It has been repeatedly shown that higher procedure volumes, by hospital and by physician, are associated with better outcomes. Buttressed by large-scale selective service purchasing, surgical care for many volume-sensitive operations has been regionalized. However, the implications of outcome disparity data for obtaining valid patient consent remain uncertain. When the first large scale outcome study appeared showing empirically that outcomes are worse when performed by lower-volume providers, bioethicists promptly insisted that such information was material to a reasonable patient’s decision whether and where to have a volume-sensitive operation and that surgeons at low-volume hospitals should disclose it. More recently, two surgical oncologists have reiterated that argument, most especially when patients are making decisions about pancreatic or esophageal resections. This proposal tantalizingly appeals to the concept of patient empowerment, supposedly showing appropriate respect for the patient’s interest in self-determination by having his surgeon (or physician), rather than others, outline for him personally the risks and benefits associated with surgical care delivered at different hospitals. But on the contrary, a surgeon’s conducting a truthful, non-misleading, non-confusing informed consent discussion of statistical outcome disparities in the relentlessly shrinking time typically allowed for this conversation is unrealistic as a general requirement. The traditional approach to informed consent is simpler, less fraught and preferable. By law, a surgeon who is licensed to practice independently and who evidences willingness to examine and to offer an operation to a patient conveys (1) an implicit standard of care assurance, and (2) a fiduciary assurance. In other words, it goes without saying that the surgeon holds it out to the patient that s/he possesses the training and skill necessary to perform the offered service with reasonable skill and safety as measured by the applicable standard of care; that s/he will act in good faith and use his/her best medical judgment on the patient’s behalf. Liability attaches when patient harm results from a surgeon’s failure on either count. It also goes without saying that the surgeon extends similar assurances for care provided by trainees who are under his or her direct supervisory authority and control. The traditional theory of informed consent forefosts requiring desultory discussions of volume-outcome disparities and will be defended here.


**INTRODUCTION: PRACTICE MAKES PERFECT?—QUESTIONS**

It is a commonplace in medicine that care quality tends to improve with greater provider experience [1]. How much experience should be required to claim competence? Which numbers tell the story? For example, ERCP is technically difficult. The learning curve for individuals varies remarkably. And it is recognized that some trainees are untrainable. Therefore, it has been proposed that an 80% cannulation rate of the desired duct be demonstrated, before a minimally invasive gastroenterologist should claim and be recognized for competence. Estimates for the number of procedures generally required to attain this level of skill varies from 35 to 180–200 [2].

Clinicians who have placed more than 50 CVCs (central venous catheters) have less than half the complication rate of clinicians with fewer than 50 attempts. With reduced work hours, what percentage of trainees place more than 50 central lines [3]? When obtaining consent for line-placement, should an ICU attending insist that his or her residents disclose how many CVCs they have placed and the associated complication rate? Should the trainee’s “zero previous experience” be disclosed for informed consent purposes prior to attempting his or her first central line placement in a living patient [4]? Should medical and surgical attendings discuss with patients hospitalized in July and August seasonal variation in outcomes associated with the educational cycle of trainees?

Englesbe et al. [5] found a 41% mortality increase among July–August surgical patients compared with April–June patients. However, Raval et al. [6] found an increase in morbidity but a slight decrease in mortality for surgeries involving residents. In a systematic review, Young et al. found that mortality increases and efficiency decreases in hospitals because of year-end changeovers. They concluded that the existing literature does not permit firm conclusions [7].

