The Policy Implications of Using Hospital and Physician Volumes as “Indicators” of Quality of Care in a Changing Health Care Environment

KATHRYN A. PHILLIPS*†‡§ and HAROLD S. LUFT†‡§

*Center for AIDS Prevention Studies, Prevention Sciences Group, University of California, UCSF, San Francisco, California, U.S.A; †Department of General Internal Medicine, UCSF, San Francisco, California, U.S.A; ‡Department of Epidemiology and Biostatistics, UCSF, San Francisco, California, U.S.A; §Institute for Health Policy Studies, UCSF, San Francisco, California, U.S.A.

There is growing interest in the quality of health care and in using quality measures to direct patients to hospitals and providers offering high quality, low cost health care. The dilemma is that, while there is an increasing need for quality indicators as a result of a changing health care environment, this changing environment has important implications for the use of some of these measures. Since the 1970s, a growing body of research in the U.S. has addressed the empirical relationship between the number of patients with a specific diagnosis of surgical procedure and their outcomes after treatment in a particular hospital or by a particular physician (“volume-outcome” studies). In this paper, we examine the policy implications of using hospital and physician volume information as an “indicator” of quality in a rapidly changing health care environment with new players and new incentives. We begin by describing the evolution of the use of volumes within both regulatory and market-oriented contexts. We then discuss policy considerations and cautions in using volumes, along with suggestions for future research. Our purpose is to point out potential problems and clarify confusions about the use of volumes, so that policymakers and practitioners can be sensitive to the potential minefields they are traversing. © 1997 Elsevier Science Ltd.

Key words: Quality, volumes, outcomes, health policies, regulation.

INTRODUCTION

There is growing interest in the quality of health care and in using quality measures to direct patients to hospitals and providers offering high quality, low cost health care [1–3]. Although tremendous strides in the development of quality measures have been made, there is still little agreement on what constitutes “quality”, how it should be measured, and how quality information should be used. Some experts predict that it may be another 10 years before highly reliable and valid quality measures will be widely available [4]. The dilemma is that, while there is an increasing need for quality indicators as a result of a changing health care environment, this new environment has important implications for the use of these measures.

The purpose of this paper is to examine the policy implications of using hospital and physician volume information as an “indicator” of quality in a rapidly changing health care environment with new players and new incentives. We begin by describing the evolution of the use of volumes within both regulatory and market-oriented contexts. We then discuss policy considerations and cautions in using volumes, along with suggestions for future research.

Our purpose is to point out potential problems and clarify confusions about the use of volumes, so that policymakers and practitioners can be sensitive to the potential minefields they are traversing. There are a number of reasons why health care actors have gravitated towards the use of volume measures as quality proxies and as a means by which to achieve quality; however, we believe that their use requires much more nuance than they are generally given. The use of volume information is controversial, and there are important concerns and debates about: the causal relationship between volumes and outcomes, whether volumes can and should be used as a proxy for quality or a means to achieve quality, the separate role of volumes as a measure of statistical reliability, and the consequences of using volumes as a policy tool. Because volume information is being used for various purposes despite these controversies, the policy implications of its use need to be examined.

Since the 1970s, a growing body of research in the U.S. has addressed the empirical relationship between the number of patients with a specific diagnosis of surgical procedure and their outcomes after treatment in a particular hospital or by a particular physician.
"volume–outcome" studies. Virtually all of the studies examining the relationship between volumes and outcomes have been conducted in the U.S. (although a few have been conducted in Canada) [5]. One reason is that European countries routinely regulate the number of facilities that are allowed to perform certain procedures, and thus there are fewer institutions with low volumes. The ready availability of routinely collected data on procedures, diagnoses, and outcomes in the U.S. also facilitates such research [5].

Volumes have been used: (1) as a proxy for quality or a means by which to achieve higher quality because of the observed relationship between higher volumes and better outcomes, and (2) as a measure of reliability and the statistical confidence one can place in direct measures of quality. For brevity, we use the phrase "volumes as an indicator of quality" to refer to all of these uses unless we are discussing a specific use of volumes. Volumes have been used as an "indicator" of quality for over 40 years [6] and an article almost 10 years ago reported that there was increasing pressure to use data on volume, price, and quality to direct patients to specific hospitals [7]. Volume information is being used in the U.S. to set minimum volume standards for hospitals and physicians, as a tool to control diffusion of technology and facilities, to direct consumers and purchasers to higher volume providers and physicians, and as a justification for centralization. Therefore, the long history of using volumes may provide insights into the current and future issues in using quality indicators.

