American Journal of Epidemiology
Published by the Johns Hopkins Bloomberg School of Public Health 2009.

Vol. 170, No. 3
DOI: 10.1093/aje/kwp137
Advance Access publication June 4, 2009

Invited Commentary

Invited Commentary: Assessing Treatment Effects by Using Observational Analyses—Opportunities and Limitations

Philip S. Wang and Michael Schoenbaum

Initially submitted March 16, 2009; accepted for publication April 9, 2009.

Making decisions about medical treatments based upon valid evidence is critical to improve health-care quality, outcomes, and value. Although such research commonly connotes the use of randomized controlled trials, experimental methods are not always feasible, and research using observational, quasi-experimental, and other nonexperimental methods may also be important. At the same time, nonexperimental methods are inherently susceptible to various types of bias and thus present special challenges in the search for valid and generalizable evidence. The study by Gardarsdottir et al. (*Am J Epidemiol*. 2009;170(3):280–285), on which this commentary is based, addresses a key potential source of bias—mismeasurement of patients' duration of treatment—in previous research on pharmacotherapy for depression. However, the authors' study is unlikely to address other potential sources of bias, which may make interpretation of their findings more difficult.

bias (epidemiology); depression; observation; research design; treatment outcome

Making decisions on medical treatments based upon valid evidence is critical to improve health-care quality, outcomes, and value. In this context, various legislative bills, most notably the American Recovery and Reinvestment Act of 2009 (H. R. 1) (http://www.rules.house.gov/111/LegText/111_hr1_text.pdf), have been introduced in the US Congress to increase federal support for efficacy and effectiveness research. Although such research commonly connotes the use of randomized controlled trials, many pressing questions are not amenable to study through randomized trials because of cost, logistics, ethics, or other barriers. For this reason, application of observational, quasi-experimental, and other nonexperimental methods may also be important. At the same time, such methods are inherently susceptible to various types of potential bias and thus present special challenges in the search for valid and generalizable evidence.

The study by Gardarsdottir et al. (1) in this issue of the *Journal* provides a useful illustration of both the opportunities and the limitations of assessing treatment effects by using observational methods. The authors examine a question of treatment efficacy, specifically the empirical relation between patients' duration of treatment with antidepressant medication and their risk of subsequent relapse/recurrence. This question had been examined in 3 prior studies cited by the authors (2–4), each of which concluded that patients who discontinued antidepressant drug treatment relatively early had a higher risk of relapse/recurrence than those who followed treatment guidelines. In turn, as the authors report, this conclusion had itself been incorporated into some subsequent depression treatment guidelines. Because suitable randomized controlled trial data were apparently not available, the 3 prior studies, and the present one, used observational rather than experimental methods.

As readers of this *Journal* are fully aware, observational analyses of the effects of medical treatment may be subject to various types of bias. In the present study, Gardarsdottir et al. (1) identify one potentially critical source of bias: mismeasurement of patients' duration of treatment. Specifically, as the authors describe, the 3 prior studies (2–4) defined ''early discontinuers'' as patients who filled 3 or fewer antidepressant prescriptions in the 6 months after filling the initial (index) prescription and/or who had no refills between 75 days and 6 months after the index fill, whereas ''continuing users'' filled 4 or more prescriptions in the 6 months after the index fill. For all patients, relapse/recurrence was defined as the patient reinitiating use of antidepressant medication after a ''clean'' period of at least 6 months without an antidepressant prescription fill. In practice, however, the 3 studies defined the ''clean'' period

Correspondence to Dr. Philip S. Wang, National Institute of Mental Health, 6001 Executive Boulevard, Bethesda, MD 20892-9669 (e-mail: wangphi@mail.nih.gov).

as starting exactly 6 months after the index fill—even for early discontinuers who actually discontinued use before this date and for continuing users whose initial treatment course actually extended past this date.

The present study (1) improves on the prior research by defining the ''clean'' period based on when patients actually ended their initial course of treatment rather than on a fixed interval that is mismeasured in general. Moreover, the authors report that this improved measurement eliminates any significant association between early discontinuation of antidepressants and subsequent relapse/recurrence.

