Ana Johnson
and
Thérèse Stukel
Medical Practice Variations
Health Services Research
10.1007/978-1-4899-7573-7_89-3

Variations in Healthcare Spending and Quality Among Institutions

Tetsuya Otsubo[1], Yuichi Imanaka[1], Toshitaka Morishima[1], Noriko Sasaki[1], Sungchul Park[1] and Jason Lee[1]
(1)Department of Healthcare Economics and Quality Management, Kyoto University Graduate School of Medicine, Yoshida Konoe-cho, 606-8501 Kyoto, Sakyo-ku, Japan

Tetsuya Otsubo (Corresponding author)
Email: otsubo.tetsuya.6m@kyoto-u.ac.jp
Email: otsubot-kyoto@umin.ac.jp

Yuichi Imanaka
Email: imanaka-y@umin.net

Toshitaka Morishima
Email: morishima.t@hotmail.co.jp

Noriko Sasaki
Email: sasaki.noriko.3z@kyoto-u.ac.jp

Sungchul Park
Email: park.sungchul.53n@st.kyoto-u.ac.jp

Abstract

Unwarranted institutional variations in healthcare spending and quality may indicate discrepancies in the quantity of services provided, management efficiency, and staff capability at the hospital level. Because not all variations are unwarranted, analyses must take into account differences in the needs and preferences of the patient groups served. Variations can be influenced by factors such as payment systems, hospital ownership, management methods, resource availability, teaching status, and practice patterns. Although institutional variations may be intertwined with variations at the regional level, some measures of care, such as nosocomial infection rates or indicators of hospital management efficiency, are more meaningful when quantified at the hospital level. The accurate identification of unwarranted variations that stem from causes at the institutional level would also help to identify the stakeholders and decision makers who have the relevant authority and jurisdiction to address the problems. In this chapter, the empirical evidence of institutional variations in healthcare spending, medical practice patterns, and outcomes are introduced. Common methodologies used for measuring variations and the possible determinants of these variations are addressed, and the methods to reduce unwarranted variations are examined.

Introduction

Unwarranted variations in institutional healthcare spending and quality may indicate hospital-level disparities in the quantity of services provided, distribution of resources, management efficiency, and the skill of clinical staff. These variations may also point to the inefficient use of existing resources or provision of suboptimal healthcare. The growing recognition of the existence and impact of unwarranted variations among hospitals can be seen in the advent of quality-incentivizing systems, such as pay for performance (P4P) and public reporting programs.

The rationale for reducing unwarranted variations among institutions is that there is a theoretical optimal delivery of healthcare that provides the highest possible quality of care at the lowest possible cost. However, deviations from this optimum may simply reflect variations in patient populations, and do not necessarily indicate poorer quality of healthcare. Research into healthcare variations at the institutional level must therefore take into account differences in the needs of the various patient groups served.

The reliability of observed variations is limited by the quality of the data and methodological approaches. Large databases comprising standardized data from numerous healthcare providers coupled with appropriate adjustment methodologies can increase the validity of any variations observed. Variations that extend beyond the hospital level are covered elsewhere in this handbook.

Unwarranted variations refer to those that are not the result of differences in patient case mix or environmental factors. They can be influenced by payment systems, hospital ownership, insurance systems, management methods, resource availability, teaching status, and practice patterns. "Optimal" healthcare can be identified using the various guidelines and standards promulgated by expert bodies within each field while taking into account the unique characteristics of each country or region.

Uneven healthcare spending and quality at the institutional level are likely to be of most interest to hospital management staff, policymakers, payers, and health services researchers, who aim to improve healthcare quality and adjust payment systems and policies. In addition to these stakeholders, the general public has become increasingly aware of hospital-level variations. The information asymmetry been patients and doctors has traditionally led to a general impression that the care provided by any medical professional is correct and necessary. However, growing patient awareness of hospital variations in the quality of healthcare processes and outcomes (including hospital scorecards and the Centers for Medicaid and Medicare Services' [CMS] Hospital Compare program) has changed this situation. This increases the burden on health services researchers to ensure that biases are controlled and appropriate risk adjustments are made, as the reporting of flawed evaluations would be inherently unfair to the hospitals involved. A previous study has, however, shown that the public reporting of mortality rates in the USA did not have a detrimental economic effect on hospitals with "poor outcomes" (Vladeck et al. *1988*). More analyses are needed to assess if this is still the case.

Institutional variations are likely to be intertwined with variations at the regional level – examined in a preceding chapter – but should be treated differently. Some measures of care, such as nosocomial infection rates or indicators of hospital management efficiency, are more meaningful when quantified at the hospital level. In fact, the analyses of such measures at the regional levels would likely gloss over meaningful variations among hospitals. Also, regional variations are likely to be more easily affected by population health factors, whereas hospital-level variations may be more directly indicative of institutional factors in the hospitals where care is provided. Institutional- and regional-level variations may arise from different underlying factors, have different implications, and may respond to different types of interventions to reduce unwarranted variations. Accurately identifying

the unwarranted variations that stem from causes at the institutional level would also help to identify the stakeholders and decision makers who have the authority to address specific problems.
In this chapter, the following questions are considered:

1.
What are the causes of institutional variations in spending, utilization, and quality?

2.
What are the consequences of higher spending?

3.
What can be done to reduce unwarranted institutional variations?

To answer these questions, first, empirical evidence of institutional variations in healthcare spending and quality is introduced; variations in spending are considered in the context of various payment systems, and quality is discussed according to practice patterns and outcomes. Common methodologies used for measuring variations and the possible determinants of variations are addressed. This is followed by an examination of the consequences of higher spending and possible methods to reduce unwarranted variations.

Causes of Institutional Variations in Spending, Utilization, and Quality

With the rapid rise of healthcare costs worldwide, there is a greater need for more research into the causes of unwarranted variations in medical spending in order to control costs and ensure that providers deliver care more efficiently. Health services researchers can also assess provider efficiency by considering spending variations in relation to outcomes.
Hospital payment systems should motivate providers to treat all patients in need of care and to deliver an adequate number of necessary services while taking into account the appropriateness of the services and patient outcomes. There are various inpatient payment systems, such as fee-for-service (per procedure/service), bundled payments (per admission or per diem), and capitation. Many OECD countries use a mix of payment arrangements to finance hospital acute care, and a few examples are provided in the Table 1 below.**Table 1**
Hospital inpatient payment systems in major industrialized countries

	Per procedure/service	**Per admission**	**Per diem**	**Capitation**
Australia		X		X
Canada		X	X	X
France		X		

Germany		X		
Italy		X		
Japan	X		X	
United Kingdom		X		X
United States	X	X		

Adapted from Paris et al. *2010*

Fee-for-service payment systems are process-based and reimburse providers according to the number and types of activities that they perform. At a system-wide level, however, this form of payment can create incentives for overprovision as providers seek to maximize revenues that depend on the volume and intensity of services delivered.

In bundled payment systems, payers do not reimburse providers for each service, but instead pay a fixed amount to physicians and hospitals for a specific set of services. In these systems, payment is linked to the type and severity of each individual case. Patients are classified in a specific diagnostic group according to their principal diagnosis and a corresponding fixed reimbursement is given to the hospital for treating the patient. In a per admission payment system (a type of bundled payment system), a fixed amount is paid to cover a predetermined set of services during a particular hospital stay regardless of the actual services provided; the widely used diagnosis-related groups (DRG) payment method and its derivatives are examples of this approach. In contrast, a per-diem payment system utilizes a fixed rate paid per day of hospitalization irrespective of the actual services delivered or the costs incurred.

Capitation payment systems involve the reimbursement of specified amounts to providers for each enrolled case in a given group, regardless of the actual number or nature of services delivered over a set period. In theory, capitation payment systems should reduce variations in hospital spending, but they may also reduce access and quality of care.

This section looks at the variations in hospital inpatient payment that result from different payment systems, and summarizes possible causes of variations in spending.

Variations in Payment Systems

Fee-for-Service Systems

Fee-for -service is the predominant mode of payment for primary care providers in numerous countries, and is sometimes used in combination with a bundled payment system for inpatient services.

The Medicare Private Fee-for-Service Plan in the USA involves regular payments of a predetermined amount to private insurance companies for the provision of healthcare to Medicare beneficiaries on a

fee-for-service arrangement. However, considerable variations in practice patterns and medical spending in the US healthcare system have been shown to bear little or no relation to quality (Fisher et al. *2003a*, *b*). These variations are not explained by differences in the health statuses of patients or by regional differences in input costs, but instead suggest that a substantial amount of unnecessary or ineffective care is being offered. Furthermore, these variations suggest that some providers may enjoy unusual market power and pricing leverage in local markets. Provider variations in hospital spending have been shown to be explained in part by differences in services provided during end-of-life care (Wennberg et al. *2004*) and hospital ownership (Reschovsky et al. *2011*). Variations in the quantity of healthcare services provided have been observed even among academic medical centers (Fisher et al. *2004*). These studies may indicate that individual physicians have dramatically different interpretations of the appropriateness of various medical procedures, which can give rise to widely divergent costs and outcomes under a system that incentivizes the provision of more healthcare services.

Diagnosis-Related Group (DRG)/Per-Admission Payment Systems

An increasing number of countries have moved to prospective per-admission payment schemes. These schemes involve the use of DRGs to set payments based on the estimated cost of hospital care for a particular disease before services are provided. This encourages hospitals to increase the volume of cases, and has the added advantage of dissuading providers from offering extraneous services.

The DRGs are used to categorize patients intogroups that are as medically homogeneous as possible in order to ensure that payment rates accurately reflect the resources and costs of treating patients within each group. The “fineness” of the patient classification system can also affect provider incentives: a finer grouping system tends to create fewer incentives to select low-cost patients, thereby reducing cost variations within a payment category. However, a greater number of classification groups results in greater monitoring problems and more opportunities to game the system.

While fee-for-service systems provide little motivation to minimize healthcare provision and costs, DRG-based payment systems incentivize providers to pursue more efficient behavior by encouraging them to consider the prices and amount of resources used. The DRGs are used to categorize patients into groups that are as medically homogeneous as possible in order to ensure that payment rates accurately reflect the resources and costs of treating patients within each group (Busse *2012*). The “fineness” of the patient classification system can also affect provider incentives: a finer grouping system tends to create fewer incentives to select low-cost patients, thereby reducing cost variations within a payment category. However, a greater number of classification groups can result in increased monitoring problems and more opportunities to game the system. The influence of the coarseness/fineness and breadth/narrowness of the grouping system on incentives is particularly apparent when payments shift from physician fee-for-service schedules to DRG payments. A shift from fee-for-service payment to DRG-based payment may therefore weaken the incentive to over-provide care and strengthen the incentive to control costs at the provider level.

