2

Cutting Surgical Practices at the Joints: Individuating and Assessing Surgical Procedures

Alex John London

Unlike pharmaceuticals or medical devices, the development and introduction of new surgical procedures often occurs with relatively little formal oversight or regulation. In some cases, surgical procedures evolve gradually over time as practitioners adapt to new challenges, develop new techniques, or integrate new technologies into clinical practice.[1] In other cases, revolutionary new techniques and technologies are introduced directly into clinical practice without rigorous evaluation in controlled clinical trials and without the formal oversight of an institutional review board (IRB) or similar mechanism.[2] These discrepancies between innovation in surgery and other areas of medicine have been underscored in recent years by the successful completion of some notable clinical trials involving surgical procedures.[3] Nevertheless, the prospect of imposing more formal mechanisms for evaluation and oversight on surgical innovation has been met with widely divergent responses.[4] On a sociological level, it may be that these discrepancies, and some of the reaction to them, reflect the relative autonomy of the surgical profession and the extent to which independence and individual initiative are prized elements of its distinctive professional culture.[5] However inter-

esting such hypotheses may be, though, they should not obscure some of the deeper and more fundamental conceptual issues that are at work here. In particular, at the heart of these divergent attitudes lies a basic conflict over whether and to what extent such proposed reforms are conceptually and practically feasible and whether their imposition would produce more harm than good.

The degree of variation endemic to the surgical context plays a crucial role in disagreements over the feasibility and desirability of implementing more formal methods for evaluating or regulating surgical innovation. In particular, concepts such as innovation, experimentation, standard treatment, and standard of care depend for their meaningful application on our ability to make discriminations of sameness and difference within the relevant domain.[6] If the degree of variation endemic to surgery makes it more difficult to individuate and re-identify surgical procedures than drug therapies, then it may be similarly difficult to meaningfully distinguish innovative procedures from standard interventions. This may make it more difficult to determine which types of activity require more formal oversight, and when, if ever, controlled clinical trials are feasible or appropriate. Additionally, if controlled clinical trials require a greater degree of uniformity and standardization than can feasibly be achieved in the surgical context, it is reasonable to resist requiring such methods to validate innovations.[7] If, in contrast, these demands for regularity and uniformity fall within feasible bounds, then resistance to the use of clinical trials to evaluate innovative therapies might rightly be viewed as unreasonable and imprudent.[8] In either case, one's attitude toward such practical questions about whether and how to evaluate or test surgical innovations depends heavily on one's conceptual model of surgery and surgical practice.

The discussion that follows begins by articulating three versions of a position that I refer to as "surgical exceptionalism." Strong, moderate, and weak formulations of this position claim, in different ways, that the exceptional ethical or regulatory status of surgery is justified by the unique nature of surgery and the complexities of surgical practice. In the second section, I argue that disagreements between the surgical

exceptionalist and the reform-minded critic take place against a shared framework of concepts and that a careful analysis of this conceptual background reveals a significant area of common ground. As a result of clarifying this common ground, I argue that extreme and moderate versions of surgical exceptionalism rest on important confusions or overstate the difficulties associated with individuating and assessing surgical procedures.

In light of this analysis, I argue for an analytical model of foundational issues concerning the individuation of assessment of surgical interventions. This model provides a problem-solving framework within which the legitimate concerns of weak surgical exceptionalism can be evaluated on a case-by-case basis, according to a dual concern for the health interests of patients and the need to generate reliable data about the performance of treatments and procedures. Rather than simply asking whether an activity is best classified as treatment, research, or innovation, the view I articulate here asks whether the activity can be carried out in a way that is adequately responsive to this dual set of concerns against the relevant background of feasible alternatives. The result is a framework that seeks to reconcile: a dedication to individual needs in clinical practice, the clinician's epistemic responsibility to utilize modalities that offer a reasonable chance of meeting clinical objectives, and the goal of advancing scientific knowledge and thereby improving the standard of care.

REGULARITY, STANDARDIZATION, AND SURGICAL EXCEPTIONALISM

I refer to "surgical exceptionalism" as the view that the somewhat exceptional ethical or regulatory status of surgery and surgical practice is justified by the exceptional differences between surgeries and pharmaceutical interventions. It is worth distinguishing three distinct, but nevertheless complementary, versions of this more general stance.

The strongest form of surgical exceptionalism holds that every surgery is a unique response to a unique set of challenges, and, therefore, that every surgical procedure is inherently innovative, novel, and uncertain. In particular, this po-

sition holds that the crisp distinctions between standard medical practice and innovation that can be applied to the use of pharmaceuticals cannot be meaningfully applied in the surgical domain. Because of the dynamic nature of surgery itself, all surgery is innovation and standard practice just is a constantly changing process of adaptation and change.

Strong surgical exceptionalism is an extreme position that is best thought of as the position that results from connecting in an explicit manner a set of views that are normally enunciated as unconnected, individual claims. Such claims often emerge, for example, when debate turns to the ontological differences between surgeries and pharmaceuticals.

For example, in comparison with pharmaceuticals and medical devices, surgical procedures seem to occupy a fairly nebulous ontological status. Pharmaceutical interventions and implantable medical devices are both classes of physical objects. As such, the variation that exists between different instances of the same drug or device can be kept within very narrow parameters by the use of rigorous methods of quality control in their manufacture and storage. In fact, one of the central goals of this manufacturing process is to ensure that whatever differences do exist in the eventual product are sufficiently small that they don't impact treatment outcomes or performance. This precision in mass manufacturing thus plays a central role in evaluating pharmaceutical interventions, because it enables clinicians to control the degree of variation that exists in the therapeutic agent as it is delivered to large numbers of recipients. This makes it possible to identify a static set of clinically relevant properties of a pharmaceutical (for example) and to study its effects as it is administered under controlled conditions. It also enables the profession to control the degree of variability that exists in the administration of the resulting therapy. The standard dosage, and known indications and contraindications for a particular drug, can then be formalized, allowing the profession to determine a set of treatments that are admissible for a particular condition and to articulate a set of best treatment practices, when possible.

In contrast, surgeries are not physical objects. As such, they are not produced by a process of manufacture and production that can be subjected to rigorous methods of quality control.

As Love puts it, "surgical techniques, unlike drugs, do not have chemical compositions, physical properties, routes and rates of excretion, or other qualities that can be measured precisely."[9] In comparison with drugs, which "come packaged as preparations to be given by dosage," operations are fundamentally dynamic "conceptual plans that require execution, and the details of a given operation change with time among surgeons and from patient to patient."[10]

Unlike pharmaceuticals, the clinically relevant properties of surgical procedures emerge directly from the manual interaction of the surgical team with the particular anatomy of the individual patient. This means that differences in the problem to be ameliorated and the specific anatomy of the individual patient constitute unique problems to be solved by the surgical intervention. Unlike pharmaceuticals, therefore, "surgical procedures are rarely introduced as fully defined, easily reproducible techniques. Rather, they come as principles for solving particular problems, sometimes of an urgent nature."[11] As new techniques or technologies are introduced into the clinical setting, and as surgeons gain experience with a particular procedure, these conceptual plans are modified and evolve.

