. Bloor and colleagues studied surgeons from two districts in Scotland with high and low tonsillectomy rates, observing surgeons as they examined patients and interviewing them afterwards about why they did or did not recommend surgery in specific cases.[20, 21] More "aggressive" surgeons tended to put more weight on the physical examination, though they differed about which specific findings—hyperemia of the anterior pillars, swollen lymph nodes, etc.—were most important in decisions to operate. In contrast, surgeons from districts with lower tonsillectomy rates tended to discount physical findings, placing greater emphasis on the patient's medical history, particularly on the number of previous tonsillitis episodes.

It is commonly assumed that high rates of procedure use mean that procedures are being used "inappropriately", but a recent systematic review found little evidence to support this belief.[22] In one study of 13 surgical procedures performed in Colorado hospitals, for example, 97% of the 4,850 cases reviewed either met published indications for surgery or were considered reasonable by external physician review.[23] In research focusing more directly on variation in procedure rates, Chassin et al. evaluated appropriateness in regions with low, average, and high rates of coronary angiography, upper endoscopy, and carotid endarterectomy. Overall, a substantial proportion of cases (17–32%, depending on the procedure) were judged inappropriate. However, regional differences in the proportion of inappropriate procedures were small and did not explain larger variations in overall rates.[24] A follow-up study of the same procedures in 23 adjacent counties in one state confirmed these results. Inappropriate procedures accounted for only 28% of the variation in the county rates of angiography, and none of the variation in endoscopy or carotid endarterectomy.[25]

In interpreting these paradoxical findings, it is worth noting that many decisions about surgery reside within a large grey area of clinical discretion, bounded by comparatively small "tails" of clearly appropriate and inappropriate indications for intervention, as defined by high level medical evidence. Moreover, regions can have comparatively high rates of certain procedures without providing unnecessary care because the reservoir of surgically eligible patients is often extremely large compared to operation rates. In the Ontario study by Hawker et al. described previously, estimates of the population need for hip or knee replacement—defined as clinically eligible by symptom scores and x-ray findings—exceeded surgical rates by several-fold, even in the high rate region.[17] For these reasons, procedure rates can vary markedly without surgeons "breaking the rules" of scientific evidence or infringing upon clinical guidelines.

Variation in the degree to which patient preferences are incorporated into surgical decisions

For some surgically treatable conditions, the "right" answer about whether to intervene should depend as much on patient preferences as on scientific evidence about clinical effectiveness. The value of any given surgical intervention depends on trade-offs between its benefits and risks. When there are large imbalances between the two, surgical decisions are clear-cut and procedure rates tend to vary little. Regional variation becomes more pronounced when such trade-offs are a close call, or when they require the value judgments of individual patients. The value of abdominal aortic aneurysm repair, for example, depends on the extent to which the upfront risk of perioperative mortality and morbidity is offset by the survival benefit associated with the lower risk of death from aneurysm rupture. In addition to scientific evidence about the comparative magnitude of risks and benefits in different patient subgroups, optimal decisions should account for how patients value or "discount" immediate versus deferred risks of death.

The most "preference sensitive" conditions are those for which surgical decisions involve more heterogeneous trade-offs.[26] With early stage prostate cancer, for example, radical prostatectomy minimizes a patient's risks of developing metastases and ultimately dying from the disease. In return for this small longevity benefit and the psychological benefits associated with cancer removal, patients need to consider the prospect of some degree of sexual dysfunction and the risk of permanent urinary incontinence. Experts and scientific studies may argue about the probabilities of different outcomes with surgery versus other treatment options, but few would disagree that such trade-offs are highly personal, and the "right answer" is supplied by each individual patient.

Differences in the extent to which physicians incorporate the preferences of individual patients into treatment decisions may be an important factor underlying regional variation in surgery rates. Although there is comparatively little empirical information on this topic, this hypothesis is supported by observations that procedure rates for preference-sensitive conditions, such as back pain and benign prostatism, tend to vary considerably more than rates of other interventions (Figure 1).

Field studies of the surgical treatment of early stage breast cancer illustrate how surgeon beliefs can trump the values of individual patients and foster variation in practice. Multiple clinical trials have confirmed the survival equivalence of breast conservation therapy (lumpectomy and radiation) and modified radical mastectomy.[27–30] While a patient retains her breast with the former, she achieves a slightly lower risk of local recurrence and can (in many cases) avoid the need for adjuvant radiation therapy with the latter. There is little evidence of significant differences across geographical areas in the way that patients view this basic tradeoff, but numerous studies from several different countries have documented significant regional variation in rates of the two procedures.[31–36] In the mid-1990's, a New York writer decided to investigate this phenomenon by interviewing surgeons in Rapid City, South Dakota, where the Dartmouth Atlas reported virtually nonexistent rates of lumpectomy. Her interviews uncovered strong opinions against lumpectomy among multiple local surgeons, with one surgeon stating, "It is my personal bias that mastectomy does better."[37]

Studies evaluating the effects of decision aids suggest that rates of surgical interventions are strongly influenced by the extent to which patients are involved in their own treatment decisions. Decision aids and interactive "shared decision making" tools provide patients with detailed information about the risks and benefits of alternative treatments for specific conditions. A recent systematic review of randomized-controlled trials evaluating such tools confirmed that they improve patient knowledge, promote a more accurate perception of risks and benefits, and increase consistency between the patient's informed values and the treatment chosen.[38] Although no studies have assessed whether decision aids reduce regional variation in procedure rates, numerous randomized trials from Europe and North America have shown significant decreases in overall surgery rates following their implementation.[39–41] In a recent study by Arterburn and colleagues, for example, the introduction of decision aids was associated with 26% fewer hip replacement surgeries and 38% fewer knee replacement surgeries performed in patients with osteoarthritis over a six-month period.[42]

