Epistemic Humility and Medical Practice: Translating Epistemic Categories into Ethical Obligations

Abraham P. Schwab
Associate Professor, IPFW

Abstract

Physicians and other medical practitioners make untold numbers of judgments about patient care on a daily, weekly, and monthly basis. These judgments fall along a number of spectrums, from the mundane to the tragic, from the obvious to the challenging. Under the rubric of evidence-based medicine, these judgments will be informed by the robust conclusions of medical research. In the ideal circumstance medical research makes the best decision obvious to the trained professional. Even when practice approximates this ideal, it does so unevenly. Judgments in medical practice are always accompanied by uncertainty, and this uncertainty is a fickle companion—constant in its presence but inconstant in its expression. This feature of medical judgments gives rise to the moral responsibility of medical practitioners to be epistemically humble. This requires the recognition and communication of the uncertainty that accompanies their judgment as well as a commitment to avoiding intuitive innovations.

Keywords: epistemic humility, uncertainty, medical practice

Running Head: Epistemic Humility and Medical Practice

An Unexceptional Case

She was two months pregnant when she arrived at Crown Street Hospital Emergency Room. The longer her uncontrollable vomiting continued, the greater the risk of miscarriage. Having tried all other reasonable medications, Dr. William McBride considered using a new medication. Intended to be used as a sedative during labor, he consulted the packaging to confirm that the medication was safe earlier in pregnancy. On finding confirmation, he administered it and, almost miraculously, the woman stopped vomiting. Little did he know that giving her this particular medication carried substantial risk of harm. Little did he know that the
successful use of Thalidomide would lead to the morphological disfiguration of thousands of children while in utero (Stephen and Brynner, 2001, 22).

Thalidomide was never officially approved for use in the United States (William McBride was a physician in Australia), and yet, the case of William McBride, while exceptional in its ominous foreboding, represents two issues that are in the background of every medical encounter. First, medical judgments are structured by uncertainty. Although it takes a number of forms and arises in a number of spaces, uncertainty is medical judgment’s constant companion. Second, the uncertainty of medical judgments imports a complementary set of professional risks. Poorly handled, uncertainty can keep medical professionals from fulfilling their responsibilities to make sound recommendations to patients and to secure each patient’s informed consent.

Discussing the appropriate means for addressing uncertainty in general will lay the groundwork for addressing the particularities of uncertainty in medical practice. The following is split into three sections. First, a few words about epistemic humility as the appropriate response to uncertainty followed by a recommendation of reliabilism as the appropriate standard for determining what epistemic humility means for a particular instance. In the second section, identifying the guideposts along the spectrum from the wildest of guesses to the most reliable predictions will be the primary task. In the third section, the risks that this uncertainty carries are described and recommendations (strangely absent from professional standards) for handling these risks are made.

**Epistemic Humility**

Epistemic humility is a characteristic of claims that accurately portray the quality of evidence for believing the claim to be an accurate one. An internist likely holds two independent beliefs: that penicillin will cure the syphilitic patient and that a diuretic will lower the blood pressure of the hypertensive patient. In both cases, there is support for the belief and it might even be said that the physician has good reason to believe both. And yet, the levels of support for

---

1 My reason for favoring the term “epistemic humility” over the term “epistemic confidence” arises from the literature on overconfidence in judgment. As decision-makers make judgments about difficult tasks, they tend to be overconfident and when they make judgments about relatively easy tasks they tend to be underconfident. Because I take the judgments that physicians make in the clinic to be quite complex, it seems the bias will most often be one of overconfidence. Nominal evidence for this view can be found in Baumann (1994). Accordingly, the appropriate corrective is to urge for more humility not more confidence.
these beliefs differ. Penicillin will cure the syphilis, but a diuretic by itself is not effective in lowering blood pressure in most cases (ALLHAT, 2002). Epistemically humble claims are those that accurately portray this difference.

Accordingly, epistemically humble claims are inconsistent in their modesty. On the one hand, any claim about the effectiveness of a diuretic to lower blood pressure should be recognized to have nominal warranted confidence. To claim that the diuretic will always lower the blood pressure is to overreach. On the other hand, epistemically humble claims are not always particularly modest. To doubt that the penicillin will cure the syphilis is to underreach. In some cases a claim could hardly be overconfident and in others it would be hard for it to be too modest.

Up to now, I have referred to “epistemically humble claims.” And yet, humility tends to be discussed as a characteristic of individuals (Louden, 2007; Grenberg, 2007; Whittle, 2006)\(^2\) and there is a general tendency to discuss the epistemic virtues as a characteristic of individuals (Flood, 2008; Corlett, 2008; Brady and Pritchard, 2006). And yet, I contend it is better to think of epistemic humility as a characteristic of claims.\(^3\) It is not necessary here to countermand the emphasis on the characteristics of individuals as a fruitful one and one that has been beneficially applied by medical practitioners.\(^4\) And yet, three aspects of attributions of character to individuals indicate the need to emphasize aspects of particular claims instead: attributions of character are difficult to make, do not do all of the necessary work, and individuals are regularly pushed to do things that are “out of character.”

Attributing the characteristic of epistemic humility to individuals faces substantial obstacles. Looking again to Dr. McBride, how would we determine if he was epistemically humble?

---

\(^2\) Much of the recent conversation about humility concerns the humility that is required to adopt an appropriate Kantian perspective about what we can know about things-in-themselves. This discussion is structured around “our” humility or the humility that attaches to the appropriate world-view, both of which emphasize humility as a generalized characteristic and not something attributable to particular claims.