Finally, each surgical procedure poses its own challenges to the dexterity and skill of the individual surgeon. Not only may results vary between different surgeons of different abilities, they may vary across the same surgeon at different times. As a result, not only may each individual surgery require different modifications in the general procedure, but the technique of the same surgeon may differ from patient to patient. With each surgery the surgeon acquires new experience and proficiencies that may lead to additional changes in approach or execution.

When these claims are combined with relatively rigid requirements for identity, strong surgical exceptionalism is the result. A more moderate form of surgical exceptionalism focuses on some of these same claims, but leaves controversial philosophical positions on identity to the side. It holds simply that actual surgical practice is sufficiently unique that it cannot be governed at the practical level by the same norms that are applied to other areas of medicine.[12] On this view,

surgical procedures are constantly in flux because each surgeon is constantly engaged in a process of learning, updating, and adapting. The regulatory paradigm that has developed in the context of pharmaceutical research and development, however, should not be applied to surgery and surgical research, because it would be practically infeasible to do so without damaging the integrity of the discipline itself. For example, because many surgical procedures are constantly evolving in response to new techniques, technologies, and challenges, there is no bright line at which a procedure crosses from standard treatment to innovation. However, the current regulatory and oversight mechanisms presuppose the existence of such bright lines, because such boundaries are easier to discern in the case of pharmaceutical interventions.

As a result, the more moderate position holds that imposing such requirements on the surgical profession would create bureaucratic delays, stifle innovation, and, ultimately, work against the interests of patients. Although the claims of moderate surgical exceptionalism do not necessarily presuppose the more robust philosophical claims of strong exceptionalism, the strong exceptionalist position entails the practical consequences of the moderate position. In actual practice, these positions may not be carefully distinguished.

Finally, weak surgical exceptionalism accepts the practical claims of the more moderate position, but resists making such sweeping generalizations about the entire domain of surgical practice. Instead, it limits the scope of its claims to areas in which known interventions are of questionable value or where effective interventions do not exist. Established areas of surgical practice in which several established interventions exist to treat a particular condition thus fall outside the scope of weak surgical exceptionalism. Given this restricted scope, weak surgical exceptionalism claims that the current regulatory paradigm could not be applied to the innovative activities at the frontiers of surgical practice without adversely impacting the prospects for advancing the state of the art.

As I have formulated these positions, weak surgical exceptionalism is closest to the reform-minded critic, in that it limits its claims about the exceptional status of surgery to the frontiers of innovation and development. Although I have

presented them as distinct positions, in practice, elements of these views overlap in complex ways. For instance, in response to the concerns about individuating surgical procedures that are emphasized in strong surgical exceptionalism, the reform-minded critic will rightly point out that sophisticated randomized controlled clinical trials have been carried out in the surgical domain. While this claim is true, it may not defuse similar concerns whose scope is restricted to the frontier regions of surgical development that are the primary focus of the weak position.

Similarly, reform-minded critics will point out that they can simply grant that there are interesting ontological differences between drugs and operations, that surgeries are susceptible to greater kinds and degrees of variation than most pharmaceutical therapies, and even that it may not, therefore, be possible to achieve the kind and degree of regularity and consistency across different instances of an operation as it is possible to achieve across different instances of a pharmaceutical. From these concessions alone it does not follow that meaningful distinctions cannot be made between approaches and techniques that represent the standard of care for a particular condition and those that represent significant departures from such accepted practices. At best, it shows only that our use of concepts such as standard of care and innovation will have to reflect different expectations about the kind of regularity and consistency that we can expect of surgical interventions, and the fact that we may need to be concerned with a different set of factors that are salient in the surgical context from those that are salient in the case of pharmaceutical therapy.

Although these claims are sound, they do not themselves constitute an analysis of the basic conceptual and practical issues that underwrite and motivate surgical exceptionalism. A more systematic focus on these basic issues may help to demarcate areas of agreement and disagreement so that genuine differences can be more clearly stated and scrutinized. In particular, although it may appear at a practical level that the reform-minded critic and the surgical exceptionalist differ widely in their basic assumptions and fundamental judgments, there is, in fact, a fairly basic framework of concepts and dis-

tinctions that provide the background against which their differences can be articulated. Moreover, by considering the following dilemma, we can see more precisely where the surgical exceptionalist agrees and disagrees with the reform-minded critic.

THE DILEMMA OF DYNAMISM

At the most general level, surgical exceptionalism is motivated by two central claims. First, there is a descriptive claim about the degree of variation and change endemic to surgery and surgical practice. Second, there is a normative claim that the dynamic nature of surgery must be protected, because it provides a powerful engine for progress and improvement. However, these claims are in tension with one another, and, even in relatively moderate cases, they give rise to what I will call the dilemma of dynamism.

This dilemma can be expressed as follows. For the dynamic nature of surgical practice to merit protection or preservation, one must be able to show that all of these changes are more than rampant directionless variation. That is, one must be able to show that they represent an engine of progress or advancement, rather than simply an engine of variation and change. So, one horn of the dilemma is this: if new surgeries vary so radically that they cannot be compared to past instances, then the claim that the dynamic nature of surgical practice is valuable and merits protection or preservation is undermined.

Radical claims about the degree of variation endemic to surgical practice threaten to undermine the normative claim that the dynamic nature of surgical practice merits protection, because the idea that surgical variation is an engine of progress presumes that surgeons have the ability to learn from their past experience and to adapt accordingly. To learn from the past, however, requires that ability to identify salient regularities across different instances of past surgeries and to compare previous approaches with proposed alternatives. To make such comparisons, though, surgeons must have some baseline against which changes that are improvements can be distinguished from changes that are mere variations. The other horn of the dilemma, therefore, is this: if different surgeries or varia-

tions in surgeries can be compared to past instances, then the value of dynamism is retained, but the exceptional status of surgery is undermined.

As this dilemma reveals, it is not at all clear that strong surgical exceptionalism can consistently maintain both its radical view of the dynamic nature of surgery and its position that this state of affairs is in any way desirable from a clinical standpoint. The same tension faces moderate and weak forms of surgical exceptionalism to the extent that they accept these claims but try to limit the scope of their applicability to a more limited set of surgical practices. Here again, however, to substantiate their evaluative stance, moderate and weak exceptionalists must be able to show that variations that occur within this limited domain represent genuine learning and result in actual progress. To substantiate the exceptional status of surgery, they must then show that these benefits cannot co-exist with more formal methods of evaluating surgical practice.

This simple dilemma reveals that more radical versions of surgical exceptionalism that disagree in principle with the reform-minded critic may not be internally consistent or fully coherent. As a result, the most plausible versions of surgical exceptionalism will not disagree in principle with the reform-minded critic. Rather, each side presumes that it is possible to identify some regularities as salient across different surgical procedures, so that different approaches can be evaluated and changes can be evaluated to see whether they represent improvements or mere variations. They diverge over a more practical problem, namely: What are the most reliable methods of evaluation that can feasibly be employed to make such evaluations?

The dilemma of dynamism, therefore, provides a strong indication that the most plausible forms of surgical exceptionalism also share significant common ground with reform-minded critics. We can do more, however, to clarity this common ground. In particular, the analysis in the following section reveals that the descriptive and normative aspects of surgical innovation cannot be neatly separated. As a result, the very idea that something represents a change over some past instance of a surgical procedure presupposes the very

type of causal claims that determine whether we regard a change as beneficial or not.