The use of volume information in a regulatory context emerged from a health care market in the U.S. where cost-based reimbursement was still the norm, hospitals were the center of health care delivery, and the "medical arms race" drove hospitals to compete for physicians and their patients by offering more services and possibly better quality, although the latter was rarely measured explicitly. Today, the driving market factors are a focus on consumer decision-making, the increasing dominance of managed care organizations, the consolidation of health plans and facilities, and the use of selective contracting through negotiations with purchasers. This changing environment introduces new complexities in using volume information, as well as other quality indicators, within both regulatory and market-oriented environments.

**HOW AND WHY VOLUME INFORMATION IS USED**

We begin by briefly reviewing how volume information has been used in regulatory and market contexts, and then the rationales that have been given for its use. The first study to document the volume–outcome relationship across a series of surgical procedures was published in 1979 [8]. Since that time, many other studies have found an association between higher volumes and better outcomes for a variety of procedures in the U.S., although the evidence of a relationship is strong for some types of procedures and weak for others (see [9] for a review). There is also evidence that the association between higher volumes and better outcomes holds at the individual surgeon level, e.g. [10]. Most volume–outcome studies have examined specific surgical procedures, e.g. coronary bypass graft surgery, and hospital rather than physician volumes.

The volume of procedures performed has been used in several ways: as a minimum threshold in guidelines for hospitals and physicians, as a tool in hospital regulation, as an indicator of quality for consumers and purchasers, as justification for centralization and a criterion in selection of sites, and as a criterion for selective contracting decisions.

First, professional groups and other organizations have set minimum volume criteria for hospitals and physicians. As early as the 1950s, minimum volume standards for births per maternity unit were proposed [6] and recommendations regarding minimal hospital volumes for open heart surgery were published as early as 1972 [11].

Second, regulatory approaches in the U.S. such as Certificate-of-Need, use volumes and evidence of a volume–outcome relationship as criteria for determining whether new hospital facilities and services will be approved, for example, whether new open heart surgery units will be allowed [12]. Certificate-of-Need (CON) regulation, which was one of the earliest attempts to regulate hospitals, grew out of the evidence that the "medical arms race" was leading to an oversupply of beds and facilities and higher costs: "a built bed is a filled bed" [13]. States in the U.S. started implementing CON laws in the late 1960s, culminating in U.S. legislation in 1974 mandating CON in all states (PL 93-641). Although the value of CON in constraining expansion and overall costs is questionable [14–17] there is some evidence that CON regulations which include the use of volume information can slow the diffusion of new services. For example, CON appears to have controlled the diffusion of cardiac services in Pennsylvania and Ohio compared to states with more limited or repealed CON regulations [18,19]. Although many states let their CON laws lapse or weakened them after U.S. federal legislation expired in 1986, as of 1995, CON or similar regulation was in force in 40 states although with wide variation in the types of services and facilities reviewed, dollar thresholds for capital review, and the stringency with which regulations are enforced [20]. Regulatory efforts are currently not in vogue, but there has been some renewed interest in CON approaches which incorporate budget constraints [21–24]. This newer form of regulation may increase the need for quality measures in order to balance quality and cost.

Third, another approach to using volume criteria is the dissemination of volume information to consumers and purchasers. Recommendations have been made to include the volumes of procedures in "report cards" for
evaluating the quality of care for AIDS patients [25] and cardiovascular procedures [26]. Consumers are being urged to choose hospitals and providers with higher volumes [27], e.g. a book for the popular press advises consumers to examine the number of procedures performed at hospitals they are considering [28]. The New York State Department of Health disseminated cardiac surgeons' caseloads and risk-adjusted mortality, which may have contributed to a reduction in the number of low-volume surgeons and in risk-adjusted mortality [29,30].