\(^3\) Nonetheless, to describe a claim as epistemically humble seems counter-intuitive. It might be better to describe such claims as either scientific or pseudo-scientific. That is, the claims are governed by the principle of epistemic humility when they are scientific and they are not when they are pseudo-scientific. Pseudoscience is not simply claiming something to be supported by science that is not. For example, someone who runs a clinical trial haphazardly or mismanages the data from a registry will produce conclusions that are not well supported. Much as it is a mistake to claim these conclusions are well-supported, these are mistakes in execution that produce a mistaken belief. Concerns about pseudoscience, however, focus on different kinds of mistakes: mistakes in categorization. That is, pseudoscience fails to take seriously the distinctions between the different bases for knowledge.

\(^4\) Indeed, this project is complimentary to such a view. Imagine for example a physician who aims to be epistemically humble. This physician would try to avoid the extremes of overreaching and underreaching and find the median between these extremes. The remainder of this article provides guidance for how such a physician could achieve that mean.
humble in the first place? What percentage of Dr. McBride’s utterances would need to be epistemically humble to characterize him as epistemically humble? Every other one? Three out of every four? All of them? The most exacting standard is tidy, but it has the distinct disadvantage of being unattainable. That is, individuals are unable to always make perfectly calibrated claims. But if any other standard is used, there is the additional problem of what to make of problematic utterances. If three out of four utterances are epistemically humble, what sense can be made of the fourth utterance? Such standards for epistemic humility seem to face the dilemma of being either unattainable because they demand perfection or unhelpful because they do not indicate anything about this utterance.

As such, attributing epistemic humility to individuals does very little work. Characterizing Dr. McBride as epistemically humble will not be helpful if the standard is anything less than perfection. And yet, if it is not perfection, the standard would not help in any particular case. A judgment would still need to be made about the claim being made in a particular case. In the unlikely scenario that this dilemma between attainability and usefulness can be resolved, attributions of character will face additional obstacles on the individual and the social level.

On the individual level, character bends when put under conditions of stress. Studies have been illustrating this point for half a century, from the famous Milgram experiments (1963) where someone in a position of authority pushes an average person to torture a stranger to Darley and Batson’s (1973) study where the pressure of time kept seminary students from helping someone who appeared to be in cardiac distress. Robust evidence continues to illustrate that individual characterizations do not predict action. While these examples do not preclude the possibility that character has some effect on the actions of an individual, it is enough to show that character will be an unstable basis for attributions of epistemic humility. Pressures of time, accidents of fortune, and the other conditions of human activity undermine character as a predictor of actions across time and circumstance.

**Standardizing Epistemic Humility**

For epistemic humility to be a part of routine medical practice, it needs a tractable method for determining warranted confidence and identifying the tipping point between

---

5 For summaries of this work see (Bishop, 2000; Schwab, 2008a).
overreaching and underreaching. As argued elsewhere (Schwab 2008a) reliabilism is an attractive epistemic perspective for medical practice. In short, reliabilism distinguishes between beliefs by evaluating the reliability of the process used to produce those beliefs. For example, individuals may have a belief about whether or not they are HIV+ based upon on a number of processes. The Western Blot or the Enzyme-linked immunosorbent assay (ELISA) methods are very (though not perfectly) reliable belief-forming processes about one’s HIV status—these methods reliably produce claims that are very likely to be later confirmed as accurate. The raw facts that one has had unprotected sex or that one has a friend who is seropositive are much less reliable belief forming processes—the claim that an individual is seropositive using these methods will be much less frequently confirmed as accurate. Accordingly, beliefs based upon the reliable processes are to be preferred. Most important for concerns about epistemic humility, identifying the reliability of the process that formed a particular belief is obligatory. Through awareness of the reliability of the processes used to form the beliefs they have about medical practice, physicians can determine the appropriate level of confidence in specific beliefs and so make epistemically humble claims. Put another way, reliabilism provides a tractable mechanism for handling the uncertainty that dogs medical practice.

A number of variations of reliabilism exist, but what all share is the view that one should hold only those beliefs arising from a process that produces reliable predictions (Goldman, 1986; Bishop and Trout, 2002). For this reason, reliabilism has a couple of obvious advantages for application to medical practice. First, it allows for systematic evaluation of multiple methods of producing predictions lacking a requirement of a single source or type of source for claims to know something. This is important because the inputs for medical practice necessarily arise from divergent sources. These range from the controlled conditions of the clinical trial to the realism of the observational study to the intuitions of the clinician’s judgment. Second, reliabilism does not require understanding of the underlying processes that produce beliefs. Such a permissive epistemic perspective matches neatly with some of medicine’s existing epistemological techniques. Take, for example, the prescribing of bupropion for smoking cessation. Here, the mechanism by which bupropion works remains unknown, but there is robust evidence that it is substantially more effective than nothing at all or even a physician’s imploring (Roddy, 2004).

---

6 For a more robust explanation and analysis of the use of reliabilism for the epistemology of medical practice, see Schwab (2008a).
Although it may be preferable to know the mechanism for a process as well, it is enough for medical practice to have the reliable prediction.

Types of Uncertainty in Medical Practice

The uncertainty that dogs medical practice has two heads: uncertain values and uncertain applications of medical practice. The remainder of this essay will focus primarily on the latter, but first a few words about the former. Uncertain values are exemplified by the varying preferences of many patients and physicians. Jehovah’s Witnesses are a classic example of a group with divergent preferences from the rest of the population. Informed consent is clearly intended to resolve uncertainty about values in favor of the patient. That is, rather than attempt to definitively settle questions of financial value or worthwhile risks and benefits, medicine’s practitioners offer a set of recommendations, and patients are then allowed to choose among these or walk away (Schwab, 2008b). I take the process and forms of informed consent, controversial as they may be (Fisher, 2006; Brody, 1989; Veatch, 2009; Katz, 1994), to be a reasonable attempt to resolve the uncertainty of value; these issues will be left aside for this paper.

