In-training Examination (ITE) standard scales, eg, composite, pediatrics, surgery. These properties allow the user to make informed decisions regarding the proper use and interpretation of the ABFP ITE standard scale scores. However, Web-based applications have been developed by non-ABFP organizations that allow residency programs to evaluate new nonstandard scales within detailed curricular areas. For instance, one Web site\(^1\) categorizes each ITE question and reports the number of questions a resident missed in each content area. The Web site suggests that the results may be used to modify existing curricula. Approximately 22% of family medicine Accreditation Council for Graduate Medical Education (ACGME)-accredited residencies are registered users of this site.

For the 2003 ITE examination, the Web site listed 191 content areas that were represented by only one ITE test question, eg, irritable bowel syndrome. Twenty-two other content areas were represented by two to 17 questions, eg, there were 17 questions regarding HIV. Thus, content areas were assessed by “scales” containing from one to 17 ITE questions. Whereas the psychometric properties of the standard ABFP ITE scales have been reported,\(^3\) the psychometric properties of these new nonstandard scales are unknown. This study assessed the reliability and positive predictive value (PPV) of these newly offered nonstandard ABFP ITE scales with small numbers of items such as those reported by this Web site.

**Methods**

We know from psychometric theory that test reliability is greatly influenced by the number of items on the test. In fact, the Spearman-Brown prophecy formula is often used to estimate the change in test reliability as the number of test items increases or decreases.\(^5\) To estimate the reliability of scales reported by the aforementioned Web site, we used the known reliability coefficient \((0.87)^{1/3}\) of the 340-item composite score of the ABFP ITE and the Spearman-Brown prophecy formula. Thus, we estimated the change in the reliability coefficient as the number of items on the ITE decreases from the 340 to 17 or less given that the reliability of the 340-item composite score is 0.87. We then used these reliability estimates and Monte Carlo methodology\(^5\) to estimate the PPV.

**Results**

Decreasing the ITE to 17 items results in an estimated reliability coefficient of 0.25. Decreasing the ITE to one item results in an estimated reliability coefficient of 0.02 for each of the 191 one-item scales. Results indicate that for a 17-item scale with a reliability coefficient of 0.25, the PPV is approximately 27%; for the 172 one-item scales, the PPV is approximately 11%.

**Conclusions**

If we make inferences about a resident’s fund of knowledge based on an incorrect response to a single item, about 89% of our inferences will be false. Even in the best-case scenario of the 17-item scale, 73% of residents identified as having knowledge deficit in HIV will be false positives. Although it is reasonable to explain the correct and incorrect answers to individual test items, it is entirely unreasonable to make inferences about large domains of knowledge based on scales with small numbers of items, low reliability, and low PPV. Educators should not try to retrofit the standard ABFP ITE and use the results in a way that was not intended. Making changes in curriculum or lecture content based on responses to a single item or small “multi-item” scales is not appropriate.

**Medical Decision-making Risk Management Based on Aeronautical Model**

To the Editor:

In the field of aviation, it was decided a number of years ago that something different was needed to reduce the number of accidents. Through various studies, it was determined that 85% of all accidents and 52% of fatal general aviation accidents were due to pilot error or poor decision-making risk management. Therefore, the Federal Aviation Administration (FAA) embarked on a process of determining and evaluating the attitudes, behavioral traps, stresses, and other items that enter into decision making by pilots. After 12 years of research, development, and testing, the result was publication in 1987 of manuals oriented to the decision-making needs of pilots. The effectiveness of the process was validated in six independent studies. After pilots received the training, the errors in decision making decreased—leading to a 10% to 50% drop in accidents.

Pilots and physicians have a similar or the same type of makeup, ie, goal oriented, self confident, sense of invulnerability, and macho. Being a physician and a pilot, I thought maybe the malpractice problem could be attacked in the same fashion.

The scope of the malpractice problem in the United States is impressive. According to statistics that
Letters to the Editor

Volume 36, No. 6

I think that most of the checklists and techniques used in aviation for risk management in decision making can be directly applied to medicine. We should incorporate medical decision-making risk management into our residencies and provide it to already-practicing physicians. We are embarking on a trial of risk management training in medical decision making in our program. This will evolve over time and include periodic evaluation of the progress of our residents. If we can accomplish the objective of physicians incorporating medical risk management into their practices, I believe that we may in the long run start to get a handle on the malpractice crisis in this country.

Stephen C. Stripe, MD

Minot Center for Family Medicine

University of North Dakota

can be found on the Internet from the Institute of Medicine, it is estimated that 44,000 to 98,000 people die per year due to medical error. It is further estimated that 1 million people are injured due to medical error. The cost to society estimates run between $17 to $29 billion per year. The RAND Corporation has done a study that showed that 35%—40% of autopsies were missed diagnoses that directly resulted in the cause of death of those patients. Also, 5% of US physicians are responsible for 54% of all malpractice claims. And, these estimates are based on hospital statistics only. The malpractice problem in this country is complex and multifaceted. Besides physicians’ involvement in the malpractice climate, lawyers and insurance companies are also involved. We as physicians cannot do anything directly about the lawyers or the insurance companies, but we can take responsibility for our part and try and do something about that.

Physicians must be proactive, not reactive, to risk management. Many of the risks are insidious and are not perceived as being intuitive or obvious. Medical training almost exclusively focuses on skill development (mental and physical), not risk management. And, like aviation, most of the medical malpractice is probably due to poor risk management in medical decision making. And, like aviation, with training in risk management of medical decision making, we may have a similar impact on malpractice in the United States. That is to say an improvement of 10% to 50% could represent 9,800 to 49,000 lives saved per year and $2.9 to $14.5 billion per year saved.