
Enhancing the Use of Clinical Guidelines: A Social Norms Perspective

The McDonnell Norms Group

Advances in clinical investigation, data analysis, rapid dissemination, and rigorous evaluation of the findings led to the accumulation of medical “evidence.” This evidence now forms the basis of thousands of guidelines developed and promulgated by professional societies, safety and outcomes organizations, provider institutions, and regulators. With rare exception, these guidelines are inconsistently implemented or used.

This article reviews the history of guideline development and use, assesses the current state of implementation, identifies obstacles to adoption, and suggests strategies to overcome these obstacles. The major finding is that the current approach to development, dissemination, and encouraged use of guidelines is inconsistent with knowledge of psychology.

The major recommendations are that the approach to translation of evidence into practice be revised to address convenience; to respond to public and peer demand; to provide immediate feedback concerning performance; and to leverage existing incentives, transparencies, and accountabilities. Implementation of these recommendations is expected to augment safety and best practice in the health-care community.

Background

Clinicians have long recognized that there is a benefit to sharing knowledge. The tradition of clinicians synthesizing individual clinical experience into shared practice recommendations dates back at least 3,300 years. The Edwin Smith Surgical Papyrus was written around 1700 BC; it describes wounds and therapies. The Ebers Papyrus (Fig. 1), circa 1552 BC, not only describes diverse medical conditions but also recommends treatments. Still, until the mid-1960s, the basis for “standard practice” was transmitted through medical training, shared common wisdom, local customs, personal experiences,

and collegial opinion, forming the basis for “eminence-based medicine.”

After the US Great Society legislation, which guaranteed health care to the elderly (Medicare), the Congressional Office of Technology Assessment and the Institute of Medicine increasingly voiced the need for studies that would evaluate proliferating technologies in a meaningful and unbiased framework. The first efforts included creation of “clinical algorithms,” which were intended to guide both physicians and their “extenders” (nurse practitioners and physician assistants) in the proper triage and treatment of individuals with common medical disorders.¹ Such algorithms typically were expressed as paper flowcharts to which clinicians could refer when seeking guidance, but they were never well accepted by physicians and were largely relegated to triage use by nonphysicians in emergency or walk-in settings.

By the mid-1980s, the focus of academic evaluative studies began to shift from specific technologies to their effect on disease outcomes. This shift was less altruistic than pragmatic; physicians and surgeons were becoming fearful of the increasing attention and power of federal regulators. The federal attention to health services research was conveyed through the National Center for Health Services Research (NCHSR), which was authorized in 1967, and by 1980, was paying increasing attention to outcomes. In 1989, spurred by Wennberg and Gittelsohn’s² landmark research on national practice variations (that began in 1973 and continues through present day) and by RAND Corporation studies suggesting widespread inappropriate performance of common surgical procedures, Congress passed legislation with bipartisan support authorizing the successor Agency for Healthcare Policy and Research (AHCPR). This legislation more or less doubled the budget overnight for an agency that now had statutory responsibility for outcomes research and practice guideline development. In 1999, the Agency changed its name to the Agency for Healthcare Research and Quality. (Congress thought policy was its business, not an agency’s. Once the Agency relinquished its policy aspirations, Congress

Members of the McDonnell Norms Group are listed in the Appendix.

Received January 26, 2006; Revised March 15, 2006; Accepted March 15, 2006.

Correspondence address: Timothy G Buchman, PhD, MD, Washington University in St Louis, 4566 Scott Ave, Mailstop 8109, Saint Louis, MO 63110.

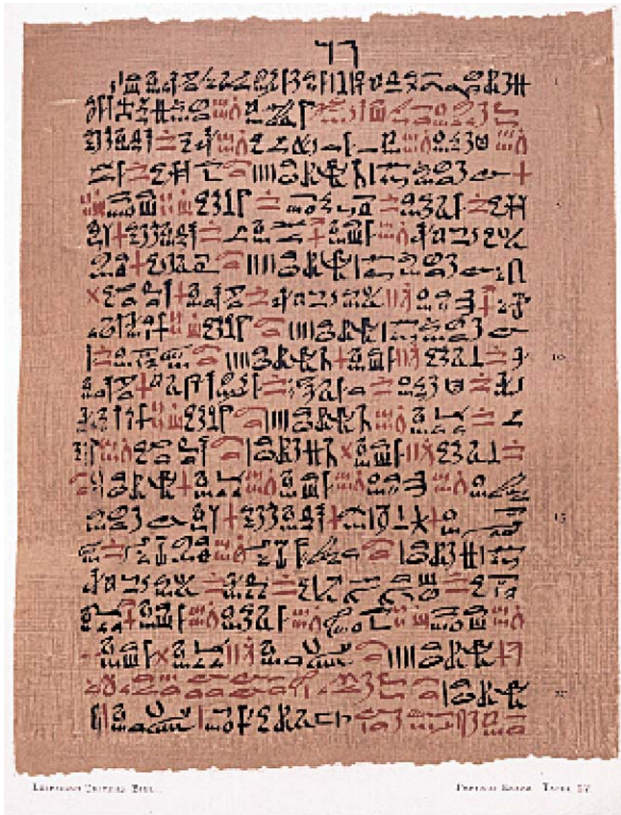


Figure 1. The Georg Ebers Papyrus. Found in Egypt in the 1870s, the Ebers Papyrus contains prescriptions written in hieroglyphics for more than 700 remedies, including this one for an acute asthma attack. From: University of Leipzig, with permission.

adjusted a previously flat budget to begin a meteoric rise.) Additional detail of the federal activities can be found in a lively and readable account given by Gray and colleagues.³

Given the accumulating data that variations in practice patterns could not (and still cannot) be justified by formal evidence,⁴ and also given the looming threat of federal practice guidelines, clinicians and their medical societies responded by creating and issuing their own clinical practice guidelines based on either expert consensus or, increasingly, on evidence from prospective studies in the literature.