Uncertain Applications of Medical Research

Uncertainty in the application of medical research arises from the nature of medical science, the features of patient populations small and large, and the methods employed in medical research. Some diagnoses, prognoses, and treatment recommendations will be supported by the most robust of evidence, including clinical trials, observational studies, and corresponding meta-analyses. Others will be supported only by the currency of intuitions and the extrapolations of received knowledge. Both kinds of evidence are limited, but intuitions and received knowledge are more dangerous guides.

---

7 In this context, medical research is defined quite broadly. Images of the clinical trial and the lab come quickly to mind, and the application of statistical methods to observational studies and registries is increasingly important. Meta-analyses of these methods are also typical. Stopping there, though, ignores what I will call informal research. Informal medical research includes those methods that sidestep the formal requirements of clinical trials (e.g., rigid protocols) and the formal requirements of analyses of trials and practice (e.g., rigid methodological analysis). Informal research is exemplified when physicians or health care teams must use their theoretical understanding of medicine to use an untested treatment on an unresearched disease. On my view, then, informal research occurs when physicians leave the quiet confidence of clinical trials and statistical analysis. Even though it would be easy, for example, to think of off-label prescribing as going beyond the scope of medical research, on this view, off-label prescribing does not go beyond medical research; it simply uses different methods.
Take for example the state of medical practice during the mid 19th century. At the time, the theoretical underpinnings of medical practice were underdeveloped in a peculiar way—though human anatomy was quite well understood, its relation to human physiology was not. By providing structural information but failing to identify the role of that structure in function, theoretical understandings of disease and health were underinformed. The risks of infection from minute organisms and their effect on overall health went unrecognized. Though the received knowledge made “sense,” it led to widespread poor decisions, including, for example, the failure of physicians to wash their hands prior to assisting in childbirth.\(^8\) Think also of the case of William McBride. One of the features that may have contributed to his willingness to take the package insert at face value was an underlying belief in “placental protection”. On this view, the placenta acted as a barrier of protection around the developing embryo and fetus—its development insulated it from toxins in the maternal body. Though the idea of a protective sac protecting the fetus is intuitively plausible, it turns out to be patently false. Had the epistemic standards for 19th century medical practice or for Dr. McBride been a rigorously followed reliabilism untold numbers of poor decisions and outcomes would have been avoided.

A Taxonomy of Uncertain Applications of Medical Research to Medical Practice

The clinical practitioner finds him or herself in a position typified not only by uncertainty, but varying degrees of it. Uncertainty in the applications of medical research fall along a spectrum from the robustly supported to the unsupported. In what follows, I outline the general features of three areas along this spectrum: its limited precision, its limited scope, and its undisciplined application. Together, they provide a thumbnail sketch of the varying degrees of uncertainty that attend the application of medical research to medical practice.

Limited Precision: Uneven Application in the Defined Domain

---

8 The shift to include the simple step of washing hands was actually ignored for nearly 20 years. This may be because there was no readily available explanation for why washing hands would improve the mother’s health post delivery. As Trout (2002) has explained, this may be due to the undue weight that scientists and lay people alike put on the “sense” that an explanation makes.
As a fixture of contemporary medicine, Evidence Based Medicine (EBM) continues to be critiqued, but its basic thrust remains unabated. Medical practice based on systematically gathered, organized, and analyzed data is better than the alternatives. To achieve this goal formal medical research must be restricted in ways foreign to medical practice: protocols are usually inflexible and at times patients with the diseases/symptoms/conditions in question must be turned away. The payoff: conclusions reached within the defined domain of formal research can be held with higher levels of warranted confidence. Predictions of treatment success for patients who fall within the defined domain (that is, patients in the clinic who meet the same criteria as subjects/participants in the trial) will be more reliable.

And yet, even within the defined domain, medical research makes uncertain predictions about medical care. Take for example the proven treatments for pre-hypertensive or stage 1 hypertensive patients. As noted above, despite the robust evidence for its inexpensive effectiveness (ALLHAT, 2002), a diuretic simply will not be effective for some patients. That is, some patients will not receive any (or not enough) benefit from the treatment.

Such uncertainty has two layers. In the first layer, the research predicts the effectiveness of the treatment (i.e., the percentage of patients who respond to diuretics). In the second layer, the research’s predictions carry a degree of uncertainty that is represented as a p-value and a confidence interval. Put roughly, these two aspects of statistical analysis indicate a range of doubt about a treatment’s stated effectiveness. What this means in medical practice is that any time a physician considers a treatment’s effectiveness and discusses it with colleagues or patients, the stated effectiveness carries with it an additional range of uncertainty. Diuretics will be effective for a certain percentage of patients, or at least, that’s apparently true.

Limited Scope: Uncertain Application at the Ragged Edge of the Defined Domain

For example, Straus and Haynes (2009) provide an excellent critique of the glut of research that is propagated upon physicians. They draw attention to the difference between the existence of an evidence base for practice and a practice that uses that evidence base.

