

To Replicate or Not to Replicate: The Case of Pharmacogenetic Studies

Establishing Validity of Pharmacogenomic Findings: From Replication to Triangulation

Stella Aslibekyan, PhD; Steven A. Claas, MS; Donna K. Arnett, MSPH, PhD

Replication of positive findings from genetic association studies has long served as the gold standard for establishing validity of observed genotype–phenotype associations.^{1–3} In addition to adequate statistical power and sound epidemiological design aimed at minimizing biases, suggested criteria for successful replication studies include: (1) same genetic variant, (2) same direction of association, (3) same definition of the phenotype, and (4) same ethnic group as reported in the discovery study.¹ However, in the context of pharmacogenetics, the phenotype requirement is further constrained by the drug and intervention strategy of the initial investigation. As a result, and because many pharmacogenetic and pharmacogenomic studies arise from randomized clinical trials, replication of gene–drug response associations has proven to be a costly and logistically difficult endeavor.^{4,5} In light of these constraints, this review presents the case for rethinking the replication requirement and proposes several alternative approaches to validating pharmacogenomic findings.

Response by Ioannidis on p 412

The prevailing view is that the requirement of replication, currently imposed or strongly suggested by most peer-reviewed journals, has improved the quality of published studies by reducing false-positive findings and overcoming the winner's curse.⁶ At first glance, replication rates for findings from both candidate gene studies and agnostic approaches, such as genome-wide association studies (GWAS) and more commonly whole exome or whole genome scans, appear low. For example, of 70 previously reported candidate loci for common diseases,

only 22 had P values $<10^{-7}$ across 100 subsequently conducted GWAS.⁷ Similarly, a recent analysis of large-scale genetic studies including at least 50 variants and published between 2007 and 2010 suggests that only 1.2% of initial findings are replicated.⁸ In pharmacogenetic studies of cardiovascular outcomes, where the phenotype is defined as response to drug treatment, the replication rate (3.4%) is comparable to that of genetic association studies in general (Table 1).^{9–13} However, such crude replication rates may be misleading and unflattering indicators of the validity of most findings in the field, as false-positives represent only 1 of many reasons for nonreplication. Below we summarize the factors that play a role in reproducibility of pharmacogenomics findings.

One factor contributing to single-digit replication rates is the low ratio of false-positive to false-negative findings inherent to genome-wide and other large-scale genetic studies.⁸ In pharmacogenetics, other factors include the low minor allele frequency of functional variants, small effect sizes, limited sample size, and often incomparable treatment regimens or phenotype definitions.¹⁴ However, a more serious challenge may lie in our definition of successful replication, which may be unnecessarily limiting, reflecting a crude grasp of the underlying biological mechanisms rather than failures of study design. For example, in some cases replication failure may actually be due to the underlying genetic architecture, ie, varying allele frequencies, locus heterogeneity or epistatic interactions between functional polymorphisms.¹⁵ Such differences are likely to become more prominent as the field shifts away from the common disease, common variant to place more emphasis on the contributions

From the Department of Epidemiology, School of Public Health, University of Alabama at Birmingham.

Correspondence to Donna K. Arnett, MSPH, PhD, Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, RPHB 220E, 1530 3rd Ave S, Birmingham, AL 35294-0022. E-mail arnett@uab.edu
(*Circ Cardiovasc Genet*. 2013;6:409-412.)

© 2013 American Heart Association, Inc.

Circ Cardiovasc Genet is available at <http://circgenetics.ahajournals.org>

DOI: 10.1161/CIRCGENETICS.112.000010

**Table 1. Published Replication Studies of Pharmacogenetic Findings in Cardiovascular Disease***

First Author, y	Drug	Cardiovascular Phenotype	N Discovery/N Replication	N Variants: Replication Attempted	N Variants: Replicated
Peters (2011) ⁹	Statins	Myocardial infarction	1885/2446	5	0
Nordestgaard (2010) ¹⁰	Losartan, Atenolol	Systolic blood pressure, diastolic blood pressure, pulse pressure, mean arterial pressure, heart rate	Varied†/3503	13	0
Thompson (2009) ¹¹	Statins	LDL cholesterol, HDL cholesterol, triglycerides	1984/3761	50	0
Link (2008) ¹²	Statins	Statin-induced myopathy	175/16664	2	1
Thompson (2005) ¹³	Statins	LDL cholesterol, HDL cholesterol, triglycerides	Varied†/2735	18	2 (LDL-C)

HDL indicates high-density lipoprotein; and LDL-C, low-density lipoprotein cholesterol.