Fourth, volumes and their association with outcomes are also used as justification for centralization and as a criterion for selection of sites. Enthoven and Singer [31], in discussing how managed competition can be improved, state that spending on hospitals could be reduced and outcomes improved by concentrat- ing procedures such as open-heart surgery in efficient regional centers. They note, however, that closing departments and hospitals is proving to be difficult politically. Another recent example is provided by Jollis et al., who found that higher volumes of percutaneous transluminal coronary angioplasty (PTCA) were associated with better outcomes, stating that these data provide evidence in support of regionalization of PTCA procedures [32]. Examples of services which have been centralized in some areas include cardiac procedures, perinatal services, pediatric and adult intensive care, burn care, trauma care, and organ transplantation [33,34].

Lastly, there are indications that volume information is used for selective contracting of purchasers with providers. Selective contracting requires information on the quality and costs of care so that consumers and businesses can choose among health plans, and insurers and administrators can decide which hospitals to contract with. Although price often plays the key role in selective contracting negotiations, some health plans consider minimum levels of volume when contracting with delivery systems as a means to avoid poor quality providers [35] (R Miller, pers. commun., 7/1/96). As an example of the use of volumes for selective contracting, managed care networks look not only at price but also at volume, patient access, and survival rates; “and if you don’t meet their minimums, they won’t even ask you to the dance” [34]. As another example, U.S. Medicare reimbursement for organ transplantation is limited to hospitals performing at least 12 transplants annually [36,37].

Volumes have been used by health care actors as an “indicator” of quality for several reasons: the documented association between volumes and outcomes, their usefulness as a statistical indicator of the confidence one can place in direct measures of quality, the evidence of an association between volumes and costs, and expediency.

First, there has long been a documented association between volumes and outcomes [9]. Although many structural indicators of quality such as volumes [38,39] have fallen into disfavor because they are not always readily linkable to outcomes, a number of prior studies have demonstrated that low volumes are associated with worse outcomes. (While volume is not clearly as immutable an “indicator” of quality as is, for example, teaching hospital status, in Donabedian’s categorization volumes is most appropriately labeled as a structural measure.)

A second rationale for using volumes is that they serve as a measure of reliability, that is, as a statistical indicator of the confidence one can place in direct measures of quality. Low volumes imply statistical uncertainty when measuring outcomes using infrequently occurring “hard” endpoints such as mortality, whereas high volumes allow for the measurement of outcomes with less statistical ‘noise’. Although outcomes are a more direct measure of quality [40] the science of measuring and using outcome data is still in its infancy and hospitals or providers with low volumes may have too few cases to reliably measure quality by focusing on rare events, such as deaths [9]. Low volumes lead to wide confidence intervals, and it is statistically impossible to detect better than average performance when zero deaths is within the confidence interval surrounding “average” or expected outcomes [41].

The association of volumes and costs is another rationale for the use of volume information. In this era of cost containment, means by which to reduce costs that also maintain or increase quality are attractive. Although there is less research linking volumes and costs, it is commonly assumed that hospitals with higher volumes have lower costs, e.g. [34,42]. Economic theory predicts that hospitals with higher volumes will have lower costs if economies of scale are present. For example, it has been argued that regionalization of heart surgery would be cost-effective since average costs would fall as volume increased (due to spreading out fixed costs among more patients) [43]. It is also likely, to the extent that complications and deaths are both aspects of poor quality, that low-quality (perhaps low-volume) hospitals will be more costly per successfully treated patient because patients with complications generate more costs (the overall average cost may be misleading if deaths occur early in the stay and therefore lower the average cost per case). A few studies have suggested that hospitals with high volumes of heart surgery have lower charges or costs [44,45], lower lengths of stay [46,47], or shorter procedure times [48] than hospitals with low volumes, which may translate into lower costs.

Lastly, another reason why health care actors may use volumes is expediency: they are more readily obtainable than, for example, process measures that require extensive chart review. Volume data are also relatively easy to disseminate, e.g. regulators and purchasers of care can more easily understand data on numbers of procedures rather than more complex measures of deviations from expected mortality rates.
THE POLICY IMPLICATIONS OF USING VOLUMES IN A CHANGING HEALTH CARE ENVIRONMENT

We discuss three recommendations and considerations in using volume information in the context of health system changes:

(1) In the absence of other quality measures, low volumes can serve as an “indicator” of possible quality problems, particularly if research has demonstrated a relationship between volumes and quality. Low volumes, however, do not necessarily indicate low quality, and the changing health care environment introduces new uncertainties about the causal mechanisms linking volumes, outcomes, and prices.

