

# The Need for Effectiveness Trials to Inform Evidence-Based Psychiatric Practice

Isabel T. Lagomasino, M.D., M.S.H.S.

Megan Dwight-Johnson, M.D., M.P.H.

George M. Simpson, M.D.

Recent reports have highlighted significant variations in the quality and safety of medical care provided in the United States (1). Such reports, along with the need to deliver cost-effective care, have driven efforts to organize and disseminate evidence-based medical practices. Within psychiatry, there has been a dramatic increase in the number of treatment algorithms and expert consensus guidelines for major mental disorders. However, their usefulness and potential for implementation in real-world clinical settings is limited by significant gaps in existing evidence for common clinical problems. This column reviews the key features and unique challenges of effectiveness trials and describes how they might contribute to our current knowledge base.

## Effectiveness trials and efficacy trials

A major limitation of the current evidence base in psychiatry is its almost exclusive reliance on results from randomized clinical trials conducted by expert practitioners in academic or commercial centers with patients

who go through a highly selective screening process. These trials, which remain the gold standard for establishing the efficacy of new treatments, determine “the maximum potential benefit to be derived from an intervention in ideal circumstances and under a controlled environment” (2). However, often an “efficacy-effectiveness gap” exists between the results achieved in efficacy trials and those observed by usual practitioners treating real patients in common settings. This finding has led to recent calls to conduct mental health effectiveness trials, or practical clinical trials, whose results would complement those of efficacy trials and increase the practical applicability of the current evidence base for clinicians and health care policy makers (3). Effectiveness trials differ from efficacy trials in several key features: they compare clinically relevant treatments; include less homogeneous study samples; are conducted in real-world practice settings, often with usual providers; and examine a broad range of outcomes (4).

## Treatments

Although efficacy trials often compare a single active treatment to a placebo, effectiveness trials reflect real-world clinical situations by comparing two or more active treatments, which may include nonpharmacologic alternatives. For example, in the ongoing Sequenced Treatment Alternatives to Relieve Depression (STAR\*D) trial for nonpsychotic major depressive disorder, patients with-

out a satisfactory response to citalopram are randomly assigned to four switch options (sertraline, bupropion, venlafaxine, and cognitive therapy) or three augmentation strategies (bupropion, buspirone, and cognitive therapy) (5). The study algorithm leads those who do not show significant improvement through a sequence of subsequent relevant switch or augmentation strategies. The STAR\*D trial will thus address the effectiveness of commonly used treatment strategies that currently lack an evidence base. Similarly, as part of the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), several effectiveness trials will be conducted to compare commonly used but unproven treatments for bipolar depression and relapse prevention (6). In the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project, a series of blinded and open-label randomized trials of active treatments will help answer questions about the comparability of commonly used medications for schizophrenia and strategies for addressing treatment resistance (7).

## Samples

Whereas efficacy trials carefully select highly motivated, homogeneous populations without significant psychiatric or medical comorbid conditions, patients from real-world settings are frequently ambivalent about treatment, come from a variety of ethnic and socioeconomic backgrounds, and have significant comor-

---

*Dr. Lagomasino and Dr. Simpson are editors of this column and are affiliated with the department of psychiatry and behavioral sciences at the University of Southern California Keck School of Medicine, 2020 Zonal Avenue, IRD 218, Los Angeles, California 90033 (e-mail, lagomasi@usc.edu). Dr. Dwight-Johnson is with the department of psychiatry and behavioral sciences at the University of Washington School of Medicine in Seattle.*

bid conditions. For example, compared with patients with Alzheimer's disease in community samples, similar patients enrolled in clinical trials are better educated, wealthier, younger, less medically ill, and less behaviorally disturbed; they deteriorate more slowly and experience lower mortality rates (8). Only 4 to 8 percent of the patients with Alzheimer's who were listed in a large clinical database would have met eligibility criteria for enrollment in two clinical trials of cognitive enhancers (8).

Similarly, Keitner and colleagues (9) evaluated the generalizability of recent efficacy trials of antidepressants for major depressive disorder and determined that 86 percent of study applicants were ineligible for the trials: 17 percent had bipolar disorder, 16 percent abused substances, 14 percent had mild depression, 13 percent had a medical contraindication, and 12 percent were using other psychotropics that were not permitted in the trial. These trials thus do not address the common clinical presentations seen in practice.