The defined domain should be read as directly as possible. That is, rather than use any other guidepost the defined domain should be determined on a study-by-study basis. For individual studies this will be quite straightforward, but for any analysis that incorporates studies with divergent inclusion and exclusion criteria or other substantial differences, the definition will be, well, less defined. In these cases, whether the research provides an adequately defined domain or produces a more ragged edge (as described below) will have to made on a case-by-case basis. This, however, is not a problem for this taxonomy or the thrust of this manuscript. The point here is not to produce a comprehensive taxonomy to neatly categorize all types of medical research, but to provide a rough structure within which medical research can be located.
Within the defined domain, independent variables are eliminated or attenuated. When patients in practice no longer fall neatly within this defined domain, when they have not only the focal diagnosis but also complicating conditions, they fall along the ragged edge. Here there is a mingling of confident application of the conclusions of formal research with extrapolations beyond the parameters of research protocols. This mingling leaves practitioners in a position of middling confidence. Yes, the patient has the focal diagnosis, but the patient also has characteristics placing them outside the defined domain of application.

Two components of the diminishing confidence along the ragged edge are illustrated by a disease management program (DMP) investigated by Sindaco et al (2007). Sindaco et al (2007) analyze the effects of a structured DMP on elderly patients following heart failure. The particular DMP they evaluated includes "discharge planning, education, therapy optimization, improved communication, early attention to signs and symptoms." (Sindaco et al, 2007, 325) Usual care (the control) included whatever various treatments and services were ordered by the patients’/research subjects’ primary care physician and/or personal cardiologist. They hypothesized that their DMP would work more effectively than usual care at breaking the cycle of admission and discharge, would improve overall health, and would save money. The study lasted two years with earlier end points for either death or hospital admission for heart failure. As it turned out, subjects treated with the DMP had a higher quality of life and functional status than subjects treated with usual care. Further, the DMP also proved to be more cost-effective at a rate of about 1,000 euros per subject.

Despite the data suggesting that the DMP is both clinically and cost-effective, the application of this study to medical practice has two significant limitations. First and foremost the exclusion criteria of this study, as in all protocol analyses, produce certain obstacles to application. In this case, patients with "coexisting non-cardiac illness likely to reduce life expectancy" were excluded. (Sindaco et al, 2007, 325) As much as this exclusion increases warranted confidence in conclusions about cost (and even clinical) effectiveness, the population of the study (the elderly population) is more likely than average to have such coexisting and complicating conditions. Accordingly, this restriction handcuffs the application of the study—DMPs limited to elderly patients who have suffered a heart attack but have no complicating conditions will have a narrow domain for confident application. This leaves open the question of effectiveness for patients at the ragged edge of this domain.
Second, the two-year time frame of the study obfuscates evaluations of cost-effectiveness. Sindaco et al illustrate a substantial cost-savings per patient, but this was likely a corresponding result of the DMP’s clinical effectiveness. Because the DMP decreased the chance that subjects would be readmitted to the hospital for any reason and maintained higher functional status (Sindaco et al, 2007, 327), these subjects were less likely to spend resources during the two years of the study. This does not mean that these individuals not return to the hospital at a time after the study has concluded. The money may not actually be saved, but simply deferred to a later time. To put this point another way, the study cannot tell us about treatments that occur beyond its temporal parameters.

Similar limits on confident application exist in other areas of medical research as well. For example, it is not uncommon for research into behavioral and psychological treatments for depression and other psychiatric disorders to exclude potential subjects who have suicidal ideation or major depression in the recent past—common characteristics of patients to whom the research would apply.

Putting this point more generally, the parameters of a study limit the confidence of its application in practice. For the present paper, it is enough to recognize the substantial difference between applications of formal research within a neatly defined domain and the applications of research at the ragged edge.

Before moving on to the third guidepost in this outline of medical uncertainty, it should be noted that some of the uncertainty of the ragged edge can be eliminated by expanding the defined domain. Specifically, the use of observational studies and registries to supplement and, in some cases, supplant, clinical trials continues to extinguish some uncertainty in practitioner judgment by expanding the variables accounted for within medical research’s confident conclusions. More will be said about this in the final section of the paper.

Undisciplined Application: Medical Practice beyond the Ragged Edge

Medical practice entirely beyond the domain of clinical trials and observational studies returns us to the uncertainty of medical practice from an earlier time. In such cases, medical
practice is much less the application of proven techniques, and more the intuitive and innovative use of known principles or theories.\(^{11}\)

Nonetheless, there are two ways the decisions made outside the defined domain enjoy an advantage over the medical practice of centuries past. First, the background theories are more robust and more accurate today. This limits the kinds of mistakes that can be made through responsible application of these theories. Second, the set of proven therapies is much larger. This limits the areas in which mistakes can be made. All that said, when practitioners apply medical research outside the defined domain warranted confidence all but disappears.

Atul Gawande’s “Case of the Red Leg” (2002) is instructive here. In short, a young, athletic woman presents with what may be a severe case of cellulitis, but Gawande, called in as a surgical consultant, worries that it might be necrotizing fasciitis (“flesh-eating bacteria”). Gawande explicitly recognizes the likely bias of availability in this worry—he saw a case of necrotizing fasciitis just a few weeks before. Nonetheless, it turns out that he is right.

The standard practice in such cases is amputation—this young woman should have lost her leg. But, as Gawande (2002, 244) puts it, she was just so young and there may have been “a purely emotional unwillingness to cut off the limb of a pretty twenty-three-year-old.” So instead, they attempted an untested treatment: first, to “debride” the wound by taking out the diseased tissue and second, to use a hyperbaric oxygen chamber, in which patients are put into a pressure chamber to increase the oxygen in their tissue and thereby assist the body as it fights off infection. To recommend such a treatment is a clear example of an undisciplined application of medical research.\(^{12}\) As Gawande (2002, 244) describes the thought process: “it was partly instinct again, an instinct that her youth and fundamentally good health might allow him [the surgeon] to get by with just removing the most infested tissue.”