The PubMed search strategy, adapted from Ioannidis et al,⁹ included the following terms: (replication OR collaborative OR genome-wide) AND (genetic) AND (candidate [All Fields] OR previously reported [All Fields] OR previously proposed [All Fields]) AND (cardiovascular system [MeSH Terms] OR (cardiovascular [All Fields] AND system [All Fields]) OR cardiovascular system [All Fields] OR cardiovascular [All Fields]) AND (pharmacogen OR drug OR treatment). Additional studies were identified using reference lists of extracted articles.

†Initial findings were taken from previously published studies.

of many rare, and thus presumably more likely to be causal, genetic variants. Furthermore, the recently published results of the Encyclopedia Of DNA Elements (ENCODE) projects stress the significance of the still poorly understood role of intergenic variants, previously considered junk DNA.¹⁶ As high-resolution genotyping becomes more accessible, the number of potentially relevant genetic markers will increase exponentially, making the same-genetic-variant component of the replication requirement not only unfeasible but also meaningless, as the functional role of most newly identified variants is yet unknown.

In this review, we discuss alternative approaches that integrate different methods of testing kindred hypotheses, such as functional validation, combined analysis of several populations, and simulations. These could jointly provide a more comprehensive and rigorous avenue for validating genetic association findings (Table 2).

Functional Validation

Functional validation of genetic variants showing significant associations with drug response in the initial screen is the most widely accepted substitute for replication.⁴ If the mechanisms of drug action are well known, it is often possible

to design cell line or animal model experiments that could aid in establishing causality of the observed associations.^{5,8} One such study used a zebrafish model to screen for genes associated with response to QT-prolonging agents and successfully confirmed a hit from 2 human GWAS, located in the zebrafish ortholog of the *GINS3* gene.¹⁷ In another study, lymphoblastoid cell line models were used to generate evidence supporting the causal association between a haplotype within *HMCGR* and decreased response to statin therapy.¹⁸ Despite the compelling nature of the results that stem from well-controlled experimental design, the effectiveness of the functional validation approach may be limited by imperfect translation between animal and human organisms, lack of appropriate assays, and ethical considerations.⁵

Joint Analysis of Several Populations

Combining the discovery and validation studies into a meta-population has distinct advantages over the classic replication paradigm. Specifically, the 2-stage GWAS design, where the second stage is used as a replication study, has been shown to have consistently lower statistical power than joint analysis, even when more stringent significance levels are applied.¹⁹

Table 2. Alternative Approaches to Validation of Genetic Association Findings

Method	Strengths	Weaknesses
Functional validation	Biologically informed Experimental design	Does not always translate to human populations Limited by crude understanding of underlying biology and availability of appropriate assays Ethical issues
Combined analysis	Improved statistical power Possible with existing data sources	Likely to further decrease false-positive:false-negative ratio Threats to validity include heterogeneity between populations and publication bias
Simulation	Can incorporate real data features such as interactions and/or population substructure Does not require new data collection	Strain on computational resources Methods are still under development

Nevertheless, the increases in efficiency due to joint analysis do not obviate the concerns surrounding validity, as using lower significance levels is only likely to decrease the already low false-positive/false-negative ratio in large-scale studies.⁸ Additionally, heterogeneity in allele frequencies between populations may result in biased summary estimates of association, thus sometimes threatening rather than confirming the validity of the findings.²⁰

Simulation-Based Methods

The most novel group of methods that could constitute viable alternatives to replication of genetic studies comprises computational approaches based on simulation. Briefly, simulation-based methods consist of identifying candidate single-nucleotide polymorphisms using an agnostic, usually genome-wide approach and subsequently comparing the predictive ability of models including the preselected single-nucleotide polymorphisms with those constructed using a large number of random selections of genomic single-nucleotide polymorphisms.²¹ Although the simulation strategy has not been widely applied to studies of genetic predictors of drug response, it has been used in etiological studies, eg, to evaluate susceptibility genes in sporadic Parkinson disease.^{21,22} Although the use of this approach may still be limited by the availability of computational resources, it remains a promising avenue of investigation and could present a feasible alternative to replication in pharmacogenetics.