Interpreting geographical patterns of surgical variation

The importance of clinical uncertainty and patient preferences are illustrated in the "variation profiles" of different procedures. The relative degree to which surgical interventions vary is commonly measured using the systematic component of variation (SCV). An SCV is more robust than other measures of variation because it distinguishes variation between areas (systematic variation) from variation within areas (random variation), and as a result reflects the true, non-random part of observed variation in surgery rates.[10, 43] These calculations are usually standardized for age and sex, and are calculated based on location of residence, regardless of where a surgery is performed.[43] In general, SCVs greater than 5 are considered indicative of high variation, while SCVs greater than 10 are considered indicative of very high variation.[44] At one end of the spectrum, regional rates of hip fracture repair and colectomy for colon cancer vary little across the 306 hospital referral regions of the United States (Figure 1). With the latter, for example, rates in the highest and lowest rate regions varied less than three-fold and the SCV was only 3·5 in 2008–10. For both procedures, there is wide professional consensus that surgery is the

preferred treatment strategy for almost all patients, and differences in patient preferences generally play a relatively small role in decisions about surgery.

Rates of hip replacement (SCV 5·8) and coronary artery bypass surgery (CABG, SCV 7·2) vary to an intermediate degree. For CABG, effectiveness has been examined in numerous large randomized clinical trials, but many of these trials have conflicting findings and do not provide evidence about all indications for the operation. CABG is moderately preference-sensitive. In some subsets of patients it is performed exclusively to prolong life expectancy, but in others more complex tradeoffs are required. Patients must balance the upfront risks of procedure-related mortality and cognitive disability against the downstream benefits of reduced angina and improved quality of life.

Radical prostatectomy occupies the far end of the variation spectrum. This procedure is associated with an SCV of 13·5; rates between the lowest and highest regions vary more than ten-fold. With early stage prostate cancer, patients have a wide range of treatment options, from continued surveillance, to many different types of radiation therapy, to surgical resection by various approaches. The comparative effectiveness of these competing strategies is incompletely studied and continues to be hotly debated among both individual physicians and specialty organizations. As described earlier, few procedures are as preference-sensitive.

Broader influences on regional variation in surgery rates

Although this review has focused on understanding regional variation in surgery rates in a clinical decision making paradigm, it is important to note that treatment decisions are influenced by many broader environmental factors (Figure 2).

Technology diffusion

Surgical innovation and new technologies can amplify regional variation in practice patterns. New procedures or new approaches to established procedures can contribute to variation simply by increasing the number of therapeutic alternatives available to clinicians. They are often disseminated in advance of rigorous evidence of their comparative effectiveness in specific populations. Variation can arise not only because of differences in physician opinions about whether the new procedure should substitute for existing interventions, but also because physicians disagree about whether the new procedure should change indications for surgery in the first place. For example, despite the absence of new evidence to inform thresholds for operating, overall rates of cholecystectomy in both the US and the UK increased substantially after the diffusion of laparoscopic surgery in the early 1990s.[45, 46]

Training and regional supply of surgeons

A region's use of a specific procedure—its "surgical signature"—may be influenced in large part by how surgical trainees are taught. Recent evidence suggests that training programs in the United States vary widely with regards to many aspects of clinical practice.[47, 48] Since physicians tend to establish their practices close to where they train, a region's use of

specific procedures may be perpetuated by the practice styles instilled by the training programs that supply its surgeon workforce.

Although the total number of physicians in a region is a major determinant of overall health spending, physician workforce is weakly correlated with surgery rates. In most instances, the regional supply of specialists is a weak predictor of procedure rates in the corresponding specialty area. In the US, for example, the regional supply of orthopedists has little association with rates of elective hip replacement, or most other procedures in that specialty. Such data highlight that surgeons in most specialties can do many different types of procedures, but tend to favor some over others. As noted by Chassin et al. with carotid endarterectomy 20 years ago, high procedure rates were explained not by differences in the number of surgeons within a given area, but rather by "enthusiasm" for that procedure among a small number of high volume surgeons in that region.[49] In a more recent survey study from Canada, spine surgeons reporting enthusiasm for a particular procedure had significantly higher rates of that procedure.[50]

Financial incentives and regulatory environment

Reimbursement models and physician incentives may influence not only overall utilization of specific procedures, but also regional variation in rates. For example, regional rates of common outpatient procedures, like endoscopy, cataract surgery, and arthroscopy, vary considerably. US regions in which those procedures are performed in physician-owned ambulatory surgery centers have overall rates twice as high or higher than those in which outpatient surgery is conducted in hospitals, by physicians that have no direct financial interest.[51, 52] Similarly, physician-owned hospitals specializing in cardiac care have significantly higher rates of CABG and PCI.[53] As delivery models and associated financial incentives vary geographically, so too will procedure rates.

