Best Research Evidence for Physical Medicine and Rehabilitation

Eugene Komaroff, PhD^{1,2}; Joel A. DeLisa, MD, MS^{1,2}

¹Kessler Foundation Research Center, West Orange, New Jersey; ²UMDNJ–New Jersey Medical School, Newark, New Jersey

Received September 9, 2008; accepted September 14, 2008

Abstract

Summary: Evidence-based medicine is a strong movement in this century, and randomized clinical trials continue to be the best level of evidence for establishing cause–effect relationships between treatment interventions and outcomes. The field of physical medicine and rehabilitation has many excellent research questions on the effects of treatment but seems to rely mostly on weak observational methods (eg, chart review, case series, and single-group designs) for answers. This paper highlights 3 basic and relatively simple principles of good experimental design: control, randomization, and replication that were developed by R. A. Fisher for large, complex, agricultural field trials. The principles diffused readily into many scientific arenas, and Fisher even applied the principles in his backyard studies into Mendelian genetics. The thoughts of R. A. Fisher, arguably the most influential statistician of the 20th century, on the promise and challenges of randomized clinical trials in medicine should motivate rehabilitation clinicians to do their own small-scale controlled trials, with Institutional Review Board approval, when faced with equally realistic and plausible treatment options for their patients.

J Spinal Cord Med. Feb 2009;32(1):3–5

Key Words: Evidence-based medicine; Clinical trials; Randomization; Cause–effect; Rehabilitation, physical

Evidence-based medicine (EBM) is “the integration of best research evidence with clinical expertise and patient values” (1). Clinical expertise and patient values are certainly not new to medicine, but “best research evidence” has grown in both coinage and controversy in the last 2 decades. According to a recent newspaper article: “The sad truth is that less than half of all medical care in the United States is supported by good evidence that it works” (2). The Congressional Budget Office is cited as the source of the “sad truth,” but there is no elaboration on the meaning of “good evidence.” Another leading US newspaper declared that politicians are “ramping up federal funding for so-called evidence-based medicine,” but lamented that “evidence-based programs are largely driven by the political imperative to cut costs—not the medical imperative to give patients the best care possible” (3). Regardless of the forces motivating EBM, the idea that physicians should rely on best evidence in their treatment decisions is clearly

reaching widespread public awareness. Although best evidence may be intuitively appealing, its vitality in EBM depends on fundamental scientific research principles.

EBM has promulgated a hierarchy of best research evidence with systematic reviews of homogenous, high-quality, randomized controlled trials at the top of the pyramid. Below that are individual, high-quality, randomized controlled trials, followed by nonrandomized controlled trials, cohort and case-control observational studies, and at the very bottom (worst evidence) are descriptive, nonconsecutive case series and expert opinion (eg, http://www.mcw.edu/FileLibrary/User/fvastalo/Oxford_Levels.pdf). Although the hierarchy is controversial in medicine (4), it is comforting that EBM has gained a foothold in physical medicine and rehabilitation (5–8). However, only a few rehabilitation studies reach the top of the evidence pyramid because of apparent difficulties in using randomization to allocate subjects to treatments (9). This is unfortunate because, without randomization, interventional trials and observational studies can produce strong evidence of association, but as is well known, correlation is not causation. In the absence of randomization, every effort has to be made to exclude every other possible explanation before a cause–effect argument becomes convincing. Neverthe-

---

Please address correspondence to Joel DeLisa, MD, MS, Kessler Foundation Research Center, 1199 Pleasant Valley Way, West Orange, NJ 07052; p: 973 243 6905; f: 973 324 6969 (e-mail: jdelisa@kesslerfoundation.net).

less, it should be clearly understood that randomized and nonrandomized studies are not at opposite poles of a quality continuum—they are essentially different research paradigms and both can suffer equally from lack of quality (10).

The Consolidated Standards of Reporting Trials (CONSORT) has been recently extended to clinical trials with nonpharmacologic treatments that are prevalent in “rehabilitation, surgery, technical interventions, psychotherapy, behavioral interventions, implantable and non-implantable devices, and complementary medicine” (11). This new standard offers guidance on problematic issues that plague rehabilitation researchers (12) such as blinding, complexity of interventions, and nonequivalence in expertise of care providers and centers.

