The Intersection of Evidence and Values in Clinical Guidelines: Who Decides What Constitutes Acceptable Risk in the Care of Children?

Risk is implicit in all clinical decision-making. Whether a clinician refrains from ordering a head computed tomography scan for a patient with a headache or decides to obtain blood work on a child with a fever, each decision involves a balance between accepting and limiting risk. Ideally, clinicians strive to achieve this balance. Clinicians accept some risk (eg, the risk of missing a rare disease) to avoid placing undue burden on the patient and the health care system with invasive, expensive, and/or potentially unnecessary testing. Yet, clinicians also strive not to assume too much risk so that timely diagnoses are made and morbidity and mortality prevented.

Achieving this balance is difficult for a number of reasons. First, clinicians often lack precise estimates of the risks and harms for a given clinical scenario. There is simply an absence of or paucity of data to help quantify many of the risks, benefits, and outcomes associated with different possible therapeutic options. This confounds the ability to know whether pursuing 1 particular option corresponds to accepting too much or too little risk. As a result, decision-making is often intrinsically imperfect.

Second, even if precise estimates of the involved risks are known, determining the threshold constituting acceptable risk (ie, the level above which too much risk would be assumed) is largely subjective. The issue of acceptable risk is inherently a matter of values, which, in pediatrics, includes not only the clinician’s values but also those of the parent and sometimes the child. Although there is general agreement that it is unacceptable for a parent to assume high levels of preventable risk and harm on behalf of his or her child, what constitutes a high level of risk and harm? Is it a 1 in 100 or 1 in 100,000 risk of the child contracting a treatable disease with no chronic sequelae? A disease with established chronic sequelae? A life-threatening disease? There is no

KEY WORDS
bioethics, evidence-based medicine, practice guidelines, shared decision-making

ABBREVIATIONS
NICE: National Institute for Health and Clinical Excellence
NNT: number needed to treat

Douglas J. Opel, MD, MPH,1,2 James A. Taylor, MD,2
Carrie A. Phillipi, MD, PhD,3 Douglas S. Diekema, MD, MPH1,2

1Treuman Katz Center for Pediatric Bioethics, Seattle Children’s Research Institute, Seattle, Washington;
2Department of Pediatrics, University of Washington School of Medicine, Seattle, Washington; and
3Department of Pediatrics, Oregon Health and Science University, Portland, Oregon

Address correspondence to Douglas J. Opel, MD, MPH, Treuman Katz Center for Pediatric Bioethics, Center for Clinical and Translational Research, Seattle Children’s Research Institute, 1900 Ninth Ave, M/S: C9S–6, Seattle, WA 98101. E-mail: djopel@u.washington.edu

doi:10.1542/hpeds.2012-0090

HOSPITAL PEDIATRICS (ISSN Numbers: Print, 2154 - 1663; Online, 2154 - 1671).
Copyright © 2013 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: No external funding.
consensus regarding where thresholds of acceptable risk lay in particular clinical situations.

Lastly, determinations of acceptable risk in pediatrics are also challenged by the need to account for the risk paradox that exists in the care of children. In general, there is a low risk of a life-threatening illness in children relative to older patients, but the consequences of being wrong (ie, assuming too much risk) are greater, at least as measured by years of quality life lost. Given these constraints, how are we making these determinations, and how should we be?

**CLINICAL GUIDELINES**
The emergence of shared decision-making as the standard in medical decision-making shifted value-laden risk determinations away from being predominantly clinician-driven toward a process in which patient or parent preferences are the focal point. The degree of decisional responsibility accepted by the clinician and patient falls along a shared decision-making continuum and is tailored to the preferences of each individual patient. Therefore, determinations of what constitutes acceptable risk in pediatrics are ideally made by using an individualized shared decision-making process incorporating parental values and preferences. This approach is becoming increasingly undermined, however, by the codification of risk through the introduction and dissemination of an expanding number of clinical guidelines.

Guidelines can be influential to clinicians because they summarize the best available evidence and are usually published by authoritative groups in the field, such as a professional medical society. Although their intent is to standardize care, reduce variability, and improve the overall quality of care to patients, they also facilitate the navigation of risk inherent to a specific clinical situation by offering clinical recommendations that incorporate a preset determination of how much risk the clinician should accept. This preset determination, however, poses several problems. First, guidelines may obviate the need for clinicians to make their own determinations of what amount of risk to accept in a specific clinical situation. Some argue that this may encourage complacency and discourage critical thinking. Second, these determinations remain subject to bias. Many guidelines are largely developed by disease experts who have a natural tendency to focus primarily on the susceptibility to, and severity of, the condition that is the subject of the guideline.