DUAL ASPECTS OF INNOVATION AND STANDARD PRACTICE

To clarify the relationship between normative and descriptive elements of innovation, three distinct but interrelated aspects of the concept must be carefully distinguished. The first aspect deals with change or alteration. An innovation is something new, either because it represents a distinctive modification of something that already exists, or because it is a first of its kind. The second aspect of innovation involves the evaluative status of this change. That is, the term "innovation" implies or connotes that the change represents an improvement or an advantage of some kind. This term is frequently used, therefore, as an honorific to describe a modification or invention that succeeds in advancing some interest or in improving some process or outcome. Finally, as these previous points illustrate, innovation is a relational or comparative concept. Something is only an innovation relative to some baseline state of affairs or set of admissible alternatives. This relational quality is perhaps most salient in the evaluative aspect of innovation, where it is more natural to speak of something as an advantage over something else, or an improvement relative to a previous set of alternatives. However, the aspect of innovation that relates simply to novelty or change is also relational or comparative in nature.

This relational aspect of the concept of innovation helps to highlight one of the central roles played in the medical context by concepts such as standard medical practice and standard of care. Namely, these concepts refer to the practices that provide the baseline against which alternative means of treating a particular condition may be measured and evaluated. This point also helps to underscore the important evaluative aspect of these latter concepts as well. On the one hand, to say that something is standard medical practice or the standard of care for some condition is frequently a way of indicating the *popularity* of the treatment in the field. In addition to this *descriptive* use, however, these terms too function as honorif-

ics. That something is standard medical practice or the standard of care frequently implies or connotes that it has been validated as an effective means of achieving a valuable end. As such, these terms are also used in a *prescriptive* sense to indicate that a certain intervention or treatment modality ought to be used to treat a particular condition.

The dual descriptive and evaluative aspects of these concepts makes it possible to construct the following somewhat oversimplified matrix (see figure 2.1). The rows in the matrix represent the need to determine whether an intervention constitutes a deviation from current practice, either because it is a modification of an existing practice or intervention, or an entirely new invention. The columns represent the need to determine whether the practice or intervention is advantageous or disadvantageous, better or worse than the alternatives for treating the same medical condition.

Although this matrix is crude in many respects, it provides a useful representation of several important relationships. First, important epistemological issues are involved in placing an intervention within this simple matrix. In particular, even when it is relatively easy to place an intervention in one of the above rows, it may be significantly more difficult to determine the box in the row to which it belongs. There are many instances in the medical literature, for example, of treatments that were regarded as effective and used on a widespread basis only to be shown to be either ineffective or posi-

Simplified Matrix for Classifying Treatments or Procedures that Target a Specific Medical Condition

	Advantageous	Disadvantageous
Change	Genuine innovation/ New admissible treatment	Failed innovation
Current intervention	Admissible treatments/ Prescriptive Standard of care	Inadmissible/ Aspect of practice in need of reform

Figure 2.1.

tively harmful. In such cases, the belief that a common intervention falls into the box marked "admissible treatment" turns out to be false. No change in practice has occurred, but careful evaluation, often carried out within a controlled clinical trial, reveals that an existing intervention belongs instead in the box marked "inadmissible treatment." The same is true for many innovative drugs, devices, and surgical procedures. Conscious deviations from standard practice are almost always undertaken under the belief that the changes or modifications will produce a treatment or procedure that fits into the box marked "genuine innovation." In many cases, though, careful study reveals that the change does not represent a significant benefit and the new intervention is subsequently discarded or withdrawn as what I have labeled a "failed innovation."

Concern about whether something is a standard activity or an innovation, therefore, almost always functions as a proxy for a network of more basic concerns that are brought into relief in the above analysis. The value placed on care that is customary or standard practice hinges on the presumption that this status has some probative value; either its effectiveness has been validated in some experimental setting, or there is some other form of warrant for the belief that it represents a causally efficacious means of achieving a desired clinical objective. When the evaluative status of a practice or intervention is cast into doubt, the question of whether it is customary or novel loses much, if not all, of its probative value. Similarly, new practices are attractive when we believe that they represent genuine innovations — advances over existing alternatives — and they are objects of suspicion when the fact that they are new practices whose evaluative status has yet to be clarified is made salient.[13]

This brief analysis enables us to see that, although they may disagree about the usefulness of terms such as "innovation" and "standard of care," both surgical exceptionalism and the reform-minded critic presuppose the need to make informed judgments about the therapeutic merits of surgical interventions. In this respect, they ultimately agree about the need to assess the merits of surgical procedures. They disagree, however, about the proper methods of assessment. Surgical exceptionalism favors informal methods of evaluation

and a reliance on the judgment of individual surgeons, whereas the reform-minded critic claims that these informal methods are highly susceptible to serious epistemic shortcomings.[14] Adopting a problem-solving perspective, we can thus reframe the central issue in this aspect of the dispute as follows: To what extent is it possible to use more reliable scientific and statistical methods to advance the assessment of surgical practices, including innovative practices, beyond the informal methods accepted by the surgical exceptionalist?

As this problem-solving perspective illustrates, this aspect of the debate can be reduced to a well-defined issue for which there are relatively settled means of resolving disputes that exist in particular cases. I will return to this issue in more detail below. For now, I want to emphasize that this reduction is made possible by the additional shared presumption that it is possible to detect sufficient regularities across individual surgeries to reliably distinguish instances of similar surgical procedures. The conceptual and practical issues that are involved in this presupposition can be clarified further by noting a second feature of the above matrix.

The above matrix illustrates the tight interrelation that exists between the dual descriptive and evaluative aspects of our concepts of innovation and standard practice. In particular, there are unaccountably many differences between any set of things that are not numerically identical. From the clinical or therapeutic standpoint, however, not all of these differences are relevant. To place an intervention into one of the above rows, therefore, we must determine whether a difference between any two interventions is significant or meaningful. For example, wearing blue rather than transparent gloves would not constitute a significant change to a procedure, unless it could be shown that the color of the gloves had a *causal impact on some relevant process or outcome.* One might claim that wearing blue gloves is a change to the procedure, but deny that it is significant because it has no causal impact on a relevant process or outcome. Alternatively, one might simply deny that wearing blue gloves is a change in procedure. However one chooses to *describe* such a case, the fundamental point is that *the underlying judgment about what constitutes a significant or meaningful change* is itself bound

up with judgments about important causal questions concerning the mechanisms or pathways for influencing relevant processes or outcomes.

In light of these interrelationships, a framework for identifying and evaluating innovation in the surgical domain should focus on two key elements. First, it must identify the defining features of surgical procedures that constitute the focal points for questions about individuation and assessment. Second, it must articulate the set of factors that provide the cannons of relevance against which variations among these foci can be identified and assessed. After clarifying these points, we may then return to the above mentioned approach to issues of assessment.

INDIVIDUATING SURGICAL PROCEDURES

Surgical interventions can be analyzed into formal and material components. At a formal or abstract level, surgeries are conceptual plans that represent as salient a set of treatment goals or endpoints and a set of steps or procedures for achieving or bringing them about. To effectuate change in the world, however, the procedures laid out in these conceptual plans must be implemented under controlled conditions by particular agents. Implementations of the same conceptual plan may thus differ in a variety of respects, as the procedures of the more abstract conceptual plan are tailored to the particularities of the individual case and executed by agents of different abilities. Because the actual causal properties of a surgical intervention are the result of this interaction between *information, agent,* and *environment,* each of these factors is relevant to the assessment of a surgical intervention. However, the formal or informational component provides the appropriate focus for individuating surgical procedures.