Nevertheless, does informed consent “reasonableness” require telling patients about the unsettled state of the literature and then invite them to decide how to evaluate and apply the information to their own situations? Is that consistent with fiduciary responsibility? Should a patient who has sought care at a teaching hospital nevertheless have the privilege to demand that the attending (and not a resident) do his...
operation or place his central line? What if the attending who teaches line placement hasn’t herself placed one in several years? Is it fair for patients at academic medical centers to free-ride, to demand the highest level of skill that medical education has produced but refuse to support the medical education system by participating in it? Should a surgeon be required to discuss the morbidity and mortality statistics of his or her own hospital for any proposed volume-sensitive procedure and compare the numbers with outcomes at other (teaching and non-teaching) hospitals in town or reasonably nearby if there is any published data to cite when doing so? But if so, why stop there? Surgical tourism evidences some patients’ willingness to travel in pursuit of better outcomes at lower prices. Shouldn’t disclosing surgical options empower those interests? And beyond site of care, should surgeons initiate discussion of their own personal outcome statistics for volume-sensitive procedures, comparing them with those of other surgeons, near and far [4,8,9]? No consensus has emerged about this [10–12].

THE ERA OF BIG DATA BEGINS

Until the study by Luft et al. [13] it had not been previously demonstrated empirically how much experience is required to obtain a noteworthy benefit—a measurable reduction mortality for surgical inpatients. Based on a review of average results for 1 year (1974 and 1975), focusing on data for 12 operations at 1,498 hospitals, the authors found that mortality for open heart surgery, vascular surgery, transurethral prostatectomy, and coronary bypass was inversely related to hospital volume. In hospitals where 200 or more of these operations were done annually, death rates were 25–41% lower than in hospitals with lower volumes.

However, for some operations (e.g., hip replacement), in hospitals where 50 to 100 operations were done, the same mortality resulted as where 200 or more were done. Indeed, the mortality/volume curve for some procedures tended to flatten out at a relatively low volume (between 10 and 50 operations/year). These included cholecystectomy, total hip replacement, resection and graft for abdominal aortic aneurysm, vagotomy, and/or pyloroplasty for ulcer of the duodenum, and cholecystectomy with bile duct incision.

The study did not enable identifying individual hospitals in any way. Nor were the data sufficiently fine-grained to discern an “experience effect” for individual surgeons, anesthesiologists, operating room teams, and nurses. After all, it is the surgeons, anesthesiologists, and their support teams who work in hospitals (not hospitals), that actually provide surgical care to patients.

The authors acknowledged significant study limitations: they could not define precisely the optimal volume for each operation, nor could they specify a defensible minimum at the hospital level. Nevertheless, they recommended a staged-process of policy development to encourage regionalizing operations whose outcomes were “markedly less satisfactory” in low-volume hospitals. They envisioned practically meaningful results—lives saved and reduced costs.

The Luft paper launched a substantial literature on the volume-outcome relationship. It also prompted regionalization for certain operative procedures [14–41]. Regionalization of adenectomies to high volume centers has been documented and was associated with reduced inpatient mortality [15]. Studies have examined the volume-outcome effect for non-oncologic as well as oncologic operations. The outcome and cost benefits of regionalizing oncologic resection of the pancreas and esophasus has been repeatedly shown. These latter are also the only oncologic surgeries for which the Agency for Healthcare Research and Quality lists hospital volume as a measure in its Inpatient Quality Indicators [42]. However, regionalization has not always been associated with improved outcomes. Colavita and colleagues found that regionalizing ventral hernia repair was not associated with reduced mortality, while costs actually increased [43]. Zemlyak et al. [41] found no regionalization for non-trauma related splenectomy so no basis can be claimed for a benefit in doing so.

VOLUME-SENSITIVE OPERATIONS AND INFORMED CONSENT

Ethicists Culver and Gert (1980) found the Luft study so compelling that they claimed physicians at low volume hospitals henceforth had a moral obligation to disclose to all candidates for volume-sensitive operations that they would be at somewhat greater risk for mortality than they might be at a high-volume hospital. Why? Because, they claimed, the reasonable patient standard for measuring disclosure adequacy would require it: “surely any rational person who is advised to undergo a major operation...would want to know that traveling 50 miles (or five blocks) might lessen his or her chance of dying during or after surgery.” A physician’s withholding such information would be a culpable omission—“self-serving at worst and unjustifiably paternalistic at best” [44]. The ethicists’ signal hedge was “might,” which of course implies “might not.”