However, almost all countries that have introduced DRG-based payments have done so by implementing a “mixed” payment system where hospital funding is not determined solely byDRG payments. Hospitals may also receive funding in other forms, including funding for teaching and research, financial support to compensate for different costs related to geographical location, and payments to cover some elements of the fixed costs of providing services. The composition of these other payment forms varies among countries and over time.

Diagnosis Procedure Combination/Per-Diem Payment Systems

This type of system gives hospitals incentives to reduce costs and unnecessary procedures, but can also encourage longer average length of stay (LOS). Japan, which has implemented a Diagnosis Procedure Combination/per-diem payment system (DPC/PDPS), has the longest average hospital LOS among the OECD countries (OECD *2011*). The Japanese payment system does not have a cap on the total amount to be reimbursed, nor does it cover outpatients regardless of whether the patients used to be hospitalized or are to be hospitalized at the same institution. Thus, hospitals have a financial incentive to keep patients in beds for as long as possible during each hospitalization.

Japan has few arrangements for evaluating the performance of hospitals because there is currently no systematic collection of treatment or outcome data, and therefore no means of implementing mechanisms promoting best-practice care (such as P4P programs). Research studies examining variations at the institutional level under this system have employed administrative claims data; topics include the varying rates of transferring patients with cerebral infarction to alternative facilities (Sekimoto et al. *2008*), different resource utilization and charges (Kuwabara et al. *2006*), and end-of-life spending (Morishima et al. *2014*).

Causes of Variations in Hospital Spending

Many countries have introduced some form of bundled payment for hospital care, whereby hospitals are paid according to the number and type of patients they treat. Although variations in patient case mix are likely to be a major determinant of spending variations, the classification systems used to determine case mix can never fully account for all cost variations; several additional hospital characteristics that determine hospital spending levels have been identified from empirical studies.

Patient Case Mix

Observed differences in spending are largely due to variations in patient characteristics, as different types of patients invariably lead to different levels of spending. Hospitals may treat different levels of disease severities, with some hospitals treating higher proportions of sicker or older patients. Before statistical methods are used to investigate hospital-level cost differences, the following patient-level variables are frequently adjusted for, regardless of main diagnosis: sex, age, socioeconomic status, diagnosis, disease severity, and accompanying comorbidities at admission. In addition, other disease-specific patient characteristics and procedures are also important and are likely to vary heavily according to each specific condition.

Hospital Characteristics

Street et al. (*2010*) have reviewed the literature to examine the international evidence for existing instruments, systems, or concepts for measuring and explaining differences in hospital spending. Characteristics that may contribute to institutional variations among providers are summarized below. The underlying reasons for their influence are also discussed, with further details available in Street et al. (*2012*).

Teaching Status

Teaching hospitals often incur higher spending than nonteaching hospitals. There have been numerous studies addressing the reasons underlying this cost difference, which can be summarized by the following two explanations: Firstly, teaching hospitals may generally treat patients with more serious illnesses, and therefore incur higher costs. Secondly, teaching hospitals require more funding to support their teaching activities (Newhouse *2003*).

Ownership and for-Profit Status

The ownership of a hospital can have direct implications on the level of healthcare spending, as private for-profit hospitals are expected to exhibit higher spending compared with publically-owned or nonprofit hospitals. Furthermore, it has been shown that hospital ownership within an area can affect spending at the hospital level, with a higher presence of "for-profit" type hospitals associated with higher costs (Reschovsky et al. *2011*).

There are two opposing viewpoints with regard to cost differences between for-profit and nonprofit hospitals: The first argues that for-profit hospitals tend to operate at lower costs and are therefore more efficient than nonprofit ones, while the second argues that there is no detectable difference in the decision behavior and efficiency of for-profit and nonprofit hospitals (Hollingsworth *2008*). A review of the literature has shown that private ownership is not necessarily associated with higher efficiency when compared with public ownership (Tiemann et al. *2012*). Thus, insofar as providing cost-effective healthcare is concerned, no consistent differences have been shown between the performances of for-profit and nonprofit hospitals.

Volume of Activity

Larger hospitals are thought to operate at lower costs than smaller hospitals (Posnett *2002*), the underlying reasoning being that larger hospitals benefit from economies of scale and that staff experience results in lower costs as case volume increases. However, diseconomies of scale may arise beyond a particular size, perhaps because a larger size can lead to overly high expenditures for overhead, increasingly bureaucratic forms of organization and complex interdependencies giving rise to problems of coordination and cooperation. Econometric analyses of hospital costs generally include some measure of hospital size as a variable for adjustments.

Provider Prices

In the USA, providers receive different compensation for beneficiaries of commercial insurance than for patients with Medicaid coverage. Furthermore, transacted prices for healthcare vary across hospitals depending on market structure and provider-payer – concentration in non-Medicare spending (Gaynor and Town *2011*). Hospital and provider price differences have been reported to account for one-third of overall healthcare spending variation among insured US residents (White *2012*). Most of these variations in provider prices were unexplained and could not be attributed to the cost of doing business.

Quality

Another possible explanation of differential spending lies in the differences in the quality of healthcare provided at the hospital level. The relationship between quality and costs, however,

remains unclear; and is addressed later in this chapter. Higher quality of healthcare may require higher resource consumption, but higher quality healthcare and efficient delivery of care are not mutually exclusive. The association between healthcare spending and quality is stronger when they are directly linked through P4P programs or reimbursement systems with punitive withholding of payment for reduced performance. However, evidence of the effectiveness of P4P in improving quality and reducing costs is mixed and inconclusive (Van Herck et al. *2010*).

Geographical Location

Whether a hospital is based in a rural or urban location can impact the level of spending when treating similar patients. As there are many confounding factors at play – including environmental, sociodemographic, or socioeconomic variables – there is no simple relationship between location and spending. On one hand, hospitals in rural locations would have higher transport costs and are also less likely to treat sufficient numbers of patients to achieve effective economies of scale. On the other hand, urban-based hospitals would likely have higher running costs and treat patients with more complex conditions.

Others

Other characteristics that may influence institutional variations in spending include specialization (Street et al. *2012*), interhospital competition (Dranove *2011*), and technological equipment (Street et al. *2010*). However, there is considerable residual variation in spending that cannot be explained by these factors, although it is the objective of health services research to reduce this unexplained portion to a minimum.

Institutional Variations in Medical Practice Patterns

With increasing demands to improve quality in healthcare and curtail rising healthcare costs, the use of quality indicators (QIs) to monitor medical practice patterns is becoming standard practice. Since the 1990s, report cards have been published on patient care and outcomes associated with hospitals, physicians, and managed care plans as measurements of quality in healthcare. Despite the possible ramifications that these report cards may have on the business operations and profits of hospitals that are deemed to perform poorly, it has been shown that there were no differences between highly-rated hospitals and nonrated hospitals with regard to risk-adjusted mortality rates or readmission rates in patients with acute myocardial infarction (AMI) (Chen et al. *1999*). However, highly-rated hospitals showed shorter LOS durations and hospitalization costs, leading to the conclusion that the report cards may have been measuring operating efficiency rather than superiority in clinical outcomes.
Ratings of hospital performances in medical practice continue to be published in many countries, although their presentation and interpretation should be approached with caution. The various stakeholders in healthcare may have different expectations or requirements from such reports, and the targets for improvement and the methodologiesmay be largely influenced by the type of healthcare system in each country. Healthcare systems can be categorized into three major types: (1) the Beveridge “public” model, in which funding is based mainly on taxation (adopted by countries such as the UK, Sweden, Norway and Canada); (2) the Bismarck “mixed” model, funded mainly by a premium-financed social/mandatory insurance system (Germany, France and Japan); and (3) a

“private” insurance model in which funding is based solely on premiums paid to private insurance companies (this system is essentially limited to the USA) (Lameire et al. *1999*).
The private insurance model, in theory, encourages hospitals to compete to provide high quality of care and conduct operations efficiently. In contrast, hospitals in countries that employ either the Beveridge or Bismarck models tend to aim for a standardization of medical care characterized by an equal opportunity for patients to receive the best medical practice available based on objective criteria or guidelines. Access and timeliness may take precedence over the efficiency and financial health of individual healthcare institutions, and the pursuit of “best medical practice” is usually a major concern in the evaluation of hospital performance.
Variations in medical practice and healthcare quality have been recorded between countries, regions, and hospitals. The Dartmouth Atlas of Health Care in the US and the National Health Service (NHS) in the UK have presented regional variations in their respective atlases. However, when multiple organizations independently finance, manage, and deliver the health services in a particular region (mainly in countries under the Bismarckand private insurance models),the analysis of variations in individual hospital performance may be more useful than the analysis of regional variations in elucidating underlying problems. Although the existence of variations is not necessarily indicative of poor quality per se, unwarranted variations may reflect the limits of professional knowledge and failures in its application, whereas good variations can result from a patient-centered approach to healthcare (Mulley *2010*).
As patients become more aware of their healthcare options and the variations in healthcare quality, they are more likely to make educated choices in where they obtain care. Insurers and other payers have also shown high levels of interest in the quality of healthcare services provided at each hospital. Services based on clinical evidence can indicate the general quality of a hospital, as well as the efficiency of the management system. Payers have also employed financial incentives and withholding of reimbursements based on how well a hospital performs instipulated QIs.
The analysis of variations in healthcare processes at the hospital level may therefore be of interest to policymakers, providers, patients, and payers. The following section looks at examples of institutional variations in medical practice patterns (i.e., processes of care) and their causes.

Evaluating Hospital Performance

Various indicators have been developed to measure quality in healthcare since the mid-1990s. There are many QIs representing different aspects of care currently used for major medical conditions. Examples of QIs have been published from various organizations and agencies in different countries, such as the Agency for Healthcare Research and Quality (AHRQ), the Centers for Medicare & Medicaid Services (CMS) and the Joint Commission in the USA; the NHS in UK, the Canadian Institute for Health Information (CIHI) in Canada, and *Bundesgeschäftsstelle Qualitätssicherung* (BQS) in Germany.
QIs that reflect processes of care are designed to measure whether recommended services are provided to eligible patients. These include indicators that measure the use of medical diagnostics, the use of medication, and surgical procedures. Timeliness of care can also be included in process indicators (e.g., whether AMI patients are given aspirin upon arrival to hospital or whether pneumonia patients are administered antibiotics within 6 h of arrival) (www.hospitalcompare.hhs.gov).