What lessons does this study (1) hold? With regard to its main methodological finding, that results can be sensitive to measurement error, we are persuaded, but hardly surprised (which should not diminish the value of the authors' demonstrating it empirically in this case). With regard to its main clinical claim, that—counter to prior findings—early discontinuation of antidepressant medication is not actually associated with elevated risk of relapse/recurrence, we remain skeptical, for several reasons.

Most obviously, mismeasurement of the initial treatment course is not the only potential source of bias in this analysis (1). For instance, as the authors point out in the Discussion section of their paper, they lacked data on depression severity. More generally, the authors lacked definitive information on why patients received antidepressant medication. The current study does focus on patients whose medical record includes a depression diagnosis, a plausible indicator that patients have depression. On the other hand, in general practice, many antidepressant prescriptions—whether or not actually written for depression—have no associated diagnosis in the medical record, so focusing on only those patients with a recorded depression diagnosis may limit generalizability (5, 6). Other potential sources of bias include loss to follow-up, for example, if some patients switch from the provider or delivery system from which the data come; incomplete capture of prescription drug fills, for example, if claims data miss prescriptions filled in some pharmacies and/or that patients pay for out-of-pocket; and imperfect correspondence between prescription drug fills and patients' actual medication use, for example, if patients hoard pills or take them only sporadically (5).

We recognize that these issues may apply to varying degrees in the current and various prior studies, each of which was conducted in a different setting. We also recognize that the current study (1) is likely to have high internal consistency, since both the mismeasured and ''correct'' analyses were conducted with the same study sample and data; on the other hand, any special aspects of this particular study setting may limit generalizability. (In this context, while we agree with the authors that many randomized controlled trials may have limited external validity, we suggest that it is not inherent to the experimental method. It is certainly possible to conduct randomized trials by using naturalistic patient populations, delivery systems, and even financing mechanisms, as various ''practical'' clinical trials—including some focusing on depression—have demonstrated (7).)

Although information on the effects of various treatments is needed throughout medicine, it may be particularly important for mental health care, where prevailing patterns of care are relatively poor while expenditures are rising rapidly, especially with respect to the prescribing of psychotropics. Observational analyses, conducted—and interpreted—carefully, can and should play a major role in helping to improve treatment, by giving clinicians—and policy makers—additional evidence on which to base their choices.

ACKNOWLEDGMENTS

Author affiliation: Division of Services and Intervention Research, National Institute of Mental Health, Bethesda, Maryland (Philip S. Wang, Michael Schoenbaum).

The views expressed in this commentary do not necessarily represent the views of the National Institute of Mental Health, the National Institutes of Health, the Department of Health and Human Services, or the US government.

Conflict of interest: none declared.

REFERENCES

1. Gardarsdottir H, Egberts TC, Stolker JJ, et al. Duration of antidepressant drug treatment and its influence on risk of relapse/recurrence: immortal and neglected time bias. *Am J Epidemiol*. 2009;170(3):280–285.
2. Melfi CA, Chawla AJ, Croghan TW, et al. The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression. *Arch Gen Psychiatry.* 1998;55(12):1128–1132.
3. Claxton AJ, Li Z, McKendrick J. Selective serotonin reuptake inhibitor treatment in the UK: risk of relapse or recurrence of depression. *Br J Psychiatry.* 2000;177:163–168.
4. Sood N, Treglia M, Obenchain RL, et al. Determinants of antidepressant treatment outcome. *Am J Manag Care*. 2000;6(12):1327–1336.
5. Croghan TW, Schoenbaum M, Sherbourne CD, et al. A framework to improve the quality of treatment for depression in primary care. *Psychiatr Serv.* 2006;57(5):623–630.
6. Pincus HA, Tanielian TL, Marcus SC, et al. Prescribing trends in psychotropic medications: primary care, psychiatry, and other medical specialties. *JAMA*. 1998;279(7):526–531.
7. Wang PC, Ulbricht CM, Schoenbaum M. Improving mental health treatments through comparative effectiveness research. *Health Aff (Millwood).* 2009;28(3):783–791.