Paternalistic Surgeons?

The ethicists concern that surgeons might manipulate information and patients in questionable ways was not unfounded. The medico-legal literature is replete with cases of surgeons withholding intentionally information about the seriousness of a proposed operation, misrepresenting its urgency and its attendant risks. Surgeons as a whole have never embraced discussing the comparative extent of their own training, experience with a proposed operation and the skill level of and the role to be played by their assistants [45].

In the famous case of Pratt v Davis (1905), the surgeon, who proposed curing a patient’s epilepsy by a gynecologic operation, testified how he bamboozled his patient, getting her to go to surgery voluntarily: “I thought...that her mental condition was such that she could be influenced by what I told her to the extent of walking to the operating room instead of being carried. I worked her deliberately, systematically, taking chances which she did not realize the full aspect of, deliberately and calmly deceived the woman. That is, I did not tell her the whole truth” [46].

And, only 25 years prior to Culver and Gert warning of surgeons’ paternalism, North Carolina’s Supreme Court (Hunt v Bradshaw) had vindicated a surgeon’s downplaying an operation’s risks as follows: “It is understandable that the surgeon wanted to reassure the patient so that he would not go to the operating room unduly apprehensive. Failure to explain the risk involved, therefore, may be considered a mistake on the part of the surgeon, but under the facts cannot be deemed such want of ordinary care as to import liability” [47].

However, the ethicists were unrealistic to expect that physicians would, on the basis of the Luft study, suddenly and straightaway amend the content of their informed consent discussions to disclose and explain the statistical importance of volume-outcome disparities. Indeed, studies of “clinical inertia” have found that physicians resist changing their behavior, for example, to intensify medical therapy, even when the evidence is unambiguous that their doing so will be patient-beneficial [48].

Disclosing outcome disparities may also be self-serving or paternalistic, for example, suppose a surgeon at a low volume hospital recognized that the Luft study would prompt increasing scrutiny of outcomes and the prospect of economic profiling. From jealous regard of his own record for good outcomes, mightn’t he adjust downward his risk-tolerance, tempted to use outcome disparities data to persuade frail elderly patients with multiple co-morbidities to have their pancreatic resections at a high volume center [49]. Alternatively, denouncing outcome disparities might prompt the surgeon’s leaving the low-volume hospital for a high volume academic center—better for
him perhaps, although he might wind up doing fewer operations personally, but it would have an unfortunate consequence for his community hospital’s hopes to build a meritorious surgical program.

Biased framing of medical statistics has been well-documented [50–52]. Halvorsen and colleagues, found that treatment effects expressed in terms of NNT yielded higher consent rates than did those expressed as equivalent postponements, suggesting that anticipated outcome-framing may influence the patient’s willingness to accept a recommended intervention [53]. Young and Oppenheimer found that semantic characterizations of risk (e.g., “at increased risk”) are associated with patients’ systematically overestimating risk. The effect was so strong that “[t]he decision to present semantic versus probabilistic information is tantamount to a decision about whether to encourage risk acceptance versus risk avoidance” [54,55]. Disclosing and as well as omitting to disclose outcome disparities may be self-serving, paternalistic and disserve patient self-determination.

Professional resistance to changing customs alone would predict that primary care physicians (PCPs), oncologists, and surgical oncologist would not embrace the ethicists’ argument for incorporating a discussion of outcome statistics in the informed consent process. And with one notable exception (Johnson v Kokemoor, discussed below) [56], courts have refused to require that physicians discuss comparative outcome statistics—by hospital or by surgeon. One State’s Supreme Court has refused to see an informed consent violation even when the surgeon had frankly lied about his experience with oncologic resection of the esophagus [57].