Empirical Evidence

QIs that target direct processes of care are relatively easy to interpret, given that the majority of these indicators reflect guidelines based on clinical evidence or expert consensus. These indicators not only allow interhospital comparisons of quality, but the temporal monitoring of indicators can also track the effects of policy or systemic changes. Indicators in this category generally do not require risk adjustment, because the denominators in these indicators exclude patients for whom the use of the process is precluded. For example, angiotensin converting enzyme (ACE) inhibitors and angiotensin receptor blockers (ARB) are frequently indicated in the treatment of patients with left ventricular systolic dysfunction, but not all patients are suitable for prescription due to renal dysfunction or other adverse effects. Failure to exclude such patients may result in the observation of "excessive" variations in the provision of these drugs.

Several examples of hospital-level variations observed in process indicators are presented in the following section, and variations in specific aspects of healthcare (e.g., elective surgery, hospitalizations, and emergency services) are addressed elsewhere in this volume.

Diagnostic tests

The use of appropriate diagnostic tests can shape subsequent treatment approaches and are indicated for conditions such as head trauma or suspected cerebrovascular accidents. Considerable variations in the rates of computed tomography (CT) scans were shown in a large-scale comparison of 34 academic trauma centers in the USA (Bulger et al. *2002*). In this study, centers that took a more "aggressive" approach to treatment tended to provide more scans. This increased use of scans may have resulted in subsequent therapy variations, and may also be associated with the lower mortality rates observed in centers providing aggressive treatments.

A survey of over 1,000 doctors in mostly urban hospitals in China regarding acute ischemic stroke care found that only 69 % of hospitals had CT scanners, and only 88 % of the doctors in those hospitals would routinely scan stroke patients (Chen et al. *1997*). This may also point to an issue of resource availability at the hospital level and its association with quality of care.

Medications

The correct use of antibiotics to prevent and treat infections has been strongly encouraged for several decades. The overuse of antibiotics has very real implications in promoting antibiotic resistance, and most guidelines aim to curtail the rampant use of these drugs. A survey of 100 Japanese teaching hospitals showed wide variations in prophylactic antibiotic use, with prophylaxis durations for gastrectomy ranging from 2.3 days to 7 days at the hospital level (Sekimoto et al. *2004*). In the Netherlands, an analysis of nine antibiotic utilization indicators in community-acquired pneumonia patients from eight hospitals also demonstrated wide hospital-level variations. These indicators looked at the timely initiation of antibiotic therapy (within 4 h after presentation), the provision of antibiotic regimens according to national guidelines, and the adaptation of dose and dose intervals of antibiotics according to renal function; performance in these three indicators ranged from 36 % to 87 %, 5 % to 59 %, and 40 % to 100 %, respectively (Schouten et al. *2005*). The Pneumonia PORT cohort study conducted on healthcare institutions located in the USA and Canada also documented large variations in antimicrobial prescribing practices for outpatients and inpatients for the treatment of community-acquired pneumonia. These practice variations were also observed to result in antimicrobial cost variations, despite no differences in medical outcomes (Gilbert et al. *1998*).

Cancer treatments can vary among hospitals, although observed variations in chemotherapy may subsume patient preferences in treatment approach. Despite recommendations of adjuvant

chemotherapy in stage III colon cancer patients within Dutch treatment guidelines, large variations in adjuvant chemotherapy use have been demonstrated at the hospital level in the Netherlands, ranging from 82 % to 96 % in patients younger than 65 years of age, 59–78 % in patients aged between 65 and 74, and 9–25 % in patients aged over 75 years (Van Steenbergen et al. *2009*).

Variations in drug prescriptions can extend even to the use of generics and brand-name drugs. A cross-sectional analysis of Veteran Affairs Medical Centers in the USA found as much as a two-fold difference in dispensing rates of brand-name oral drugs, despite a closely-managed formulary (Gellad et al. *2010*). These findings suggest a correlation between practice variations and healthcare spending.

Surgery

In an analysis of 306 Hospital Referral Regions (designated by the Dartmouth Atlas Project) in the USA, Wennberg investigated the underlying causes of practice variations in surgery for three common cancers: colon cancer, breast cancer and prostate cancer (Wennberg *2010*). Colectomy can considerably improve prognoses and quality of life in colon cancer patients, and there are few other surgical or nonsurgical treatment options available for these patients; the study revealed that there were relatively small regional practice variations for these surgeries. In contrast, mastectomy in breast cancer patients and prostatectomy in prostate cancer patients had a higher degree of variations, which were associated with the increase in availability of other treatment options depending on each patient's cancer stage or personal preference (Wennberg *2010*). These findings corroborate the results from a previous study of breast conserving surgery (BCS) in the initial management of breast cancer in 182 Ontario-based hospitals, which showed substantial variations (6 %–84 %) at the hospital level (Iscoe et al. *1994*).

In recent years, there has been an increase in the online availability of information regarding surgery-related clinical performance at the hospital level, such as the frequency of specific surgeries performed and outcomes including mortality, LOS, readmission rate, and total costs. Examples of these online resources are the Annual Hospital Performance Report from the Pennsylvania Health Care Cost Containment Council ("PHC4"; http://www.phc4.org/), hospital-specific data from the Dartmouth Atlas Project (http://www.dartmouthatlas.org/data/hospital/), and results from the Canadian Hospital Reporting Project (CHRP; http://www.cihi.ca/CIHI-extportal/internet/en/document/health+system+performance/indicators/performance/chrp_ireport_about).How-ever, there is a need to be cautious when interpreting the results of each hospital because of variations in the disease complexity of patients treated and available resources. These results may be used for interhospital comparisons, tracking and reviewing quality improvement in each hospital, and to support the implementation of effective interventions from a policy-making perspective.

Blood Transfusions

Patients with severe anemia due to hemorrhage, cancer, or other critical conditions generally receive blood transfusions. However, large variations in rates of blood transfusions have been observed in both adult (Hasley et al. *1994*) and pediatric patients (Ringer et al. *1998*; Bednarek et al. *1998*) despite the widespread availability of guidelines. Transfusions are dependent on patient casemix, physiological factors (such as hematocrit levels (Rao et al. *2004*)), and medical opinion. Furthermore, variations in providers' definitions of bleeding can also impact the decisions to use transfusions (Steinhubl et al. *2007*). For these reasons, the simple analysis of utilization rates may be insufficient.

The use of risk adjustment methods to assess variations in blood product use has been previously reported (Sekimoto et al. *2010*).

Causes of Variations in Medical Practice Patterns

There are three major categories of factors that can affect institutional variations in healthcare quality, the first of which involves patient characteristics. These can include basic demographics such as age and sex, as well as clinical factors such as disease severity, comorbidities, and clinical indication for a particular healthcare service. Other factors such as patient ethnicity, socioeconomic status, smoking history, and medical history may also be important contributing factors. Patients with contraindications to a healthcare service should not be included as potential candidates to receive the service, as this can bias the indicator to reflect an excessively poor performance. Process indicators that are based on general clinical practice, such as the appropriate use of antibiotics in infections or diagnostics in determining treatment approaches, generally do not require risk adjustments. However, other process indicators that require certain clinical criteria to be fulfilled may need more complex adjustment methodologies. Risk adjustment can account for the differences in intrinsic patient health risks and suitability at admission, thereby making hospital-level comparisons more meaningful (Iezzoni *2003*). Additionally, there are several types of care that are heavily reliant on patient preference, such as the use of life-sustaining drugs or palliative treatments in end-of-life care. These types of care present more complex challenges to the accurate definition of what constitutes good quality.

Secondly, healthcare provider characteristics can affect institutional variations. Hospital resource availability (e.g., numbers of acute care beds, special care units, diagnostic or therapeutic equipment, and staff), case volume, hospital ownership, teaching status, and residency/fellowship program involvement can affect the types of patients a hospital treats and treatment approaches. The presence of integrated multidisciplinary teams for specific diseases (such as cancer or stroke) may also affect the quality of processes for these conditions. An impact of patient volume on lymph node detection after colorectal carcinoma resection patients has been demonstrated, with low-volume hospitals showing a proclivity to recovering fewer lymph nodes when compared with medium- and high-volume hospitals (Miller et al. *2004*). As in the example of blood transfusions (Steinhubl et al. *2007*), variations in medical opinion among the clinicians at the point of care, rather than patient preference, may in fact dominate treatment choice (Wennberg *2002*).

Thirdly, external circumstances may be important factors underlying the observed variations in healthcare quality. These factors include regional differences, resource allocation differences, proximity and competition with other hospitals, and overall healthcare delivery and payment systems. Henderson (*2002*) has described possible confounding factors in the quality of healthcare with the acronym SALT (S: Sociocultural considerations, A: The aging of population, L: The legal system and medical malpractice, T: Technology in medicine).

Institutional Variations in Outcomes

The major paradigm for evaluating healthcare quality is the structure-process-outcome model developed by Avedis Donabedian (*1980*, *1988*). In this model, “structure” quality measures refer to those that evaluate the attributes of the settings in which care is provided, while “process” quality measures evaluate the actions conducted when giving and receiving care. The “outcomes” facet of the Donabedian triad of quality analysis refers to the effects of healthcare on the health statuses of

patients and populations, and may be considered the most direct reflection of healthcare effectiveness and quality. Outcome measures can indicate the cumulative results of various aspects of healthcare, including those that are difficult to quantify, such as the expertise and skill of healthcare staff (Mant *2001*). Although arguably not definitive measures of quality, outcomes are generally easier to interpret than structure and process measures.

Outcomes also tend to garner the most interest from the general public and policymakers. However, there are obstacles to producing fair and comparable outcomes, such as those stemming from variations in confounding factors such as patient case mix. Additionally, outcome measures may be indicative of both the quality of healthcare provided at a hospital and the general health of a population, and must therefore be approached cautiously by analysts. Despite this caveat, hospital variations in outcomes should be analyzed regardless of whether they result from quality of care variations, as such variations may indicate other factors that warrant further investigation.

Empirical Evidence

Outcome measures frequently include mortality rates (case fatality rates), readmission rates, LOS durations and patient satisfaction at discharge. The interpretation of these measures is dependent on national and regional context. For example, mortality rates may not be a direct indicator of healthcare quality in Japan, where the hospital is the most common place of death.