When quality measures or clinical data to perform detailed risk-adjustments are not routinely available, for procedures exhibiting a volume-outcome relationship, it is more reasonable to assume that a low-volume hospital will have worse than average outcomes than it is to assume that all hospitals have equal outcomes. An example will illustrate. Picture what is likely to happen if you attempt to choose a high-quality obstetrician for prenatal care. First, you call potential sources of information about the quality of different obstetricians. The information you get is: (1) where they went to school and did their residencies (which may be a proxy for quality, but you know that some individuals do poorly even at the best schools and some excellent clinicians trained at less prestigious schools), (2) board certification (but you know that there is little evidence linking certification to outcomes, [4]) and (3) office hours (which may be a factor in your access and satisfaction but is not a measure of quality). Although you want an obstetrician who has good outcomes delivering high-risk babies, the particular risk you are concerned about is too rare to measure outcomes reliably for individual physicians so there is no available information. In desperation, you do what other people do: you choose a doctor and hospital that deliver a lot of babies. You assume that, given the lack of better information, you will get better care from high volume providers and facilities.

There are several concerns, however, about the use of volumes: low volumes cannot be used as an overall “indicator” of poor quality; the causal mechanisms linking volumes, outcomes, and prices are unclear; and selective contracting changes the process by which hospital selection is made.

One caution in using volumes as “indicators” of quality is that studies of the association between volumes and outcomes examine patterns across many hospitals, but the inference may not be true for individual hospitals or providers. There can be a number of reasons, other than poor quality, why specific hospitals or providers may have low volumes, such as the start-up of new services, rural location, or a procedure performed by a high-volume surgeon in several low-volume hospitals. Furthermore, hospitals may have high volumes and quality for some procedures but low volumes and quality for others, or volumes and quality may fluctuate over time. Therefore, low volumes cannot be used as an overall “indicator” of poor quality, volume standards will vary by procedure and disease, and it will be useful to have multiple measures and longitudinal data. However, in the past one could reasonably say that in the absence of other quality measures, one would probably have a higher likelihood of better outcomes with a high-volume provider than a low-volume provider.

A key concern about using volumes as “indicators” of quality in today’s health care environment is that the causal mechanisms linking volumes, outcomes, and prices are unclear. One of the controversial issues about volumes is whether an observed association between higher volumes and better outcomes is a result of more experience leading to better outcomes, the “Practice-Makes-Perfect” hypothesis, or whether patients are attracted to hospitals with better outcomes, thereby increasing their volumes, the “Selective-Referral” hypothesis [9]. That is, if increased volume is sufficient to achieve better outcomes, then increases in volumes at selected sites will improve outcomes (ignoring other problems such as access to care). However, if the observed volume-outcome relationship reflects selective referrals to better quality providers, regulatory and other “steering” strategies would need at least as effective indicators of quality and means to choose the best providers. There is some evidence that both hypotheses may be true to varying degrees for different procedures [49].

The introduction of selective contracting with health care providers by third-party payers dramatically changes the process by which hospital selection is made. In physician-dominated markets under the traditional model of competition, hospitals competed for patients largely through efforts to attract physicians by providing high technology services, which tends to increase the number of low-volume hospitals and increase costs. In insurer-dominated markets under the model of selective contracting, price is a more important factor [35,50] although there are some indications that measures of quality such as outcomes and infrastructure are also being used for contracting decisions [51].

To the extent that higher volumes are associated with lower costs and prices, these newer forms of competition based on price may result in higher quality and lower costs. Studies in California, which in 1982 was the first state in the U.S. to enact legislation to encourage selective contracting, have found that the introduction of selective contracting led to a decrease in costs, particularly in high-competition markets [50]. However, since high-quality (and perhaps high-volume) hospitals and providers could have higher prices or charges, selective contracting which emphasizes price without consideration of quality may result in a reduction in quality. That is, in a managed care environment where contracting decisions are often based primarily on price, hospitals may have higher volumes...
Policy implications of using volumes

not because they have higher quality but because they have lower prices. In fact, some low-volume hospitals, either believing that increased volume alone could improve quality or lower costs or that low-volumes might exclude them from future competitive bids, might have submitted very low bids in order to "buy market share". Unless the higher volume which resulted did lead to better outcomes, this could lead to worse outcomes overall.

