



The silver lining of recent effectiveness trials

ALAN F. SCHATZBERG

Department of Psychiatry and Behavioral Sciences,
Stanford University School of Medicine, Stanford,
CA, USA

Effectiveness trials have played an increasingly prominent position in the evaluation of the comparative benefits or adverse events of various psychotropic agents. Fleischhacker and Goodwin review results of several recent trials in

psychiatry and discuss their pros and cons. They argue for the need for such effectiveness studies in late Phase III or post approval Phase IV of the drug development cycle, but with some key caveats, emphasizing the need for incorporating key elements of traditional randomized clinical trials (RCT): randomization and concealment allocation. These would have strengthened many of the recent studies that have been wanting in terms





of what they have really taught us about optimal treatment. This was an unfortunate consequence of STAR*D, where Phase II inpatients could themselves choose to switch to a new agent or augment with the addition of a second one (1,2). While Fleischhacker and Goodwin's observations and suggestions are warranted, perhaps it might be wise to pause a moment and reflect on what we are trying to accomplish, where we have come from, and where should we be heading.

Clinical drug development in psychiatry has become largely focused on demonstrating efficacy by achieving two or more positive, pivotal trials in which an investigational agent is shown to be statistically more effective than placebo in alleviating a specific disorder (e.g., major depression) or less commonly specific symptoms across several syndromes (e.g., agitation in dementia, depression, etc.). Companies generally conduct 4-8 studies to yield at least two positive trials. A filing with the regulatory agencies often includes one or more failed or negative trials but, with enough positive trials and a side effect profile that was not severe or dangerous, an agent would likely be approved. A typical development program trial might include 1,000-1,500 patients exposed to the agent. Phase III trials may include active comparators, but rarely in sufficiently large numbers to allow for enough power to demonstrate superiority of an investigational agent over an available therapy. They are used largely for so-called assay sensitivity to assess the reasons for a so-called failed trial.

This type of efficacy approach has provided a pathway to approval but not for helping the clinician decide when to use a drug, particularly in relationship to older, available compounds that are frequently less expensive. Hence, this divide between efficacy and effectiveness.

To remedy this problem, the field in recent years has embarked on a number of effectiveness trials that many hoped would answer key questions and justify their costs. Unfortunately, many have argued that the studies have taught us little we had not known already and that the cost has not been justified. These trials

have not answered the key question for us: which medication strategy is best for a particular patient's disorder.

Yet the studies may have been useful for other reasons. They have allowed us as a field for perhaps the first time to conduct relatively large-scale trials. These have been common in treatment of patients with cancer or cardiovascular disease but have been rare in psychiatry. Indeed, several of these studies did recruit relatively large samples of subjects. Protocol development, study implementation, data collection and analysis were well coordinated and well implemented. And, while the lack of randomization and at times blinding has prevented many important questions to be answered, the foundation has been laid to conduct large-scale, true comparison studies in the future. They have also demonstrated that DNA samples can be collected to assess for genetic predictors of response.

A few comments on where we should be heading. Future studies can build on these new infrastructures but still require that we build in key features. As Fleischhacker and Goodwin point out, randomization is key to make real comparisons. This needs to be routinely built into trials. Comparing two known active strategies should make it easier for both patients and investigators to feel comfortable building in this essential feature in the trial.

Setting entry criteria to allow inclusion of as many representative patients is also essential. Here we may need to do some research and have active discussion regarding risks and ethics involved in re-exposure to a specific agent or class of agents. There has been a tendency to exclude subjects based on a past history of adverse event or lack of response to a specific agent or class. Not uncommonly we may exclude patients whose previous response – particularly if it were long ago – had been complicated by other factors (e.g., concomitant flu-like symptoms), or in whom adverse events were not particularly severe. All too often with studies in more chronic conditions we are confronted with the issue of previous treatment response. If we do include such patients, we may need to stratify in

the randomization on the basis of positive or negative response.

With more experience, large scale studies need to be reassessed as to optimal design. Should we be nesting substudies that compare two agents A vs. B, C vs. D within larger trials? Should we be employing adjustments of the randomization based on results to date (e.g., so-called play the winner strategies) (3)? These could all add power to the study design.

DNA samples for pharmacogenetics should be routinely collected in large scale studies or even in smaller studies within large scale drug development programs. Indeed, this is a plus of the recent studies, with several of them having already reported interesting genetic prediction data (4-6). Unfortunately, some DNA samples were not collected at baseline, prior to drug initiation, such that important data on dropouts or drug intolerance may have been lost. Thus, genetic samples should truly be routine parts of the design – i.e., collected in all subjects at baseline (7). These types of data can then be combined with clinical measures to develop moderators or predictors of response (8). As we collect more data, the field can develop criteria for assessing the utility of such predictors for drug selection and for determining when to adopt them clinically. This will help move us beyond a conclusion that the newer agents are not more beneficial than are older ones to a recommendation that a particular patient would best be treated with one or another drug. That is ultimately what we want out of these studies: greater beneficial effects for the patient in our office or in our waiting room. We then will need to all be prepared for a more individual-based practice. The recent effectiveness trials need to be seen as a step in the evolution of our clinical specialty, both in terms of research and treatment applications.

References

1. Rush AJ, Trivedi MH, Wisniewski SR et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med* 2000;354:1231-42.
2. Trivedi MH, Fava M, Wisniewski SR et al.





Medication augmentation after the failure of SSRIs for depression. *N Engl J Med* 2006; 354:1243-52.

3. Krishnan KRR. Efficient trial designs to reduce placebo requirements. *Biol Psychiatry* 2000;47:724-6.
4. Perlis RH, Purcell S, Fava M et al. Association between treatment-emergent suicidal ideation with citalopram and polymorphisms near cyclic adenosine monophosphate response element binding protein in the STAR*D study. *Arch Gen Psychiatry* 2007; 64:689-97.
5. Lekman M, Laje G, Charney D et al. The FKBP5-gene in depression and treatment response – an association study in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Cohort. *Biol Psychiatry* 2008;63:1103-10.
6. Campbell DB, Ebert PJ, Skelly T et al. Ethnic stratification of the association of RGS4 variants with antipsychotic treatment response in schizophrenia. *Biol Psychiatry* 2008;63:2-41.
7. Murphy GM Jr, Kremer C, Rodrigues HE et al. Pharmacogenetics of antidepressant medication intolerance. *Am J Psychiatry* 2003;160:1830-5.
8. Kraemer HC, Frank E, Kupfer DJ. Moderators of treatment outcomes: clinical, research, and policy importance. *JAMA* 2006; 296:1286-9.