Through the early 1990s, organized medicine began to embrace the connection among measured outcomes, guidelines based on those outcomes, and, more generally, the notion of evidence-based care.⁵ Development and dissemination of clinical practice guidelines created by organized medicine then began in earnest. By 1999, the National Guideline Clearinghouse (www.guidelines.gov), sponsored by the Agency for Healthcare

Research and Quality, the American Medical Association, and the American Association of Health Plans, boasted about 650 clinical practice guidelines. (Some of these guidelines had, in fact, been created by the federal government; through 1996, clinical practice guidelines were still being developed in-house at AHCPR. The decision to stop AHCPR's guideline development was political and pragmatic; there was an ongoing concern that dissemination of practice guidelines would anger affected providers, who, in turn, would attack the AHCPR's funding in the annual appropriations process. Given the near-disastrous budget experience in 1995 [recall the Contract with America, a Congressional budget-reduction agenda that culminated in several memorable government shutdowns at the end of that calendar year], AHCPR made a strategic move to replace its internal guideline development and dissemination program with support of external "evidence-based practice centers," which would develop and disseminate guidelines.)

By 1999, it was also clear that generating a guideline was no guarantee of its being widely adopted. In their landmark review, "Why don't physicians follow clinical practice guidelines?" Cabana and colleagues⁶ identified seven general categories of "barriers" affecting physician knowledge (Fig. 2). The Cabana categorization was adopted as the "differential diagnosis" most commonly applied to guideline implementation failures.⁷ We would add an eighth barrier, namely, failing to make guideline-based advice available at the point of care, rather than relying on the ability of clinicians to read, remember, and properly apply the guidelines encountered (typically) in lectures, journals, or flyers.

Guideline production accelerated even as implementation failures accumulated. Professional societies charged one panel after another with describing optimal care, and then writing, reviewing, approving, and disseminating guidelines without much thought to implementation barriers. Between 1999 and 2005, the number of guidelines deposited in the National Guideline Clearinghouse tripled. Despite ready accessibility and increasing scientific rigor, guideline use by busy clinicians could most charitably be described as lackluster. Flying in the face of both the exhortation ("practice consistent with current medical knowledge") and the threat related to deviation ("errors so costly in terms of ability to conduct business in the marketplace, market share and reputation that the organization must take action")

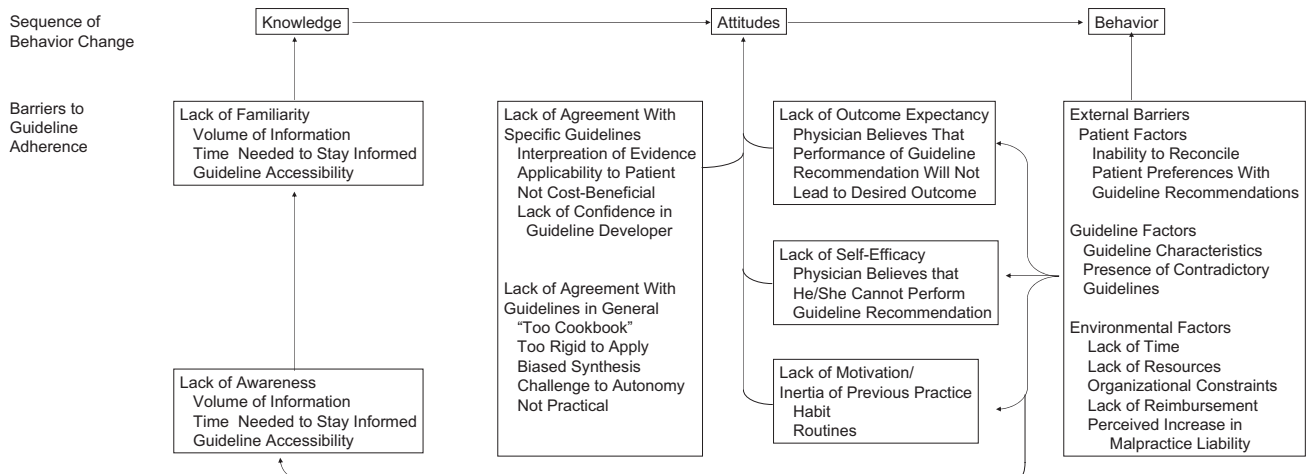


Figure 2. Cabana and colleagues' description of barriers to physician adherence to practice guidelines in relation to behavior change. From: Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;282:1458–1465, with permission.⁶

embedded in the 2000 report of the Institute of Medicine, "To err is human,"⁸ individual practice continues to be discrepant with community recommendations.^{9,10}

This perplexing variance currently sustains scholars, institutes, and legions of consultants. Although a comprehensive review of the explanatory literature is beyond the scope of this article, foci have included physicians (variously caricatured as ignorant, disinterested, disaffected, or inadequately compensated); guidelines (described as inconvenient, impenetrable, unavailable at the point of care, underperforming, and—very recently—unsafe when applied to high-risk groups¹¹); and patients (thought not to comprehend the significance of the sound advice offered to them by health-care workers).