---

\(^{11}\) Another way to describe the difference between undisciplined application and application within the defined domain is the difference between human fallibility and reliable processes. That is, neither the robust conclusions of medical science nor the ineffable judgment of medical professionals will be perfectly reliable—the expected result will not always come about. The question, then, is which strategy warrants greater confidence in general. Adler (2005) has taken fallibility to be compatible with knowledge. He goes further to argue that the use of fallible human judgment is preferable to a process that has a higher truth ratio. Without the burden of a detailed textual analysis, the conclusion he reaches is that fallible human judgment at least has the possibility of producing knowledge because of its understanding of the underlying process. As noted in the early section on reliabilism, however, medical practice is unencumbered by the need to understand the underlying process and so, this argument is unpersuasive in this context. Medicine foregoes a world of knowers (or at least relegates them to the background) for a world of practitioners whose recommendations produce better health outcomes.

\(^{12}\) To be undisciplined in this context is less to be unruly than simply be uncontrolled. Here discipline is not punishment but a controlling of activities. To be undisciplined is to ignore the controlling influence of medical research as a guide for medical practice.
This case illustrates undisciplined application of the conclusions of medical research beyond the ragged edge of the defined domain. As an intuitive innovation, this judgment, this recommendation goes beyond the confident conclusions of medical research. Not a blind guess, it represents that which appears intuitively plausible. And yet, by making the decisions they did, they put the girl at risk for even greater harm. At the time of the treatment they would have needed to amputate either just below or just above the knee. If the innovative treatment had failed to stop the progression of the disease, the amputation may have cost the patient her entire leg, perhaps even her life.

Moreover, the use of ‘young’ and ‘healthy’ as criteria for judgment in this case highlights some of the particular dangers of intuitive innovations in medical practice. Gawande and the rest of the health care team could not have known if these criteria would be relevant to the patient’s recovery or the risks to which the patient would be subject. Intuitive criteria can be profoundly misleading. What if she had been obese? There is some evidence that surgeons are too quick to use obesity as an excluding criteria for surgery for some patients (Everett et al, 2003). Given Gawande’s characterizations, as the case presented itself, the bias may have pushed the other way—that youth and vitality would provide a certain kind of protection, perhaps like the mythical placental protection of Dr. McBride’s time.

So long as there are expansive spaces beyond the reach of medical research, areas where medical practice must practice with limited guidance, these judgments beyond the ragged edge are unavoidable. For the foreseeable future, medical practice will include judgments with the minimal confidence that accompanies undisciplined application. In the following sections, this taxonomy’s implications for medical practice will be teased out. Specifically, risks associated with these uncertainties will be identified and recommendations for avoiding attenuating these risks will be identified.

**Risks of Uncertainty**

13 Although there is not space to pursue a full treatment here, it should be noted that the biases of human judgment can compound the uncertainties of the application of medical research to medical practice. Specifically, within the defined domain, predictable overconfidence (Moore and Healy, 2008) may lead practitioners to have confidence in the effectiveness of treatments that outstrips the evidential support. At the ragged edge, overconfidence, the confirmation trap, (Wason, 1960) and hindsight bias (Fischhoff, 1975) all skew judgment. Outside the defined domain, the already mentioned biases are exacerbated by the absence of well-grounded guideposts for practitioner judgment.
This taxonomy of uncertainty in medical practice is not idle. The varying degrees of confidence impregnate the judgments of medical practice with risks of poor medical advice and failures to procure patients’ informed consent. Only through accurate diagnosis and careful extraction of uncertainty can these risks be eliminated or attenuated.

Professional Judgment

Physicians’ fiduciary responsibilities to their patients require them to work on their patient’s behalf. These responsibilities are often codified as beneficence, non-maleficence, justice and autonomy. If uncertainty hides in the shadows of medical judgment, it can make of physicians irresponsible fiduciaries. One cannot reliably do good or avoid harm if their confidence in their judgments is uncalibrated. One cannot distribute resources reasonably or equitably if there is unacknowledged or unknown uncertainty. Those moments when warranted confidence and actual confidence diverge, when uncertainty is compounded by ignorance, medical practitioners can utterly fail their patients in their judgments about risks, benefits, and justice.

This possibility of failure is compounded when the lone practitioner is part of a health care team. On the group level, when a treatment plan is organized, the only issues that will be considered are those issues explicitly raised during any group discussion. As a generalist or specialist presents his or her thoughts about a case, others on the team will use these thoughts to produce their own conclusions. Just in case these physicians do not recognize or communicate the levels of uncertainty in any conclusion they present, the uncertainty of the first specialist will be compounded by the uncertainty of the second physician. In such circumstances, the sum may become less than its parts. If the uncertainty of these initial judgments is not adequately communicated, all future judgments for which these initial judgments are foundational will be inadequately informed.

Informed Consent

If physicians blithely make judgments covered in the oil of uncertainty, their judgments and recommendations may run counter to their fiduciary responsibilities. Further, when patients are uninformed about a practitioner’s uncertainty, they are kept from pursuing their own

---

14 Patient autonomy is protected through informed consent procedures, which will be discussed in the next section.
interests. There are many conceptions of informed consent and many standards that are recommended for evaluating whether or not informed consent has occurred (Waller and Repko, 2008; Jones, McCullough and Richman, 2005). Despite this competition among conceptions, they share the minimum requirement that physicians provide some accurate information regarding the risks and benefits associated with a treatment recommendation. If a practitioner fails to acknowledge and communicate the epistemic status or uncertainty of his or her recommendations, he or she fails to meet these minimum conditions. Importantly, the risks arising from such failures increase as the treatment recommendations stretch beyond the ragged edge of the defined domain.