Regulatory constraints on capacity may have the opposite effect and serve to dampen variation in surgery rates. For example, for many years New York State has limited the diffusion of new cardiac surgery programs with so-called "Certificate of Need" requirements, while California has had no such restrictions. While other factors may also be at work, New York has historically had much lower overall rates of CABG,[54] and less geographical variation within its boundaries.[18]

Nonetheless, it is important not to overstate the importance of financial incentives and regulatory factors as determinants of regional variation in surgery rates. Although such factors may help explain international differences in surgery rates (as described next), they cannot explain regional variation within countries sharing a common reimbursement model and healthcare policies. The fact that variation is often idiosyncratic and specialty-specific (regions have high rates for some procedures and low rates for others) also speaks against financial incentives and the regulatory environment as major determinants of regional variation in the use of surgery.

Regional intensity of medical care

Although many may assume that the use of surgery is a marker of the general "aggressiveness" in the approach to medical care by physicians serving a given region,

current evidence does not support this belief. At least in the US, there is very little correlation between a region's use of inpatient surgery and its overall healthcare utilization, as reflected by regional expenditures.[26, 55] Moreover, there is little evidence that some regions are simply more inclined toward surgical intervention than others. Instead, procedure rates tend to be idiosyncratic and highly condition- and specialty-specific. For example, data from the *Dartmouth Atlas* indicate that rates of cardiac procedures are largely uncorrelated with those of orthopedic procedures. More surprisingly, there is sometimes little correlation in regional rates of procedures of the same specialty or subspecialty. In Harlingen, Texas, for example, the rate of total hip replacement is among the lowest in the US, while the rate of total knee replacement is among the highest.[56]

International comparisons

Although this article reviews the literature on small area analysis (and thus regional variation within countries), it is worth noting that rates of common surgical procedures also vary markedly across international boundaries. Based on data from the 1990s, for example, rates of back surgery were 164% higher in the US than in Canada.[57] In more recent comparisons, rates of carotid interventions (either endarterectomy or stenting) were markedly higher in the US (300 per 100,000 Medicare beneficiaries in 2006) than across 10 European countries (9.6 per 100,000 in 2008).[58, 59] Such disparities no doubt reflect international differences in healthcare capacity, reimbursement policy and financial incentives, and other factors discussed in the previous section.

Like variation across small areas within countries, however, variation in surgical rates across countries also reflects differences in physician beliefs about indications for surgery. In the 1980s, for example, overall rates of inguinal hernia repair in the UK (137 per 100,000) were only half those in the US (276 per 100,000).[3, 60] Given the perceived risks of bowel strangulation and related complications (concerns largely debunked in later clinical trials), surgical repair was considered the "right" answer by American surgeons for almost all patients. In the UK, by contrast, patients with hernias that were small or minimally symptomatic often received a truss or were left untreated.[43, 60] Contemporary studies of carotid endarterectomy show that 26% of patients undergoing surgery in Europe have asymptomatic disease, a subgroup in whom the benefits of surgery are lower.[61] In the US, by contrast, patients without symptoms account for 48% to 92% of individuals undergoing this procedure.[62, 63]

Despite international differences in overall surgical rates, the extent to which specific procedures vary within countries is remarkably consistent.[9, 10] In one international study, the United States had substantially higher overall rates of hysterectomy, tonsillectomy, and hernia repair than Norway and the United Kingdom.[43] However, the relative degree of regional variation within countries was consistent across all three settings. Analysis of more recent data further confirms that variation "signatures" for most procedures are similar across national boundaries. As seen in Figure 3, the SCVs of six common procedures in the UK, as reported by *The King's Fund,* are almost linearly correlated with SCVs of the same procedures in the US, as published in the *Dartmouth Atlas*. Such data underscore that the

clinical decision making paradigm—not environmental and other external factors—is central to understanding and ultimately reducing regional variation in the use of surgery.

Conclusions

Practice variation is of course not unique to surgical care. Similar regional disparities have been described in medication prescribing patterns,[64–66] the use of radiological imaging,[67] and admission rates for medical conditions.[10, 68, 69] Nonetheless, because the consequences of an operation can be so much more dramatic for a patient than a prescription or an x-ray, it is not surprising that regional variation in surgery—and its implication that many patients are getting procedures they do not want or need — has garnered a disproportionate share of attention from researchers, policy makers, and patient advocacy groups.

As considered more fully elsewhere in this issue, optimal strategies for reducing regional variation in surgical rates remain uncertain. The persistence of geographical disparities, despite 80 years of research and growing professional awareness of this issue, suggests that natural history alone will be insufficient. One obvious answer is better evidence about the comparative effectiveness of competing therapies and thus an answer to the basic question, “Which rate is right?” Still, the stubbornness of regional variation in surgery rates, despite major advances in clinical science over time, raises questions about the power of scientific evidence alone in reducing variation. After all, some of the most widely studied surgical procedures continue to be the most variable, while other procedures vary little in the absence of high-level evidence.

Because high variation procedures tend to be preference-sensitive, reducing variation will instead require more systematic efforts to ensure that treatment decisions are driven by the well-informed preferences of individual patients. With growing evidence of their effectiveness for many surgically-treated conditions (see the companion article in this issue), decision aids and shared decision making programs are becoming increasingly popular among payers, policy makers, and health system leaders in the U.S and UK.[70, 71]

There are of course other important types of variation in surgical practice. Surgical technique and operative approach vary considerably among surgeons, driven more by surgeon training and preference than evidence about comparative effectiveness. More importantly, surgical quality and patient outcomes after surgery also vary considerably across surgeons, hospitals, and regions. Such variation has prompted wide-ranging efforts aimed at accelerating quality improvement in surgery, from centralization of complex procedures, to pay for performance initiatives, to operative checklists, to clinical registries and performance feedback.[72] Recent studies suggest that such efforts have begun to pay dividends for patients, at least as reflected by population-based studies of surgical mortality.[73] Nonetheless, improving the quality of decisions to operate in the first place is perhaps even more important to the well-being of patients, and will require sustained efforts aimed at reducing unwanted variation in regional rates of surgery

Acknowledgments

Dr. Birkmeyer is supported by grant from the National Institute on Aging (5-P01-AG-019783-10).

