Scientific knowledge is advanced through hypothesis testing and rigorous evaluation. We can’t be true believers that a new treatment works before we have tested it, or we are likely to subject patients to ineffective and potentially harmful treatments. Ethical concerns are present in clinical trials when the investigator believes that one arm of the trial is more beneficial than another when in fact there is no actual proof that the benefit exists. Thus, a central principle of clinical research is that there be genuine uncertainty about how a new treatment will compare with the standard of care. This principle is called equipoise.

Innovations in medical education should also adhere to the principle of equipoise and be subjected to rigorous evaluation, but often there is tension between the desire for evaluation and the desire for educational equivalency. Educational equivalency requires that educationally equivalent experiences should be provided for all of the learners in each class. Ensuring educational equivalency means that when we implement new curriculum, either in terms of what we teach or how we teach it, the new methods should whenever possible be shown to be at least as effective as the old curriculum.

This issue of educational equivalency is a possible barrier to evaluation when one considers that students are actually customers paying many thousands of dollars per year for an education. Recruiting medical students into a trial of a new curricular innovation when the impact of the new curriculum could affect their grade point average or their National Board of Medical Examiners (NBME) score is problematic. Medical students are wary of interventional but also may be reluctant to be enrolled in the placebo or “usual care” group. Consequently, many curricular changes are implemented across the entire class or institution because they are seen as novel and innovative but are not evaluated rigorously because students/instructors all want the new and improved version. However, it is important to remember that either improvement or non-inferiority are only hypotheses and that testing our hypotheses through research is required to prove equivalency or better. Without rigorous evaluation it is unclear that the new curriculum is actually improved. Consequently, with significant dollars and careers at stake it is difficult to undertake trials of new educational programs in a way that manifests equipoise.

An example of the introduction of an innovative educational curriculum and the value of evaluation was the introduction of the problem-based learning (PBL) curriculum. When PBL was first developed by Harold Barrows at McMaster Medical School in 1969, it was thought that learning in small groups using clinical cases might be interesting to medical students. Many educators were drawn to the concept of teaching cases in small groups instead of teaching in large lecture halls. However, there was an obvious potential downside to such a substantial change in providing educational content to medical students. Concerns about increased costs with PBL (principally for more faculty time and more small classrooms) needed to be offset by improved educational outcomes. Being educationally equivalent wasn’t enough. PBL had to be better than the traditional way to justify the increased costs. It was important that rigorous evaluations were undertaken with equipoise rather than just assuming that new meant improved.

Almost 25 years after PBL was introduced, two large meta-analyses were conducted, examining more than 40 studies to assess whether PBL was equivalent or better to traditional strategies in teaching medical students. The results were mixed. In terms of academic achievement, which included NBME I, “other knowledge tests” and incidence of academic failure or attrition, the traditional students were found to be statistically favored. The PBL curriculum, however, was found to be
significantly superior with respect to students' program evaluation (ie, students attitudes and opinions about their curriculum). Using most of the other measures, there has been no statistically significant difference between PBL and “traditional” curricula. Recently, a new meta-analysis was performed looking at the effects of PBL on physician competency. This systematic review found that PBL teaching had moderate to strong levels of evidence in benefiting physician competencies in coping with uncertainties, appreciation of legal and ethical aspects of health care, communication skills, and self-directed continuing learning. The arrival of this evidence in support of PBL may, however, have come too late to save the PBL movement. One important lesson of the PBL experiment to date is that if we can agree on the important educational outcomes that we want to have changed by the intervention, and we can perform well-organized research on an curricular change before we make large-scale change, we may advance educational innovation in a more efficient and sustainable way.

New technology and distance learning continue to offer opportunities to develop innovative ways of teaching medicine. It is important that evaluations of these strategies are undertaken and with equipoise and scientific rigor. Assessing the value of these innovations has led to some exciting findings for medical educators. Research has shown that standardized patients (SPs) can teach physical exam skills to medical students just as effectively as physician faculty at a reduced cost to the medical school. Research has shown that a highly immersive virtual human (VH) is a better assessment tool than a standardized patient for consistent assessment of the learners' presentation and physical examination skills. Further, one evaluation found that simulation-based training was superior to problem-based learning in attainment of critical assessment skills. However, not all new innovations are better than the current standard. One evaluation found in the case of adding ultrasound to a second-year medical school clinical skills course that this new addition to the curriculum did not improve student scores on an objective structured clinical exam (OSCE).

Are there strategies that might help us conduct evaluations and still provide educational equivalency? One strategy is to use a cross-over design where at some point in the study all students received the intervention. Another method that has been used in quality improvement research is a pragmatic cluster-randomized trial design. Two theoretically equivalent educational interventions could be given at the same time, and outcomes measures at the end of the study could test for equivalence. Such a design would avoid randomizing a student to no intervention, potentially preserving their enthusiasm about their education. Finally, educators should go back after the trial and offer the new curriculum to any students who did not participate in the intervention if the intervention has been found beneficial.

It is important that we keep the creative juices flowing while proposing novel and innovative ways of teaching. At the same time, we have to keep in mind that new doesn’t always mean improved. The real value to students is to teach them through strategies that work, and it is incumbent upon us to be objective in our evaluations to provide that evidence. Educational equivalency and equipoise can coexist if we are careful in our study design and rigorous in our study methods. Improving medical education requires a scientific approach no less rigorous than clinical research.

CORRESPONDING AUTHOR: Address correspondence to Dr Egan, University of Chicago Pritzker School of Medicine, Department of Family Medicine. 773-834-2850. megan@medicine.bsd.uchicago.edu.

References