At the formal level, surgical interventions are defined by two primary features: the target endpoints that the procedure is designed to bring about most directly, and the hierarchy of instrumental goals that constitute the means through which the desired target endpoints are effectuated. For any medical condition, therefore, alternative surgical procedures can be distinguished at the most general level by their pursuit of dif-

ferent target endpoints. In response to breast cancer, for example, radical mastectomy is defined by the target endpoints of removing the entire breast, pectoral major and minor muscles, and lymph nodes. In contrast, the modified radical mastectomy is defined by the target endpoints of removing the entire breast and frequently the axillary lymph nodes. Similarly, lumpectomy is defined by the target endpoints of excising only the breast cancer tumor as well as a surrounding border of normal breast tissue. Each set of target endpoints defines a genus of surgical intervention for breast cancer.

Surgical procedures may then be further individuated into species by specifying the hierarchy of instrumental goals that constitute the means that are used to realize or bring about the target endpoints that define the relevant genus. Here too, for any particular set of target endpoints there may be numerous alternative species. For example, one hierarchy of subordinate goals for achieving a target endpoint might involve a traditional, open procedure while an alternative employs a newer, minimally invasive laparoscopic procedure. If the genus of the surgical procedure is defined by its target endpoints, then different hierarchies of instrumental goals define the various species of the same genus.

In theory, however, every hierarchy of instrumental goals can itself be broken down into a further hierarchy of subsubgoals. For example, it may be possible to make a particular incision with either a traditional scalpel or a laser and to repair a particular rupture with a variety of sutures, staples, or additional means. Analytically speaking, then, there may be multiple ways of realizing or achieving any particular goal in a surgical procedure. In response to this analytical point, strong surgical exceptionalism adopts the most rigid means of individuating surgical procedures according to which there are as many possible surgical procedures of the same species as the product of all the possible hierarchies of subgoals, subsubgoals, and so on. From a clinical or therapeutic standpoint, however, not all of the very fine differences that are recognized by the most rigid means of individuating surgical procedures are relevant.

How do we determine, therefore, when variations in these focal points count as genuine differences? The cannons of rel-

evance against which variations in these foci can be identified and assessed are provided by the set of factors that have the greatest potential impact on the procedure's net therapeutic advantage. Following Freedman, I treat an intervention's "net therapeutic advantage" as "a compendious measure of a treatment's attractiveness."[15] It thus includes considerations such as an intervention's direct impact on disease reduction, symptomatology, and ability to function, discounted by its particular side effect profile. So, for example, differences in degree of invasiveness or in the particular pathways used by the intervention may be relevant to several of these dimensions of concern. From a fairly fastidious viewpoint, these factors may be regarded as properties of the method of intervention that constitute its efficacy.

The boundary between efficacy and effectiveness, however, is difficult to draw in the surgical domain, given that the causal properties of surgical procedures emerge from the interaction of information, agent, and environment. As a result, the net therapeutic advantage of a surgical intervention must also include broader, more practical considerations such as the efficiency and reliability of a procedure, especially when it is implemented in the broader surgical community. I will refer to this last feature as "projectability."

Like many of these factors, projectability is itself a function of the cognitive and manual demands that a particular procedure places on the surgical team, given the existing skill set of team members. These aspects of a procedure are relevant precisely because surgical interventions must be repeatable interventions that can be disseminated to and implemented by competently trained personnel. If all else is equal, therefore, one procedure might be preferred to another because it is less complex, or requires maneuvers that are easier to execute or which deviate less from the core of a team's existing skill set in comparison with alternatives.

The very rigid conditions for identity emphasized by the strong exceptionalist position emphasize the possibilities for variation that arise from this blurred boundary between efficacy and effectiveness in the surgical domain. However, this position itself represents a failure to think systematically about the relationship between the formal or informational aspect

of surgeries and the cognitive and physical abilities of individual surgeons. In particular, at a formal or informational level, the instrumental goals of a surgical procedure are modules of steps to be accomplished. To ensure that the causal properties of surgical procedures can be retained as they are implemented by different individuals or on different occasions, it is necessary to ensure that individual members of the profession have a common mastery of the basic set of skills, techniques, and maneuvers necessary to execute the steps of these modules. A fundamental commitment to quality control in the surgical context must therefore involve setting and enforcing high, uniform standards for education, training, and licensure, to ensure that members of the profession possess the requisite basket of skills and abilities (see figure 2.2).

However, the key issue concerns whether such a process of quality control on education and training will be sufficient to ensure that agents are capable of implementing the surgical procedure in a way that preserves its effectiveness. In certain relatively simple cases, therefore, it may be necessary only to individuate instances of a surgical procedure down to the hierarchy of subordinate goals that define the species of that relevant genus. Beyond that, the tolerance on allowable variation in the means that are used to execute these subordinate goals might be as expansive as the set of techniques that are recognized in the surgical profession as reasonable examples of the competences that surgeons in that area are expected to possess. Speaking of the variety of approaches to fixing an inguinal hernia, for example, Meakins suggests, "the large number of operations testify to the reality that in many situations, the technique does not matter despite strenuous and occasionally acrimonious discussion in support of one operation or another."[16] In some cases, therefore, it may not matter which specific approach is employed, as long as it falls within a family of acceptable implementations and that it is competently executed.

In more complicated cases, it may be necessary to individuate instances of a surgical procedure at a finer level, specifying more precisely the particular steps that must be taken and techniques employed in order to effectuate the desired target endpoints. Ultimately, however, the descriptive ques-

tion of where to fix the bound on individuating surgical procedure hinges on our beliefs about the tolerances that must be maintained to bring about the relevant target endpoints in a way that is most likely to result in a positive net therapeutic advantage to the patient.