Mortality Rates

Mortality rates garner more attention from the general public than other types of indicators. While variations in the quality of healthcare processes are more indicative of a hospital's adherence to recommended processes of care, the interpretation of mortality rates as QIs can be problematic. Efforts must be made to ensure completeness of data, analysis methods that include appropriate risk adjustments, and reporting of pertinent caveats to aid in their interpretation.

Mortality is generally categorized as in-hospital mortality, which is limited to deaths that occur within a healthcare institution; or general mortality, in which post-discharge mortalities are included. Access to the latter type of information may be more useful to researchers as this can shed light on other factors such as a patient's choice in their place of death or attempts to seek healthcare elsewhere. Measures can then be further categorized into crude mortality, standardized mortality, and risk-adjusted mortality. The time length of the mortality indicator can also be adjusted according to study objective and data availability: short-term mortality can be measured with 3-day or 7-day mortality, while longer mortality can include the entire admission period or extend beyond hospital discharge. In an analysis of over 700 acute stroke patients from eight hospitals in the UK; crude 30-day mortality rates for 1 year was observed to range from 25 % to 38 % at the hospital level (Mohammed et al. *2005*). The authors compare these rates with those reported by an independent company, and note the smaller breadth in variations in their own data. However, this difference in variations was likely affected to a large degree by the fact that the study used data spanning 1 year, whereas the independent company used a sample from over 6 years. This highlights the importance of temporal effects when comparing mortality rates. The effects of variations in the types of patients treated can be reduced by the appropriate application of risk adjustment methodologies; stringent study design can reduce the effects of variations in data collection methods and chance. Ideally, a comparison of the variations in mortality rates should reveal the existence and extent of variations, if any, in the quality of care provided.

In addition to the relatively orthodox explanatory variables of patient characteristics, some studies have explored the association of hospital factors on variations in mortality rates. A study of 34 academic trauma centers in the USA reported variations in care and their influence on outcome (Bulger et al. *2002*); specifically, the study showed the effect of quality of care on variations in observed mortality rates for patients with closed head injuries and long bone fractures. Of particular interest was the finding that management at "aggressive" centers (in which intracranial pressure monitors were used in more than 50 % of patients) was associated with significant reductions in the risk of mortality. Similarly, a study conducted by Brooke et al. (*2012*) compared 658 hospitals in the US that were either partially or fully compliant with National Quality Forum (NQF) safe practices standards; the results showed that full compliance with these standards was associated with decreased mortality.

An alternative to the mortality rate, whether crude or adjusted, is the use of ratios or differences to emphasize variations from a population or sample mean. For example, Krakauer et al. (*1995*) have addressed variations in medical practices and outcomes using "added days of life", defined as "the cumulative difference per patient between the observed probability of death over the period of observation (180 days) and that predicted from the national experience for the patients of the hospital", as an indicator. In this indicator, mortality is interpreted in the context of population mortality rates, with positive quantities indicating that observed probabilities of death are lower than predicted. In addition, observed (O) and expected (E) mortality can also be analyzed using the O/E ratio or O-E difference (Iezzoni *2003*) after calculating E values by regression modeling. These indicators can be applied to performance profiling in the comparison of outcomes across providers. The widely used standardized mortality rate (SMR) is an example of an O/E ratio, with E values derived from indirect standardization. Indirect standardization involves the application of stratum-specific mortality rates from all patients to the number of cases in each stratum within a hospital. In contrast, direct standardization methods involve the calculation of mortality rates for each risk stratum within a hospital, which are then applied to a standard population case mix.

Readmission Rates

Readmission rates can refer to short-term or long-term readmissions, and can include all-cause readmissions or be restricted to disease-specific readmissions. These rates may indicate the quality of healthcare with the assumption that appropriate and sufficient healthcare reduces the necessity for patient readmissions. This indicator may therefore be more appropriate for analyzing unplanned readmissions, as the inclusion of planned readmissions in staged procedures would introduce bias. Reasons for unplanned readmissions include relapses in the original primary diagnosis and the development of a separate condition, such asan infection (iatrogenic or otherwise). Ideally, variations in readmission rates should focus on unplanned and *nonpreventable* readmissions, as not all unplanned readmissions are preventable.

Payers may view high readmission rates as a source of unnecessary spending, and in some cases variations in these rates have been linked to hospital reimbursement systems. For example, the CMS has introduced financial penalties to hospitals in the USA for excessive read missions in AMI, heart failure, and pneumonia patients. Readmissions are calculated to be in excess through the use of a hospital's "excess readmission ratio", which measures a single hospital's performance compared with the national average for each condition. The ratio includes adjustments for patient demographics, comorbidities and patient frailty.

Variations in readmission rates in acute asthma patients admitted to two NHS hospitals in the UK have been shown to be associated with variations in clinical practice (Slack and Bucknall *1997*).

These clinical practice variations include the use and timing of corticosteroids. The identification of other factors would benefit future analyses by allowing analysts to employ better risk adjustment and to have a better frame of context for interpretations.

Length of Hospital Stay

LOS indicates the period of time in days spent by a patient in a healthcare institution, and is one of the main factors influencing hospitalization costs. Variations in durations can be very country specific: for example, LOS in Japan is very high compared with other OECD countries (OECD *2011*); this can be explained in part by an increasingly aged population and the provision of subacute care and chronic care in acute care hospitals. The per diem payment system in Japan may also play a part in this protracted LOS.

Prolonged LOS durations can indicate a lack of efficiency in a healthcare system, resulting in higher hospitalization costs, reductions in availability of empty beds for new admissions, and an increased risk of iatrogenic conditions. Quantifying variations in LOS durations may therefore help identify institutions where patients are hospitalized for longer periods than necessary. These potentially-avoidable costs are a major concern for stakeholders, and a reduction in LOS is a frequent goal of policymakers and insurers. Factors such as faster and more frequent transfers to postacute care providers or the increased use of day surgeries are possible LOS-reducing and cost-saving options. However, any incentives or penalties introduced to reduce LOS should not encourage hospitals to discharge or transfer patients too quickly.

An analysis of five teaching hospitals in Japan showed that mean unadjusted LOS durations for breast cancer surgery patients ranged from 16.7 days to 30 days for those who had undergone BCS and from 25.1 days to 28.9 days for patients who had undergone mastectomies (Ishizaki et al. *2002*). This study used univariate analyses to investigate unadjusted LOS variations, as well as linear mixed models to investigate the underlying factors explaining these variations. The results showed a much shorter average LOS but a much greater degree of variation in BCS patients compared with mastectomy patients. The hospitals used in the analysis were voluntary participants in a program with the expressed objective of monitoring QIs in order to improve quality. As such, these hospitals may be more indicative of institutions that are more conscientious toward quality improvement, and the variations observed may be an underestimation of the variations existing among all Japanese hospitals.

Morgan and Beech (*1990*) noted that the organization of care at the hospital level and clinical practice variations may underlie the observed variations in LOS. The authors observe that median LOS durations in NHS hospitals ranged from 4.4 to 6.0 days for inguinal hernia repair and 9.7 days to 10.7 days in cholecystectomy. Organizational factors that may influence LOS include different forms of health service funding, variations in the recording and coding of specific data items, bed allocation procedures, management and availability of surgical theaters, turnaround time for laboratories and x-ray departments, and organization of ancillary services. Clinical practice variations that affect LOS are likely to be more apparent in innovative practices with greater leeway for individual clinical judgment; these clinical practices include differences in preoperative investigations and postoperative management of patients, approaches to day surgeries or subsequent outpatient care, and rates of adoption of new forms of surgical management.

Using a combination of patient medical records and surveys, Cleary et al. (*1991*) analyzed the variations in LOS for patients admitted to six teaching hospitals for AMI, coronary artery bypass graft surgery, total hip replacement, cholecystectomy, or transurethral prostatectomy. The study showed significant interinstitutional differences in LOS for almost all of these conditions. Patient case mix

was able to account for the majority of variations for total hip replacement patients, but not for the differences in the other conditions. Three of the hospitals were located in California, while the other three were located in Massachusetts; it was observed that the hospitals in Massachusetts had longer average LOS durations than the California hospitals for AMI patients, suggesting the influence of regional factors in institutional variations.

Patient Satisfaction

Patient satisfaction with a healthcare provider can be divided into satisfaction with the technical elements of healthcare and satisfaction with the interpersonal elements of healthcare (Donabedian *1988*). Patient satisfaction takes into account the more intangible aspects of patients' preferences and choices, and the ability of a system to respond to these choices. Socioeconomic and cultural factors play a large role in patient satisfaction, which must be taken into account when conducting comparisons across providers, regions, and countries.

As with any survey-based qualitative approach, patient satisfaction as an outcome measure is vulnerable to researcher bias (particularly in the design stage), weak generalizability of results from small sample volumes, and varying patient expectations. However, these problems may be addressed using rigorous study design, ensuring acceptable statistical power and sample sizes, and employing the use of anchoring vignettes to correct for individual differences in expectations and perceptions.

Causes of Variations in Outcomes

Variations in outcomes may arise due to differences in practice variations, resource availability, payment systems, demographics, and random variations.

Medical practice variations may include the overutilization and underutilization of specific health services, and indicate inefficiencies in healthcare provision. For conditions where evidence-based standards of care are available, variations can be interpreted with respect to these standards. However, there are difficulties in establishing standards for all aspects of care, as well as getting clinicians to accept and adhere to these standards. When the target outcome is improvement to the quality of a patient's life, then the patient's preferences and values become critical factors in determining whether a procedure is indicated or not. Accordingly, practice variations may also reflect differences in valuations at the individual level, and such variability must be considered desirable (Mulley *1990*).

Decisions made at the medical frontlines are a combination of a clinician's experience, the intrinsic conditions of a specific patient, and extrinsic factors such as resource availability or patient density. Therefore, decision analysis (referring to the systematic approach to decision making under conditions of uncertainty) can also be used when investigating practice variations.

A lack of resources can affect a provider's ability to provide sufficient care. For example, lack of medical technologies such as CT scanners or the absence of specialized staff would render a provider unable to provide some types of services. Similarly, an excess in resources may be an incentive for clinicians to overuse healthcare services. Despite the fact that these variations can impact the quality of care at the institutional level, the effective allocation of resources to hospitals is better conducted at a regional level. Although it is possible for each individual hospital to independently provide comprehensive care, this level of redundancy makes little financial sense. However, systems that allow market forces to dictate resource distribution may result in wide variations in the distribution of resources such as high-cost medical technologies (Otsubo et al. *2011*).