An analogy may clarify why high volumes can give misleading signals to consumers and purchasers about quality. You are trying to choose a restaurant for dinner in a town in which you have no restaurant reviews or recommendations, so you look for restaurants with long lines of people waiting in the hope that this signals high quality. On the one hand, if higher volumes lead to better food because the kitchen becomes more experienced and reliable, then using a long line as an indicator of quality is likely to help you choose a better quality restaurant. On the other hand, if the long lines are the result of tour operators who have chosen the restaurant because of its low prices, they may provide little information about the quality of the restaurant.

Our discussion suggests that it is quite a leap from identifying quality indicators to using those indicators as criteria in selecting facilities or providers. The process by which contracting decisions are made is still largely a "black box" due to its nature, e.g. health plans may not want to publicize their selection criteria and negotiating strategies. It is therefore difficult to determine how quality indicators such as volumes are being used and the consequences. More research in this area is needed, possibly using in-depth interviews rather than published reports.

(2) The use of strict volume thresholds and the centralization of services as a result of the perceived problem of low-volume facilities may have unintended consequences that need to be anticipated and addressed.

Although volumes are a tool that can be used to achieve some regulatory or market goals, they are a relatively blunt tool that may distort the market, provide perverse incentives, and incur large transaction costs. In particular, it is more problematic to set strict volume thresholds as the criterion for regulatory approval, purchasing decisions, or an acceptable quality rating in "report cards", than it is to give weight to higher volumes as one of several criteria. Few studies have examined the functional form of the volume-outcome relationship and therefore a strict threshold for minimum volumes may be inappropriate. For example, although studies of coronary angioplasty (PTCA) have found an inverse relationship between volumes and outcomes, it is not certain whether the current recommended minimum volume of 200 procedures is the most appropriate threshold. Incorporating minimum volume thresholds in CON regulations can keep out of the market new entrants which may be able to provide services more efficiently. Thresholds can also give hospitals an incentive to raise volumes by relaxing standards of appropriateness if their license could be revoked for volumes below the minimum. (While many CON laws have implied such penalties, we know of no instances of its being enforced.) The regulation of volumes can also contribute to large information and administrative costs.

The centralization of services because of the perceived problem of low volume facilities may also have unintended consequences. Some efforts to centralize or regionalize services have been successful. However, centralization efforts have often focused on services which are less desired by hospitals because of factors such as low reimbursement rates, e.g. trauma care. In comparison, cardiac procedures such as CABG and PTCA are desired by hospitals because they increase hospitals' prestige and profit potential, and therefore centralization efforts have been less successful. Although centralization is often advocated as a solution to poor outcomes and high costs, the relative lack of successful centralization suggests that such policies face a number of difficulties and need to consider several caveats. Concentrating procedures in fewer hospitals may require longer travel distances and increase the need to transport patients, which could increase the chances of complications. An appropriately regionalized system might have to tolerate low volumes in isolated rural areas, particularly for emergency procedures although the benefits of expedient handling of emergencies have to be weighed against possible decrements in quality for the elective cases treated at those facilities. Policies to increase volumes will vary not only by geographic area but also by procedure, since redirection of some types of patients would require large shifts with little potential improvement in outcomes. For some procedures and locations, however, closing low-volume hospitals may have little effect on travel distances. For example, eliminating hospitals with very low volumes of CABG (<100 cases annually) in California would increase travel distances to a CABG hospital for a small number of residents only slightly. Lastly, it is difficult to determine the impact of centralization on outcomes and quality of care, and existing evidence is mixed. For example, although volume and level-of-care standards for perinatal care have been advocated since the 1930s, observational and controlled studies have not demonstrated that centralization per se improves outcomes.