Such failures to procure informed consent due to a failure of epistemic humility may be worse than failures of moral humility. That is, paternalism—the circumventing of patient decision-making for the patient's benefit—often results from a failure of moral humility. Physicians or some other medical practitioner has failed to recognize the limits of his or her own view of a situation. Unchecked paternalism is proscribed in medical ethics because it so easily fails to take seriously the values of the patient, the pluralism of society, and the right to control one’s body. And yet, this fundamental mistake may be less undesirable than a failure to make epistemically humble claims. While medical paternalism should generally be avoided, failure to make epistemically humble claims compound paternalism’s failure to get consent by subjecting patients to unknown risks for unknown benefits. This is not failure to get consent in the hopes of serving the patient’s interests—this is a failure to get consent in the ignorance of uncertain risks of benefits. The goals of medicine and good medical practice may remain controversial, but not a single set of recommended goals endorses subjecting patients to unknown risks for unknown benefits without their consent.

A Lack of Professional Guidance

Despite the substantial risks that come with unknown and unacknowledged uncertainty, the advice of professional associations is surprisingly sparse and apparently misguided. The guidance that is offered includes no cautionary note about variable levels of uncertainty or what such uncertainty might mean for professional judgment or informed consent. Instead, the opinions offered regarding the use of intuitive innovations fall into two basic camps: the protective and the infective.
Professional associations are protective of innovation in medical practice, reminding members of medical community, policy makers, and the broader public that a cornerstone of good medical practice is the freedom of medical practitioners to exercise their intuitive judgment. The American Medical Association argues that: “Physicians should prescribe drugs, devices, and other treatments based solely upon medical considerations and patient need and reasonable expectations of the effectiveness of the drug, device or other treatment for the particular patient.” (AMA CEJA Opinion 8.06 Prescribing and Dispensing Drugs and Devices, emphasis added) They do not go on, however, to define what counts as a reasonable expectation, and as we shall see below Great Britian’s General Medical Council takes a similar position of protection without caution.

The American College of Surgeons has a “Statement on Emerging Surgical Technologies and the Evaluation of Credentials,” [ST-18] which includes the following: “The development of a new technology must be accompanied by a scientific assessment of safety, efficacy, and need. Both the rigor and scope of an assessment will depend, to some extent, on the novelty and complexity of the technology. The assessment process may range from carefully monitored observational studies with evaluation to controlled clinical trials.” (ACS, 2008) Admirable as it may be to identify that the levels of assessment should vary with the type of innovation in play, what is lacking here and in the rest of the statement is any mention of communication to patient or any recognition of varying levels of uncertainty. While it may be true that some innovative surgeries do not include substantial changes from well-established practice, this does not eliminate uncertainty or the possibility of diminished effectiveness.

Professional associations also encourage the infective spread of practitioners’ intuitive and innovative judgments. These organizations provide reminders to members of the medical community (and its sub-communities) that they have a responsibility to share their novel ideas with others. The motivations here are likely reasonable—it would not do to have a community of professionals dithering over worries about intellectual properties while their patients go unmended. As the AMA puts it: “Physicians have an obligation to share their knowledge and skills and to report the results of clinical and laboratory research. . . This tradition enhances patient care, leads to the early evaluation of new technologies, and permits the rapid
dissemination of improved techniques.”\textsuperscript{15} (AMA CEJA Opinion 9.08 New Medical Procedures, emphasis added) They go on: “The intentional withholding of new medical knowledge, skills, and techniques from colleagues for reasons of personal gain is detrimental to the medical profession and to society and is to be condemned.” (AMA CEJA Opinion 9.08 New Medical Procedures)

Here the problem is the definition of research, and so the determination of what should be shared, goes unarticulated.

A lack of attention to research about the predictable biases of human judgment runs through both the protective and the infective recommendations of professional associations. The advice of Great Britain’s General Medical Council is illustrative. In their public statement on ‘Good Practice in Prescribing Medicines’ they state:

“When prescribing a medicine for use outside the terms of its licence you must:

- \textit{Be satisfied} that it would better serve the patient's needs than an appropriately licensed alternative
- \textit{Be satisfied} that there is a \textit{sufficient evidence base and/or experience} of using the medicine to demonstrate its safety and efficacy.” (GMC, 2008, emphasis added)

To ‘be satisfied’ is not defined further. At face value, this guidance admonishes practitioners to be intuitively satisfied with the recommendation. That is, the recommendation is to be sure that the prescription satisfies informal research standards (despite their incumbent uncertainty). As noted above, part of the value of evidence-based medicine is its move away from individual judgments with the inherent limits on sample size and memory and the biases of overconfidence and confirmation. Invoking the value of intuitive insight is only plausible because so much of medical practice happens outside the defined domain of medical research. Accordingly, one might expect to find matching exhortations to avoid such judgments when possible or to limit the

\textsuperscript{15} Best guesses are important to spur on research, but they must be acknowledged as just that: a guess. Because they spur on research and carry the promise of improved medical care, it is not surprising that existing professional guidance works to protect and foster such guesses. Nonetheless, it is surprising that there is no guidance on how these guesses should be used in practice.
use of such judgments, and to communicate the level of uncertainty such judgments include. But there are none.