Payment systems include prospective payment systems and fee-for-service systems. Prospective payment systems may encourage providers to under provide services in order to maximize profits. On

the other hand, fee-for-service systems tend to provide more services to patients who are covered under the system, and lower levels of care to patients without coverage (McPherson *1990*).
Although arguably more germane to regional variations, population and demographic variations areimportant determinants of case mix, as well as of intrinsic variations in morbidity. Variations in age composition, sex, ethnicity, diet, socioeconomic factors and environmental exposure can affect disease rates, thereby affecting demand for specific services, which in turn can influence variations in healthcare.
Random variations are those caused by chance factors that by definition cannot be anticipated, detected, or eliminated. As the rates of outcomes are measured by dividing the number of events by the population at risk, the contribution of random variations is heavily influenced by the number of cases and the frequency with which outcome events occur. It may be prudent for analysts to exclude hospitals with extremely low sample sizes from analyses in order to reduce the effects of chance.
Another possible factor influencing variations at the institutional level is the volume–outcome relationship. This relationship is based on the concept that "practice makes perfect", and that hospitals that treat many patients with a particular disease are able to develop the skill sets to efficiently and effectively perform these services with few complications (Luft et al. *1979*). Although studies have shown a significant association between volumes and outcomes (Birkmeyer et al. *2002*), magnitudes of association vary greatly and causations have yet to be clarified (Halm et al. *2002*). It is possible that good outcomes have attracted more patients to a particular hospital, resulting in higher patient volumes; or that hospitals with high volumes tend to be larger institutions with numerous specialist staff and technologies.
Additionally, shifts in the prevalence of risk factors have been shown to result in shifts in mortality (Green and Wintfeld *1995*). This emphasizes the importance of consistent monitoring of measures, and modifying the standards of care and risk factors when appropriate.

Methodologies for Analyzing Institutional Variations

Development and Validation of Quality Indicators

The AHRQ has described QIs as consisting of the following four elements, all of which should be present:

1.
A concept, the specific aspect of quality captured by the measure

2.
A perspective, the point of view from which the measure it taken

3.
A method, how the actual concept is measured

4.
An application, how is the measure actually used

The first step in developing QIs is to identify candidate measures. This consists primarily of three steps: a literature review, development of conceptual model, and expert engagement (AHRQ *2011*).

By reviewing the existing literature, it is possible to idxentify articles that address potential indicators, data sources, validation efforts, and risk adjustment methods. Development of a conceptual model is likely to help identify candidate measures. Expert opinions are a critical contribution to the development process – especially in lieu of scientific evidence – as they can improve the acceptability and feasibility of the QI.

Candidate measures obtained from existing sources have initial specifications for the inclusion/exclusion criteria, as well as defined numerators and denominators for calculation. Analysts can further modify these specifications to adapt the indicator to available data, improve the indicator based on new evidence, or incorporate changes to the data source.

QIs should be assessed for acceptability, feasibility, reliability, sensitivity to change, and validity. The acceptability of each QI to both the providers being assessed and their assessors depends on the clinical merit of the indicator as well as possible privacy considerations. Feasibility refers to whether a QI can be practically applied in healthcare services research, and is influenced by the availability and quality of data. Reliability is the extent to which a measurement with an indicator is reproducible. Sensitivity to change is critical to the success of a QI, as quality measurements need to detect changes in the quality of care in order to discriminate within and across providers. Finally, indicators developed by consensus techniques have face validity, whereas those based on rigorous evidence (such as those based on randomized controlled trials) possess content validity (Campbell et al. *2003*; McGlynn and Asch *1998*).

Risk adjustments can help identify and account for relevant explanatory variables. Empirical analyses also make it possible to determine the relative bias of the candidate measures, the precision and reliability of each indicator, and variation in rates among institutions.

Importance of Local Context

Many QIs have been developed and used to evaluate quality of care, and the wealth of information regarding currently used QIs allows analysts to choose from a wide variety of existing measures. However, preexisting QIs may need to be modified to allow for local attributes and specific objectives because they are frequently developed and validated using data or motives that are different from those of the current user. Adaptation to the local context is necessary for two reasons (Delnoij and Westert *2012*): First, the information infrastructure that is used to calculate indicator scores may be different between healthcare systems. Second, healthcare systems vary across regions and countries. For instance, hospital SMRs should be adjusted for regional variations in the organization and performance of healthcare institutions adjacent to the hospital of interest.

In additional to data and systemic differences, there may also be social and cultural factors that influence specific QIs. The effects of these determinants may be more noticeable in qualitative measures (such as patient satisfaction) or indicators that involve end-of-life treatments, due to cultural differences in attitudes toward mortality (Fletcher et al. *1992*). Without considering these and other local factors, it may be impossible to accurately evaluate the quality of care.

Methodological Approaches to Institutional-Level Variations

Consider a hypothetical situation showing different unadjusted mortality rates in AMI patients in two hospitals, with Hospital A at 10 % and Hospital B at 8 %. At first glance, Hospital B has better quality of care due to better survival. However, a separate analysis may show that average resource utilization at Hospital A is much higher than that of Hospital B, leading observers to conclude that Hospital A is

providing more, and therefore better, care. With increasing availability of healthcare data, questions like these confront physicians, hospitals management staff, and policymakers in their quest to understand and improve quality of care. It is the responsibility of the researchers providing information to these decision makers to ensure that the information is accurate and reflects the influence of all relevant factors. The conclusions to the hypothetical situations above are simplistic and possibly misleading, and appropriate steps are required to account for possible differences between treatment groups. Although detailed research methodology is addressed in another chapter of this handbook, we briefly address some methodological aspects of analysis specific for variations at the institutional level.

Risk Adjustment Methods

Evaluations of quality of care (both process and outcome measures) are traditionally sought to compare new treatments or program with existing ones. Recently, QIs are increasingly being used to identify variations in the quality of care across different providers. However, comparator groups are often different with respect to important baseline risk factors for the quality of care. The challenge is to render these comparisons fair, meaningful, and ultimately, useful.

As stated earlier, analyses should adjust for intrinsic variations such as patient demographics to ensure that the variations observed are not simply a result of differential case mix. Different types of hospitals treat different types of patients, and some hospitals (such as teaching hospitals) may treat sicker patients. Analysts must be aware of these intrinsic differences, adjust for them when possible, and take caution in interpreting the results.

Adjustment methodologies include simple matched comparisons, as well as more sophisticated methods to account for multiple covariates. Adjustments can be applied at the design stage or analytical stage.

Accurate risk adjustment starts with attention to design issues, particularly with respect to the data source. There are three main categories of data sources: administrative data, clinical data, and survey data. As each of these data sources has its own unique characteristics, the data source should ideally be selected for the specific purpose of each QI. However, there are feasibility limitations in database availability for analysts.

The next step is to ensure that the appropriate subjects are included in the study sample. For example, AMI patients are unlikely to be discharged home within a short period of time. Therefore, administrative data that identify a patient as being admitted for an AMI with a LOS of only 2 days (with no mortality) may represent a data entry mistake, and their inclusion in analysis may result in misleading conclusions.

Risk adjustments at the analytical stage are generally more focused on outcome measures than process measures. Regression modeling is central to most analytic strategies. In multivariable regression models, risk adjustment factors can include age, sex and gender, race and ethnicity, acute clinical stability (referring to the physiologic functioning of a patient), principal diagnosis, comorbidities, functional status, health behaviors, sociocultural and environmental attributes, overall health statuses, and patient preferences (Iezzoni *2003*).

Advanced statistical and econometric methods have been developed and used in health service research. For example, instrumental variables or differences in differences have been employed to address unmeasured or unknown confounders (Newhouse and McClellan *1998*; Yew *2009*).

Furthermore, multilevel regression models can be used to adjust for patient and hospital characteristics while controlling for the hierarchical structure of the data. Patients treated in the same hospital are more likely to be similar to each other than with randomly chosen patients, but traditional

multivariate and logistic regression models generally ignore this correlation, and may therefore overestimate the precision of hospital-level associations (Grieve et al. *2005*). Propensity scores are increasingly used in observational studies due to an ability to explicitly determine the degree to which measured characteristics are balanced across treatment groups (D'Agostino *2007*). Even if the degree of risk adjustment varies across several methods, they can reduce the unexplained variations in the types of patients being treated; however, overly-complicated systems are harder to reproduce and more difficult to interpret for nonexpert stakeholders.

There are several limitations to risk adjustment methods that should be noted (Yew *2009*). First, the accuracy and usability of outcome measures is reliant on the completeness and accuracy of the data available. Second, there may be no perfect methods to adjust for risk factors (Iezzoni *2003*). Third, risk adjustments cannot be used to directly compare the outcome indicators when comparators are very different (Shahian and Normand *2008*).

Consequently, risk-adjusted indicators should be interpreted cautiously in light of the limitations mentioned above, and analysts should report the methodologies used. This transparency in QI development and analytical methodology can support the interpretation of the QIs, as well as better allow other researchers to use them appropriately.

Additionally, the simultaneous monitoring of structure, process, and outcome measures would be beneficial for understanding the variations in healthcare quality at the institutional level, as it would allow analysts to investigate the relationships between the various measures. For example, in the case of nosocomial infections, an analysis of (1) the presence/absence of infection control teams, (2) antibiotic utilization or infection reduction activities, and (3) nosocomial infection rates would provide much more usable information for decision makers than an analysis of infection rates alone.

Consequences of Higher Spending

The consequences of higher spending can include both direct and indirect effects on the quality of healthcare. This relationship has important implications for healthcare provision because sweeping reductions in healthcare utilization as a means of controlling costs may inadvertently compromise the quality of care to patients. Therefore, simply limiting healthcare utilization may not necessarily be the optimal solution to curtail rising costs (Bodenheimer and Fernandez *2005*). In 1980, Avedis Donabedian described a hypothetical relationship between the volume of services provided and patient outcomes (Donabedian *1980*). He proposed that the provision of additional services would not always lead to better outcomes due to the increase in risk that accompanies medical care. Under this concept, quality of care is therefore unlikely to exhibit monotonic improvements brought about by increases in the volume of services provided; instead, there will be a point where further inputwould result in no change or even reductions in the quality of care.