Comprehensive papers on clinical trial methodology have been published for rehabilitation research (13,14); however, their thoroughness may have discouraged or intimidated clinically oriented physicians from designing and conducting their own randomized clinical trials. It is our contention that with proper Institutional Review Board (IRB) oversight, clinical trials can be designed, conducted, analyzed, and results can properly be interpreted by practicing physiatrists and other rehabilitation professionals. Toward that end, we provide a few simple guidelines for designing small clinical trials based on the ideas of R.A. Fisher, arguably the most influential statistician of the 20th century (15). Fisher invented ANOVA, the exact test, and method of maximum likelihood among many other noteworthy contributions to mathematical statistics.

Fisher (16) tested treatment effects by accommodating nuisance factors in his designs for agricultural research. His ideas for simple experimentation diffused readily into other research arenas; however, Fisher observed that they met with resistance in medicine (17):

“Medical research has had to rely a good deal on uncontrolled experiments, uncontrolled observations. Dr. Snow, who studied and in the end quelled the occurrence of cholera in London, used a very large number of different types of inquiry in order to gain sufficient confirmation of his important conclusion, namely, that it was fecal contamination in the water supply that was responsible for the cholera.... Consequently, when inconclusive evidence is criticized on the grounds that it is inconclusive, it is not uncommon for medical men to defend it, perhaps with a certain indignation, on the ground that in the past medical science has made notable advances primarily—not solely, never only, but primarily—by the observational method” (pp 156–157).

Fisher (16) identified 3 first principles that permit causal inferences from simple experiments: control, randomization, and replication. Fisher applied these principles not only to large, complex agricultural field trials but also in his own small-scale studies on Mendelian genetics that he performed in his backyard (18). Next,

we next consider each principle in the context of clinical research.

There have to be at least 2 treatment groups in a clinical trial. A patient in the control group gets the standard of care, which can be no treatment (placebo), whereas a patient in the treatment group receives the promising, but unproven experimental intervention or treatment. For example, in a pharmacologic study, a control subject would get a pill that feels, looks, tastes, and smells exactly like the experimental pill but lacks the active ingredients. Besides this difference, all subjects must receive exactly the same care, attention, and follow-up from study physicians and staff. Cross-over designs in which subjects serve as their own controls are legitimate; however, these designs are feasible mainly for chronic conditions and interventions with relatively short-lasting treatment effects (19).

Fisher (16) was well aware of the challenges that randomization posed for clinical trials:

“Deliberate experimentation has not been very widely used in the medical field. There is a movement at the present time to organize clinical trials, let us say, of new drugs or of new antibiotics in such a way that an impartial judgment of comparing the new with the old may be obtained by hospital staffs. And that would involve applying the new and the old at random to some of the hospital patients. So long as no body of medical opinion can say with confidence that one is better than the other, or perhaps that in matters usually as complicated as this, for what cases one drug is the better and for what cases the other—so long as the state of ignorance remains, it would be perfectly fair, I think, to clear the air by such simple experimentation” (p 156).

In the above, Fisher foreshadowed what is now called equipoise to justify randomization in clinical trials. Equipoise does not mean that a clinician/investigator has to believe that all treatment interventions are equal but only that there exists “an honest, professional disagreement among expert clinicians about the preferred treatment” (20).

Fisher (16) is clear that randomization does not mean assigning the next intervention to the next study participant in whatever order that occurs to the investigator. Proper randomization requires the use of an external process or mechanism that is known to operate randomly, such as tossing coins or dice, drawing cards from a shuffled deck, selecting the next digit from a random numbers table, or generating the next number with a validated computer random numbers generator. In addition, proper randomization requires an unpredictable allocation sequence so that investigators who have a stake in the outcomes can not predict the next treatment assignment. Otherwise, the study drops a notch on the EBM “strength of causal evidence” pyramid. Undoubtedly, physicians loath to surrender important patient treatment decisions to chance, but randomization remains the best method for achieving comparability

between experimental groups and is the scientific foundation for statistical inference (19).

