Letters to the Editor

The Impact of Underpowered Studies on Clinical Trial Results

TO THE EDITOR: The commentary by Marder et al. in the September 2017 issue of the *Journal*, "Why Are Innovative Drugs Failing in Phase III?" (1), brings up a critical topic: the frequent failure of phase III trials for CNS programs. The authors effectively summarized investigations into this issue. However, we believe that a possible explanation for phase III trial failure was minimized, as noted in the commentary: "that the phase II results were misleading" (1, p. 829).

It is perhaps an underappreciated truism that statistically underpowered studies with small sample sizes (often early-phase studies) are highly likely to both under- and overestimate treatment effect sizes (2, 3). By considering large effects from small studies to be true effects rather than overestimation errors, some researchers have concluded that smaller trials are more advantageous.

Take, for example, the citation of Undurraga and Baldessarini, who suggest limiting phase III antidepressant trials to 30–75 patients (4, p. 860). Testing this suggestion using data from Food and Drug Administration (FDA) antidepressant phase III trials, we found that treatment arms with Ns of ≤75 had both the highest (0.75) and lowest (−0.29) effect sizes of the entire group of 115 arms and that they achieved statistical significance only 50% of the time. In contrast, the largest treatment arms (with Ns of ≥350) had a 100% rate of statistical significance and little variation in effect size (from 0.24 to 0.33).

These smaller trials empirically demonstrate the chance findings of underpowering—they are no more reliable than tossing a coin. These findings are unsurprising, given that a two-arm trial with Ns of ≤75 would be powered only at 50% for an effect size of 0.5 (2). Underreporting negative studies in published research may mask the downside of the underpowering coin, which can be seen in the FDA data; that is, those "unlucky" underpowered trials resulting in failure and underestimations of drug effect.

The increase in size of phase III trials should not be viewed as a design flaw because this rejects the statistical principles of the scientific method. If phase II results are a reflection of the true magnitude of treatment effects, by rule they should be replicated in a larger sample. If they are not, then the phase II results were likely misleading. Seeking to replicate such lucky results from underpowered studies by underpowering future studies only perpetuates scientifically unsound methods and high failure rates.

REFERENCES

1. Marder SR, Laughren T, Romano SJ: Why are innovative drugs failing in phase III? Am J Psychiatry 2017; 174:829–831
2. Kraemer HC, Mintz J, Noda A, et al: Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry 2006; 63:484–489
3. Gibertini M, Nations KR, Whitaker JA: Obtained effect size as a function of sample size in approved antidepressants: a real-world illustration in support of better trial design. Int Clin Psychopharmacol 2012; 27:100–106
4. Undurraga J, Baldessarini RJ: Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. Neuropsychopharmacology 2012; 37:851–864

Arif Khan, M.D.
Kaysee Fahl Mar, M.A.
Walter A. Brown, M.D.

From the Northwest Clinical Research Center, Bellevue, Wash.; the Department of Psychiatry, Duke University School of Medicine, Durham, N.C.; and the Department of Psychiatry and Human Behavior, Brown University, Providence, R.I.

Address correspondence to Dr. Khan (akhan@nwcrc.net).

The authors report no financial relationships with commercial interests.

This letter was accepted for publication in December 2017.

Am J Psychiatry 2018; 175:188; doi: 10.1176/appi.ajp.2017.17091016

The Importance of Adequately Powered Clinical Studies: Response to Khan et al.

TO THE EDITOR: Dr. Khan and coauthors emphasize that statistically underpowering phase III trials can lead to misleading results. We do not disagree with them. Our concern is that the pressures to meet aggressive study timelines may lead sponsors to include lower quality study sites and to use recruitment incentives that lead to the inclusion of marginal subjects.

We agree that phase III trials should be appropriately powered based on the effect size that has been established in earlier proof of concept and other trials. We also emphasize that the powering of phase III trials should take into account the increase in variability from the greater heterogeneity of the population enrolled in larger registration trials and the accompanying expansion of sites and geographies. Conducting smaller, underpowered studies as a way to manage or mitigate the challenges we have highlighted in our commentary would only add to, not subtract from, the probability of failures.

Stephen R. Marder, M.D.
Thomas Laughren, M.D.
Steven J. Romano, M.D.

From the Semel Institute for Neuroscience and Human Behavior, UCLA, Los Angeles; the VA Desert Pacific Mental Illness Research, Education, and Clinical Center, Los Angeles; Laughren Psychopharmacology Consulting, Rockville, Md.; and Mallinckrodt Pharmaceuticals, St. Louis.