Several studies have previously examined whether higher healthcare spending is associated with better quality of care at the hospital level. The bulk of these studies are from the US, and results thus far appear to be mixed: although some studies have reported a lack of association (or inconsistent associations) between the quality of care and spending (Chen et al. *2010*; Hvenegaard et al. *2011*; Jha et al. *2009*; Kaestner and Silber *2010*; Yasaitis et al. *2009*), others have shown associations between lower mortality rates and higher healthcare spending for several diseases including AMI (Barnato et al. *2010*; Ong et al. *2009*; Romley et al. *2011*; Schreyögg and Stargardt *2010*; Stukel et al. *2012*). Asystematic review published in 2013 found that over 80 % (10/12) of studies conducted on regional-level associations reported negative or no associations between healthcare spending and quality (Hussey et al. *2013*). Furthermore, the Institute of Medicine (IOM) reported that there is no evidence

supporting the existence of a consistent regional-level relationship between spending and quality (IOM *2013*). In contrast, positive associations between spending and quality appear to be more frequent at the hospital level than at the regional level (Hussey et al. *2013*).

Incremental costs in healthcare are often influenced by the volume of medical services provided. The Dartmouth Atlas has defined "aggressive treatment style" as the level of intensity of medical treatment according to the quantity of medical resources utilized (Fisher et al. *2003a*, *b*). Although the potential benefits from aggressive treatment styles may seemingly be outweighed by the accompanying risks of increased iatrogenic complications and mortality, there is some evidence that this assumption may not necessarily hold true: in an investigation of the relationship between aggressive treatment style and postsurgical patient outcomes, Silber and colleagues (*2010*) found that aggressive treatment for surgical procedures was associated with significantly reduced mortality. Therefore, the type of medical treatment should be taken into consideration when examining the relationships between spending and overall outcomes.

Wennberg and colleagues (Wennberg et al. *2002*) introduced a typology of healthcare comprising three categories: effective care, preference-sensitive care, and supply-sensitive care (See other chapter). Skinner (*2012*) further described the conceptual relationship between benefits and costs in each category. In effective (Type I category) care, the net value of care is comparatively high against the risk of side effects. In contrast, the overall benefits of preference-sensitive (Type II category) care tend to be less dominant in relation to total cost. Supply-sensitive (Type III category) care provides few incremental benefits in relation to costs, but is likely to account for a majority of the differences in regional-level costs. In this way, the type of healthcare may also play an important role in the relationship between spending and outcomes (Fisher and Skinner *2010*).

Figure 1 shows a hypothetical model of the relationship between healthcare spending and quality; the circles labeled A through F represent hospitals exhibiting different spending–outcome relationships. The functions f(x) and g(x) indicate conventional production functions. A hospital on the f(x) function may be able to achieve more efficient and effective healthcare delivery than a hospital on the g(x) function by actively adhering to measures such as the use of surgical checklists or ensuring careful hand hygiene. If the intrinsic differences in production functions between hospitals are not taken into account, there is a risk that the associations between quality and spending may be misinterpreted. For example, an analysis of Hospitals A and B without taking into consideration the different production functions may erroneously indicate the existence of a positive correlation between spending and outcome. Similarly, by not acknowledging the differences in production functions, an analysis of Hospitals C and D may imply a lack of correlation between spending and outcomes while an analysis of Hospitals E and F would indicate a negative correlation between the two factors. Therefore, the efficiency and effectiveness of healthcare delivery is an important factor when examining the relationship between quality and spending.

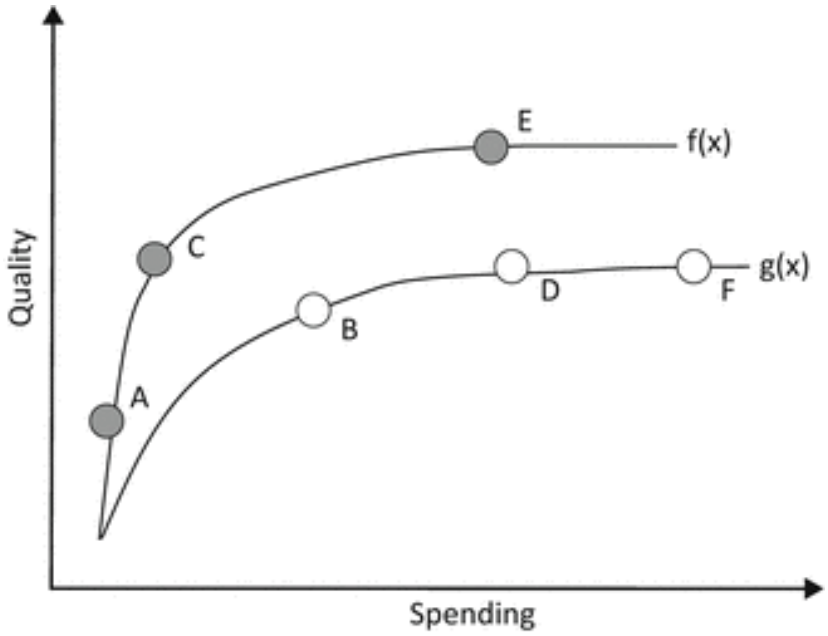

Fig. 1
Hypothetical institutional-level relationship between spending and quality

When interpreting observed associations between spending and quality or when developing strategies to reduce spending or improve quality, it is important to take into account national and institutional contexts such as payment systems, insurance systems, care coordination systems, and hospital teaching status. Also, the associations between spending and quality can be influenced by hospital structural factors that dictate the treatment capability of each hospital (Donabedian *2005*; Spertus et al. *2003*). Examples of such factors include the presence and volume of human resources, medical equipment, and facilities.

With the aim of understanding the consequences of higher healthcare spending, the relationship between spending and quality has been examined with increasing intensity in recent years. As described above, current evidence suggests that these two factors are not always correlated. The influence of each country's unique healthcare system on this relationship means that different countries cannot reply on each other's results to guide health policy, and analyses must be conducted in the context of each country. Analysts need to take into account these and other factors in their study designs and consider potential biases to reduce misinterpretations.

Reducing Unwarranted Institutional Variations

Studies in a number of countries have noted widespread variations in healthcare spending and quality. In cases where spending appears to be intrinsically linked to quality, efforts to reduce variations in one may translate to reductions in the other. Poor quality in healthcare, such as that resulting from numerous medical errors, is also likely to have a substantial impact on healthcare spending because of the ensuing costs associated with rectifying errors. Here, improving the quality of care may reduce extraneous costs. Achieving improved quality of care also requires systems that reward good performance. In addition, providers need to show that the prices they charge consumers are commensurate with the services provided. Stakeholders in healthcare must agree upon – and report – standardized outcome measures for the value of healthcare, such as risk-adjusted mortality rates.

These changes will make it easier to move toward a hospital payment system based on standardized care, and allow patients to judge the value and quality that providers offer.

For unwarranted variations in medical practice, the key to reducing variations lies in their accurate identification after taking into account factors that do not reflect quality of care, such as patient case mix, suitability of treatment, and random error. In addition to the factors highlighted above, practice variations can also be affected by the fundamental and inevitable presence of uncertainty in clinical decision making. The decision-making process is a complicated interaction of scientific evidence, patient desire, doctor preference, local culture, and various external influences, not all of which may be reasonable. Additionally, improvements to the depth, quality, and sharing of medical information (between providers, payers, and patients) may also help to reduce variations in spending and quality by supporting better planning and monitoring.

Once the variations and underlying causes have been identified, hospital management staff, and policymakers may decide if improvements are necessary and feasible, and accordingly direct resources to do so. Elucidating variations at the institutional level allows better identification of the appropriate people in authority, which ultimately supports the implementation of necessary improvements at the correct level. It is possible that the variations at the institutional level are intertwined with those at regional levels, and addressing these variations may therefore require collaborative efforts between regional- and institutional-level stakeholders. Wennberg has proposed the following four steps in order to improve healthcare quality across the board and reduce unwarranted variations (Wennberg *2010*):

1.

Promoting organized systems of healthcare delivery

2.

Establishing informed patient choice as the ethical and legal standard for decisions surrounding elective surgeries, drugs, tests, and procedures, and at the end of life

3.

Improving the science of healthcare delivery

4.

Constraining undisciplined growth in healthcare capacity and spending

Conclusions

Readers should note that there is no optimal scenario in which healthcare is provided with a complete lack of variation at the institutional level. Variations can and should arise in response to differing patient values, expectations, and preferences. Also, epicenters of new and effective knowledge, methodologies, and technologies would be associated with better healthcare, and would manifest as a change in the status quo rippling through providers and regions as these innovations are adopted at varying rates.

It is, instead, the unwarranted variations in healthcare that are the focus of researchers and policymakers, but therein lies another problem: how can we distinguish “unwarranted” variations from those that occur as an inevitable consequence of intrinsic differences in populations and environment, and would exist regardless of the actions of healthcare providers? Figure 2 shows a

conceptual summary of the distinction between observed variations and unwarranted variations at the institutional level and possible factors leading to unwarranted variations in spending and quality.

