

Hopes and Dreams: Healthcare Reform and Comparative Effectiveness Research for Personalized Treatment in Psychiatry and Neurology

Andrew A. Nierenberg, MD

President Barack Obama has advocated for healthcare reform that results in affordable high quality care with increased safety and decreased cost. To achieve this goal, President Obama and Congress have championed comparative effectiveness research (CER) backed by American Recovery and Reinvestment Act stimulus monies and perhaps more enduring sources of funding. Through this funding, the administration hopes that the results from CER will start to fill important gaps in evidence and provide data to physicians and healthcare workers so that they can make better and more rational decisions for their patients.

CER, however, presents a wide variety of challenges to researchers and clinicians. If designed correctly, CER can achieve President Obama's goals. Researchers will compare competing treatments and, if a winner emerges, the superior treatment will be recommended (and perhaps preferentially paid for by insurance). If designed incorrectly, CER can produce results that can be misinterpreted and misused by regulators and clinicians to favor less expensive but inferior treatments.^{1,2} The purpose of this editorial is to briefly review aspects of CER that will lead to more good

than harm in psychiatry and neurology.

First, what is CER and how does it differ from trials designed to obtain FDA approval (ie, registration studies)? A CER study includes:

- a) a broad range of patients similar to those who would be treated in the community in a standard clinical setting. Such patients have been systematically excluded from efficacy studies that compared active treatments to placebo or other active comparators;
- b) outcomes that integrate sustained response or remission along with tolerability and attrition. Registration studies have principally presented single time point outcomes on last observation carried forward data;
- c) flexible additional treatments to manage specific symptoms commonly associated with the disorder, side effects, and comorbid conditions;
- d) continuation treatment for a sufficient period to determine persistence of benefit and enduring or late developing adverse effects; and
- e) follow up of all participants to study completion regardless of compliance or adherence with taking their medication.

Dr. Nierenberg is professor of psychiatry at Harvard Medical School, co-director of the Bipolar Clinic and Research Program, and associate director of the Depression Clinical and Research Program at Massachusetts General Hospital (MGH) in Boston.

Faculty Disclosures: Dr. Nierenberg consulted to or served on the advisory boards of Abbott, Appliance Computing, Inc., Brain Cells, Inc., Bristol-Myers Squibb, Eli Lilly, EpiQ, Forest, GlaxoSmithKline, Janssen, Jazz, Merck, Novartis, Pamlab, Pfizer, PGx Health, Pharmaceutica, Schering-Plough, Sepracor, Shire, Somerset, Takeda, and Targacept; he has received research support from Cederroth, Cyberonics, Forest, Medtronic, NARSAD, the NIMH, Ortho-McNeil-Janssen, Pamlab, Pfizer, Shire, and the Stanley Foundation through the Broad Institute; he has received past support from Bristol-Myers Squibb, Cederroth, Eli Lilly, Forest, GlaxoSmithKline, Janssen, Pfizer, Lictwer Pharma, and Wyeth; he has received honoraria from the MGH Psychiatry Academy (MGHPA activities are supported through Independent Medical Education grants from AstraZeneca, Eli Lilly, and Janssen); he earns fees for editorial functions for *CNS Spectrums* through MBL Communications, Inc., and *Psychiatric Annals* through Slack, Inc.; he receives honoraria as a CME Executive Director for the *Journal of Clinical Psychiatry* through Physicians Postgraduate Press; he has been on the speaker's bureaus of Bristol-Myers Squibb, Cyberonics, Eli Lilly, Forest, GlaxoSmithKline, and Wyeth; he has received royalties from Cambridge University Press and Belvoir Publishing; he owns stock options in Appliance Computing, Inc.; and owns the copyrights to the Clinical Positive Affect Scale and the MGH Structured Clinical Interview for the Montgomery Asberg Depression Scale, exclusively licensed to the MGH Clinical Trials Network and Institute.