Recently, Housri and Koniaris [10] have reuscitated the Culver and Gert proposal, also on ethical grounds. They argue that, outcome disparity data is increasingly available, and from multiple sources. They cite one study finding that 25% of patients were aware of online hospital ratings by Leapfrog or Kaiser [58]. Few seemed to be basing their decisions about site of care on such evidence, however. Outcome data is often misinterpreted or presented misleadingly, sometimes for medical center marketing purposes. So, Housri and Koniaris argue, physicians should take responsibility for providing a de-biased presentation of outcome disparity data during the informed consent process when the differences are consistent and “substantial.” They do not define “substantial” but they do say that it clearly includes oncologic resections of the esophagus and pancreas.

The Question Again: Should Physicians Be Required to Disclose and Discuss Outcome Disparities When Obtaining Consent?

When Culver and Gert weighed in on the matter, standards for assessing a physician’s disclosure for adequacy included:

- the customary, reasonable practitioner standard, ± a locality rule. This has been the most commonly recognized standard and remains in force in approximately half of jurisdictions in the United States [59–63].
- The material risk or “reasonable patient” standard which measures disclosure adequacy regarding treatment or procedure choices by asking what a reasonable patient would want to know in order to make an intelligent decision in the circumstances rather than by what (supposedly) paternalistic doctors tend to disclose [64,65].
- Finally, there are hybrid statutory standards that include elements of reasonable physician and reasonable patient, ± a locality provision [66].

Whatever their nuances, all the standards have been procedure-focused, not provider-focused. None require discussion of disparities among providers.

Were Culver and Gert basically correct 1980? Are the arguments of Housri and Koniaris valid today? I think the answers are “no” and “no.” Culver and Gert over-interpreted the Luft paper. The data were not risk-stratified, so could not provide useful guidance to individuals. And despite their laudable goal of dispensing de-biased statistical information to patients, there are multiple reasons to reject Housri and Koniaris proposal to make it an informed consent duty.

Understanding statistical concepts and critiquing statistics-based arguments is difficult [67–69]—for example, that significant association is not causation, that the positive predictive value of a positive test may be low, how to interpret an odds ratio, that studies discussed on TV news shows may be worthless, what number needed to treat and intent to treat mean, the difference between reductions in relative risk and absolute risk and why large reductions in relative risk may not be clinically meaningful, the limitation of retrospective studies, even very large ones, and so on.

Physicians themselves struggle with such statistics [70]. Studies of physician communication reveal widespread use of vague, undefined qualitative terms [71]. Even though surgeons are the most knowledgeable experts about the studies in their field, their ability to communicate that knowledge to laypersons is highly variable.

A doctor’s well-meaning effort to de-bias a media report “study shows that you are best advised to have your operation at a high-volume center” risks introducing new biases and different, but not necessarily more benign misunderstandings. And the effort may be infected with non-eliminable bias, and perhaps self-serving or paternalistic factors [72].

Producing a non-misleading summary of statistical arguments that enhances rather details the decision making of a fearful layperson who needs major surgery and who may labor under twin disabilities of innumeracy and low health literacy compounds the difficulties—well beyond what it is reasonable to expect a surgeon to resolve during an informed consent discussion, an information exchange that gets ever briefer. Fitten and Waite found that sicker patients are at greater risk for misunderstanding the complexities of risk-communications than healthier patients [73].

MISINTERPRETING OUTCOME STUDIES: THE ECOLOGICAL FALLACY

The Luft study averaged inpatient mortality and compared it with expected mortality among hospitals performing a variety of operations at various volumes—an analysis of an average of averages. It was not possible to identify an individual hospital. A staff physician at a low-volume community hospital could not know whether mortality data from his hospital was even included in the Study’s analysis. He could not know whether the mortality/experience curve at his own hospital flattened at remarkably low volumes for any specific volume-sensitive operation.

Were a physician to have inferred that mortality at his own hospital must be “about average” among low volume hospitals, he would have reasoned fallaciously. To infer a conclusion about an individual (e.g., a hospital or a patient) solely on the basis of an analysis of group data commits “the ecological fallacy.”