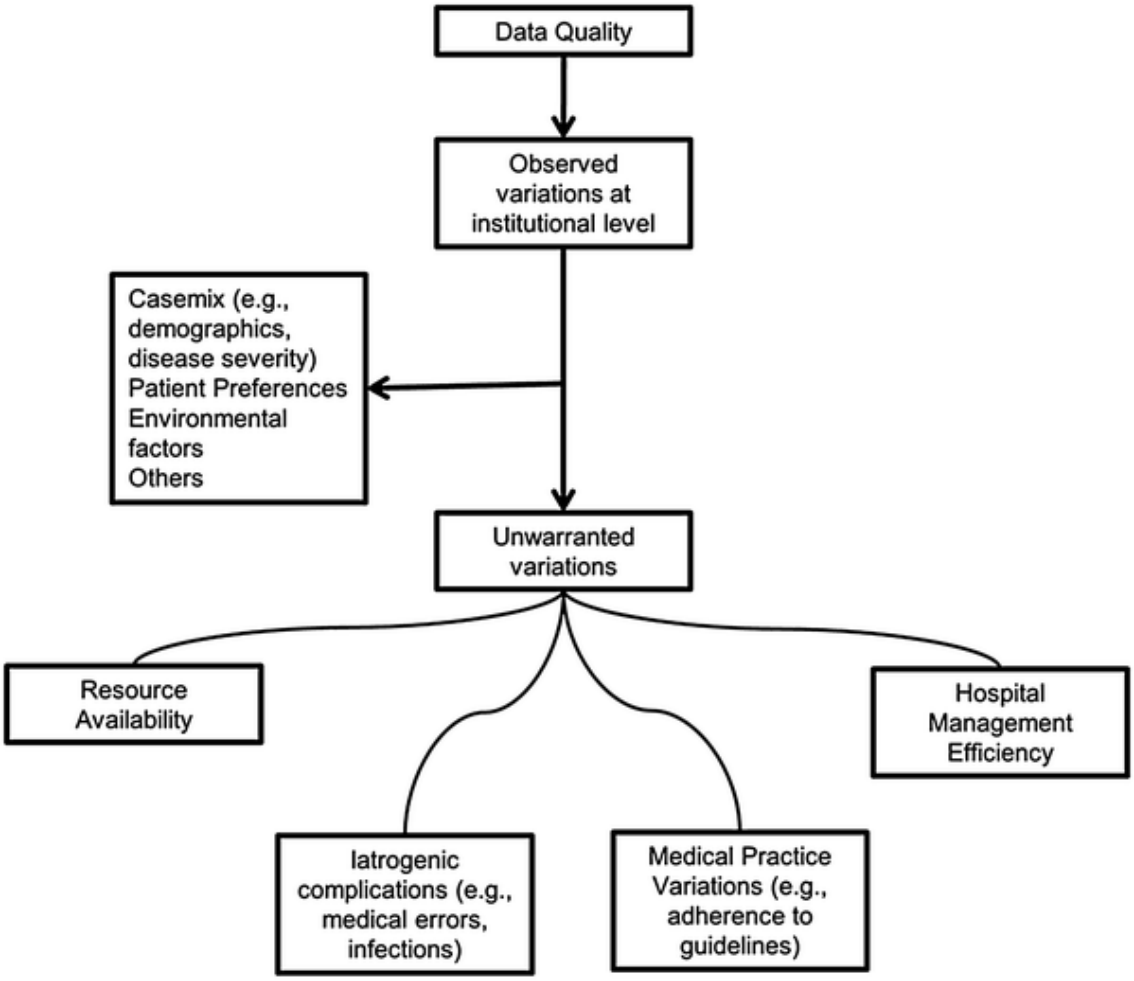

Fig. 2
The conceptual distinction between observed variations and unwarranted variations, and the possible factors that influence unwarranted variations in healthcare at the institutional level

Firstly, the extent of the observed variations is dependent on the quality of the data analyzed, which can be affected by missing data or systematic error. Researchers should control for patient case mix (such as patient demographics and disease severity), patient preferences, and environmental factors. These factors represent intrinsic variables affecting the demand of types and quantity of healthcare and therefore cannot be considered “unwarranted” or indicative of poor quality. Factors that can influence unwarranted variations include uneven resource availability, hospital management efficiency, medical practice variations, and iatrogenic complications. Variations in resource availability may arise in countries that rely solely on free markets, which can lead to the maldistribution of resources. In such cases, it may be beneficial for policy implementation to ensure better distribution of resources and care. Hospital management efficiency could give rise to spending variations as hospitals differ in their utilization of existing resources and proper training of management and clinical staff may reduce inefficiencies. Medical practice variations occur when there are no standards agreed upon by consensus and evidence or when clinicians provide care differently from standards and guidelines. If clinicians disagree with these standards and actively provide healthcare contrary to the recommendations, this may indicate problems in the formulation of these standards or failures to present them convincingly to frontline healthcare providers. If the issues

lie in differences in patient preferences, cultural differences, or other such extenuating circumstances, then these variations must be interpreted accordingly. Also, differential rates of iatrogenic complications and errors may reflect differences in the training of clinicians and other staff or the general safety culture within each hospital.

Additionally, the simultaneous monitoring of structure, process, and outcome measures would be beneficial for understanding the variations in healthcare quality at the institutional level, as it would allow analysts to investigate the relationships between the various measures.

The health services researcher must also understand that all variations are products of the surrounding healthcare systems at the regional and national levels and must therefore be interpreted within that context. For example, the variations observed in a healthcare system with a uniform reimbursement system would be different with those observed in a system that utilizes several reimbursement methods. Additionally, an overall lack of any particular resource within a region would still have an influence on institutional-level variations.

The identification of unwarranted variations at the institutional level is a stepping stone to the greater objective of identifying and explaining the underlying reasons for the variations. This ensures that stakeholders with the appropriate jurisdiction and authority are made aware of the problem and are able to divert resources to address the issue. Stakeholders must work to reach a consensus on what types of care and spending are appropriate, accurately identify causation, and support the design and implementation of interventions to reduce unwarranted variations at the institutional level.

References

AHRQ Quality Indicator Measure Development, Implementation, Maintenance and Retirement. Quality indicator development. Agency for Healthcare Research and Quality; 2011. http://www.qualityindicators.ahrq.gov/modules/Default.aspx. Accessed 29 June 2014.

Barnato AE, Chang C-H, Farrell MH, et al. Is survival better at hospitals with higher “end-of-life” treatment intensity? Med Care. 2010;48:125–32.
PubMedCentral CrossRef PubMed

Bednarek FJ, Weisberger S, Richardson DK, et al. Variations in blood transfusions among newborn intensive care units. J Pediatr. 1998;133:601–7.
CrossRef PubMed

Birkmeyer JD, Siewers AE, Finlayson EVA, et al. Hospital volume and surgical mortality in the united states. N Engl J Med. 2002;346:1128–37.
CrossRef PubMed

Bodenheimer T, Fernandez A. High and rising health care costs. Part 4: can costs be controlled while preserving quality? Ann Intern Med. 2005;143:26–31.
CrossRef PubMed

Brooke BS, Dominici F, Pronovost PJ, et al. Variations in surgical outcomes associated with hospital compliance with safety practices. Surgery. 2012;151:651–9.
PubMedCentral CrossRef PubMed

Bulger EM, Nathens AB, Rivara FP, et al. Management of severe head injury: institutional variations in care and effect on outcome. Crit Care Med. 2002;30:1870–6.
CrossRef PubMed

Busse R. Do diagnosis-related groups explain variations in hospital costs and length of stay? – Analyses from the EURODRG project for 10 episodes of care across 10 European countries. Health Econ. 2012;21:1–5.
CrossRef PubMed

Campbell SM, Braspenning J, Hutchinson A, et al. Improving the quality of health care: research methods used in developing and applying quality indicators in primary care. Br Med J. 2003;326:816–9.
CrossRef

Chen Z, Sandercock P, Xie J, et al. Hospital management of acute ischemic stroke in china. J Stroke Cerebrovasc Dis. 1997;6:361–7.
CrossRef PubMed

Chen J, Radford MJ, Wang Y, et al. Performance of the “100 top hospitals”: what does the report card report? Health Aff. 1999;18:53–68.
CrossRef

Chen LM, Jha AK, Guterman S, et al. Hospital cost of care, quality of care, and readmission rates: penny wise and pound foolish? Arch Intern Med. 2010;170:340–6.
CrossRef PubMed

Cleary PD, Greenfield S, Mulley AG, et al. Variations in length of stay and outcomes for six medical and surgical conditions in massachusetts and california. JAMA. 1991;266:73–9.
CrossRef PubMed

D’Agostino Jr RB. Propensity scores in cardiovascular research. Circulation. 2007;115:2340–3.
CrossRef PubMed

Delnoij DMJ, Westert GP. Assessing the validity of quality indicators: keep the context in mind! Eur J Public Health. 2012;22:452–3.
CrossRef PubMed

Donabedian A. The definition of quality and approaches to its assessment, Explorations in quality assessment and monitoring, vol. I. Ann Arbor: Health Administration Press; 1980. p. 8–11.

Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260:1743–8.
CrossRef PubMed

Donabedian A. Evaluating the quality of medical care. Milbank Q. 2005;83:691–729.
PubMedCentral CrossRef PubMed

Dranove D. Health care markets, regulators, and certifiers. In: Pauly MV, Mcguire TG, Barros PP, editors. Handbook of health economics. Amsterdam: Elsevier; 2011. p. 639–90.

Fisher E, Skinner J. Comment on Silber et al.: aggressive treatment styles and surgical outcomes. Health Serv Res. 2010;45:1893–902.
PubMedCentral CrossRef PubMed

Fisher ES, Wennberg DE, Stukel TA, et al. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Ann Intern Med. 2003a;138(4):273–87.
CrossRef PubMed

Fisher ES, Wennberg DE, Stukel TA, et al. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Ann Intern Med. 2003b;138:288–98.
CrossRef PubMed

Fisher ES, Wennberg DE, Stukel TA, et al. Variations in the longitudinal efficiency of academic medical centers. Health Aff (Millwood). 2004;Suppl Web Exclusive:VAR19-32.

Fletcher A, Gore S, Jones D, et al. Quality of life measures in health care. II: design, analysis, and interpretation. Br Med J. 1992;305:1145–8.
CrossRef

Gaynor M, Town RJ. Competition in health care markets. In: Pauly MV, Mcguire TG, Barros PP, editors. Handbook of health economics. Amsterdam: Elsevier; 2011. p. 499–637.

Gellad WF, Good CB, Lowe JC, et al. Variation in prescription use and spending for lipid-lowering and diabetes medications in the veterans affairs healthcare system. Am J Manag Care. 2010;16:741–50.
PubMedCentral PubMed

Gilbert K, Gleason PP, Singer DE, et al. Variations in antimicrobial use and cost in more than 2,000 patients with community-acquired pneumonia. Am J Med. 1998;104:17–27.
CrossRef PubMed

Green J, Wintfeld N. Report cards on cardiac surgeons: assessing New York state's approach. N Engl J Med. 1995;332:1229–32.
CrossRef PubMed

Grieve R, Nixon R, Thompson SG, et al. Using multilevel models for assessing the variability of multinational resource use and cost data. Health Econ. 2005;14:185–96.
CrossRef PubMed

Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med. 2002;137:511–20.
CrossRef PubMed

Hasley PB, Lave JR, Kapoor WN. The necessary and the unnecessary transfusion: a critical review of reported appropriateness rates and criteria for red cell transfusions. Transfusion. 1994;34:110–5.
CrossRef PubMed

Henderson JW, editor. Confounding factor – SALT. In: Health economics and policy. 2nd ed. Ohio: South-Western Publishing Co.; 2002. p. 299–332.

Hollingsworth B. The measurement of efficiency and productivity of health care delivery. Health Econ. 2008;17:1107–28.
CrossRef PubMed

Hussey PS, Wertheimer S, Mehrotra A. The association between health care quality and cost: a systematic review. Ann Intern Med. 2013;158(1):27–34.
CrossRef PubMed

Hvenegaard A, Arendt JN, Street A, et al. Exploring the relationship between costs and quality: does the joint evaluation of costs and quality alter the ranking of Danish hospital departments? Eur J Health Econ. 2011;12:541–51.
CrossRef PubMed

Iezzoni L, editor. Risk adjustment for measuring health care outcomes. 3rd ed. Chicago: Health Administration Press; 2003. p. 33–70.