Dr. Reames is supported by a grant from the National Cancer Institute (5 T32 CA-009672-22).

These funding sources had no involvement in the manuscript herein. We thank Mark P. MacEachern for his assistance in constructing the literature search.

References

1. Glover JA. The incidence of tonsillectomy in school children. 1938. Int J Epidemiol. 2008; 37(1):9–19. [PubMed: 18245048]
2. Wennberg J, Gittelsohn. Small area variations in health care delivery. Science. 1973; 182(4117): 1102–8. [PubMed: 4750608]
3. Wennberg J, Gittelsohn A. Variations in medical care among small areas. Sci Am. 1982; 246(4): 120–34. [PubMed: 7079718]
4. Wennberg JE, Freeman JL, Culp WJ. Are hospital services rationed in New Haven or over-utilised in Boston? Lancet. 1987; 1(8543):1185–9. [PubMed: 2883497]
5. Wennberg JE, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. N Engl J Med. 1989; 321(17):1168–73. [PubMed: 2677726]
6. Westert GP, Groenewegen PP, Boshuizen HC, Spreeuwenberg PMM, Steultjens MPM. Medical practice variations in hospital care; time trends of a spatial phenomenon. Health & Place. 2004; 10(3):215–20. [PubMed: 15177196]
7. Raymont A. Hospital discharges in New Zealand 1991–2005: changes over time and variation between districts. N Z Med J. 2008; 121(1279):66–74. [PubMed: 18709049]
8. Hanstede MM, Burger MJ, Timmermans A, Burger MP. Regional and temporal variation in hysterectomy rates and surgical routes for benign diseases in the Netherlands. Acta Obstet Gynecol Scand. 2012; 91(2):220–5. [PubMed: 22043840]
9. The NHS atlas of variation in healthcare : reducing unwarranted variation to increase value and improve quality. [S.l.]. Right Care. 2010
10. Appleby, J.; Raleigh, V.; Frosini, F. Variations in Health Care : The Good, the Bad, the Inexplicable. London: King's Fund : [distributor] King's Fund; 2011.
11. Westert GP, Groenewegen PP. Medical practice variations: changing the theoretical approach. Scandinavian Journal of Public Health. 1999; 27(3):173–80. [PubMed: 10482075]
12. Dartmouth Atlas Project Database. The Dartmouth Institute of Health Policy and Clinical Practice. 1996–2010.
13. Sheikh K, Bullock C. Rise and fall of radical prostatectomy rates from 1989 to 1996. Urology. 2002; 59(3):378–82. [PubMed: 11880074]
14. Lu-Yao G, Albertsen PC, Stanford JL, Stukel TA, Walker-Corkery ES, Barry MJ. Natural experiment examining impact of aggressive screening and treatment on prostate cancer mortality in two fixed cohorts from Seattle area and Connecticut. BMJ. 2002; 325(7367):740. [PubMed: 12364300]
15. Shao YH, Albertsen PC, Shih W, Roberts CB, Lu-Yao GL. The impact of PSA testing frequency on prostate cancer incidence and treatment in older men. Prostate Cancer Prostatic Dis. 2011; 14(4):332–9. [PubMed: 21709691]
16. Lucas FL, Siewers AE, Malenka DJ, Wennberg DE. Diagnostic-therapeutic cascade revisited: coronary angiography, coronary artery bypass graft surgery, and percutaneous coronary intervention in the modern era. Circulation. 2008; 118(25):2797–802. [PubMed: 19064681]
17. Hawker GA, Wright JG, Coyte PC, et al. Determining the need for hip and knee arthroplasty: the role of clinical severity and patients' preferences. Med Care. 2001; 39(3):206–16. [PubMed: 11242316]
18. Dartmouth Medical School. Center for the Evaluative Clinical Sciences., Maine Medical Center (Portland Me.). Center for Outcomes Research and Evaluation. The Dartmouth atlas of cardiovascular health care. Chicago, Ill: AHA Press; 1999.
19. Wennberg JE, Fowler FJ Jr. A test of consumer contribution to small area variations in health care delivery. J Maine Med Assoc. 1977; 68(8):275–9. [PubMed: 894131]