Recommendations

Detailed recommendations for addressing uncertainty in medical practice will best be informed by specific professional associations and the particularities of their specialties. For example, the features of uncertainty in specialties that use technical skills more substantially (e.g., surgery and orthopedics) may very well be different than the features in specialties that use cognitive skills more substantially (e.g., internal medicine or pediatrics). Nonetheless, a few general features of such standards can be described. Specifically, recognition and communications of uncertainty and the avoidance of intuitive innovations are obligatory.

Recognize Uncertainty

Given the features of medical research’s relationship to medical practice, a commitment to epistemic humility requires, first that physicians and other practitioners recognize the warranted confidence of their conclusions. It would be unreasonable to ask physicians to estimate a precise p value and confidence interval for all their diagnoses, prognoses, and treatment recommendations. Nonetheless, the rough differences in warranted confidence between robustly supported, minimally supported, and unsupported claims should be recognizable. And yet, the biases of human judgment present some predictable and substantial obstacles to this recognition. To begin, predictable overconfidence plagues decision-makers in general (Moore and Healy, 2008). Baumann et al (1991) showed that medical practitioners are not immune. Particularly insidious, overconfidence can keep the well-intentioned physician from an accurate assessment of their evidential support.

This overconfidence can be compounded by the confirmation trap. When practitioners have a history of overconfident judgment—a history of failing to accurately calibrate the confidence of their judgments, the confirmation trap comes to bear. It is well-documented that decision-makers at all levels (from naïve to expert) attend to new evidence according to the

---

16 Some have argued that this study does not show what it claims to—in fact if both sets of physicians had correct treatment recommendations then their confidence is vindicated. This misses the point that if multiple treatments would achieve the same result, the confidence a practitioner has in any of these treatments should be lessened accordingly. Their recommendation cannot be THE right recommendation.
conclusions they have previously drawn (Lord, Ross, Lepper, 1979; Nickerson, 1998; Jonas et al., 2001; Pronin, Gilovich and Ross, 2004). As a general rule, decision-makers are likely to ignore or discount evidence that runs counter to previously reached conclusions. Though there is some debate over how problematic this bias is in general (Klayman and Ha, 1987), it, along with overconfidence, can be particularly problematic in medical practice.

In the everyday decisions of everyday life, the confirmation trap may lead to limited harms. That a professor may be overconfident that he can produce a sharp analysis or provide a clear elucidation of difficult concepts during a class may be more beneficial than harmful—a lack of confidence may keep him from pushing the print button and may paralyze him in the classroom. In medical practice, however, the stakes are higher. Physicians who are too confident in their judgment or who fail to consider competing diagnoses and treatments can place patients in grave danger.

Both overconfidence and the confirmation trap can at least be attenuated. The most promising strategy for attenuating their effects is to consider the opposite (Hirt and Markman, 1995). By acknowledging and emphasizing the potential support for contradictory claims or conclusions (and, implicitly, the limited epistemic support for their claims), physicians put themselves in a better position to accurately recognize uncertainty.

Communicate Uncertainty

Part of the lingering hesitancy regarding the application of social science conclusion is that it is unclear what it means to the individual case that there is, say, a four in ten chance that a diuretic will lower blood pressure. In medical practice this is one reason to defer to the patient. Put the patient in the best circumstances available to make a good judgment about what serves the patients interest and then step back. Informed consent as a means to sidestep professional uncertainty. Any discussion of these risks without explicit acknowledgement of the underlying uncertainty cannot achieve informed consent.

To procure informed consent, the physician should, at a minimum, discuss the risks and benefits associated with the treatment. But if she discusses these alone, she fails to make epistemically humble claims. If she does not identify the epistemic status of the claim, she leaves it open to patient presumption. Such strategies can go wrong in a couple of ways. First, the skeptical patients may not believe that the treatment in question is particularly likely to lead to an
improved medical state. The patient may believe that there is less reason to be confident in the recommendation than is warranted. Second, a blindly trusting patient may be too ready to believe a treatment will be effective. Without indications from the physician that either thoroughgoing skepticism or blind trust is incorrect, patient understanding of the decision before them will be crippled. For example, many patients will be likely unaware of what an off-label prescription is and so without explicit communication, they cannot be aware of the wildly varying uncertainty that can be associated with different off-label prescriptions. Even with proven treatments, patients may be unaware that there is no certain result (risk or benefit) for the treatment that has been through multiple clinical trials. As such, attempts to achieve informed consent are undermined when the physician fails to make epistemically humble claims.

If a physician were to preface (or even interrupt) their discussions with patients with information about $p$ values and confidence intervals and an explanation of what these terms mean, it would be cumbersome and of unlikely benefit. Moreover, as discussed above, physicians will often be in positions that they don’t know these aspects of their conclusions or claims in any precise fashion. These facts, however, do not leave the physician with no means to communicate different degrees of confidence. Certainly, the best way to do this may depend on the preferred communication modalities of a physician, but some general features can be articulated.

At a minimum, differences in warranted confidence between well-established and robustly supported (i.e., clinical trials, observational studies, etc.) recommendations and relatively unsupported recommendations can be clearly communicated (e.g., “think of it as the difference between getting advice about how to fix your car from your dealership’s mechanic and getting advice from your next door neighbor who owns a car but has never done any work on it”). More nuanced distinctions could also be drawn between robustly supported treatments and single-trial supported treatments. (e.g., "think of it as the difference between a slam-dunk and a half-court shot").