In contrast, Partners in Care, a trial of quality improvement strategies for depression in managed primary care, screened patients in waiting rooms to recruit a diverse patient population (10). Of the 1,356 patients enrolled, 40 percent had subthreshold depression commonly found in primary care, 43 percent had comorbid anxiety disorders, 33 percent had problematic alcohol or drug use, and 79 percent had one or more chronic medical conditions. Additionally, 43 percent of study participants were from a racial or ethnic minority group.

### Settings

Effectiveness trials often take place in real-world settings and use usual providers. Efficacy trials are often conducted in specialty academic or commercial research settings, even though epidemiologic evidence suggests that most patients with depression and anxiety who receive treatment do so in primary care (11,12). Participants in efficacy trials also receive costly, state-of-the-art interventions and an intensity of clinical care not frequently available to patients in

usual-care settings. For example, even participants who receive a placebo are scheduled for weekly follow-up visits that involve supportive interactions and the monitoring of active symptoms. Even the process of obtaining informed consent educates and engages participants to a greater extent than may occur in usual clinical settings. Providers of medication or psychotherapy are often experts in their field, potentially influencing the quality and outcomes of treatments delivered. Thus, although efficacy trials may represent "cookies baked by an expert chef according to a recipe," real-world treatments are similar to "cookies baked at home, loosely based on a recipe, often with some substitution of ingredients" (personal communication, Duan N, 2004).

Effectiveness trials may help us understand how to best implement treatments in real-world settings to optimize outcomes. Partners in Care (10) and Project Improving Mood: Promoting Access to Collaborative Treatment (IMPACT) (13), among others, implemented depression treatment in real-world primary care settings, using usual care as the comparison group, with usual providers delivering study treatments after receiving training by study experts. In Partners in Care a significant difference was seen between the two groups: 51 percent of patients in the intervention group compared with 40 percent of patients in the usual-care group received minimally adequate depression care ( $p < .001$ ) (10). In Project IMPACT a significant difference was also seen between the two groups: 78 percent of patients in the intervention group compared with 54 percent of patients in the usual-care group received adequate depression care ( $p < .001$ ) (13).

### Outcomes

Finally, effectiveness trials tend to examine a broader range of outcomes than traditional efficacy trials. They frequently include outcomes that are highly relevant to patients, including not only symptom burden but also functioning, quality of life, satisfaction, and costs. Persons who make health care decisions often benefit from the inclusion of cost-effective-

ness and cost-utility outcomes, both of which determine the costs of clinical gains or improvements in quality of life, respectively (3).

For example, costs may be determined for the impact of treatments in terms of quality-adjusted life years, a measure that takes into account both the quantity and quality of life generated by an intervention. As a standardized measure, the quality-adjusted life year (QALY) can be used to compare the relative merits of different treatments across disease conditions and populations. For example, in Partners in Care the estimated cost per QALY gained for Latinos was \$6,100 or less for a quality improvement intervention that facilitated psychotherapy but more than \$90,000 for a quality-improvement intervention that facilitated antidepressant use. For whites, the estimated cost per QALY gained was \$30,000 for both interventions (14). A cost of \$50,000 per QALY is generally considered cost-effective and a good use of health care resources (15).

Effectiveness trials also examine outcomes over a longer, more clinically relevant time frame. In their comparison of olanzapine versus haloperidol plus benztropine for schizophrenia, Rosenheck and colleagues (16) assessed traditional outcomes of psychotic symptoms and extrapyramidal side effects, as well as quality of life and medical costs for 12 months. Longer follow-up periods also allow for evaluation of longer-term side effects, including weight gain, hyperlipidemia, and insulin resistance (17). Partners in Care is currently examining five-year outcomes that include employment status and household wealth.

Because of their differences in methods, results from efficacy and effectiveness trials sometimes differ. For example, several randomized trials have suggested that olanzapine was superior to haloperidol in the treatment of schizophrenia in terms of symptom improvement, extrapyramidal side effects, and quality of life. However, Rosenheck and colleagues' study (16), which was conducted in 17 Department of Veterans Affairs medical centers with usual providers following a study algorithm, found

no significant differences between patients assigned to haloperidol plus benztropine and patients assigned to olanzapine in terms of study retention; positive, negative, or total symptoms; extrapyramidal symptoms; or quality of life. Although olanzapine was less likely to produce akathisia and tardive dyskinesia, the drug also had greater potential for weight gain and increased cost. These findings may differ from those of previous efficacy trials at least in part because previous studies compared second-generation antipsychotics with haloperidol without prophylactic benztropine. Thus the comparator did not reflect “the next-best choice as used in optimal real-world practice” (18).