20. Bloor MJ, Venters GA, Samphier ML. Geographical variation in the incidence of operations on the tonsils and adenoids. An epidemiological and sociological investigation. Part I. J Laryngol Otol. 1978; 92(9):791–801. [PubMed: 701980]
21. Bloor MJ, Venters GA, Samphier ML. Geographical variation in the incidence of operations on the tonsils and adenoids. An epidemiological and sociological investigation (Part 2). J Laryngol Otol. 1978; 92(10):883–95. [PubMed: 712221]
22. Keyhani S, Falk R, Bishop T, Howell E, Korenstein D. The relationship between geographic variations and overuse of healthcare services: a systematic review. Med Care. 2012; 50(3):257–61. [PubMed: 22329997]
23. Elliott RV, Kahn KA, Kaye R. Physicians measure up. A study of 13 surgical procedures. JAMA. 1981; 245(6):595–600. [PubMed: 7192755]
24. Chassin MRKJ, Park RE, Winslow CM, et al. Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. JAMA. 1987; 258(18): 2533–7. [PubMed: 3312655]
25. Leape LL, Park RE, Solomon DH, Chassin MR, Kosecoff J, Brook RH. Does inappropriate use explain small-area variations in the use of health care services? JAMA. 1990; 263(5):669–72. [PubMed: 2404147]
26. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. Health Aff (Millwood). 2002; (Suppl Web Exclusives):W96–114. [PubMed: 12703563]
27. Fisher B, Bauer M, Margolese R, et al. Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. N Engl J Med. 1985; 312(11):665–73. [PubMed: 3883167]
28. Arriagada R, Le MG, Rochard F, Contesso G. Conservative treatment versus mastectomy in early breast cancer: patterns of failure with 15 years of follow-up data. Institut Gustave-Roussy Breast Cancer Group. J Clin Oncol. 1996; 14(5):1558–64. [PubMed: 8622072]
29. Fisher B, Anderson S, Bryant J, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med. 2002; 347(16):1233–41. [PubMed: 12393820]
30. Poggi MM, Danforth DN, Sciuto LC, et al. Eighteen-year results in the treatment of early breast carcinoma with mastectomy versus breast conservation therapy: the National Cancer Institute Randomized Trial. Cancer. 2003; 98(4):697–702. [PubMed: 12910512]
31. Nattinger AB, Gottlieb MS, Veum J, Yahnke D, Goodwin JS. Geographic variation in the use of breast-conserving treatment for breast cancer. N Engl J Med. 1992; 326(17):1102–7. [PubMed: 1552911]
32. Farrow DC, Hunt WC, Samet JM. Geographic variation in the treatment of localized breast cancer. N Engl J Med. 1992; 326(17):1097–101. [PubMed: 1552910]
33. Smith GL, Xu Y, Shih YC, et al. Breast-conserving surgery in older patients with invasive breast cancer: current patterns of treatment across the United States. J Am Coll Surg. 2009; 209(4):425–33. e2. [PubMed: 19801315]
34. Ridao-Lopez M, Garcia-Armesto S, Abadia-Taira B, Peiro-Moreno S, Bernal-Delgado E. Income level and regional policies, underlying factors associated with unwarranted variations in conservative breast cancer surgery in Spain. BMC Cancer. 2011; 11:145. [PubMed: 21504577]
35. Thompson B, Baade P, Coory M, Carriere P, Fritschi L. Patterns of surgical treatment for women diagnosed with early breast cancer in Queensland. Ann Surg Oncol. 2008; 15(2):443–51. [PubMed: 17909915]
36. van Steenbergen LN, van de Poll-Franse LV, Wouters MW, et al. Variation in management of early breast cancer in the Netherlands, 2003–2006. Eur J Surg Oncol. 2010; 36 (Suppl 1):S36–43. [PubMed: 20620013]
37. Green L. Geography is Destiny: Where you Live - Not the Facts of your Case - May Determine What Medical Treatment you Receive. Mirabella. 1996 Nov 1.:154–8.
38. Stacey D, Bennett CL, Barry MJ, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2011; (10):CD001431. [PubMed: 21975733]