Now suppose a hypothetical physician, per the insistence of Culver and Gert, were to disclose to a candidate for a volume-sensitive operation: “patients who have their volume-sensitive operations done at low volume hospitals increase somewhat their risk of death.” It would be natural for our hypothetical patient to reason as follows: “The doctor tells me that the operation I need is ‘volume-sensitive.’ Therefore, I will reduce somewhat my risk of in-patient death if I have my operation at the high-volume hospital 50 miles away or at the one across town.”

This reasoning would have ignored that some “volume sensitive” operations demonstrated the “experience effect” at comparatively low volumes. The surgeon’s hypothetical disclosure omitted that. The patient’s reasoning also commits the ecological fallacy, inferring his own non-stratified risk and how to reduce it from group-based data. The Luft study was not a “travel effect” study. It did not randomize risk-matched patients who “stayed home” at their low volume hospital with...
patients who traveled to high-volume centers for their volume-sensitive surgeries.

Clearly, Luft et al. aspired to supplement discussions whether or not to have a surgery with some guidance about where patients should have volume-sensitive operations. “The guiding principle should be: the greater the excess risk of low-volume surgery, the greater distance one should be willing to travel to a hospital or surgeon with high volumes and good outcomes.”

However, their principle wouldn’t help a patient’s deciding with whom he should go to surgery. Suppose a community primary care physician (somehow) knew that, within some time frame, (say, 5 years ago) a study had found greater than expected inpatient mortality for the Whipple procedure at his own low-volume hospital. It could nevertheless be true, as Luft et al. recognize, that the good outcomes of one fellowship trained hepatobiliary surgeon on staff were swamped by the predominately worse outcomes of several “occasional” surgeons. Our community physician might know by reputation this one well-trained surgeon and refer his patient on that basis—for a volume-sensitive operation at a low volume hospital. A trust-based referral based on local knowledge.

On the other hand, suppose our community physician merely referred the patient to a high-volume hospital. This could not preclude his being served by one of several “occasional” surgeons on staff whose failures were masked by the aggregate statistics of their more numerous, fellowship-trained colleagues. Disclosing hospital outcome disparity data in an informed consent discussion would not have promoted better decision making for an individual patient who wanted to know not only whether to have surgery, when to have it, and where to have it but most importantly, with whom.

The Luft study data set was very large (as are many subsequent outcome studies) but the data came from only 1 year and was almost 5 years old when the paper appeared in the New England Journal. In other words, it was a “freeze-frame” in a still running video. It focused exclusively on inpatient mortality, did not consider longer-term mortality (at 30 and 90 days). It ignored complications. Importantly, it ignored “process measures,” which are associated with quality care and off-set, to some extent, cherry-picking incentives that outcome measures encourage.

The study justly became a classic for providing a numerical elaboration of the “experience effect” for several operations and for recommending regionalization. The study is not commonly cited for the hedge in its final sentence hedged, “…it is possible that large segments of surgical practice may not require change.” But surgical training and practice changed, and for many reasons. Increasingly, surgeons go to fellowship, to enter practice at a community hospitals as, say, hepatobiliary specialists able and willing to build meritorious programs.

Since the Luft study appeared, outcome measures have been pushed beyond inpatients, (a measure that created an incentive to discharge patients quicker) to include 30-day and 90-day mortality. Additional measures include planned and unplanned readmissions and return to surgery. Kheri et al. [74] found an independent inverse association between post-op complications and reduced long-term survival. Complications are associated with prolonged hospitalization. Patients who endure a complicated post-op course have reduced quality of life for many years when compared with patients who escape such complications [75]. There is growing interest in measuring care quality by disability-free survival at 1 year [76].

Increasingly, outcome data are enabling individualized scrutiny of providers [77]. The Society of Cardiothoracic Surgeons in the UK, relying on a decade of data, has begun publishing risk-adjusted outcomes—surgeon-by-surgeon. In the name of “transparency,” the National Health Service recently has decided to do likewise for eight other surgical specialties [78].

Tools for better risk-stratifying patients have also been introduced. The American College of Surgeons has published online for free-use a Surgical Risk Calculator with a data base of 2.5 million surgical patients [79]. Other calculators stratify risk based on frailty.