Regarding replication, consider the following thought exercise: Assume that all subjects enrolled in a controlled clinical trial are identical in every way and thus have the same baseline scores on a perfectly reliable measure of pain. Each subject receives either the investigational treatment intervention or the usual standard of care. After a reasonable period of time during which it is assumed the intervention had a chance to have an effect, patients fill out a pain-rating scale. Now, we take any 2 subjects at random and find a difference in their pain scores. Clearly something has produced the difference, but was it the investigational treatment or merely chance (random error)? If both subjects were in the investigational treatment group, we can be certain it was chance because they were both exposed to the same treatment. Similarly, if both subjects were from the usual standard of care group, the difference in their scores must also be caused by chance. However, if they were from different treatment groups (one investigational and the other standard), it could be either a “treatment effect” or a “chance effect”—the 2 effects are hopelessly intertwined or confounded. It is replication or addition of more subjects that help us detect the true effects.

In closing, we paraphrase DeLisa (21) by emphasizing that clinical trials carry more weight than ever in health care. Randomized clinical trials are the foundation for EBM and are a standard for good medical practice. We urge rehabilitation professionals not to waste valuable time and energy attacking the EBM hierarchy, but instead focus on creative ways to design and conduct proper randomized controlled trials that accommodate both treatment and nuisance factors in the complex and challenging arena of rehabilitation research.

REFERENCES

1. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based Medicine: How to Practice and Teach EBM*. 2nd ed. Churchill Livingstone: Edinburgh; 2000.
2. Rosenthal A. The high cost of health care (editorial). *New York Times* November 25, 2007.
3. Pitts P. ‘Evidence-based’ Rx miscues. (commentary). *The Washington Times* April 15, 2008.
4. Concato J, Shah N, Horwitz RJ. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25):1887–1892.
5. DeLisa JA. Shaping the future of medical rehabilitation research: using the interdisciplinary research model. *Arch Phys Med Rehabil*. 2004;85(4):531–537.
6. Frontera WR, Fuhrer MJ, Jette AM, et al. Rehabilitation medicine summit: building research capacity. *J Spinal Cord Med*. 2006;29(1):70–81.
7. Ebenbichler G, Kerschman-Schindl K, Brockow T, Resch KL, Ludwig K. The future of physical & rehabilitation medicine as a medical specialty in the era of evidence-based medicine. *Am J Phys Med Rehabil*. 2008;87:1–3.
8. Heinemann A. State-of-the-science on postacute rehabilitation: setting a research agenda and developing an evidence base for practice and public policy. An introduction. *J Spinal Cord Med*. 2007;30(5):452–457.
9. Johnston M, Sherer M, Whyte J. Applying evidence standards to rehabilitation research. *Am J Phys Med Rehabil*. 2006;85:292–309.
10. Guyatt GH, Oxman AD, Vist GE, et al. for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–926.
11. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. for the CONSORT group. Extending the CONSORT Statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med*. 2008;148:295–309.
12. Whyte J. Clinical trials in rehabilitation: what are the obstacles? *Am J Phys Med Rehabil*. 2003;82(suppl 10):S16–S21.
13. Wyndaele JJ, ed. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel [theme issue]. *Spinal Cord*. 2007;45(3):190–242.
14. DeLisa JA, Johnston MV, eds. Clinical trials in medical rehabilitation: enhancing rigor and relevance [supplement]. *Am J Phys Med Rehabil*. 2003;82(10):S3–S57.
15. Efron BR. Fisher in the 21st century. *Stat Sci*. 1998;13(2):5–122.
16. Fisher RA. *The Design of Experiments*. 7th ed. New York: Hafner Publishing; 1960.
17. Fisher RA. Cigarettes, cancer, and statistics. *Centen Rev*. 1958;2:151–166.
18. Box JF. R.A. Fisher: *The Life of a Scientist*. New York: Wiley; 1978.
19. Friedman LM, Furbur CD, DeMets DL. *Fundamentals of Clinical Trials*. 3rd ed. New York: Springer-Verlag; 1998.
20. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317(3):141–145.
21. DeLisa JA. Clinical trials: the cornerstone of medical rehabilitation. *Am J Phys Med Rehabil*. 2003;82(suppl):S1–S2.