Surgery: An Interaction Between Information, Agents, and Environment

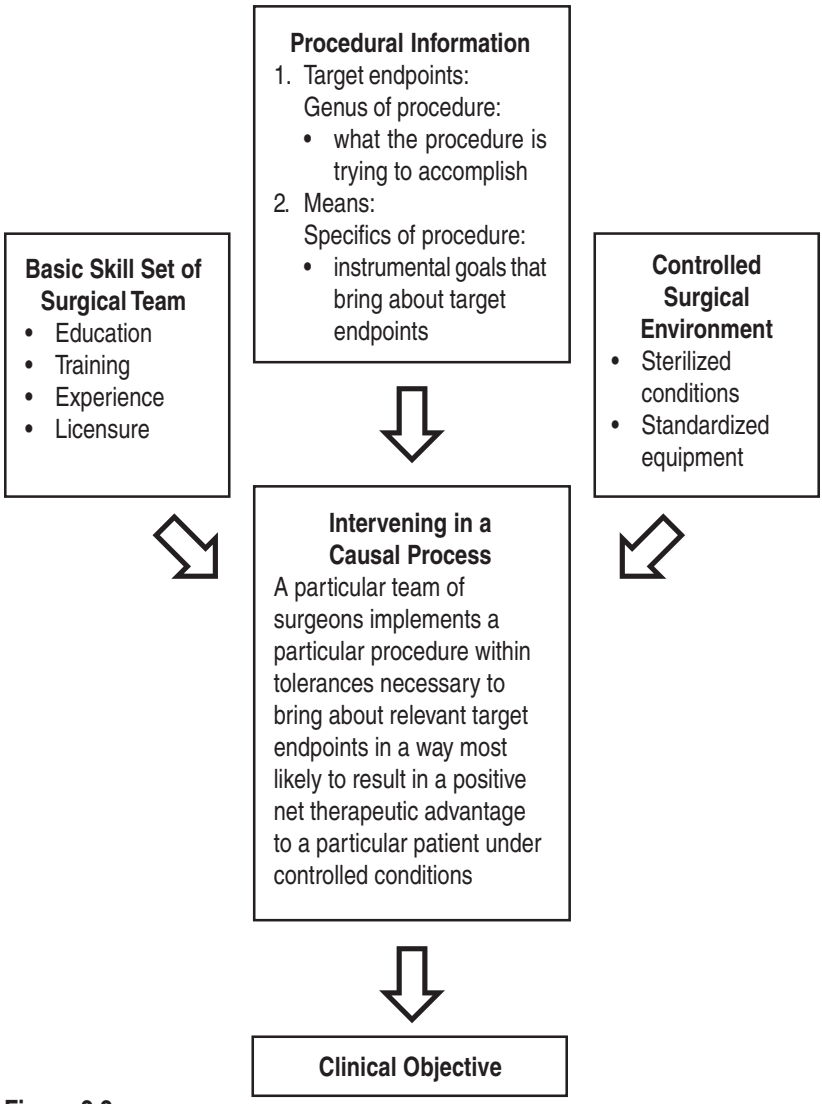

Figure 2.2.

As a result, procedures that fall into the row in the above matrix for commonly accepted surgical procedures can be defined fairly broadly as the set of interventions that are commonly used to treat the medical condition in question. This set will include genera of surgical procedures that are defined by different target endpoints. Whether two sets of endpoints count as different will depend on our beliefs about whether whatever variations that do exist are likely to exert a causal influence on the intervention's net therapeutic advantage. For each of these genera, the species are defined by different hierarchies of instrumental subordinate goals that utilize techniques that are recognized in the surgical profession as reasonable examples of the competences that surgeons in that area can be expected to possess as qualified members of the profession.

The row in the matrix for new surgical procedures will thus include procedures that represent two general types of deviation from common practice. First, the procedure can represent a new genus of intervention for a particular condition, either because no surgical intervention currently exists for that condition or because it is defined by a set of target endpoints that deviate from those that are commonly used. Alternatively, the procedure may represent a new species of an accepted genus, in that it employs a hierarchy of instrumental subordinate goals that fall outside the bound on commonly accepted variation. Again, whether a variation at either of these levels should be counted as a genuine difference will depend on our beliefs about whether they are likely to exert a causal influence on the intervention's net therapeutic advantage.

As I noted earlier, however, sorting procedures into the appropriate row is of secondary importance to the process of sorting procedures into the appropriate column. In other words, classifying a procedure as an accepted practice or an innovation is often an indirect way of trying to determine the evaluative status of the intervention. Additionally, directing our attention primarily at techniques that deviate from common practices is only warranted to the extent that the therapeutic advantages of the practices that are already in place have been verified by an appropriate means. When this is not

the case, the primary focus of concern should be on determining the evaluative status of options in question.

Within either of the above rows, therefore, the set of admissible interventions will consist of a subset of surgical procedures from these larger classes. This subset will include the genera of procedures that are most likely to provide patients with a positive net therapeutic advantage when executed by a competent surgical team. For each genera, admissible species may then be defined by the hierarchies of instrumental subordinate goals that fall within the bound on acceptable variation. Here again, the bound on acceptable variation will be set by the tolerances that must be maintained to bring about the relevant target endpoints in a way that is most likely to result in a positive net therapeutic advantage to the patient.

EPISTEMOLOGICAL AND ETHICAL ISSUES IN SURGICAL ASSESSMENT

As the above analysis emphasizes, important epistemological issues are involved in our judgments of similarity, difference, and relative value in the surgical domain. Because these judgments are themselves grounded on beliefs about causation, our ability to refine and apply the categories described in the previous section will hinge on the extent to which we directly investigate these underlying causal claims. Moreover, these epistemological issues and the imperative to investigate them directly are intimately connected to a network of basic moral issues. The exceptionalist's resistance to the use of more formal methods of assessment in the surgical domain may therefore receive specious support from confusion about each of these issues.

For example, given the conceptual interrelationships discussed above, the exceptionalist's commitment to largely informal and often unreliable methods of assessment may itself be causally responsible for the exceptionalist's beliefs about both the degree of variation that exists in the surgical domain, and its significance. Individual judgment and informal case series are susceptible to error from a variety of sources. For example, retrospective case series may be subject to significant bias in the selection of the cases that are reported and to

bias in the selection of surgical candidates. Additionally, such reports often lack the methodological mechanisms necessary to distinguish the effects of the surgical intervention from variations in the natural history of the disease. Because individual judgment and such informal case series often lack the capacity to distinguish the effects of the procedure from such confounding variables, and because of a reporting bias that favors studies with positive results, this informal approach to inquiry can make it appear that subtle differences in surgical procedures have significant causal consequences. The misguided acceptance of the reliability of such approaches, combined with professional pressures and potential financial incentives, may then encourage additional efforts at innovation on the part of a larger number of individuals. At both a cognitive and a practical level, therefore, surgical exceptionalism may itself be a causal link in a self-reinforcing process that encourages the proliferation of variation in surgery.

Such faulty epistemic assumptions exaggerate the capacity of individual surgeons to disentangle the actual causal properties of a change in surgical procedure from background noise and confounding variables. In doing so, they also provide the background against which the imposition of more formal methods of assessment may appear to raise special moral problems for surgeons. In particular, some may believe that as long as innovative activities do not follow a formal protocol with a well-defined set of hypotheses, they do not fall under the heading of medical research, and do not therefore generate special moral requirements, such as the need to secure external oversight or outside approval. Such a view combines the faulty epistemic assumptions discussed above with the faulty moral assumption that special moral requirements that are imposed on research involving human subjects are generated by the use of formal protocols or methods of assessment.

The special moral requirements that attach to research involving human subjects, however, do so not because a formal protocol or method of assessment is being employed. They arise from two sources: the potential tension that exists between the goals of sound clinical practice and the goals of sound clinical research, and the potential impact of this tension on the interests of patient/participants. These same ten-

sions arise in the case of what is sometimes referred to as "informal research" in the surgical domain. In the context of "informal research," however, they are addressed in a manner that deviates in problematic ways from the proper norms of both sound clinical practice and sound clinical research.

What I am calling sound clinical practice combines two elements. First, this activity is guided by an *individual focus,* in which the primary goal is to minister directly to the health needs of the individual patient. Values such as respect for autonomy and informed consent serve, in part, to encourage a process of effective communication between patients and physicians, so that this individual focus can be effectuated in a way that is consistent with the patient's broader goals and values. A central goal of this relationship, therefore, is to select from the set of admissible treatments the option that is the most attractive in light of this larger set of values and objectives.