Address correspondence to Dr. Marder (marder@ucla.edu).

The authors' disclosures accompany the original article.

This reply was accepted for publication in December 2017.

Am J Psychiatry 2018; 175:188–189; doi: 10.1176/appi.ajp.2017.17091016r

Pharmacogenetic Tests in Psychiatry

TO THE EDITOR: We feel it necessary to highlight several overlooked and important considerations related to the Clinical Case Conference by Rahman et al., "Misleading Guidance from Pharmacogenomic Testing," published in the October 2017 issue of the *Journal* (1).

Pharmacogenetic testing is unstandardized. More than 20 companies in the United States offer testing with highly variable gene content, results reporting, and evidence base (2). Thus, the selection of a test is not trivial, and no two tests should be assumed to be equivalent or interchangeable. The test ordered for Mr. A included the *DRD2* gene. However, none of the major pharmacogenetic resource hubs (the Pharmacogenetics Knowledgebase, the Clinical Pharmacogenomics Implementation Consortium, and the Food and Drug Administration) indicate an actionable interaction between *DRD2* and clozapine. In fact, less than 20% of pharmacogenetic tests relevant to psychiatry include *DRD2* on their testing panels (2), suggesting that the case of Mr. A is an outlier that should not be extrapolated to all pharmacogenetic testing. Unfortunately, the authors do not report which company performed the testing. Such transparency is required to discourage irresponsible testing of unvalidated genetic markers that can cause a setback in the clinician-patient relationship.

There are major differences in the subfields of psychotropic pharmacogenetics. Antipsychotic pharmacogenetics, though promising, has not yet been evaluated in a randomized controlled trial. In contrast, antidepressant pharmacogenetic testing has been supported by multiple randomized controlled trials (3–5), and at least six more are under way.

Providers who order pharmacogenetic testing have a responsibility to educate themselves and their patients about such testing and its limitations. Given the complex biopsychosocial context in which drug response occurs, it would be naive to believe pharmacogenetic testing will ever provide definitive prescribing advice for psychiatric drugs. Pharmacogenetic testing provides an additional tool for clinicians to assist in thoughtful implementation of evidence-based treatment methodologies based on the unique characteristics of the individual, enhancing, rather than replacing, treatment guidelines.

REFERENCES

1. Rahman T, Ash DM, Lauriello J, et al: Misleading guidance from pharmacogenomic testing. Am J Psychiatry 2017; 174:922–924
2. Bousman CA, Hopwood M: Commercial pharmacogenetic-based decision-support tools in psychiatry. Lancet Psychiatry 2016; 3: 585–590
3. Singh AB: Improved antidepressant remission in major depression via a pharmacokinetic pathway polygene pharmacogenetic report. Clin Psychopharmacol Neurosci 2015; 13:150–156
4. Pérez V, Salavert A, Espadaler J, et al: Efficacy of prospective pharmacogenetic testing in the treatment of major depressive disorder: results of a randomized, double-blind clinical trial. BMC Psychiatry 2017; 17:250
5. Bradley P, Shiekh M, Mehra V, et al: Improved efficacy with targeted pharmacogenetic-guided treatment of patients with depression and anxiety: a randomized clinical trial demonstrating clinical utility. J Psychiatr Res 2017; 96:100–107

Chad Bousman, M.P.H., Ph.D.
Josiah Allen, B.A.
Harris A. Eyre, M.D., Ph.D.

From the Departments of Medical Genetics, Psychiatry, and Physiology and Pharmacology, University of Calgary, Calgary, Alta., Canada; Medigenics Consulting, Rochester, Minn.; the Innovation Institute, Texas Medical Center, Houston; the Center for Innovation in Mental and Physical Health and Clinical Treatment, School of Medicine, Deakin University, Geelong, Victoria, Australia; the Department of Psychiatry, University of Melbourne, Melbourne; and the Discipline of Psychiatry, University of Adelaide, Adelaide, South Australia, Australia.

Address correspondence to Dr. Bousman (chad.bousman@ucalgary.ca).

Dr. Eyre is an employee of and holds equity in CNSDose. Mr. Allen holds equity in Assurex Health, a wholly owned subsidiary of Myriad Genetics. Dr. Bousman reports no financial relationships with commercial interests.

This letter was accepted for publication in November 2017.

Am J Psychiatry 2018; 175:189; doi: 10.1176/appi.ajp.2017.17101086