IOM (Institute of Medicine). Variation in health care spending: target decision making, not geography. Washington, DC: The National Academies Press; 2013.

Iscoe NA, Goel V, Wu K, et al. Variation in breast cancer surgery in ontario. CMAJ. 1994;150:345–52.
PubMedCentral PubMed

Ishizaki T, Imanaka Y, Hirose M, et al. A first look at variations in use of breast conserving surgery at five teaching hospitals in Japan. International J Qual Health Care. 2002;14:411–8.
CrossRef

Jha AK, Orav EJ, Dobson A, et al. Measuring efficiency: the association of hospital costs and quality of care. Health Aff. 2009;28:897–906.
CrossRef

Kaestner R, Silber JH. Evidence on the efficacy of inpatient spending on Medicare patients. Milbank Q. 2010;88:560–94.
PubMedCentral CrossRef PubMed

Krakauer H, Bailey RC, Cooper H, et al. The systematic assessment of variations in medical practices and their outcomes. Public Health Rep. 1995;110:2–12.
PubMedCentral PubMed

Kuwabara K, Imanaka Y, Matsuda S, et al. Profiling of resource use variation among six diseases treated at 82 Japanese special functioning hospitals, based on administrative data. Health Policy. 2006;78:306–18.
CrossRef PubMed

Lameire N, Joffe P, Wiedemann M. Healthcare systems – an international review: an overview. Nephrol Dial Transplant. 1999;14:3–9.
CrossRef PubMed

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.
CrossRef PubMed

Mant J. Process versus outcome indicators in the assessment of quality of health care. International J Qual Health Care. 2001;13:475–80.
CrossRef

McGlynn EA, Asch SM. Developing a clinical performance measure. Am J Prev Med. 1998;14:14–21.
CrossRef PubMed

McPherson K. International differences in medical care practices. In: Health care systems in transition, OECD social policy studies, vol. 7. Paris: OECD; 1990. p. 17–28.

Miller EA, Woosley J, Martin CF, et al. Hospital-to-hospital variation in lymph node detection after colorectal resection. Cancer. 2004;101:1065–71.
CrossRef PubMed

Mohammed MA, Mant J, Bentham L, et al. Comparing processes of stroke care in high- and low-mortality hospitals in the west midlands, UK. International J Qual Health Care. 2005;17:31–6.
CrossRef

Morgan M, Beech R. Variations in lengths of stay and rates of day case surgery: implications for the efficiency of surgical management. J Epidemiol Community Health. 1990;44:90–105.
PubMedCentral CrossRef PubMed

Morishima T, Lee J, Otsubo T, et al. Association of healthcare expenditures with aggressive versus palliative care for cancer patients at the end of life: a cross-sectional study using claims data in Japan. International J Qual Health Care. 2014;26:79–86.
CrossRef

Mulley Jr A. Medical decision making and practice variation. In: Andersen T, Mooney G, editors. The challenges of medical practice variations. London: Macmillan Press; 1990. p. 59–75.

Mulley AG. Improving productivity in the NHS: reducing practice variation thorough better decision making is key. Br Med J. 2010;341:c3965.
CrossRef

Newhouse J. Accounting for teaching hospitals' higher costs and what to do about them. Health Aff. 2003;22:126–9.
CrossRef

Newhouse J, McClellan M. Econometrics in outcomes research: the use of instrumental variables. Annu Rev Public Health. 1998;19:17–34.
CrossRef PubMed

OECD. Average length of stay in hospitals. In: Health at a Glance. 6th ed. Paris: OECD Publishing; 2011. doi:10.1787/health_glance-2011-en. Accessed 29 June 2014.

Ong MK, Mangione CM, Romano PS, et al. Looking forward, looking back: assessing variations in hospital resource use and outcomes for elderly patients with heart failure. Circ Cardiovasc Qual Outcomes. 2009;2:548–57.
PubMedCentral CrossRef PubMed

Otsubo T, Imanaka Y, Lee J, et al. Evaluation of resource allocation and supply – demand balance in clinical practice with high-cost technologies. J Eval Clin Pract. 2011;17:1114–21.
CrossRef PubMed

Paris V, Devaux M, Wei L. Health care delivery. In: Health systems institutional characteristics: a survey of 29 organisation for economic co-operation and development (OECD) countries. OECD Publishing; 2010. doi:10.1787/5kmfxfq9qbnr-en. Accessed 29 June 2014.

Posnett J. Are bigger hospitals better? In: McKee M, Judith H, editors. Hospitals in a changing Europe. Philadelphia: Open University Press; 2002. p. 100–18.

Rao SV, Jollis JG, Harrington RA, et al. Relationship of blood transfusion and clinical outcomes in patients with acute coronary syndromes. JAMA. 2004;292:1555–62.
CrossRef PubMed

Reschovsky JD, Hadley J, Saiontz-Martinez CB, et al. Following the money: factors associated with the cost of treating high-cost Medicare beneficiaries. Health Serv Res. 2011;46:997–1021.
PubMedCentral CrossRef PubMed

Ringer SA, Richardson DK, Sacher RA, et al. Variations in transfusion practice in neonatal intensive care. Pediatrics. 1998;101:194–200.
CrossRef PubMed

Romley JA, Jena AB, Goldman DP. Hospital spending and inpatient mortality: evidence from California. Ann Intern Med. 2011;154:160–7.
PubMedCentral CrossRef PubMed

Schouten JA, Hulscher ME, Kullberg B, et al. Understanding variation in quality of antibiotic use for community-acquired pneumonia: effect of patient, professional and hospital factors. J Antimicrob Chemother. 2005;56:575–82.
CrossRef PubMed

Schreyögg J, Stargardt T. The trade-off between costs and outcomes: the case of acute myocardial infarction. Health Serv Res. 2010;45:1585–601.
PubMedCentral CrossRef PubMed

Sekimoto M, Imanaka Y, Evans E, et al. Practice variation in perioperative antibiotic use in Japan. International J Qual Health Care. 2004;16:367–73.
CrossRef

Sekimoto M, Kakutani C, Inoue I, et al. Management patterns and healthcare costs for hospitalized patients with cerebral infarction. Health Policy. 2008;88:100–9.
CrossRef PubMed

Sekimoto M, Imanaka Y, Shirai T, et al. Risk-adjusted assessment of incidence and quantity of blood use in acute-care hospitals in Japan: an analysis using administrative data. Vox Sang. 2010;98:538–46.
CrossRef PubMed

Shahian DM, Normand S-T. Comparison of "risk-adjusted" hospital outcomes. Circulation. 2008;117:1955–63.
CrossRef PubMed

Silber JH, Kaestner R, Even-Shoshan O, et al. Aggressive treatment style and surgical outcomes. Health Serv Res. 2010;45:1872–92.
PubMedCentral CrossRef PubMed

Skinner J. Causes and consequences of regional variations in health care. In: Pauly M, Mcguire T, Barros P, editors. Handbook of health economics, vol. 2. Waltham: Elsevier; 2012. p. 45–93.

Slack R, Bucknall CE. Readmission rates are associated with differences in the process of care in acute asthma. Qual Saf Health Care. 1997;6:194–8.
CrossRef

Spertus JA, Radford MJ, Every NR, et al. Challenges and opportunities in quantifying the quality of care for acute myocardial infarction: summary from the acute myocardial infarction working group of the American Heart Association/American college of cardiology first scientific forum on quality of care and outcomes research in cardiovascular disease and stroke. Circulation. 2003;107:1681–91.
CrossRef PubMed

Steinhubl SR, Kastrati A, Berger PB. Variation in the definitions of bleeding in clinical trials of patients with acute coronary syndromes and undergoing percutaneous coronary interventions and its impact on the apparent safety of antithrombotic drugs. Am Heart J. 2007;154:3–11.
CrossRef PubMed

Street A, Scheller-Kreinsen D, Geissler A, et al. Determinants of hospital costs and performance variation. In: Methods, models and variables for the EuroDRG project. 2010. http://opus.kobv.de/tuberlin/volltexte/2010/2661/pdf/10_05_10_WPplaintext_MiG_Band3_mg.pdf. Accessed 29 June 2014.

Street A, Kobel C, Renaud T, et al. How well do diagnosis-related groups explain variations in costs or length of stay among patients and across hospitals? Methods for analysing routine patient data. Health Econ. 2012;21:6–18.
CrossRef PubMed

Stukel TA, Fisher ES, Alter DA, et al. Association of hospital spending intensity with mortality and readmission rates in Ontario hospitals. JAMA. 2012;307:1037–45.
PubMedCentral CrossRef PubMed

Tiemann O, Schreyögg J, Busse R. Hospital ownership and efficiency: a review of studies with particular focus on Germany. Health Policy. 2012;104:163–71.
CrossRef PubMed

Van Herck P, De Smedt D, Annemans L, et al. Systematic review: effects, design choices, and context of pay-for-performance in health care. BMC Health Serv Res. 2010;10:247–59.
PubMedCentral CrossRef PubMed

Van Steenbergen LN, Rutten HJT, Creemers GJ, et al. Large age and hospital-dependent variation in administration of adjuvant chemotherapy for stage III colon cancer in southern Netherlands. Ann Oncol. 2009;21:1273–8.
CrossRef PubMed

Vladeck BC, Goodwin EJ, Myers LP, et al. Consumers and hospital use: the HCFA “death list”. Health Aff. 1988;7:122–5.
CrossRef

Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. Br Med J. 2002;325:961–4.
CrossRef

Wennberg JE, editor. In health care, geography is destiny. In: Tracking Medicine: a researcher’s quest to understand health care. 1st ed. New York: Oxford University Press; 2010. p. 3–13

Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. Health Affairs. 2002; Suppl Web Exclusives:W96–114

Wennberg JE, Fisher ES, Stukel TA, et al. Use of Medicare claims data to monitor provider-specific performance among patients with severe chronic illness. Health Aff (Millwood). 2004;Suppl Web Exclusive:VAR5-18.

White C. Health status and hospital prices key to regional variation in private health care spending. National Institute for Health Care Reform Research Brief; 2012. http://www.nihcr.org/Spending_Variation.html. Accessed 29 June 2014.

Yasaitis L, Fisher ES, Skinner JS, et al. Hospital quality and intensity of spending: is there an association? Health Aff. 2009;28:w566–72.
CrossRef

Yew YD. Risk adjustment: towards achieving meaningful comparison of health outcomes in the real world. Ann Acad Med Singapore. 2009;38:552–8.