Published in final edited form as:

Am J Psychiatry. 2011 March ; 168(3): 234–236. doi:10.1176/appi.ajp.2010.10121738.

Betting on Biomarkers

ROY H. PERLIS, M.D., M.Sc.

In this issue, Johnson and colleagues (1) report on a study of ondansetron, a serotonin receptor antagonist, for alcohol-dependent patients. Had they conducted a standard randomized placebo-controlled trial, they would have found no statistically significant difference between treatment groups on the primary outcome measure and stopped there. That is, on average, their study participants were not significantly less likely to drink if they received ondansetron than if they received placebo.

Instead, however, the authors made an important and risky bet, choosing to stratify randomization according to a common genetic variation and then to analyze their results within genetically defined subgroups. With the use of a biological marker, an otherwise murky picture of treatment response became clearer, pointing to a potential novel intervention for drinking—but only in one patient subgroup.

Biomarkers may be employed in randomized trials in a variety of ways, and study designs doing so fall into several broad categories. The first and most traditional is the post hoc analysis (sometimes referred to as a retrospective-prospective analysis [2]), in which the biomarker is simply analyzed as a covariate or moderator. The risk of false positive findings in post hoc analyses is high, and it increases with the number of potential moderators examined, but this approach requires the fewest assumptions a priori—only the foresight to collect the marker to be studied. The potential power of this approach was recently demonstrated in a phase 2 study of bapineuzumab in mild to moderate Alzheimer's disease (3). The study as a whole failed to separate drug from placebo on its primary endpoint, and in the past this result might have led the compound to be shelved. However, post hoc analysis identified larger effects in the subgroup of APOE4 epsilon-4 noncarriers, leading to next-step studies focusing on this patient subgroup.

A second category is the biomarker-enriched design, in which the researchers make the “strong” assumption of larger effect in a particular subgroup and elect to enroll subjects only from that subgroup. While the use of a biomarker for this purpose is relatively novel, the concept of an enriched design is not; indeed, a generation's worth of antidepressant trials have explored severity thresholds, with mixed results. Enriched designs should be more efficient, allowing smaller sample sizes to demonstrate a given effect size. On the other hand, such enriched designs may not be feasible when the group of interest is less common or when identification of the biomarkers is more labor intensive. Also, from a regulatory as well as a scientific perspective, they will almost always entail a follow-up study to establish specificity of effect. That is, if the treatment works in the marker-positive group, does it work in the marker-negative group?

A third category, which addresses these limitations, is the biomarker-stratified design, which is the one used by Johnson and colleagues here. In this approach, the investigators again assume that a marker will be associated with a differential response. However, rather than

excluding marker-negative subjects, the researchers simply stratify randomization to ensure a balanced distribution of the marker in question. Analysis can then proceed sequentially (i.e., first examining one group, then the other), keeping in mind the need to adjust the significance threshold for the number of tests conducted. The strength of this approach is that it allows investigation of gene-by-treatment effects directly—that is, one can examine whether associations are truly marker specific.

In the case of ondansetron in the treatment of alcohol dependence, the presence of a marker-negative group greatly facilitated the interpretation of the results. In the latter group, no drug-placebo separation was observed, supporting the notion that ondansetron truly exhibits genotype-specific effects. Had Johnson et al. conducted a simple biomarker-enriched study, the same effect would have been observed in the patients who were entered, but we would have no notion of whether it was relevant to the excluded group.

How will these kinds of findings translate to clinical practice? After some resistance to the notion of reducing their potential market share by looking for more responsive subgroups, drug developers have come around to the notion that a smaller piece of a big pie is better than no pie at all. The notion of drug-diagnostic codevelopment—the simultaneous validation of a new treatment and a new test to guide its use—has become an area of intense interest within the pharmaceutical world. With the forthcoming publication of new guidelines for drug-diagnostic codevelopment from the U.S. Food and Drug Administration, this process should only accelerate.