Let's start with the simplest of questions. Which is better, treatment A or treatment B? Both have proven efficacy (superiority over placebo for the outcome of interest), but researchers have never compared the two. Treatment A has one set of benefits and harms while treatment B has another. CER randomizes appropriate patients to receive A or B and the researchers compare overall outcomes. Even with this simple example, the problem is that the average outcomes are compared and important differences for relevant subgroups can be obscured.³ Subgroups can show the opposite from overall average results in terms of benefit or harm as well as cost effectiveness. The challenge is to identify the relevant subgroups.

Now let's refine the clinical question. Not "which treatment is better" but, instead, "who should receive treatment A and who should receive treatment B"? Or, more precisely, who should receive treatment A and not treatment B and who should receive treatment B and not treatment A? These questions cannot be answered by a simple CER design, but instead by a more complicated individualized CER study or *i*-CER.⁴ The challenge of *i*-CER is to not only choose the most relevant variables to describe patients, but also to perform the most clinically relevant analyses that can actually be used by clinicians. These analyses use a moderator approach, with criteria set forth by Kraemer and colleagues.⁵ A moderator variable must precede treatment, be uncorrelated with treatment, and have interactive effect with treatment on outcome. To be clinically useful, a moderator must be prevalent and easy enough to measure in the treated population, and have a large enough effect size on differential outcome (those with the moderator variable who respond differently to treatment A compared to treatment B) so that the study can have sufficient statistical power to show a clinically important difference in outcomes.

Moderator variables can be demographic/clinical (eg, age, sex, severity or duration of illness) or biomarkers (eg, quantitative electroencephalogram, cytokines, growth factors, genes). With several biomarkers, a "biosignature" can result that can profile patients so they can be matched to the best treatments to maximize benefit and minimize harm. But moderator studies that generate clinically

useful biosignatures are just the first stage in developing personalized medicine. Once researchers identify these biosignatures, the study needs to be replicated, and once replicated, the next generation of studies need to be conducted.

Let's go back to *i*-CER for treatment A and treatment B. Let's say that the *i*-CER shows that 70% of patients with biosignature X respond to treatment A but only 40% respond to treatment B. It also shows that 70% of patients with biosignature Y respond to treatment B but only 40% respond to treatment A. Using just this study risks having these results apply to the group that was included in the study and is not necessarily generalizable to most patients. To show generalizability, the study must be replicated. Again, patients are randomized to treatment A and treatment B and biosignature X and Y are analyzed as moderators. Let's say that the study is replicated. Good, but the job is not finished. Note that patients with each biosignature were not actually randomized directly to each treatment. The final study takes patients at baseline, stratifies them by biosignature X and Y, and then randomizes all of the patients with biosignature X to treatment A or B and all of the patients with biosignature Y to A or B. Only then will we know if those with biosignature X truly respond better to A and those with biosignature Y truly respond preferentially to treatment B.

Successful programs of CER that progress to *i*-CER and personalized treatment depend on many factors including large enough studies that are powered on the relevant moderators to find clinically relevant differences. The hope is that President Obama and Congress provide sufficient sustained funding to do this difficult work. The dream is that we have real personalized medicine for our patients with psychiatric and neurological disorders. **CNS**

REFERENCES

1. Kraemer HC, Glick ID, Klein DF. Clinical Trials Design Lessons From the CATIE Study. *Am J Psychiatry*. 2009;166:1222-1228.
2. Meltzer DO, Basu A, Meltzer HY. Comparative Effectiveness Research For Antipsychotic Medications: How Much Is Enough? *Health Aff*. 2009;28:w794-w808.
3. Fava GA. The intellectual crisis of psychiatric research. *Psychother Psychosom*. 2006;75:202-208.
4. Basu A. Individualization at the Heart of Comparative Effectiveness Research: The Time for *i*-CER Has Come. *Med Decis Making*. 2009;29:N9-N11.
5. Kraemer H, Wilson T, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry*. 2002;59:877-883.