The second element of sound clinical practice relates to the medical professional's *epistemic responsibility* to pursue this individual focus by using diagnostic and therapeutic modalities that he or she has good reason to believe offer a reasonable chance of achieving the desired clinical objectives. This epistemic responsibility provides the basic motivation for trying to determine a set of admissible treatment options and a prescriptive standard of care for a particular medical condition. That is, it helps the medical profession to ensure that, in routine medical practice, patients can make informed decisions and receive an intervention whose relative net therapeutic advantage over the available alternatives has been established by some credible means.

This epistemic responsibility therefore provides a key element in the division of labor between clinical medicine and clinical research: a basic role of medical research is to provide data that will resolve uncertainty that exists about the evaluative status of an intervention or set of interventions. The defining goal of sound clinical research, therefore, is to generate reliable information or data that will resolve uncertainty about a clinically relevant question. However, to achieve this goal, the activities of researchers often must be fixed or regulated by terms that are articulated in the research protocol. These

terms, however, often reflect, not the individual focus of clinical medicine, but the scientific and statistical requirements that must be met for the results of the study to be methodologically reliable or significant. The use of a formalized protocol, therefore, is simply one palpable indication that a sound method of inquiry is being used to address a clinically relevant scientific question.

Special moral obligations attach to the latter activity because of the potential tensions that exist between the individual focus of clinical practice and the methodological requirements that enable clinical research to generate generalizable scientific data. Often, for instance, trade-offs must be made between these methodological requirements and the extent to which an individual's treatment can be tailored or adjusted to his or her specific needs. When such tensions exist, they may affect the subject's interests in ensuring that his or her current health interests are not sacrificed or adversely impacted by the goal of generating information that will primarily benefit future patients. The special moral requirements that attach to clinical research, therefore, are intended to ensure, among other things, that these trade-offs do not unfairly disadvantage research subjects and that subjects who participate in research are aware of the nature of the trade-offs that a particular research project entails.

Informal surgical experimentation or innovation deviates in problematic ways from both sound clinical practice and sound clinical research. Altering a surgical procedure out of a desire to produce a genuine innovation creates a significant tension between the individual focus of clinical practice and the epistemic responsibilities of the medical professional. It also creates a significant tension between the goal of assessing the relative therapeutic merits of the innovative practice and employing the scientific and statistical methods that are necessary to carry out this process. Because informal surgical research is not fully responsive to the norms of either of these regulative ideals, there will almost always be an alternative way of proceeding that will be preferable from both a moral and an epistemic point of view.

Grasping this point is essential to adopting a more productive problem-solving perspective on this issue. It suggests

that we reframe the central issue here as: For any particular case, to what extent would it be feasible to pursue an alternative approach that is more adequately responsive to the norms of these twin regulative ideals?

A PROBLEM-SOLVING PERSPECTIVE

I have contrasted treatment and research by highlighting the extent to which these activities are organized around different goals. The fact that these goals are conceptually distinct, however, should not be confused with the very different question of whether they are mutually exclusive or incompatible.[17] Figure 2.3 represents the misguided *a priori* assumption that these goals are mutually exclusive and that the tradeoffs between them are, therefore, zero sum.

In truth, the potential for integrating these conceptually distinct goals in actual practice is more robust than is recognized by the zero sum model. On the one hand, something like this recognition is latent in the exceptionalist's own efforts to combine treatment of individual patients with so-called informal research. The central deficiency of the exceptionalist approach is that it suffers from significant ethical and methodological flaws. However, other approaches exist that are more responsive to each of these goals, which do not suffer from these ethical shortcomings.

Figure 2.4 below provides a more accurate representation of the conceptual and practical potential for achieving both of these goals on the same occasion. Each axis should be understood as a dimension of value that can be used to rank a set of options, when the options are activities in which individuals with a specific medical condition might participate. Ranking

Relationship of Clinical Practice and Clinical Research Represented as Zero Sum

Figure 2.3

each option according to its responsiveness to the interests of individual participants determines the relative place of the option on the horizontal axis. Ranking each option according to its capacity to generate generalizable scientific information determines the relative place of the option on the vertical axis. Mapping alternative options onto this space simply provides a visual representation of various possible dominance relationships. In particular, some option (O1) weakly dominates another (O2) just in case there is no dimension of value on which O2 is preferred to O1 and O1 is preferred to O2 on at least one dimension of value.[18]

For the purposes of the present chapter, figure 2.4 simply presents a more realistic account of the space of possible options for carrying out research activities. The northwestern quadrant represents activities that are designed to generate generalizable data without also attempting to meet the particular health needs of individual research participants. Phase I clinical trials provide a nice illustration of the kind of activity that would be mapped onto this quadrant. The southeastern quadrant represents the contrasting case of standard clinical practice in which treatments are tailored to the needs of

Possible Relationships between Practice and Research, Highlighting Integrative Potential of the Two Activities

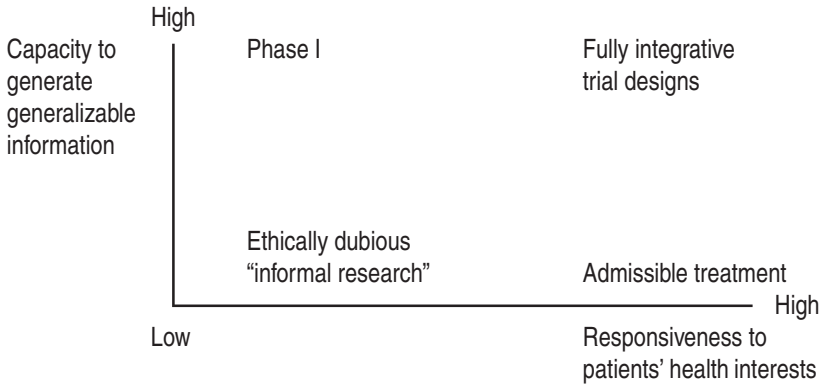

Figure 2.4. Labels in the figure are intended simply to illustrate the type of activity that might fall into the designated quadrant.

individual patients without also attempting to generate generalizable data that might answer a well-formed hypothesis of some kind.

Activities that fall into the southwestern quadrant raise both ethical and scientific concerns. Informal research activities will frequently fall into this quadrant for several reasons. First, they are used to minister to the individual needs of particular patients, but they do so through the use of a means or pathway whose evaluative status is unknown. Additionally, the way the treatment is administered may or may not be constrained by the goal of generating informal research data. Although these goals are in competition, there is no clear sense of how they are to be reconciled.