### Challenges of effectiveness trials

The gains to be made from effectiveness trials are not without costs or challenges. The most relevant and reasonable clinical treatments, including psychosocial therapies, must be identified for testing. Rational treatment algorithms should be developed on the basis of efficacy trials and, where evidence is lacking, expert consensus used to identify relevant comparison treatments. Such consensus might help eliminate the inclusion of inappropriate treatments sometimes used in clinical practice—for example, mega-doses of antipsychotics or off-label use of gabapentin.

Psychosocial interventions need to be appropriate for the level of training and supervision available in typical community settings. Because effectiveness studies use more diverse samples and treatment settings, they will require a larger sample to overcome “noise” and detect treatment effects. Settings as well as a diverse population of providers and patients, who may be naive to research, must be recruited. Also, providers, clinics and administrators must be able to implement a common intervention strategy across sites. Researchers may

be unfamiliar with the concerns and needs of community practices. For example, clinical practices may be wary of randomly assigning patients to treatment groups, especially to a placebo treatment group. Thus effective partnerships between research institutions and community practices must be established to share knowledge and staff, incorporating the values and priorities of community partners in the research design.

Relevant outcomes for diverse stakeholders must also be determined by consulting with patients, providers, and policy makers. Finally, because effectiveness trials require large numbers of participants and alternative settings, the costs of these trials far exceed those of smaller efficacy trials, and the mechanisms for funding research and community partners (subcontracting) may be quite cumbersome. A combination of funding streams, from both public and private agencies, is often required. Despite these challenges, the evidence gained from effectiveness trials is critical to promoting high-quality, cost-effective mental health care. ♦

### References

1. To Err is Human: Building a Safer Health System: Committee on Quality of Health Care in America. Edited by Kohn LT, Corrigan JM, Donaldson MS. Washington DC, Institute of Medicine National Academy Press, 1999
2. Bausewein C, Higginson IJ: Appropriate methods to assess the effectiveness and efficacy of treatments. *Journal of Palliative Medicine* 7:423–430, 2004
3. Wells KB: Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. *American Journal of Psychiatry* 156:5–10, 1999
4. Tunis S, Stryer D, Clancy C: Practical clinical trials: increasing the value of clinical research for decision making in clinical health policy. *JAMA* 290:1624–1632, 2003
5. Rush AJ, Fava M, Wisniewski SR, et al: Sequenced Treatment Alternatives to Relieve Depression (STAR\*D): rationale and design. *Controlled Clinical Trials* 25:118–141, 2004

6. Sachs GS, Thase ME, Otto MW, et al: Rationale, design, and methods of the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Biological Psychiatry* 53:1028–1042, 2003
7. Stroup TS, McEvoy JP, Swartz MS, et al: The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophrenia Bulletin* 29:15–31, 2003
8. Schneider LS, Olin JT, Lyness SA, et al: Eligibility of Alzheimer's disease clinic patients for clinical trials. *Journal of American Geriatrics Society* 45:923–928, 1997
9. Keitner GI, Posternak M, Ryan C: How many subjects with major depressive disorder meet eligibility requirements of an antidepressant efficacy trial? *Journal of Clinical Psychiatry* 64:1091–1093, 2003
10. Wells K, Sherbourne C, Schoenbaum M, et al: Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA* 283:212–220, 2000
11. Regier DA, Goldberg ID, Taube CA: The de facto US mental health services system: a public health perspective. *Archives of General Psychiatry* 35:685–693, 1978
12. Young AS, Klap R, Sherborne CD, et al: The quality of care for depressive and anxiety disorders in the United States. *Archives of General Psychiatry* 58:55–61, 2001
13. Unützer J, Katon W, Callahan CM, et al: Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* 288:2836–2845, 2002
14. Schoenbaum M, Miranda J, Sherbourne C, et al: Cost-effectiveness of interventions for depressed Latinos. *Journal of Mental Health Policy Economics* 7:69–76, 2004
15. Laupacis A, Feeny D, Detsky AS, et al: How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Canadian Medical Association Journal* 146:473–481, 1992
16. Rosenheck R, Perlick D, Bingham S, et al: Effectiveness and cost of olanzapine and haloperidol in the treatment of schizophrenia: a randomized controlled trial. *JAMA* 290:2693–2702, 2003
17. Meltzer HY: Putting metabolic side effects into perspective: risks versus benefits of atypical antipsychotics. *Journal of Clinical Psychiatry* 62(suppl 27):35–39, 2001
18. Rosenheck RA: Effectiveness versus efficacy of second-generation antipsychotics: haloperidol without anticholinergics as a comparator. *Psychiatric Services* 56:85–92, 2005