Such diversity of opinion has led to interesting commentary and recommendations. So the redoubtable Joint Commission on Accreditation of Healthcare Organizations published "Lessons from experienced guideline implementers;"¹² the most important two of which were telegraphed in the subtitle: "Attend to many factors and use multiple strategies." If one is unsure why a guideline implementation has failed, a GLIA (Guideline Implementability Appraisal) has been proposed that assesses nine specific dimensions, including "decidability, executability, effect on process of care, presentation and formatting, measurable outcomes, apparent validity, novelty/innovation, flexibility, and computability" of the guideline in question.¹³ Failures have been attributed to violation of one of "The ten commandments of effective clinical decision support:" ie, No. 9, which states, "Monitor impact, get feedback, and re-

spond."¹⁴ Even the military, which historically encouraged standard behavior by court martialing surgeons who failed to follow prescribed practice, seems perplexed. (During World War II, failure to treat penetrating colon injuries with diversion could result in court martial. Based on this wartime experience, colostomy for civilian colon wounds became the standard of care for the next 4 decades. This is now known to be unnecessary in most patients.¹⁵) The 2001 Guide for Action entitled "Putting practice guidelines to work in the Department of Defense medical system," prepared by the RAND Corporation for the US military, provides a detailed method to identify barriers, yet no advice about how to overcome them.

The McDonnell Norms Group suggests that the general failure of clinical guideline implementation might rest in erroneous assumptions about what motivates human behavioral changes, in failures to account for the dynamic evolution of social norms, and in illusions about the physician-patient dyad. These recommended changes in practice detailed in guidelines are often not integrated into clinicians' work environments, making the recommended changes challenging to accomplish in daily practice. In addition, the guidelines are not presented in a clearly understandable or decipherable form. Even when guidelines are in an easily understood form, implementation often requires new elements not standard at the point of care. Each of these factors contributes to difficulty in changing clinicians' behavior and allowing evidence-based guidelines to guide their decisions. More simply, guidelines designed and dissemi-

nated without explicit consideration of the day-to-day realities of practice environment cannot succeed.

Rational (?) behavior

Reluctance to adopt clinical guidelines recommendations is pervasive, affecting both caregivers and patients. Regardless of recommendations from trusted and trustworthy sources, patients continue to act in ways contrary to their best interests or the best interests of the community at large. Familiar and common-sense advice about (for example) smoking cessation and moderation in alcohol often go unheeded until a catastrophic event occurs.

Clinicians now recognize that simply demonizing a risky behavior will not alter it. The current alternative strategy is to “medicalize” the consequence of the behavior. For example, a new diagnosis is “metabolic syndrome,” including at least glucose intolerance, obesity, hypertension, and dyslipidemia.¹⁶ One can treat the components piecemeal with medicines, and then attempt an end run to correct the “predisposing factors,” namely, the underlying behaviors (ie, overeating and lack of exercise). There are many other medical and non-medical examples of sound advice falling on seemingly deaf ears. Understanding why plentiful and reliable information fails to drive important, even essential, behavioral change is as critical to society as it is to the individual.¹⁷

Paradoxically, people follow fashion—trends in music, art, and clothing—that spread rapidly and seemingly without any perceived effort. (There is often a significant marketing strategy propelling the spread. Sponsorships at sports arenas, product placements in movies, and ad hoc events surrounding product releases are three among many marketing strategies in wide use.) Emotional and cultural factors that affect behavioral change and facilitate “overnight” transformations are not completely understood, but fundamental processes, such as imitation, are now recognized to drive consumption, sustain belief systems, and even promote undesirable practices such as teenage drinking and risky driving, both of which contribute to trauma morbidity and mortality. The mechanisms that drive nonmedical fads can be redirected toward at least some aspects of health care.

Mass media have been used with great effectiveness to communicate and even create demand for nonessential treatments, recently including various “fashionable” drugs. Nowhere is this more evident than in the direct-

to-public marketing of drugs aimed at treating erectile dysfunction.¹⁸ Understanding how information, ideas, and behaviors, in general, spread through populations could inform the problem of poor compliance with recommendations. Improving the ability of communities to efficiently and effectively spread information and specific practices could result from active collaboration with researchers whose scholarly focus is studying the spread of ideas. Declaration of “what is best practice” is inadequate. The guidelines must be marketed. As a foundation for change, experts in this domain are in general agreement on at least three points.

1. People are functionalists, so they “go with what works.” In addition, immediacy is highly valued. Unless behavior can be clearly seen to cause the desired outcomes, people are generally skeptical of value and are less likely to follow up on the behavior.
2. Natural validity—by this we mean widespread acceptance and use of a practice irrespective of a rationale—plays an important role in sustaining and reinforcing behaviors. If the environment confirms and condones a particular behavior, that behavior, regardless if it is desirable or not, most likely will be sustained. Such confirmation is more dependent on culture than evidence. The hierarchy of residents on a teaching surgical service solicits behavior from the most senior personnel and expects junior trainees to adhere to and ultimately emulate that behavior. Deviation from the exemplar constitutes, in Bosk’s terminology, a quasi-normative error.¹⁹ So strong is the hierarchical structure of a traditional surgical service that natural validity is effectively the only validity, and attempts to introduce alternative behaviors, even those that appear preferable based on high-quality evidence, are commonly derided.
3. Adaptation alters perspective.²⁰ This is familiar to every surgeon who has cared for patients with sudden, life-changing illnesses such as acute quadriplegia. Initially, patients insist that they do not want to live with such a severe disability and often ask to be permitted to die. When polled months later, nearly all have adapted physically and emotionally and are grateful to be alive. Clinicians (and patients) evaluate a new clinical guideline based on their initial response (too cumbersome, too costly, and so on), believing that they will continue to have the same response months or years later. In fact, successful guidelines become “part of the landscape” and are hardly noticed, even while they are being regularly used.