39. Kennedy AD, Sculpher MJ, Coulter A, et al. Effects of decision aids for menorrhagia on treatment choices, health outcomes, and costs: a randomized controlled trial. JAMA. 2002; 288(21):2701–8. [PubMed: 12460093]
40. Auvinen A, Hakama M, Ala-Opas M, et al. A randomized trial of choice of treatment in prostate cancer: the effect of intervention on the treatment chosen. BJU Int. 2004; 93(1):52–6. discussion 6. [PubMed: 14678367]
41. Whelan T, Levine M, Willan A, et al. Effect of a decision aid on knowledge and treatment decision making for breast cancer surgery: a randomized trial. JAMA. 2004; 292(4):435–41. [PubMed: 15280341]
42. Arterburn D, Wellman R, Westbrook E, et al. Introducing decision aids at Group Health was linked to sharply lower hip and knee surgery rates and costs. Health Aff (Millwood). 2012; 31(9):2094–104. [PubMed: 22949460]
43. McPherson K, Wennberg JE, Hovind OB, Clifford P. Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway. N Engl J Med. 1982; 307(21):1310–4. [PubMed: 7133068]
44. McPherson, K. Systematic variation in surgical procedures and hospital admission rates : a methodological study : report II, April 1996, for the Department of Health. London: London School of Hygiene & Tropical Medicine, Dept. of Public Health & Policy, Health Promotion Sciences Unit; 1996.
45. McMahon AJ, Fischbacher CM, Frame SH, MacLeod MC. Impact of laparoscopic cholecystectomy: a population-based study. Lancet. 2000; 356(9242):1632–7. [PubMed: 11089821]
46. Rutledge R, Fakhry SM, Baker CC, Meyer AA. The impact of laparoscopic cholecystectomy on the management and outcome of biliary tract disease in North Carolina: a statewide, population-based, time-series analysis. J Am Coll Surg. 1996; 183(1):31–45. [PubMed: 8673305]
47. Goodman DC, Stukel TA, Chang CH, Wennberg JE. End-of-life care at academic medical centers: implications for future workforce requirements. Health Aff (Millwood). 2006; 25 (2):521–31. [PubMed: 16522606]
48. Morden NE, Chang CH, Jacobson JO, et al. End-of-life care for Medicare beneficiaries with cancer is highly intensive overall and varies widely. Health Aff (Millwood). 2012; 31(4):786–96. [PubMed: 22492896]
49. Chassin MR. Explaining geographic variations. The enthusiasm hypothesis. Med Care. 1993; 31(5 Suppl):YS37–44. [PubMed: 8492584]
50. Bederman SS, Coyte PC, Kreder HJ, Mahomed NN, McIsaac WJ, Wright JG. Who's in the driver's seat? The influence of patient and physician enthusiasm on regional variation in degenerative lumbar spinal surgery: a population-based study. Spine (Phila Pa 1976). 2011; 36 (6): 481–9. [PubMed: 21488247]
51. Hollingsworth JM, Ye Z, Strope SA, Krein SL, Hollenbeck AT, Hollenbeck BK. Physician-ownership of ambulatory surgery centers linked to higher volume of surgeries. Health Aff (Millwood). 2010; 29(4):683–9. [PubMed: 20368599]
52. Hollingsworth JM, Krein SL, Ye Z, Kim HM, Hollenbeck BK. Opening of ambulatory surgery centers and procedure use in elderly patients: data from Florida. Arch Surg. 2011; 146(2):187–93. [PubMed: 21339431]
53. Nallamothu BK, Rogers MA, Chernew ME, Krumholz HM, Eagle KA, Birkmeyer JD. Opening of specialty cardiac hospitals and use of coronary revascularization in medicare beneficiaries. JAMA. 2007; 297(9):962–8. [PubMed: 17341710]
54. Anderson GM, Grumbach K, Luft HS, Roos LL, Mustard C, Brook R. Use of coronary artery bypass surgery in the United States and Canada. Influence of age and income. JAMA. 1993; 269(13):1661–6. [PubMed: 8455299]
55. Wennberg, JE. Tracking medicine : a researcher's quest to understand health care. Oxford: Oxford University Press; 2010.
56. Dartmouth Medical School. Center for the Evaluative Clinical Sciences, Maine Medical Center (Portland Me.). Center for Outcomes Research and Evaluation. The Dartmouth atlas of musculoskeletal health care. 2000.

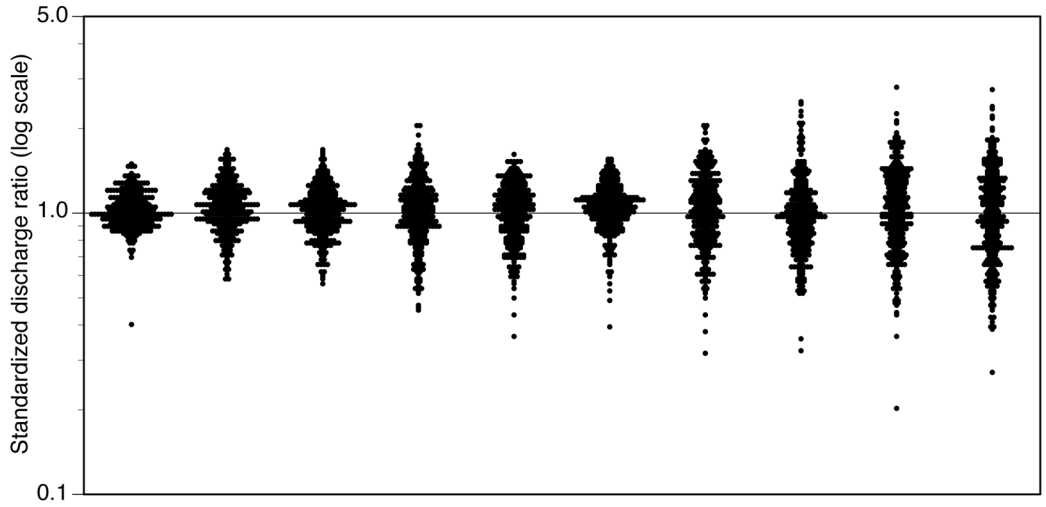
57. Lavis JN, Malter A, Anderson GM, et al. Trends in hospital use for mechanical neck and back problems in Ontario and the United States: discretionary care in different health care systems. Cmaj. 1998; 158(1):29–36. [PubMed: 9475907]
58. Patel MR, Greiner MA, DiMartino LD, et al. Geographic variation in carotid revascularization among Medicare beneficiaries, 2003–2006. Arch Intern Med. 2010; 170(14):1218–25. [PubMed: 20660840]
59. Vikatmaa P, Mitchell D, Jensen LP, et al. Variation in Clinical Practice in Carotid Surgery in Nine Countries 2005–2010. Lessons from VASCUNET and Recommendations for the Future of National Clinical Audit. European Journal of Vascular and Endovascular Surgery. 2012; 44 (1): 11–7. [PubMed: 22633072]
60. Wennberg JE, Bunker JP, Barnes B. The need for assessing the outcome of common medical practices. Annu Rev Public Health. 1980; 1:277–95. [PubMed: 6753863]
61. Menyhei G, Björck M, Beiles B, et al. Outcome Following Carotid Endarterectomy: Lessons Learned From a Large International Vascular Registry. European Journal of Vascular and Endovascular Surgery. 2011; 41(6):735–40. [PubMed: 21450496]
62. Woo K, Garg J, Hye RJ, Dilley RB. Contemporary Results of Carotid Endarterectomy for Asymptomatic Carotid Stenosis. Stroke. 2010; 41(5):975–9. [PubMed: 20339122]
63. Eslami MH, McPhee JT, Simons JP, Schanzer A, Messina LM. National trends in utilization and postprocedure outcomes for carotid artery revascularization 2005 to 2007. Journal of Vascular Surgery. 2011; 53(2):307–15. [PubMed: 21093200]
64. Zhang Y, Baicker K, Newhouse JP. Geographic variation in Medicare drug spending. N Engl J Med. 2010; 363(5):405–9. [PubMed: 20538621]
65. Zhang Y, Baicker K, Newhouse JP. Geographic variation in the quality of prescribing. N Engl J Med. 2010; 363(21):1985–8. [PubMed: 21047217]
66. Naughton C, Bennett K, Feely J. Regional variation in prescribing for chronic conditions among an elderly population using a pharmacy claims database. Ir J Med Sci. 2006; 175(3):32–9. [PubMed: 17073245]
67. Parker L, Levin DC, Frangos A, Rao VM. Geographic Variation in the Utilization of Noninvasive Diagnostic Imaging: National Medicare Data, 1998–2007. American Journal of Roentgenology. 2010; 194(4):1034–9. [PubMed: 20308507]
68. A Dartmouth Atlas Project Topic Brief: Supply Sensitive Care. Lebanon, NH: The Dartmouth Atlas of Health Care; 2007.
69. Busato A, Matter P, Künzi B, Goodman DC. Supply sensitive services in Swiss ambulatory care: An analysis of basic health insurance records for 2003–2007. BMC Health Services Research. 2010; 10(1):315. [PubMed: 21092250]
70. Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. BMJ. 2010; 341:c5146. [PubMed: 20947577]
71. Dartmouth-Hitchcock Medical Center. [accessed March 11th, 2013] Center for Shared Decision Making. 2013. http://patients.dartmouth-hitchcock.org/shared_decision_making.html/
72. Birkmeyer JD. Progress and challenges in improving surgical outcomes. Br J Surg. 2012; 99(11): 1467–9. [PubMed: 23001839]
73. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. N Engl J Med. 2011; 364(22):2128–37. [PubMed: 21631325]