Replication Versus Triangulation

The idea of demanding a simply articulated gold standard for replication is tempting—it elegantly codifies a discipline-wide, acceptable level of skepticism. It also offers journal editors and reviewers a clear and convenient metric of study quality. However, until the definition of successful replication is sufficiently informed by advances in basic and clinical science, imposing the stringent requirement of reproducing associations at the variant level is as likely to obfuscate as clarify gene–phenotype associations. Already others have argued that relaxing the replication requirement from single variant to the gene or pathway level may bring the concept closer to etiologic relevance.²³ We have suggested here that using different methods of testing the same pharmacogenetic hypothesis, such as those discussed above, as well as those yet to be developed, be considered in lieu of strict variant-centric replication. This approach is not without precedent: the idea of methodological triangulation is well established in the social sciences, another domain characterized by complex systems, nascent knowledge, and severe logistical and ethical constraints on experimental design. For many in the genomic sciences, embracing triangulation might be an uncomfortable departure from the epistemological binary afforded by a gold standard criterion; it would require reviewers to be intimately familiar with recent research in the area and assess the merits of each combination of studies individually.

Finally, it is important to remember that even when etiologic validity has been satisfactorily established, successful clinical applications of genetic association findings are far from guaranteed. A study of several chronic disease outcomes demonstrated that even replicated SNPs with highly significant measures of genotype–phenotype associations often have poor discrimination and reclassification ability, highlighting the need for better methods of identifying candidate variants.²⁴ Making the necessary albeit inconvenient step from pure replication to multipronged evaluation of evidence would help build a more capable pipeline from genomic discovery to clinical application, charting the course toward the ultimate goal of personalized therapeutic approaches.

Disclosures

None.

References

- Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, et al. Replicating genotype-phenotype associations. *Nature*. 2007;447:655–660.
- Sillanpää MJ, Auranen K. Replication in genetic studies of complex traits. *Ann Hum Genet*. 2004;68(Pt 6):646–657.
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet*. 2001;29:306–309.
- Daly AK. Genome-wide association studies in pharmacogenomics. *Nat Rev Genet*. 2010;11:241–246.
- Roden DM, Wilke RA, Kroemer HK, Stein CM. Pharmacogenomics: the genetics of variable drug responses. *Circulation*. 2011;123:1661–1670.
- Kraft P. Curses—winner's and otherwise—in genetic epidemiology. *Epidemiology*. 2008;19:649–651; discussion 657.
- Siontis KC, Patsopoulos NA, Ioannidis JP. Replication of past candidate loci for common diseases and phenotypes in 100 genome-wide association studies. *Eur J Hum Genet*. 2010;18:832–837.
- Ioannidis JP, Tarone R, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology*. 2011;22:450–456.
- Peters BJ, Pett H, Klungel OH, Stricker BH, Psaty BM, Glazer NL, et al. Genetic variability within the cholesterol lowering pathway and the effectiveness of statins in reducing the risk of MI. *Atherosclerosis*. 2011;217:458–464.
- Nordstgaard BG, Kontula K, Benn M, Dahlöf B, de Faire U, Edelman JM, et al. Effect of ACE insertion/deletion and 12 other polymorphisms on clinical outcomes and response to treatment in the LIFE study. *Pharmacogenomics*. 2010;20:77–85.
- Thompson JF, Hyde CL, Wood LS, Paciga SA, Hinds DA, Cox DR, et al. Comprehensive whole-genome and candidate gene analysis for response to statin therapy in the Treating to New Targets (TNT) cohort. *Circ Cardiovasc Genet*. 2009;2:173–181.
- Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy—a genomewide study. *N Engl J Med*. 2008;359:789–799.
- Thompson JF, Man M, Johnson KJ, Wood LS, Lira ME, Lloyd DB, et al. An association study of 43 SNPs in 16 candidate genes with atorvastatin response. *Pharmacogenomics J*. 2005;5:352–358.
- Voora D, Shah SH, Reed CR, Zhai J, Crosslin DR, Messer C, et al. Pharmacogenetic predictors of statin-mediated low-density lipoprotein cholesterol reduction and dose response. *Circ Cardiovasc Genet*. 2008;1:100–106.
- Greene CS, Penrod NM, Williams SM, Moore JH. Failure to replicate a genetic association may provide important clues about genetic architecture. *PLoS One*. 2009;4:e5639.
- Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74.