How choices are made

Physicians (and patients) are socialized to use utility-based techniques when making choices. The usual situ-

ation is that there is uncertainty about outcomes or optimal strategy, there are meaningful trade offs to be made, and the choice is nontrivial. Such situations are not only common in individual patient decision making, but also appear in guideline development, where several of the options have been used historically. Typically, a decision tree is constructed from evidence-based narrative guidelines that reflect alternatives to represent the available choices; outcomes that depict potential consequences of these choices; probabilities defining the likelihood of outcomes of each choice; and utilities, or values, assigned to outcomes. Once such a tree is constructed, the guideline is formulated to maximize a utility, and it often takes the form of an algorithm.^{21,22} Studies show that algorithm-based (flowcharts) and text-based (narrative) practice guidelines are used differently by physicians, with different preferences and adoption rates at differing levels of expertise. Patients are becoming increasingly well educated and are beginning to challenge existing protocols as inappropriate to a current and specific situation.²³ See, for example, Schneider's recent monograph.²⁴

Such stylized approaches to decision making differ substantially from another approach more commonly used by surgeons, in which decisions are embedded in a broader context and are part of a decision-action cycle. Here, the decision tree is reshaped as the case evolves, with subsequent decisions affected by earlier actions in a continuously changing environment.²⁵ Such decisions are made in the context of ill-structured problems with competing goals, where, over time, there is the potential for conflicts, shifting priorities, and tradeoffs. Actions can frequently generate problems of their own through unintended effects (complications); often it is difficult to determine whether the cause of a complication is rooted in the original problem or is a result of a particular midcourse action.

This latter view of decision making stresses that actions are taken on the basis of satisficing strategies,²⁶ in which decisions are made once a reasonable, adequate solution has been found, rather than after a systematic and exhaustive comparison of alternatives to find the one that is optimal. Such satisficing is often based on a superior ability to assess the unique characteristics of the particular situation, after which a solution may just fall into place.²³ This is the rationale often used for letting experience trump evidence. Such personal heuristics, used by clinicians under complex conditions when mak-

ing real-life decisions, provide an important source of knowledge for incorporation by guideline developers.

Heuristics and algorithms are not the same thing. The former can be thought of as "rules of thumb," the latter simply as rules. Heuristics are based on experience and do not guarantee a successful result, but are nevertheless highly effective and often are the only ways to achieve a goal. A familiar example is the varied sequence in which an experienced surgeon conducts a review of systems as part of a preoperative history; certain diagnoses prompt more or less emphasis among the various systems. Not every possible question is asked, yet the pertinent data are nearly always elicited. It is, of course, possible to conduct an exhaustive review of systems using a structured logical process—an algorithm—that ensures that the same questions are asked every time. There is no guarantee that any particular algorithm is correct, but given an established structure, the performance of a particular algorithm becomes highly predictable. Guidelines developed through consideration of heuristics may have a better chance of being used than those that rely exclusively on decision analytic theory.

Existing guidelines: Challenges and opportunities

It follows that guidelines that do not meet user requirements with regard to assumptions of their existing expertise, knowledge content, and integration with workflow may not be readily adopted. Yet there is an abundance of existing guidelines, evidence-based and endorsed by professional societies, that are, nevertheless, frequently violated. Examining the barriers to their adoption may provide insight into strategies for effective future guideline development.

The barriers are conveniently, if arbitrarily, divided into guideline-related barriers and practitioner-related barriers.

Guideline-related barriers

Fundamentals of heuristics are not considered. Most problems—computational, physical, biologic, and social—have more than one solution. The possible solutions are not equivalent; there is usually some tradeoff between efficiency and precision. A current and familiar example relates to evaluation of the abdomen after blunt trauma. Historically, diagnostic peritoneal lavage was used to evaluate such patients in the emergency department. It was invasive and required some surgical skill, but offered precise information about the presence of

visceral injury. Recently, diagnostic peritoneal lavage has been eclipsed by either of two types of imaging: focused ultrasonography examination and CT scan. There is little question that CT is more precise than ultrasonographic examination. The proximity of a CT scanner to the trauma resuscitation bay strongly influences local behavior. Where CT is convenient, ultrasonography is relegated to a minor supporting role. Where CT is inconvenient, heuristics have elevated ultrasonography to the central role in screening for abdominal injury. The role of heuristics—balancing efficiency and precision—is commonly underplayed by guideline developers. In addition, people, including health care providers, do what they do because a status quo choice has already proved economically feasible and remains cognitively less taxing than all other alternatives. An example is protocol-driven insertion of a device into the brain for monitoring of intracranial pressure after brain injury. Although in principle it is possible to perform serial neurologic examinations on many patients, the nursing and physician effort required to perform such examinations far exceeds the cost and risk associated with placing the device. Put differently, local culture and its attendant decisional inertia can trump the most persuasive data.

Guideline versus protocol. A related issue is the appellation *guideline* versus the more prescriptive *protocol*. Professional societies are loath to publish protocols, often from fear of liability for inevitable adverse outcomes, regardless of whether the outcomes had anything to do with the protocol. Ubiquitous disclaimers to the effect that clinicians must rely on their professional judgment in applying guidelines paradoxically encourage clinicians to develop personal heuristics. Further reflection suggests a triad of perverse incentives. Guideline creators indemnify themselves by adding disclaimers and stating that individual judgment should prevail. Third-party payors, including insurance companies and HMOs, expect physicians to adhere to published guidelines. Physicians have learned that “I was just following the guideline” does not preclude litigation, especially when common practice is at variance with the guideline.