This process may lead to guidelines biased toward a minimal acceptance of risk through excessive testing or treatment. Third, guidelines must draw on the same imprecise and incomplete evidence base that is available to clinicians and therefore suffer from similar imperfect calculations. The uncertainty this introduces into the guideline recommendation is not always quantified or explicitly acknowledged. The US Preventive Services Task Force’s use of the insufficient evidence (I) rating in favor of a rating category based on expert opinion or weak evidence offers a good example of acknowledging uncertainty when it exists. Not only can a recommendation based simply on expert opinion have negative consequences for patients if followed, but failing to be explicit about recommendations based on weak evidence can also contribute to the misperception that there is no reasonable clinical alternative to the guideline. This was illustrated in a recent case discussion on the management of a newborn whose mother was diagnosed with chorioamnionitis and prompted us to publish a commentary criticizing the Centers for Disease Control and Prevention’s guidelines for not explicitly acknowledging and incorporating this uncertainty into their recommendations.

**A COMPREHENSIVE ASSESSMENT OF RISK**
Further challenging the guideline development process beyond being hampered by a lack of quality evidence is that some risks may be inappropriately neglected by guideline committees. Committees must consider not only the risk of disease in the absence of the intervention but also the risks of the intervention itself. Neglected risks include those that are not directly related to the specific intervention, such as the risks of harmful effects related to the psychosocial stress of undergoing the treatment or testing and even the unanticipated risks associated with simply being treated, a random catastrophic medical error leading to death, or significant morbidity. The recent outbreak of fungal meningitis due to a contaminated steroid medication is a stark reminder of these risks.

Only through a comprehensive consideration of all the potential risks of an intervention can its projected burdens be best quantified (albeit approximately) and compared with its projected benefits. For instance, consider the recommendation that all newborns be screened for jaundice with a serum or transcutaneous bilirubin measurement before discharge from the newborn nursery to prevent kernicterus. Before the subjective determination can be made...
as to whether universal screening represents the acceptance of too much or too little risk, the entirety of risk involved needs to be understood. The most relevant risk is that of developing kernicterus. In healthy term neonates without isoimmunization, this risk is likely <5 per 1,000,000. The indirect effects of universal bilirubin screening (the risks associated with hospitalization of newborn infants, including medical error, and psychosocial effects of treatment and testing, such as disrupted maternal–infant bonding) are also important and relevant because there is more subsequent testing and treatment of hyperbilirubinemia with universal screening. Although these latter indirect risks may be understudied, their consideration is nonetheless critical because their inclusion in the comprehensive risk evaluation may tip the balance between accepting and limiting risk that universal screening seeks to achieve.

Calculations such as the number needed to treat (NNT) or the number needed to harm can help give perspective to this balance of accepting and limiting risk. For example, if a guideline regarding male infant circumcision was developed, it would be pertinent to the committee’s deliberations to consider that the NNT to prevent 1 case of HIV in the United States or in Canada through neonatal circumcision is likely >5000. This NNT could be balanced with other benefits of circumcision, such as the NNT of ~100 to prevent 1 urinary tract infection, and the risks of an adverse medical effect from the procedure and the potential psychosocial stress to the infant and parents.

If the ideal health system is safe, effective, patient-centered, timely, efficient, and equitable, comprehensive risk assessments in guideline development must also be accompanied by cost-effectiveness analyses. These analyses require quantification of how beneficial a clinical intervention is in terms of outcomes such as length and quality of life and the resources required to realize these outcomes. These assessments are complex, time-consuming, and involve additional stakeholders such as the government and health plans but are critical to policy-making. Cost-effectiveness analysis features prominently in European guideline development, such as those developed by the UK’s National Institute for Health and Clinical Excellence (NICE). Inadequate attention paid to these same assessments in the recent US National Heart, Lung, and Blood Institute’s guideline regarding universal screening and drug treatment of dyslipidemia in children has been cited as a significant limitation of the guideline.