Lack of Informed Consent or Frank Lies & Fraud? Johnson v Kokemoor (1996) [56].

A patient presented to her family physician complaining of headache. He diagnosed her with a bifurcation aneurysm in the posterior circulation of her brain. He referred her to a local neurosurgeon who had done 30 aneurysm repairs during residency—all in the anterior circulation. He never operated on a posterior-circulation aneurysm during his residency. He had done so on a bifurcation aneurysm only twice since entering practice. He had never operated on one as large as the patient’s. He offered to clip her aneurysm.

She pressed him about the extent of his experience with the proposed operation. He responded reassuringly, saying that he had done it “lots of times, dozens of times.” In bad faith and betraying his implicit commitment to use his best judgment on her behalf, Kokemoor omitted explaining that the risks of comparatively inexperienced surgeons performing the operation were increased by more than an order of magnitude. Despite that the operation is among the most difficult in neurosurgery, Kokemoor lied when saying that the operation was less risky than the pre-op angiogram she would be having, its seriousness comparable to that of tonsillectomies, appendectomies and gall bladder surgeries.

When Mrs. Johnson asked about the risk of delay, he said the risk of rupture was 2%/year, cumulative. She consented. The surgeon clipped the aneurysm successfully, performing the operation within the applicable standard of care. However, the complication was devastating. Mrs. Johnson wound up an incomplete quadriplegic. She sued him for lack of informed consent.

At the time of the events in the case (1990), Wisconsin’s informed consent statute required the physician’s disclosures to relate to the proposed treatment or procedure and alternatives, including no operation, along with the risks and benefits of each option. A disclosure’s adequacy was to be measured by “material risk,” what a reasonable patient would want to know in order to make an intelligent decision. The statute said nothing about “provider-specific” disclosures. No jurisdiction had ever required a physician to disclose the extent of his own training, his experience with the planned operation, nor that of his hospital. Whether it was permissible to give frankly dishonest and misleading answers to a patient’s direct questions had not been tested in Wisconsin.

Nevertheless, Wisconsin’s Supreme Court gave the statute an unprecedented interpretation, holding that, in this particular case, information regarding a physician’s experience in performing a particular procedure, a physician’s risk statistics as compared with those of other physicians who perform that procedure, and the availability of other centers and physicians better able to perform that procedure would have facilitated the plaintiff’s awareness of “all of the viable alternatives” available to her and thereby aided her exercise of informed consent.

The plaintiff’s expert, a highly accomplished neurosurgeon who practiced at a high-volume center, testified that he routinely disclosed his own morbidity and mortality rate for the operation. He did not claim that his practice was a professional custom among neurosurgeons.

The Court’s imposing an after-the-fact duty to disclose comparative outcome disparities, was well-aligned with the spirit of what Culver and Gert had advocated in 1980. But the Court went substantially beyond their demand. The Johnson court insisted that Kokemoor disclose comparative risk statistics for inexperienced versus experienced surgeons (implying that he should have specifically included himself in that category) and should have reported his own outcome statistics for the operation.

It apparently didn’t occur to the court that the outcome statistics of a surgeon who had recently entered practice would not be statistically meaningful. Disclosing small numbers would have further misled the
Outcome Disparities and Informed Consent

REFERENCES


12. Treatment effects expressed in terms of NNT yielded higher consent rates than did those expressed as equivalent postponements. This result suggests that the description of the anticipated outcome may influence the patient’s willingness to accept a recommended intervention.


9. In forming of actual percentage rate of adverse events was associated with less fear about adverse events and greater intent to comply with prescribed regimens. Using verbal descriptors to disclose adverse event risk information was associated with less intent to comply.


4. The T. J. Hooper, 60 F. 2d 737 (2d Cir. 1932).


79. Surgical risk calculator, ACS/NSQIP. Available at: http://riskcalculator.facs.org/
83. Shadrick v Coker, 963 S.W.2d 726, 1998.
84. 18 U.S.C. 1346.