The target audience is ill-defined, ill-chosen, or both. The complexity of guidelines intended for physician implementation and, more specifically, their application to an individual patient, varies widely. While life-saving protocols (such as Advanced Cardiac Life Support) have been intentionally streamlined (Advanced Cardiac Life Support now focuses on a single

primary antiarrhythmic drug and a handful of specific exceptions for its use), other guidelines, such as selection of empiric antibiotics for infections before identification of the organisms, allow for broad choice, not only among drugs, but even in dosing of a single drug. Although such breadth may be appropriate for the expert who can appreciate the nuances of different agents and patients, it is less suitable for novices such as medical students. The latter might be better served by one safe choice instead of four options.

Practitioner-related barriers

The recommended practice may not be seen by the individual being asked to change as truly representing a change for the better. There is ample evidence that minimizing sedation in critically ill patients leads to fewer complications and shorter stays.²⁷ Yet harried nurses who are free to titrate sedative medications will often choose doses that produce deep sedation because a somnolent patient is easier to care for and is less likely to cause self-harm. It is easy to rationalize the choice because “the patient needs rest,” building on an illusory equivalence between natural sleep and drug-induced unresponsiveness.

Individuals may find the information promoted in support of the recommended practice unconvincing. This is especially common when there is conflicting information or lack of trust in the recommending authority. Provided that current practice is not met with an intolerable level of immediate adverse outcomes, it is commonly regarded as safer than either of the newly proffered alternatives. Such reasoning is reinforced by experience subsequent to recommendations that have reversed findings of even highly touted studies appearing in the most prestigious literature.²⁸ The lay public is equally sensitive to such conflicts; recently, the Centers for Disease Control issued reports on obesity-associated mortality that differed markedly with respect to risk. Picking up on conflicting official information, many obese patients (and some “experts”²⁹) now assert that obesity carries minimal, if any, risk to health. The resulting lack of trust in the recommending authority (here, the government) makes it unlikely that related official advice such as spending an hour a day (most days) in moderately vigorous exercise will be seriously considered.³⁰ In such situations, any reinforcement of the status quo, no matter how risky, might make that choice appear safer. Marketing to the status quo, for example,

the garment industry's twin tactics of "size creep" (today's size 8 was yesterday's size 10) and expanded production of extraplus-sized garments, sends the message that obesity, if not okay, is certainly not as bad as the government is making it out to be.

The recommended practice may benefit a larger community, but have costs (or simple neutrality) for relevant decision makers. The classic example involves a pediatrician, a child with a viral respiratory infection, and an anxious mother. Although good clinical practice recommends reassurance and supportive care, mothers (and too many clinicians) are conditioned to expect an antibiotic prescription. The conditioning may not be "official," but rather as simple as "the neighbor's child had the same thing, got an antibiotic, and got better." Failure to provide the expected prescription may diminish the stature of, or the trust the mother has in, the pediatrician, a cost that is substantial in comparison to the risk of providing an antibiotic that is perceived as "safe." So, unsafe practices are propagated by social pacts and expectations. This is especially problematic in teaching environments. For example, the decision of an attending physician to enter an isolation room without donning appropriate protective garb legitimizes the practice in the eyes of the trainees. Why does an attending physician choose not to use the garb? Generally because the immediate nuisance value of the time spent is perceived as larger than some subsequent (and often invisible) risk of transmitting a pathogen to another patient.

The recommended practice may run counter to established norms or perceived norms. After a tracheostomy procedure, the established norm is to obtain a chest x-ray, even for uncomplicated patients. Clinical evidence strongly argues against this routine practice because it is costly, risky (unnecessary radiation), and without benefit.³¹ Yet the practice has proved almost impossible to eradicate because ICU staff perceive that they will be blamed for failing to identify and correct a vanishingly rare life-threatening complication. The calculus of the ICU staff undervalues (for example) the incremental risk of a future cancer versus the immediate (but much smaller) risk of an undetected life-threatening emergency consequent to tracheostomy placement. Here again, time plays a role in determining practice; immediate proof allows the clinician to go on to the next task rather than continuously monitor the patient.

Disadvantages of adopting the "best practice" may be easier to perceive than advantages. The cost of change is immediate, although benefits—especially those that are temporally delayed or that involve prevention of low-probability events—are excessively discounted. So strict adherence to barrier precautions when performing an invasive procedure, such as bedside insertion of a central venous catheter, is frequently violated because the immediate nuisance associated with gloves, gowns, hats, masks, surgical scrub, drapes, and so on looms larger than the small but real risk of an infection that will become clinically apparent only long after the line is inserted.

Approaches that enhance adoption of preferred practices

Although barriers to guideline adoption and implementation are great, they are not insurmountable. How can spread be encouraged?

