To the Editor:—While we applaud the intent of Fallis and Frické1 in their recent attempt to provide the public with indicators of accuracy for health information on the Internet, we take exception to some of their methods and conclusions. The topic they have chosen to study is certainly an important one. There exists on the Internet a vast amount of health information, and much of it may be incorrect or easy to misinterpret. The authors’ call for impartial validators of such information should be cheered. Nevertheless, we have some methodological concerns with their recent study.

The first concern involves their choice of clinical subject matter. The authors claim that managing childhood fever in the home was chosen “because there is a wide consensus among experts in this area.”1 This issue is a point of significant contention among pediatricians.2–4 Even the five topics covered in their methods are sure to spark many debates among health care professionals.5–7 We consider it an important omission that the authors did not provide the “correct” answers or cite the experts who provide the consensus to the five topics set forth in the paper. We wonder how much of the information that was labeled inaccurate by the authors was merely a reflection of differences in practice. If this same topic were brought up in a pediatric journal, we believe that reaching a consensus on what is “correct” would be difficult. Without clear consensus, there can be no gold standard. Without a gold standard, there can be no measure of accuracy.

The second concern is over the methodology used for selection of Web sites. The authors provide no criteria as to how they decided which 100 sites to analyze. Given the enormous number of possible sites, it would be critical to know who selected sites and how they decided which ones to select. Without such information, the study cannot be repeated or understood fully, nor can possible selection bias be ascertained.

The third concern we have is the definition of “accurate” and “inaccurate” sites. The authors arbitrarily set the dividing line between accurate and inaccurate at the median of the sites’ ratings. Although they did confirm their results using several different cut-offs, we are concerned with the use of these terms at all. Their accuracy scores range from 0.55 to 0.75, with the vast majority of sites in the 0.6 to 0.7 range. These numbers are not only very close together; they also have no readily apparent clinical significance. From a pragmatic point of view, the significance of a difference in scores between 0.67 and 0.63 is unclear. According to the authors, one is accurate and the other is not.

Unfortunately, these numbers provide no value to the practitioner or the consumer. A site with an accuracy rating of 0.7 may correctly label a number of sites to check for temperature but give an unsafe dose of medication to the consumer. A site with an accuracy of 0.6 may give a safe dose of medication but recommend temperature sites incorrectly. One may be more “accurate” but also less safe. This is the true measure of importance.

The fourth concern is that the authors’ discussion of their results goes beyond what their data support. They claim that “the presence of the HONcode logo on a Web site is a fairly good indication that a Web site contains accurate information on the treatment of fever in children.”1 Their own data do not support
this: 21 percent of sites with the HONcode logo were judged to be inaccurate. Furthermore, those judged to be more “accurate” cannot be said to be safer.

All this leads us to the larger problem. There is currently no good way to judge the safety, accuracy, and completeness of health information given out over the Internet. Only information that is provided by groups willing to take responsibility for the content of their sites should even be considered at this time. Such sites may be run by large professional or governmental organizations, like the American Academy of Pediatrics (http://www.aap.org/) or the Centers for Disease Control and Prevention (http://www.cdc.gov/).

We appreciate the authors’ attempts to provide us with objective criteria to judge the safety of health-related information on the Internet. We disagree with their conclusions, however. The HONcode, copyright, or organization domain cannot be used to adequately judge the validity of health information. Until indicators are found that can, we will still recommend that our patients discuss any health information that they encounter on the Internet with us before using it.—AARON E. CARROLL, MD, SUNIL SALUJA, MD, PETER TARCZY-HORNOCH, MD

References


In reply:—We appreciate this opportunity to clarify the methodology and conclusions of our paper “Indicators of Accuracy of Consumer Health Information on the Internet,”1 in response to the letter from Drs. Carroll, Saluja, and Tarczy-Hornoch.

Carroll et al. may be correct that consulting a physician is the best way for people to obtain accurate health information. However, the reality is that millions of people are using the Internet to obtain health information.2 And it is unlikely that they are restricting themselves to sites “run by large professional or governmental organizations.”3 In addition, the information that consumers find on the Internet frequently has an effect on treatment decisions.2

Given that people are using the Internet to obtain health information, it seems reasonable to provide them with guidelines for evaluating this information. Furthermore, such guidelines should be empirically grounded. A number of individuals and organizations have published guidelines for evaluating the quality of health information on the Internet.3 However, at the time we began our study, no empirical work had been done to test these guidelines. In other words, no one had checked to see whether any of the proposed indicators of accuracy really are indicators of accuracy.

Since ours was one of the first empirical studies to test these guidelines, we would not want to claim that our methodology is the ideal one. As more researchers address this important issue, better methodologies will undoubtedly be developed. Even so, we do not believe that our study is invalidated by the concerns raised by Carroll et al.

To determine whether a proposed indicator of accuracy really is an indicator of accuracy, one has to be able to measure the accuracy of information on a particular health topic. As Carroll et al. suggest, this is an extremely difficult task. In designing our study, however, we took encouragement from the fact that a number of studies (e.g., Berland et al.,4 Impicciatore et al.,3 Griffiths and Christensen,6 and Sandvik7) have undertaken to measure the accuracy of information on the Internet on a variety of health topics. In fact,