- 
17. Milan DJ, Kim AM, Winterfield JR, Jones IL, Pfeufer A, Sanna S, et al. Drug-sensitized zebrafish screen identifies multiple genes, including GINS3, as regulators of myocardial repolarization. *Circulation*. 2009;120:553–559.
 18. Medina MW, Gao F, Ruan W, Rotter JJ, Krauss RM. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. *Circulation*. 2008;118:355–362.
 19. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet*. 2006;38:209–213.
 20. Kavvoura FK, Ioannidis JP. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. *Hum Genet*. 2008;123:1–14.
 21. Lesnick TG, Sorenson EJ, Ahlskog JE, Henley JR, Shehadeh L, Papapetropoulos S, et al. Beyond Parkinson disease: amyotrophic lateral sclerosis and the axon guidance pathway. *PLoS One*. 2008;3:e1449.
 22. Li Y, Rowland C, Xiromerisiou G, Lagier RJ, Schrodi SJ, Dradiotis E, et al. Neither replication nor simulation supports a role for the axon guidance pathway in the genetics of Parkinson's disease. *PLoS One*. 2008;3:e2707.
 23. Peng G, Luo L, Siu H, Zhu Y, Hu P, Hong S, et al. Gene and pathway-based second-wave analysis of genome-wide association studies. *Eur J Hum Genet*. 2010;18:111–117.
 24. Jakobsdottir J, Gorin MB, Conley YP, Ferrell RE, Weeks DE. Interpretation of genetic association studies: markers with replicated highly significant odds ratios may be poor classifiers. *PLoS Genet*. 2009;5:e1000337.

Response to Stella Aslibekyan, PhD, Steven A. Claas, MS, and Donna K. Arnett, MSPH, PhD

John P.A. Ioannidis, MD, DSc

Aslibekyan et al discuss some interesting alternatives to current replication standards in pharmacogenomics. I sympathize with several of their ideas, which overlap with what I proposed, but disagree with some others.

Relaxing the standard replication criteria (same genetic variant, same direction of association, same phenotype) is problematic. Of course, genetic variants that are in perfect linkage disequilibrium are exchangeable. Moreover, often genetic variants in the same locus may have independent effects, and it is important to develop efficient methods to capture these incremental pieces of information residing in the same locus. However, unlinked or poorly linked variants in the same gene cannot be used for replication of associations. Associations in different direction in different populations are possible in theory but hardly ever documented; their documentation would require separate replication of the different directions in the varied postulated settings. Phenotype definitions may also vary across studies, but some core components should be shared. Otherwise, we do not really know what phenotype we are talking about. The other criteria listed by Aslibekyan et al as standards are more flexible. For example, replication does not have to entail the exact same ethnic group; in fact, replication across different ancestry groups would make an association more broadly applicable; statistical efficiency may also improve when different ancestry populations are combined after appropriate stratification. Moreover, we have limited evidence whether pharmacogenetic associations pertain to wide drug class effects and diverse therapeutic strategies thereof, or are more restricted to single drugs and strategies. Casting a wider net may allow documenting associations that have broader relevance.

I am worried about relaxing the statistical criteria of replication. We have a miserable history of false-positives in the field. Joint analysis (meta-analysis) of multiple datasets in large consortia is the rule currently, and it should help discover more associations without diluting the statistical thresholds required. Functional validation is likely to offer mostly adjunctive information that cannot replace sound statistical support. There are so many different functional models and experiments that can be conceived, and their results are often incongruent. Moreover, many functional models may simply be insufficiently sensitive to detect what are modest genetic effects. Finally, simulation approaches using pathways or other complex computational means are interesting but have been largely unsuccessful to date. In fact, the examples cited by Aslibekyan et al on Parkinson disease are statistically flawed, and the discovered pathways are simply statistical artifacts: their selection process did not properly adjust for the extent of multiplicity of analyses; not surprisingly, these complex pathway signatures have not moved forward in the translational space.

Using other lines of evidence is reasonable, but no triangulation can overcome the need for rigorous statistical documentation and replication. The notion that somehow we can make major progress by using smart tools in suboptimal studies with small sample size is equivalent to asking for a free lunch. Real progress will require large studies and stringent statistics, not more of the creative data dredging that has already led this field to uncreative stagnation.

To Replicate or Not to Replicate: The Case of Pharmacogenetic Studies: Establishing Validity of Pharmacogenomic Findings: From Replication to Triangulation

Stella Aslibekyan, Steven A. Claas and Donna K. Arnett

Circ Cardiovasc Genet. 2013;6:409-412

doi: 10.1161/CIRCGENETICS.112.000010

Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2013 American Heart Association, Inc. All rights reserved.

Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://circgenetics.ahajournals.org/content/6/4/409>

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation: Cardiovascular Genetics* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

Reprints: Information about reprints can be found online at:
<http://www.lww.com/reprints>

Subscriptions: Information about subscribing to *Circulation: Cardiovascular Genetics* is online at:
<http://circgenetics.ahajournals.org/subscriptions/>