Convenience is surprisingly important to the adoption of practice. Put differently, the preferred option needs to be easiest (least mental and physical effort), and needs to be available "just in time," meaning where and when needed. Failure to address the convenience and simplicity has led to resistance to the adoption of computerized physician order entry. At least one-third of attempts to introduce this cost-effective safety tool fail. See, for example, an account of the Cedars-Sinai failure.³² There is general agreement that computerized physician order entry is a preferred practice because it reduces errors engendered by name similarity, abbreviations, and transcription processes, and also avoids many significant complications by automatically checking for side effects and need for dose adjustment based on liver and kidney function. But there is a misconception by the users of the system. Users expect it to solve all problems and eradicate all errors. This cannot happen. Pen and paper are convenient, operate at "blink speed," do not require the clinician to remember a password, and do not require the clinician to perform activities previously assigned to the unit secretary. So the conventions of pen and paper are enormously attractive to the clinician, despite their inherent risks and inefficiencies to the medical-care system. Compounding resistance to computerized physician order entry is the fact that not all software systems function perfectly on initial release. As with any new implementations, they can cause unintended problems as they adapt to the users and the en-

vironment. Implementation failures at one institution become known at others, and the storied inconveniences and outright failures become legend. Failures are remembered much more vividly than successes of any system. Any new guideline, protocol, or system has to be at least more convenient than all alternatives. Proved economic benefit plays a close second in importance.

Public (and peer) demand is crucial. This is often overlooked in strategies that address behaviors of physicians separate from the community of caregivers in the health-care system. The notable reduction in wrong-site and wrong-patient surgery is based in part on development of team strategies (preoperative marking of the operative site is a shared responsibility) and team verification (the “time-out” just before initial incision) that the right operation is about to be performed on the right site on the right patient.

Immediate (or near-immediate) feedback is useful. Because benefits (eg, prevention of repeat heart attacks with aspirin and β -adrenergic blockers) that lie in an unspecified future are unduly discounted, immediate surrogates, such as a weekly report on how many patients did (and did not) receive recommended care, can be useful. Competitive immediate feedback that leverages public opinion, such as posting competitive compliance rates by caregiver or by functional unit, can accelerate compliance.

Probabilistic information should be translated into forms that are immediately comprehensible to the target audience. Evidence shows that physicians often have trouble accepting epidemiologic (generally probabilistic) information from the literature as the basis for making decisions about their own patients, whom they perceive as “individuals and not statistics.” This means that there is an extra step required to translate probabilistic information into clinically useful information at the point of care. Similarly, categorical data (eg, how many deaths were avoided, infections prevented, needle-stick-free days) are more readily interpreted by patients and providers than are continuous probabilistic data. Such an approach has been used to educate patients about the benefits of behavioral changes in cancer prevention and longevity.³³

Information that confronts misguided metanorms can have dramatic effects. For years, mild hyperglycemia was clinically accepted as an inevitable consequence of surgical stress; teleologic arguments were marshaled in favor of tolerating high glucose levels (“the body needs to

have the extra energy immediately available for emergencies”). A single, well-constructed, randomized controlled trial comparing current practice versus maintaining tight glycemic control was sufficient to confront the misguided metanorm and spur widespread efforts to more tightly control blood glucose in the ICU.³⁴

Tying desired changes to existing norms helps people understand and adopt practices. Pain is often inadequately managed because “as needed” orders require that a patient admit discomfort (which can be perceived as a sign of weakness) and disrupt other routine nursing tasks to receive medication. Designating pain as “the fifth vital sign” obligates the caregiver to routinely assess the patient, solicit the patient’s subjective assessment, and deliver relief as part of routine, not outside routine. This practice has now been widely embraced by nurses and mandated by regulatory agencies.

How messages are framed can be crucial. Celebrity advocacy is an increasingly common tactic used to encourage the lay public to comply with clinical practice guidelines.³⁵ So Katie Couric’s on-air colonoscopy triggered a sustained 20% increase in the colonoscopy rate among reporting gastroenterologists.³⁶

How messages are delivered is equally crucial. Phenytoin, a widely used antiseizure medication, has a long half-life, yet is commonly administered on daily, twice daily, and three times daily schedules. The use of a three times daily schedule inadvertently implies to both doctors and patients that less frequent dosing is inadequate. Safety is commonly compromised when the adequate daily dose is inadvertently administered on a more frequent basis. Yet there is no consistent message about phenytoin dosing and, specifically, the need for a standard dosing schedule to encourage safe dosing. As a result, many patients experience frequent phenytoin overdoses. Paradoxically, parents often underdose their children with medication prescribed for administration four or more times daily because it “seems too much” to give the child “so many doses.” Such undertreatment (especially of infectious diseases) can prolong illness and lead to bacterial resistance.

The calculus of professional behavior Incentives, transparencies, and accountabilities

Professional behavior is subject to diverse influences. Surgeons are as inconsistent as other physicians with respect to widely recommended practices such as hand washing, use of appropriate isolation gear (gowns and

gloves), and timely administration of specific antibiotics. Yet they are highly compliant with these recommendations in one specific context, namely, the operating room. How does context augment compliance with these three practices? It does so by changing the incentives, transparencies, and accountabilities.

Hand washing is a learned behavior that is encouraged early in life and reinforced early in medical training. The incentives in the operating suites are substantial: the sinks are deep, the water is warm, the environment is social (sinks are typically side by side and team members chat while scrubbing), and that most precious commodity, time, is not wasted looking for the necessary resources. Sinks are universally placed just outside the operating room and universally stocked with the necessary soap-impregnated scrub brushes. (More recently, alcohol/chlorhexidine foams have been approved for second and subsequent daily scrubs. Despite greater efficiency, many surgeons still prefer the traditional wet soap-and-brush process: it feels good and the social aspect of scrubbing is preserved.)

But these strong incentives pale beside the transparencies and accountabilities. It is quite obvious to the entire operative team that someone has (or has not) just scrubbed; wet hands and arms held at chest height ready to receive a drying towel signify to both the scrub nurse, and to the circulating nurse who is charged with watching for deviations, that the person is prepared to “gown up.” Each person in the operating room is expected to promptly report any violation of sterile technique, and there is no social cost to the violation or to reporting such a violation.