Activities that take place in the northeastern quadrant represent attempts to integrate the goals of science and medical research with the need to safeguard and secure the interests of trial participants. From a problem-solving perspective, the question that arises in any particular case is whether there exists an alternative way of proceeding that is a more northeasterly option. An example of a study that might fall into this quadrant would be a controlled clinical trial that is designed to disturb a state of equipoise that exists in the larger clinical community about the relative therapeutic merits of the interventions within the arms of the trial.[19] For instance, such a trial might be designed to evaluate the relative therapeutic merits of two interventions that are widely used to treat the same medical condition. All of the participants in such a trial would be guaranteed to receive a treatment that is consistent with care they might receive from reputable medical practitioners. Although they forego the choice of which intervention they receive, proper methodological controls and design features ensure that the results of the trial will provide valuable data about the relative merits of the study interventions.[20]

As a general rule, the imperative to employ more northerly options increases with the degree of uncertainty that exists about the relative therapeutic advantages of the available treatment options. Similarly, the imperative to employ more easterly options increases as the risks to the interests of subjects increases. So, for example, in cases where no treatment is an admissible therapeutic option, it may be permissible to

pursue a less easterly option that employs a placebo control over a more southeasterly option in which two interventions are compared in a head-to-head trial. When no treatment poses significant risks of suffering, morbidity, or mortality, however, and there exists an admissible therapeutic intervention, it is necessary to pursue a more southeasterly trial that compares the new intervention to the existing treatment rather than the more northwesterly option of the placebo-controlled trial.

Ultimately, the question of how to make the subtle tradeoffs between these concerns that may arise in particular cases touches on issues that are fundamental to research ethics. These are also issues about which there remains some dispute. However, for the purposes of the present inquiry, it is sufficient to reiterate simply that most efforts at informal research will be dominated by alternative activities, some of which may fall into this area of dispute.

Movements to this northeastern quadrant can be approached from either the treatment or the research perspective. For example, the practical realities of actual clinical practice often constrain the extent to which the regulative ideal of sound clinical practice can be realized in a particular case. In rare or unique cases, there may not be any recognized list of admissible or standard treatments. Similarly, sometimes the world presents the responsible clinician with a novel problem that must be solved immediately to prevent significant harm to the patient. In such cases, responsible clinical practice may itself require ingenuity and the ability to adapt intelligently to the circumstances. When such a problem could not have reasonably been foreseen, and when there either are no standard interventions or they cannot be feasibly implemented under the constraints of the exigent circumstances, the responsible clinician may be obligated to innovate and to problem-solve.

In such cases, the focus on individual needs that helps to define the regulative ideal of sound clinical practice requires the clinician to engage in what I will call "innovation as emergent problem-solving." A defining characteristic of this class of activities, however, is that it is carried out in response to the demands of an emergent situation in which it is not feasible either to utilize a technique of known efficacy or to for-

mally evaluate potential solutions to the problem before implementing them.

More difficult issues arise for innovative activities that fall within a broader zone of ambiguity. This is a zone of ambiguity, in which the individual focus of clinical medicine is pursued on a more regular basis through means whose evaluative status has not been clearly assessed. In such cases, the pursuit of a beneficent end is in tension with the professional's epistemic responsibility to ensure that the means that are used to bring about this end have a reasonable likelihood of success. A key element that distinguishes innovative activities that fall into this zone of ambiguity from those that constitute innovation as emergent problem-solving is the greater potential, in the former case, of pursuing alternative approaches that might better conform to the norms of sound clinical practice or sound clinical research.

As uncertainty about the relative therapeutic merits of an intervention increases, so does the surgeon's responsibility to implement a method of assessment that will clarify the issues that are in doubt. The exceptionalist may resist this claim out of a belief about the practical infeasibility of implementing more formal methods of assessment in the surgical domain. Using the model I have outlined in figure 2.4, this reticence can be understood as the belief that there are relatively few options for assessment that exist in the northeasterly quadrant, and that those that do exist are extremely cumbersome and expensive. However, this relatively impoverished view of the option space rests on the myopic identification of proper methods of formal evaluation with randomized controlled trials (RCTs). Although the RCT is widely regarded as the gold standard for clinical trial design, this particular method of assessment is not always necessary, feasible, or appropriate.[21] Recognizing that a variety of statistical and scientific methods exist for gathering and analyzing data about the performance of a surgical intervention reveals the attractiveness of adopting a problem-solving perspective in which the goal is to locate and define clinically relevant questions in a particular case, and then to match them with an appropriate method of data gathering and assessment.

For example, large multi-center RCTs are appropriate when uncertainty exists about the relative net therapeutic advantage of a well-defined set of treatment alternatives. These alternatives may include widely available and accepted interventions as well as innovative alternatives that seek recognition as admissible interventions for significant medical conditions. The fact that surgical procedures may need to go through a significant period of development and refinement before they are ready to be evaluated in an RCT does not mean that there is no need to conduct an RCT once this stage has been reached. Moreover, there is no reason why alternative methods of assessment should not be utilized at earlier stages of development.

In particular, a great deal can be done to address important causal questions simply by implementing the relatively simple steps necessary to gather richer sets of prospective data from clinical practice to construct a robust informational database. For example, Hlatky and colleagues report that predictions of multivariable statistical models based on data derived from the Duke Cardiovascular Disease Databank agreed well with the data from the three major randomized trials of coronary bypass surgery.[22] Generally speaking, the reliability of such databases will depend on the size of the treatment effects to be measured and our existing knowledge about the prognostic factors of the disease in question. As our understanding of these factors advances, so does the ability of statistical models to distinguish smaller treatment effects from background noise and confounders. Where treatment effects are pronounced and our understanding of the natural history and prognostic factors of the disease or condition is good, such databases provide powerful evaluative tools.

Similarly, so-called tracker trials represent an attractive means to monitor and assess surgical interventions as they progress through rapid periods of development or change.[23] Such innovative evaluative techniques represent a particular instance of a broader range of quality assessment tools that might be used to monitor and assess incremental changes in procedure, implementation, or technique.[24] These methods are particularly attractive when studying a change in a modular

component that impacts a discrete physiological system. When this is the case, such quality assessment methods may generate performance data about the impact of such changes without the need to conduct an RCT every time a surgeon changes a modular component of an existing procedure in the hope of improving the procedure's net therapeutic advantage.

In other cases, it may be possible to answer clinically relevant questions by conducting carefully designed open trials. In particular, the development of Bayesian statistical designs now makes it possible to conduct sophisticated open trials that are capable of generating important clinical data.[25] Additionally, as data from such carefully designed open trials accumulates, it may be possible to answer clinical questions through meta-analysis and proper statistical modeling.[26]

This is not, by any means, an exhaustive catalogue of the methods of technology assessment that can be utilized in the surgical domain. It is intended simply to illustrate the point that there exists a fairly rich set of options beyond the conduct of RCTs for evaluating the performance of alternative surgical procedures. Obviously, the process of matching particular clinical questions with an appropriate method of assessment is a task for biostatisticians working in close collaboration with surgeons. Only by fostering such collaborative efforts in actual practice, however, will it be possible to directly address the important epistemological issues that are required by the tasks of individuating and assessing alternative surgical interventions.

CONCLUSION

The failure to pursue a problem-solving approach to any problem often stems from the perception that there are no feasible solutions for such an approach to uncover, or from the belief that better outcomes can be achieved by some other means. Both of these factors are at work in the debate over the exceptional status of surgery and surgical innovation. The purpose of the present philosophical analysis has been to undermine and to discredit these perceptions and to indicate that a problem-solving approach to surgical innovation is not only feasible, but that it represents the only pathway by which

surgeons can integrate the individual focus of their therapeutic mission, their epistemic duty to utilize modalities that offer a reasonable chance of meeting clinical objectives, and the broader social goal of advancing scientific knowledge to improve the standard of surgical care. The scope of the present inquiry, therefore, has been limited to conceptual analysis and to clarifying some of the moral arguments that support pursuing the problem-solving approach that is articulated here. Obviously, however, pursuing such an approach will have costs of its own, and there may be inhibitions to undertaking it in earnest that stem from other sources. The question of how to provide additional motivation to undertake the problem-solving approach to surgical innovation — and whether formal regulation is an attractive means of facilitating this end — requires a careful consideration about the best ways to influence the complex culture and institutions of the surgical profession.