AN INTERSECTION OF EVIDENCE AND VALUES
Although quantitative, evidence-based processes for assessing risk in clinical guideline development are necessary to the development of recommendations, they remain insufficient for the determination of thresholds constituting acceptable risk. Just as in an individual clinical encounter, these determinations at the policy level remain fundamentally dependent on one’s values. Whether the combination of NNTs of 5000 and 100 to prevent 1 case of HIV and 1 case of urinary tract infection, respectively, does or does not justify a recommendation for universal male infant circumcision depends on the values, preferences, concerns, and expectations of those to whom this issue is posed. This process is also true of cost thresholds. For example, the £20,000 per quality-adjusted life-year gained threshold established by NICE for determining whether a new treatment is an effective use of the resources of the UK’s National Health Service is fraught with disputes regarding competing values. Therefore, risk and cost assessments are essential but not enough: genuine moral disagreement between reasonable people about the relative weights of certain risks, costs, and outcomes will still exist.

As such, guideline committees must recognize the importance of parental and societal values in the guideline development process and be vigilant that these values are not marginalized. One way to prevent the marginalization of values is by making the subjectivity underlying a committee’s risk assessments—and therefore, its recommendations—transparent. Not doing so can have a profound effect if the subjectivity goes unnoticed. For example, parents who refuse to comply with a recommendation made by a clinician who is strictly following a clinical guideline may be easily dismissed or threatened with use of state intervention because their refusal is viewed as a departure from the clinical standard of care. Clinician fear of litigation only encourages this view. Yet, given the inherent subjectivity of the guideline’s determination of an acceptable threshold of risk, it is perhaps more likely that the parents’ refusal simply reflects a different but reasonable assessment of what constitutes acceptable risk than the 1 made by the guideline committee. When guideline committees are not transparent about the values prioritized when making recommendations, or clinicians simply accept guideline
recommendations as the de facto determination of what level of risk is acceptable to tolerate, parental values and input are short-changed.

THE WAY FORWARD

Recently established standards for guideline committees from the Institute of Medicine26 and others, such as GRADE,27 offer an improved framework for developing recommendations by explicitly analyzing benefits versus harms, the quality of the available evidence, values of the population affected, and economic analyses. Some guideline committees, such as NICE and the Advisory Committee on Immunization Practices, now use these guidelines to develop recommendations.28,29 New guidelines from the American Academy of Pediatrics also include more explicit consideration of risk calculations and value judgments. The guideline for the Diagnosis and Management of the Initial Urinary Tract Infection in Febrile Infants and Children 2 to 24 months, for instance, mentions the committee’s assessment of the ratio of benefits to harms, the value judgments made in their assessments, and the role of patient preferences in their recommendations.30

Although these new standards are an improvement, we suggest further reorienting the clinical guideline development process to account for the value-laden activity of determining what constitutes acceptable risk even more comprehensively. This reorientation at the policy level should be anchored in the goal of shared decision-making at the individual level: “to make decisions in a manner consistent with the patient’s wishes.”22 This is not to claim that parents are better at determining risk and certainty than guideline committees nor is it to suggest that the concept of shared decision-making in guideline development is a panacea. Rather, it is simply an acknowledgment that parental insight is vital to the determination of the level of risk that they are willing to accept on behalf of their child.

To achieve this reorientation, guideline development needs to involve broader deliberation and increased societal input; that is, a comprehensive sampling of a population’s actual, rather than perceived, values.31,32 For any particular clinical scenario—particularly those in which there is weak or equivocal evidence to guide clinical actions—guideline committees should explicitly seek societal input regarding what constitutes an appropriate level of risk for parents to accept on behalf of their child in which medical guidelines are desired. Then, upon reaching a majority consensus, it would be appropriate to proceed with developing clinical recommendations that comply with this agreed-on level of acceptable risk. The Centers for Disease Control and Prevention recently chose this approach to determine how the Advisory Committee on Immunization Practices should introduce new vaccines to the recommended childhood immunization schedule.33 If obtaining societal consensus is logistically or financially difficult, pediatric guideline committees can better consider societal values by including nonmedical community stakeholders (such as a group of randomly selected parents) to review and suggest modifications to the recommendations of the medical experts as a final step before a guideline is promulgated. In addition, guideline committees could be more transparent about the level of consensus reached among the committee and/or include a dissenting opinion along with the published recommendations.

CONCLUSIONS

Guideline development has recently been lamented as a broken process.22 We offer some suggestions for its improvement. Critical to this improvement is increasing the evidence-based foundation on which recommendations are built over the long term, and meanwhile, better acknowledging the imperfection and uncertainty inherent to recommendations based on the evidence we have now. Superseding this step is the need to better incorporate societal and parental values in the guideline development process. Anything less than this seems unfair.

REFERENCES