Gowning and gloving is also a learned behavior, but one that bears only passing resemblance to the casual process of donning street clothes. On the hospital wards and in the office, avoidance of gowns and gloves is often rationalized (“I’m going into the isolation room not to touch the patient, only to speak”). One possibility is that transparency is minimal (getting “caught” is relatively unlikely unless one encounters another caregiver). Another is that accountability for the subsequent bad outcomes (carrying the organism to another patient) is even more opaque given the time delay between inoculation and the appearance of clinical infection. A third possibility is that gowning and gloving in the operating room are assisted and ritualized. Unlike the situation on the ward (where one must find a gown, find the right glove size, find someplace secure for the pager and other ac-

countments of medical practice), staff in the operating room not only have the appropriately sized equipment immediately available, but ritual demands that the staff assist the physicians (and student physicians) in donning the equipment. Convenience is paramount. When HIV and other transmissible diseases made double gloving de rigueur, this modification was more readily incorporated into the ritual because it could be done conveniently.

Getting the right antibiotic into the patient before the scalpel hits skin was problematic for decades. Surgical interns were assigned to write “antibiotic on call” orders; floor nurses busy trying to complete a preoperative checklist were expected to remember the order; and operating room personnel had to understand not to interrupt the infusion if it was still running. Failure and consequent infection were reportable in surgery morbidity and mortality conferences, so ownership of the problem was surgical. The problem was not solved until it was asked, “What practitioner has best access to an IV just before the incision, and who can also ensure that the appropriate antibiotic is administered before the surgical incision?” The answer, of course, is the anesthesiologist. Currently, the anesthesiologist has not only the responsibility to make sure that the selected antibiotic has been infused, but also the authority to halt the case if there is any ambiguity that the patient has received the right dose of the right drug at the right time. So accountability has shifted.

Mutual support and the underappreciated value of teams

The comparison of operating room procedures with general inpatient practices emphasizes an underappreciated aspect of social norm adoption and enforcement in medicine: the role of teams. The sanctity of the doctor-patient dyad is more apparent than real. With rare exceptions, American health care is currently delivered by a community of providers that includes physicians and nurses, a range of therapists, pharmacists, case managers and social workers, nutritionists, and others. Yet most clinical practice guidelines generally focus on physician actions and patient behaviors and tend to underplay the potential contributions of other health-care providers. Omission of these professionals from clinical practice guideline constructs not only impedes the expression of shared vision but also implies that the physician and patient are the only stakeholders in the processes of care. Sharing of incentives, transparencies, and accountabili-

ties can lead to a competency-driven division of labor, a division that seems important given that key patient actions (taking a pill, changing a habit, undergoing a screening test) are typically encouraged or supervised by nonphysician providers. There is accumulating evidence that nonphysician-driven clinical protocols are generally more successful than those that are designed to be driven by physicians. Nowhere is the success of this approach more evident than in the intensive care unit, where team members of diverse backgrounds become champions of protocols most relevant to their domains of expertise.

Toward a different future

It seems unlikely that the additional generation and dissemination of clinical guidelines using existing approaches will augment compliance or significantly improve patient outcomes. We suggest that systematic study of implementation successes and failures will illuminate new strategies for influencing clinician and patient behaviors. Partnerships among providers, health-care organizations, payors, and behavioral and social researchers aimed at understanding the dynamics of decision making may be able to generate novel guideline structures and implementation strategies that have a better chance for widespread adoption. The use of information technologies that provide just-in-time knowledge, and make the safe and preferred choices more apparent at the time of decision making, has potential to increase adherence to current recommendations.

Acknowledgment: This work was made possible by generous support from the James S McDonnell Foundation, St Louis, MO.

Appendix

The McDonnell Norms Group

The McDonnell Norms Group, sponsored by the James S McDonnell Foundation, aims to identify core principles in the behavioral, cognitive, and social sciences that enable the responsible application of information for the public good. The group intends to close the gap between gathering, synthesis, and provision of information—activities that culminate in the development of reasonable recommendations—and the adoption of new behaviors that reflect those recommendations. More simply, the group seeks to learn “Why we do not practice what we preach,” either as professionals or individuals. The group includes scientists and policy experts with

backgrounds ranging from clinical surgery to evolutionary biology. The group formed around a case studies approach and will focus next on inconsistencies between recommendations and use of antibiotics in clinical medicine.

The James S McDonnell Foundation is a not-for-profit foundation that uses its resources to catalyze the development of new knowledge and insight at the intersection of specific disciplines, such as complex systems science and neurobiology. The Foundation’s purpose is “to improve the quality of life” through the vision of aviation pioneer James S McDonnell and his family. Mr McDonnell’s accomplishments included building the Mercury spacecraft that first carried Americans into earth orbit. He was married to Mary Elizabeth Finney, whose father, John MT Finney, MD, FACS, was a founder and first president of the American College of Surgeons.

Members of the McDonnell Norms Group:

Timothy G Buchman, Department of Surgery, Washington University in St Louis, St Louis, MO

Vimla L Patel, Department of Biomedical Informatics and Psychiatry, Columbia University, New York, NY

Jonathan Dushoff, Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ

Paul R Ehrlich, Department of Biologic Sciences, Stanford University, Stanford, CA

Marc Feldman, Department of Biologic Sciences, Stanford University, Stanford, CA

Maryann Feldman, Rotman School of Management, University of Toronto, Toronto, Ontario, Canada

Bruce Levin, Department of Biology, Emory University, Atlanta, GA

Dale T Miller, Department of Psychology, Stanford University, Stanford, CA

Paul Rozin, Department of Psychology, University of Pennsylvania, Philadelphia, PA

Simon A Levin, Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ

Susan Fitzpatrick, James S McDonnell Foundation, St Louis, MO

REFERENCES

1. Wasson J, Sox H, Garcia R. A randomized study of the educational effects of an algorithm system. *J Med Educ* 1979;54:119–121.