ACKNOWLEDGMENTS

I thank Jonathan Moreno, Angelique Reitsma, Griffin Trotter, and the other members of the working group on regulating innovative surgery for their critical comments and support. I also thank my student Emily Evans for her comments on this chapter and for creating figure 2.2.

NOTES

1. G.J. Agich, "Ethics and innovation in medicine," *Journal of Medical Ethics* 27 (2001): 295-6; C.E. Margo, "When is surgery research? Towards an operational definition of human research," *Journal of Medical Ethics* 27 (2001): 40-3; J.W. Love, "Drugs and Operations: Some Important Differences," *Journal of the American Medical Association* 232, no. 1 (1975): 37-8.

2. K.V. Iserson and P.M. Chiasson, "The ethics of applying new medical technologies," *Seminars in Laparoscopic Surgery* 9, no. 4 (2002): 222-9; J.E. Frader and E. Flanagan-Klygis, "Innovation and Research in Pediatric Surgery," *Seminars in Pediatric Surgery* 10, no. 4 (2001): 198-203; D.C. Cronin, J.M.

Millis, and M. Siegler, "Transplantation of liver grafts from living donors into adults — too much, too soon," *New England Journal of Medicine* 344, no. 21 (2001): 1633-7.

3. National Emphysema Treatment Trial Research Group, "Patients at High Risk of Death after Lung-Volume-Reduction Surgery," *New England Journal of Medicine* 345, no. 15 (2001): 1075-83; National Emphysema Treatment Trial Research Group, "A Randomized Trial Comparing Lung-Volume-Reduction Surgery with Medical Therapy for Severe Emphysema," *New England Journal of Medicine* 348, no. 21 (2003): 2059-73; J.B. Moseley et al., "A controlled trial of arthroscopic surgery for osteoarthritis of the knee," *New England Journal of Medicine* 347 (2002): 81-8.

4. V. Brower, "The ethics of innovation: should innovative surgery be exempt from clinical trials and regulations?" *EMBO Reports* 4, no. 4 (2003): 338-40.

5. B.L. Bosk and J.E. Frader, "Obstacles for Change — Sociological Views," chapter seven of this volume.

6. A.J. London, "The Ambiguity and the Exigency: Clarifying 'Standard of Care' Arguments in International Research," *Journal of Medicine and Philosophy* 25, no. 4 (2000): 379-97; A.J. London, "Equipoise and International Human-Subjects Research," *Bioethics* 15, no. 4 (2001): 12-332.

7. Love, see note 1 above.

8. J.E. Frader and D.A. Caniano, "Research and Innovation in Surgery," in *Surgical Ethics,* ed. L.B. McCullough, J.W. Jones, and B.A. Brody (New York, N.Y.: Oxford University Press, 1998), 216-41.

9. Love, see note 1 above, p. 38.

10. Ibid.

11. Ibid.

12. Aspects of the moderate position can be found in Agich (see note 1 above) and Love (see note 1 above) although Agich is probably best understood as expounding weak surgical exceptionalism.

13. In order to prevent the descriptive and evaluative aspects of the term "innovation" from being confused in practice, new therapies and procedures are sometimes referred to as "experimental" or "investigational." The primary drawback of these terms, though, is that they may suggest that the inter-

vention is being studied or is part of a research program when this may not in fact be the case. Furthermore, as the matrix above indicates, the effectiveness of accepted interventions can also be called into question, creating some of the same uncertainties that arise in the case of new but untested interventions. For these reasons, I agree with Levine's suggestion that the term "nonvalidated" should be used to refer to interventions or practices whose evaluative status is uncertain or not adequately validated. R.J. Levine, *Ethics and Regulation of Clinical Research,* 2nd ed. (New Haven, Conn.: Yale University Press, 1988, pp. 4-5; M.F. McKneally, "Ethical Problems in Surgery: Innovation Leading to Unforeseen Complications," *World Journal of Surgery* 23 (1999): 786-8, 787. However, I suggest that the term "investigational" should be used only when such an intervention is the subject of a research study. This may help to forestall unnecessary linguistic confusion.

14. J.L. Meakins, "Innovation in surgery: the rules of evidence," *American Journal of Surgery* 183 (2002): 399-405; Margo, see note 1 above; Frader and Flanagan-Klygis, see note 2 above; B. Reeves, "Health-technology assessment in surgery," *Lancet* 353 (suppl. I, 1999): 3-5; see note 8 above.

15. B. Freedman, "Placebo-Controlled Trials and the Logic of Clinical Purpose," *IRB* 12, no. 6 (November-December 1990): 1-6, 5.

16. Meakins, see note 14 above, p. 401.

17. A.J. London, "Clinical Equipoise: Foundational Requirement or Fundamental Error?" in *Oxford Handbook of Bioethics,* ed. B. Steinbock (New York: Oxford University Press, 2006).

18. The axes on figure 2.4 should not be understood as requiring a particular cardinal scale on which alternatives can be ordered. Obviously, if such cardinal information is available, then it should be used. If it is not available, however, the figure should be seen as representing only ordinal information. For a detailed account of the decision theoretic machinery necessary to deal with the full range of complexities presented by actual cases, see I. Levi, *Hard Choices: Decision Making Under Unresolved Conflict* (New York, N.Y.: Cambridge University Press, 1986).

19. See note 17 above.

20. A.J. London and J.B. Kadane, "Sham surgery and genuine standards of care: Can the two be reconciled?" *American Journal of Bioethics* 3, no. 4 (2004): 61-4.

21. P. McCulloch et al., "Randomized trials in surgery: problems and possible solutions," *British Medical Journal* 324 (2002): 1448-51; Meakins, see note 14 above; Reeves, see note 14 above; M. Baum, "Reflections on randomized controlled trials in surgery," *Lancet* 353 (suppl. I, 1999): 6-8.

22. M.A. Hlatky et al., "Comparison of Predictions based on observational data with the results of randomized controlled clinical trials of coronary artery bypass surgery," *Journal of the American College of Cardiology* 11, no. 2 (1988): 237-45.

23. R.L. Lilford et al., "Trials and fast changing technologies: the case for tracker studies," *British Medical Journal* 320 (2000): 43-6.

24. McCulloch, see note 21 above, p. 1450.

25. D.J. Spiegelhalter et al., "Bayesian methods in health technology assessment: a review," *Health Technology Assessment* 4, no. 38 (2000): 1-130.

26. K. Benson and A.H. Harz, "A comparison of observational studies and randomized controlled trials," *New England Journal of Medicine* 342 (2000): 1878-86; J. Concato, N. Shah, and R.I. Horwitz, "Randomized controlled trials, observational studies and the hierarchy of research designs," *New England Journal of Medicine* 342 (2000): 1887-92.