Key messages

- Studies from many different countries indicate that the use of surgical procedures varies across geographical regions.
- Chance, patient demand, and differences in diagnostic practices play a relatively small role in explaining regional variation in surgical rates.
- Different attitudes and beliefs about the indications for surgery are the most important reasons for surgical variation.
- Discretionary, “preference-sensitive” procedures, such as radical prostatectomy, tend to vary considerably more than procedures for which clinical decisions are constrained to a narrow range of options, like colectomy for colon cancer.
- Tools to better incorporate the preferences of individual patients, such as decision aids, may help reduce variation for preference-sensitive procedures.
- Broader influences of regional variation include technology diffusion, training and the regional supply of surgeons, and financial incentives and the regulatory environment.

Search strategy and selection criteria

We searched Ovid MEDLINE and Embase for articles published between 1946 and May 1, 2013, and constructed our searches using a combination of MeSH or EMTREE controlled terms (such as, “small-area analysis”, “physician’s practice patterns”. and “clinical practice”) and title and abstract keywords. For the keyword portion of the search, we used the following concepts, plus various derivatives: “regional”, “geographic”, “population”, “treatment”, “small area”, or “care”, adjacent to “variation” or “pattern”, in combination with “rate” adjacent to “surgery”, “surgical”, or “procedure”. We also searched the bibliographies of articles identified by this search strategy, as well as references contained in *Tracking Medicine*, by John E. Wennberg. Given the necessarily broad search terms and the narrative nature of this review, articles highlighted in text were selected based on the authors’ judgments about their respective importance in describing, explaining, or reducing regional variation in the use of surgery.



	Hip fracture repair	Colectomy	Cholecystectomy	CABG	Hip replacement	Knee replacement	Back surgery	PCI	Carotid endarterectomy	Radical prostatectomy
US average (per 1000 Medicare enrollees)	6.9	1.0	3.4	3.4	3.8	8.8	4.7	8.2	2.1	1.4
Extremal ratio	3.68	2.88	3.01	4.55	4.52	3.96	6.40	7.81	14.12	10.05
Interquartile ratio	1.23	1.29	1.29	1.36	1.38	1.23	1.49	1.48	1.65	1.71
SCV (x10)	2.10	3.52	3.87	7.16	5.76	4.01	8.98	9.77	12.13	13.53

Figure 1.

Variation in rates of major surgery for ten common procedures among 306 hospital referral regions in the US, based on 2008–10 national Medicare data. The extremal ratio reflects the highest rates divided by the lowest. The interquartile ratio is the 75th percentile rate divided by the 25th percentile. The systematic component of variation (SCV) represents the true, non-random part of observed variations.[6,7]