Still, some caution is also warranted. It bears noting that variation in the much-maligned serotonin transporter has been linked with an astonishingly broad array of phenotypes—from anxiety to antidepressant responsiveness to creative dance (4). One might therefore fall back on the easy answer that many of these findings, like much of the published medical literature, simply represent type I error (5). Alternatively, many of them may be correct but represent an example of pleiotropy, which is perhaps not surprising for such a fundamental brain process as serotonergic neurotransmission. That is, there is no reason to think that a variant might not be manifested in numerous ways—some adverse, some beneficial—which could include responsiveness to ondansetron, either directly or indirectly.

In a similar vein, recent investigations point to genetic complexity in psychiatric disorders, with many common or rare variations of modest effect increasing liability (6). The limited pharmacogenomic literature in psychiatry supports this interpretation. Rather than presuming that all candidate gene associations are wrong, one might instead conclude that many of them are “right” but perhaps too small to be particularly useful as biomarkers. The question of whether such associations are of large enough effect to merit inclusion in clinical trials, as was done here, is a difficult one to answer other than empirically. For example, an investigation of mice carrying only a single copy of the calcium channel subunit coded by the *CACNA1C* gene, premised on a small but genome-wide significant effect on bipolar liability, successfully identified behavioral differences from wild-type mice, suggesting that small genetic effects may still be put to good use at the bench (7). The Johnson et al. study provides some validation for the corresponding approach in clinical investigation—that is, even genetic variations of relatively modest effect may still be put to good use in clinical/translational investigation.

More broadly, this study by Johnson and colleagues illustrates the thoughtful use of biological markers to tilt the odds in favor of finding novel therapies for psychiatric disorders. While the burgeoning number of markers to choose from may seem challenging to clinical investigators, conducting good clinical trials is itself no less challenging—so anything that improves the odds of success should be most welcome.

Acknowledgments

Dr. Perlis is supported by NIMH R01 MH086026 and the Stanley Center for Psychiatric Research; he has received royalties, consulting fees, or grant support from Concordant Rater Systems, Proteus Biomedical, and RIDVentures. Dr. Freedman has reviewed this editorial and found no evidence of influence from these relationships.

References

1. Johnson BA, Ait-Daoud N, Seneviratne C, Roache JD, Javors MA, Wang X-Q, Liu L, Penberthy JK, DiClemente CC, Li MD. Pharmacogenetic approach at the serotonin transporter gene as a method of reducing the severity of alcohol drinking. *Am J Psychiatry*. 2011; 168:265–275. [PubMed: 21247998]
2. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst*. 2009; 101:1446–1452. [PubMed: 19815849]
3. Salloway S, Sperling R, Gilman S, Fox NC, Blennow K, Raskind M, Sabbagh M, Honig LS, Doody R, van Dyck CH, Mulnard R, Barakos J, Gregg KM, Liu E, Lieberburg I, Schenk D, Black R, Grundman M. Bapineuzumab 201 Clinical Trial Investigators. A phase 2 multiple ascending dose trial of bapineuzumab in mild to moderate Alzheimer disease. *Neurology*. 2009; 73:2061–2070. [PubMed: 19923550]
4. Bachner-Melman R, Dina C, Zohar AH, Constantini N, Lerer E, Hoch S, Sella S, Nemanov L, Gritsenko I, Lichtenberg P, Granot R, Ebstein RP. AVPR1a and SLC6A4 gene polymorphisms are associated with creative dance performance. *PLoS Genet*. 2005; 1(3):e42. [PubMed: 16205790]
5. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005; 2(8):e124. [PubMed: 16060722]
6. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P. International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009; 460:748–752. [PubMed: 19571811]
7. Dao DT, Mahon PB, Cai X, Kovacsics CE, Blackwell RA, Arad M, Shi J, Zandi PP, O'Donnell P, Knowles JA, Weissman MM, Coryell W, Scheftner WA, Lawson WB, Levinson DF, Thompson SM, Potash JB, Gould TD. Bipolar Genome Study (BiGS) Consortium. Mood disorder susceptibility gene CACNA1C modifies mood-related behaviors in mice and interacts with sex to influence behavior in mice and diagnosis in humans. *Biol Psychiatry*. 2010; 68:801–810. [PubMed: 20723887]