Communicating uncertainty may be even more important for the health care team. The compounding effects of a failure to communicate uncertainty among health care workers give rise to the possibility of more substantial misjudgments. When a physician addresses the patient directly, any failure to communicate uncertainty will limit the patient’s ability to make an informed judgment. When a physician addresses other members of the health care team, the
failure to communicate uncertainty limits their judgments and in any case where those members of the team communicate with the patient, further limits the patients judgment.

Avoid Intuitive Innovations

From the overconfidence of Buamann’s research participants to Gawande’s fortunate, but unsupported, use of availability, I have provided a number of examples of the ways decision-makers can make predictably biased judgments. Given the robust successes of systematically gathered and analyzed data, intuitive innovations should be a choice of last resort. Moreover, given the various kinds of uncertainty associated with certain off-label prescriptions and innovative surgeries, these judgments require more extensive and explicit conversations with patients about what the uncertainty entails.

If physicians were to take the most conservative pose, intuitive innovations would be reserved for only those situations where a patient’s situation is such that doing nothing is worse than the extreme levels of uncertainty that attach to available alternatives. That an individual physician or even a group of physicians hold a belief is inadequate to justify its use. Indeed, it is the physician’s responsibility to refuse to offer treatments that they do not have good reason to believe are effective. As we saw in the opening sections, epistemically humble claims are those that are produced by reliable processes. Given the robust lack of reliability of intuitive judgments (because of overconfidence and the confirmation trap, among other biases), an intuitive innovation would rarely count as a good reason in medical practice.

And yet given the limited scope of its conclusions, formal medical research cannot be the sole arbiter of good medical practice. There is too much to be decided at the ragged edge and beyond the defined domain. This leaves physicians in the uncomfortable position of having limited information available from medical research and limited warranted confidence in their intuitive innovations. It seems some general guidelines for practitioners would be helpful here,

---

17 One of the difficulties with heuristics and biases imperative (Avoid Predictable Bias!) is that avoiding bias is not possible in some cases and may not be advantageous in others. For example, prospect theory has been used to illustrate that decision-makers change their choices based on the framing of information—when the choices are cast in terms of losses they are risk seeking, when they are cast in terms of gains they are risk-averse. One difficulty is that we cannot help but frame the decisions. Another difficulty is that there are competing gains (or losses) at stake in each decision. That is, framing as a gain in one situation implies a loss in the other. I might keep my full head of hair and feel okay for a while, but I might be killed by cancer 6 months earlier.
but I am in no position to provide recommendations for specific areas of medical practice. Nonetheless, I offer below some general features for these guidelines to model.

First and foremost, limit the need for innovative judgments. One way to put practitioners in good stead in this regard is to expand the defined domain. Movements in this direction are already evident from the increasing numbers and controversies surrounding observational studies and registries\(^{18}\) (Kirtane et al, 2009; Vandenbroucke, 2009), but increased funding is also needed to more comprehensively investigate effective applications of medical research. As it is currently structured the National Institutes of Health’s (NIH) budget puts more money into basic science research over applications of that research and the corresponding social science research. The current emphasis on translational research is an attempt, it seems, to address this problem without changing the structure of almost exclusive emphasis on basic science. The goals of translational research, while laudable in principle, lack the basic tools needed; namely a rosetta stone for translating basic science research into clinical applications. Accordingly, supplementing the emphasis on basic science with funding for the social side of medical practice will limit the need for individual innovation. Such funding should work to account for the effectiveness of known treatments at the ragged edge and beyond. Further, doctor-specific information about treatment effectiveness could be of substantial use in these ventures, particularly in technically oriented specializations like surgery.

Second, the standards to justify intuitive innovation are in inverse proportion to the need for intervention. Physicians should be unwilling to try innovative therapies unless the need warrants it.\(^{19}\) For example, there will be instances when intuitive innovation would be indefensible. When there is a known, effective treatment with limited side-effects, to use an off-label prescription with effectiveness that is less well-known would be an epistemic and moral failure. Intuitive innovations should not be universally proscribed, but given their different epistemic status than recommendations based on clinical trials, observational studies, etc, it should be curtailed as much as possible.

\(^{18}\) It should be noted that the increasing use of electronic medical records will make conducting observational studies and registries much easier and, in some cases, their conclusions more robust.

\(^{19}\) The definition of need remains elastic and elusive. While many physicians practice as though need is objectively determinable, Veatch (2008) has argued that need is subjective. Whichever view one adopts, the determination of need remains possible (either through evaluation of facts of elicitation of values) and should be used to determine the justification for innovative judgment.
Third and finally, when intuitive innovations are required, physicians should institute practices that will limit or attenuate the predictable biases of their judgment. There are a number of strategies discussed that relate either to the biases in general (Hirt and Markman, 1995; Vaughn and Weary, 2003; Koehler, 1991) or as they relate to medical practice specifically (Schwab, 2008c).

Conclusion

The practice of medicine is dogged by uncertainty. A constant companion, it is fickle. It varies from the limited to the unrestrained. And yet, it remains an elephant in the clinical consult. As I discuss above, this uncertainty suggest a taxonomy of medical claims, which in turn provides some clear recommendations for medical practice. The obligations to recognize and communicate uncertainty and to avoid intuitive innovation are the fruits of the pursuit of evidence-based medicine. And yet, they remain unspoken in the recommendations of professional associations of medicine. As such, this essay has shown the roots and implications of these obligations. Given the general structure of these obligations, however, this is the beginning of this analysis, and not the end of it.
BIBLIOGRAPHY