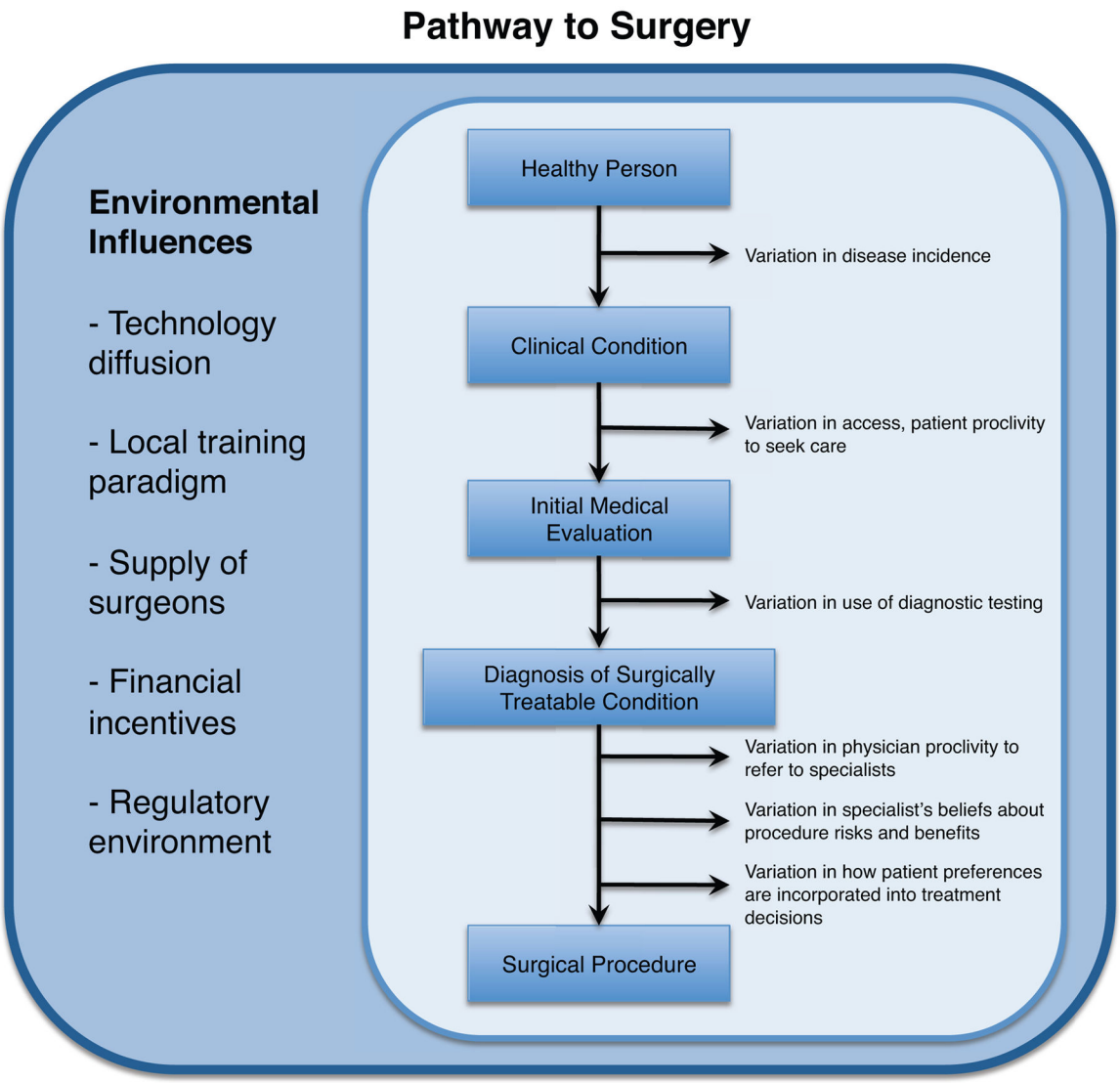


Figure 2.
Conceptual model of regional variation in the use of surgery.

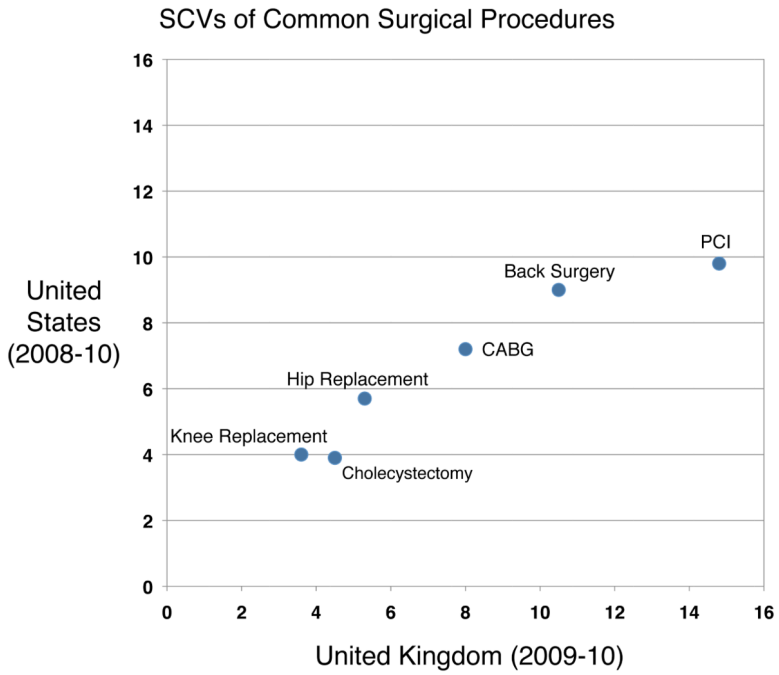


Figure 3.

The association between regional variation in the US[12] and the UK[10] for the specified period (SCV: systematic component of variation; CABG: coronary artery bypass grafting; PCI: percutaneous coronary intervention).