2. Wennberg J, Gittelsohn A. Small area variations in healthcare delivery. *Science* 1973;182:1102–1108.
3. Gray BH, Gusmano MK, Collins SR. AHCPR and the changing politics of health services research. *Health Aff (Millwood)* 2003; Jan–Jun; Suppl Web Exclusives:W3:283–307.
4. Birkmeyer JD, Sharp SM, Finlayson SR, et al. Variation profiles of common surgical procedures. *Surgery* 1998;124:917–923.
5. Guyatt GH. Evidence-based medicine. *ACP Journal Club* 1991; 114:A–16.
6. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;282:1458–1465.
7. Cabana MD. Barriers to adoption. *The Scientist* 2006;20:52.
8. Kohn LT, Corrigan JM, Donaldson MS, eds. *To err is human: Building a safer health system*. Committee on Quality of Healthcare in America. Washington, DC: Institute of Medicine; 2000.
9. Bratzler DW, Houck PM, Richards C, et al. Use of antimicrobial prophylaxis for major surgery: baseline results from the National Surgical Infection Prevention Project. *Arch Surg* 2005;140: 174–182.
10. Asch SM, Kerr EA, Keeseey J, et al. Who is at greatest risk for receiving poor-quality healthcare? *N Engl J Med* 2006;354: 1147–1156.
11. Tinetti ME, Bogardus ST Jr, Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. *N Engl J Med* 2004;351:2870–2874.
12. Solberg LI, Brekke ML, Fazio CJ, et al. Lessons from experienced guideline implementers: attend to many factors and use multiple strategies. *Jt Comm J Qual Improv* 2000; 26:171–188.
13. Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* 2005;5:23.
14. Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 2003;10:523–530.
15. Miller PR, Fabian TC, Croce MA, et al. Improving outcomes following penetrating colon wounds: application of a clinical pathway. *Ann Surg* 2002;235:775–781.
16. Reisin E, Alpert MA. Definition of the metabolic syndrome: current proposals and controversies. *Am J Med Sci* 2005;330: 269–272.
17. McDonnell Norms Group. Battling bad behavior. *The Scientist* 2006;20:51.
18. Waack KE, Ernst ME, Graber MA. Informational content of official pharmaceutical industry web sites about treatments for erectile dysfunction. *Ann Pharmacother* 2004;38:2029–2034.
19. Bosk CS. *Forgive and remember. Managing medical failure*. Chicago: University of Chicago Press; 1979.
20. Redelmeier D, Rozin P, Kahneman D. Understanding patients' decisions: Cognitive and emotional perspectives. *JAMA* 1993; 270:72–76.
21. Hadorn DC, McCormick K, Diokno A. An annotated algorithm approach to clinical guideline development. *JAMA* 1992; 267:3311–3314.
22. Lee KF, Purcell GP, Hinshaw DB, et al. Clinical palliative care for surgeons: part 1. *J Am Coll Surg* 2004;198:303–319.
23. Patel VL, Arocha JF, Diermeier M, et al. Cognitive psychological studies of representation and use of clinical practice guidelines. *Int J Med Inform* 2001;63:147–167.
24. Schneider SH. *The patient from hell: How I worked with my doctors to get the best of modern medicine and how you can too*. Cambridge, MA: Perseus Books; 2005.
25. Peleg M, Gutnik LA, Snow V, Patel VL. Interpreting procedures from descriptive guidelines. *J Biomed Inform* 2006;39:184–195.
26. Simon HA. *Models of thought*. New Haven: Yale University Press; 1989.
27. Schweickert WD, Gehlbach BK, Pohlman AS, et al. Daily interruption of sedative infusions and complications of critical illness in mechanically ventilated patients. *Crit Care Med* 2004; 32:1272–1276.
28. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005;294:218–228.
29. Campos PG. *The obesity myth: Why America's obsession with weight is hazardous to your health*. New York: Gotham; 2004.
30. *Dietary Guidelines for Americans 2005*. <http://www.health.gov/dietaryguidelines/dga2005/>. Accessed January 20, 2006.
31. Smith DK, Grillone GA, Fuleihan N. Use of postoperative chest x-ray after elective adult tracheotomy. *Otolaryngol Head Neck Surg* 1999;120:848–51.
32. Connoll C. Cedars-Sinai doctors cling to pen and paper. <http://www.washingtonpost.com/wp-dyn/articles/A52384-2005Mar20.html>. Accessed January 20, 2006.
33. Woloshin S, Schwartz LM, Welch HG. Risk charts: putting cancer in context. *J Natl Cancer Inst* 2002;94:799–804.
34. Van den Berghe G, Wouters P, Weekers F, et al. Intensive insulin therapy in the critically ill patients. *N Engl J Med* 2001;345: 1359–1367.
35. Larson RJ, Woloshin S, Schwartz LM, Welch HG. Celebrity endorsements of cancer screening. *J Natl Cancer Inst* 2005; 97:693–695.
36. Cram P, Fendrick AM, Inadomi J, et al. The impact of a celebrity promotional campaign on the use of colon cancer screening: the Katie Couric effect. *Arch Intern Med* 2003;163:1601–1605.