The trends toward managed care and selective contracting may lead to fewer providers and increased volumes, negating the need for direct centralization efforts. In an increasingly difficult market, hospitals are likely to enter into agreements with competitors to centralize services. Selective referrals to centers of excellence may increase as managed care organizations increase economic incentives to use regional centers and as they collect and disseminate data on quality. Also, if market concentration continues to increase as expected,
organizations can direct more patients to their panel of hospitals, and those hospitals may have increased patient volumes. HMOs may be more likely to direct their patients to high-volume providers, so that increased managed care market share in an area may decrease the number of low volume providers [64,65]. In these scenarios, quality of care could improve to the extent to which increased hospital volumes lead to improved outcomes [66].

Unfortunately, most studies of centralization predate the aggressive use of selective contracting. Selective contracting requires that a range of options be available in a given market and hospitals must have occupancy rates that are low enough to allow them to accept additional patients [50]. The research on the success of selective contracting in reducing costs has been conducted in California, which historically has had a large number of low-volume facilities and excess capacity. If low-volume providers and facilities are driven out of the market either through regulation or market forces, the result could be decreased competition, increased prices, and possibly lower quality.

(3) The audience and purpose has to be considered in assessing whether volume information will be useful.

The competitive-market model assumes that information will be used without the “stick” of regulation. Little is known, however, about whether or how providing information on volumes or other measures of quality changes consumers’ or purchasers’ decisions or what type of information is most useful [1,67]. As yet there is little evidence on whether consumers will use information on the technical aspects of quality such as volumes, since consumers may focus on interpersonal rather than technical aspects of quality [68]. In the U.S., information disseminated by state health data organizations (HDOs) which were created in the mid-1980s to facilitate “buy right” approaches by providing information on the price and quality of medical procedures [69,70], has been used less by the avowed target, the consumer, and more by hospitals and insurance companies [69]. Health data organizations have also become embedded in the same political- and interest-group politics that have marked direct regulation, and there is little evidence yet that the information provided necessarily leads to better control of health care costs [69]. Lastly, recommended volume standards by definition are voluntary, and overall there is little evidence that these standards have decreased the number of facilities or providers with low volumes or improved outcomes [71].

The experience with volumes suggests that simply producing and disseminating quality information is unlikely to improve quality or reduce costs, and different types of quality measures are necessary for different target audiences and for different purposes. In the “new” health care market in the U.S. where the choice of providers is commonly restricted by health plans, the primary choice for consumers becomes the health plan.

In turn, the plans choose the providers and facilities that will be covered, thus a key issue is whether plans will have incentives to choose high-quality providers. Similarly, different indicators and dissemination approaches are necessary for measures used for internal quality improvement versus those intended for patient consumption. Therefore, there is no one perfect measure of quality for all situations. Quality measures targeted to consumers might focus on interpersonal aspects of quality, while measures targeted to purchasers might focus more on technical aspects, including volumes. Since quality rankings often differ depending on which measure is used, it is prudent to combine information from several measures of quality and to examine trends over time [39,72,73].

Approaches to measuring quality will vary depending on whether the quality of hospitals, providers, or health plans is being measured. For example, most of the work on volume-outcomes has been done on hospital volumes, while physician volume has been less studied and the policy options less explored. Physician volume and quality can be harder to study since some procedures, such as CABG, are done by teams and therefore it is more difficult to identify individuals’ contributions to outcomes. Furthermore, the problems of statistical reliability become even more daunting at the physician level.

CONCLUSION

The need for quality measures within both regulatory and market-oriented contexts is likely to continue unabated, although market changes are outpacing the technology currently available to measure quality. Volumes may continue to be useful as a statistical indicator of uncertainty, but there are a number of caveats about its use as a proxy for quality, particularly as more direct quality measures become available. Nonetheless, the history of using volumes may provide insights into current and future issues in using other quality indicators.

REFERENCES

5. Banta, H., Engel, G. and Schersten, T., Volume and
Policy implications of using volumes


42. Kirklin, J., Akins, C., Blackstone, E. and Booth, D.,

43. Finkler, S., Cost-effectiveness of regionalization, the heart surgery example. Inquiry 1979; 16(Fall 1979): 264–270.


70. McClure, W., Buying right, the consequence of glut. Business and Health 1985; September: 43–46.

