Chasing—and Catching—the Wild Goose
Hypothesis-Free Post-Hoc Stratification Studies as a New Paradigm for Drug Development

Klaus Lindpaintner, MD, MPH

In this issue, Tardif et al. report on the results of a retrospective analysis of 2 phase III trials of the cholesteryl ester transfer protein inhibitor, dalcetrabip, dal-OUTCOMES, and dal-PLAQUE-2. Using a hypothesis-free whole genome screening process that could be likened to a wild-goose-chase, they identified genetically distinct subpopulations among the probands, which demonstrated markedly different clinical outcomes when treated with dalcetrabip as a secondary prevention strategy. If substantiated prospectively, the effect of this finding will be paradigm changing.

Article see p 372

Synopsis
The pivotal clinical trial, the dal-OUTCOMES study, randomized 15,871 subjects and found no reduction in cardiovascular events with an average of 31 months of follow-up on dalcetrabip compared with placebo.2 The second trial, dal-PLAQUE-2, examining the effects of dalcetrabip on intima-media thickness,3 was prematurely terminated because of media thickness, was prematurely terminated because of negative results. Changes in high-density lipoprotein (HDL) levels did not correlate with clinical outcomes. That might have been the end of the story, but instead a new question was posed: are there any subgroups that might have actually accrued benefit from the new medicine, despite the failure to find statistical significance in the large group? Tardif et al. conducted a genome-wide association study on a representative subset of the dal-OUTCOMES cohort for which the requisite blood samples and informed consent were available. This revealed a striking dichotomy in clinical outcomes according to genetic variants of the ADCY9 gene: individuals receiving dalcetrabip who were homozygous for the minor allele (the less common variant) of several single-nucleotide polymorphisms (SNPs) associated with this gene experienced ≥40% reduction in cardiovascular end points compared with placebo-treated controls. This fraction represented ≈20% of the cohort. In contrast, the subset of patients receiving dalcetrabip who were homozygous for the major allele (ie, the alternate form of the SNPs) was found to have a 30% increase in cardiovascular events compared with controls. The group of patients who were heterozygous for these SNPs showed no effect. Genotype status did not correlate with the outcomes among the placebo-treated controls. Subsequently, hypothesis-driven testing for SNPs of the ADCY9 gene in patients from the dal-PLAQUE-2 study revealed that that minor allele homozygosity was associated with reduction, and major allele homozygosity with an increase in common carotid artery intima media thickness. Although different SNPs of the ADCY9 gene yielded the statistically most significant results in the 2 studies, the authors plausibly interpret this finding as supportive evidence for the results of the dal-OUTCOMES genotyping study.

Importantly, the difference in outcomes of these subgroups was not previously hypothesized on the basis of a putative mechanistic explanation. It was discovered purely by means of the unbiased screen, followed by appropriate biostatistical analysis only recently made possible by advances in molecular biology and technology.

The authors conclude that an additional prospective trial, enrolling only patients carrying the genotype associated with favorable outcomes in this study, should be conducted to validate these retrospective results. Indeed it should!

Impact
Tardif et al. have produced what may 1 day be regarded as a landmark study. It has the potential to herald an entirely new paradigm for future drug development. Their work suggests compellingly that a heretofore hypothetical concept, that is, unbiased, hypothesis-free whole genome analysis done in pursuit of possible stratification parameters, can indeed result in the discovery of subpopulations that may derive particular benefit from a medicine that is not effective for the general patient population. This knowledge allows more targeted and, therefore safer, more efficient, and more cost-effective use of prescription medicines. Few would disagree that improved targeting of medical interventions to recipients who have a significantly enhanced likelihood of deriving benefit would be medically and ethically desirable. Indeed President Obama, in his State of the Union address this year, raised this very issue to level of public discourse by proposing dedicated funding of $215 million for a Precision Medicine Initiative, which might be expected to support exactly this kind of pioneering research.

This work raises several issues worthy of comment, both on a scientific and public health policy level.
Scientific Considerations

At the level of cardiovascular science, the observation that HDL-increases per se did not track with the presence of a beneficial or detrimental outcome in the respective genotype-defined subgroups raises intriguing questions: should HDL continue to be regarded as a meaningful surrogate marker of cardiovascular risk, or is it a poor first-approximation proxy of an underlying biology which, if better understood, could be far more informative? Is the effect of cholesteryl ester transfer protein inhibition, while superficially reflected by HDL levels, conveyed via a different mechanism, one that affects cardiovascular outcomes regardless of the effect on HDL? Is the effect of dalcetrabip on cardiovascular outcomes perhaps an off-target one, that is, not directly related to cholesteryl ester transfer protein inhibition? Both of the latter questions are raised by the present observation that genetic variants in a gene that seems unrelated to cholesteryl ester transfer protein are associated with dramatically differential response to the drug. If a prospective validation study establishes a causal relationship between biomarker and clinical phenotype, how important is a mechanistic explanation for association-based data? A biologically relevant effect must ultimately have a causative explanation and basis, and it will always be preferential to identify a mechanistic basis before acting on it. In this case, the findings on clinical end points are strengthened by the parallel observation that the phenotype of intima-media thickness seems to plausibly corroborate the former. Given the nature of the associated gene, the authors speculate that the underlying molecular mechanism may be related to β-adrenergic signaling, and additional work to further examine this should certainly be encouraged. However, given the complexity of how clinical outcomes relate to underlying biology, it may not be prudent to insist on uncovering mechanisms before embracing an epidemiologically or clinically reliable marker to guide the informed management of patients.

This study relies exclusively on genetic markers. However, if one embraces the concept of post hoc hypothesis-free screening to identify phenotypically distinct subgroups of patients, then other classes of biomarkers ought to be considered as well, as science and technology progresses and the feasibility of such screens increases. Arguably, the vastly richer information content of other classes of markers with greater complexity and dynamic range than genomics, such as transcriptomics, epigenetics, proteomics, metabolomics, etc., makes them prime candidates to be studied as well.

Public Health Policy Considerations

At the level of public policy, the effect of the study under discussion, conceptually valid today, will become compelling if the methodology used and results obtained by Tardif et al. are validated in a subsequent prospective trial. It is not too soon to begin preparing for that day. In this spirit I offer the following considerations and thought experiments, making the assumption that prospective validation has been achieved:

The unanticipated uncovering of patient subgroups with diametrically opposite outcomes hidden underneath an overall null effect of dalcetrabip provides a powerful rationale for carrying out exploratory, hypothesis-free, retrospective screens in all clinical trials of new medicines. Importantly, this should be advocated not only for trials with negative results (ie, those that show no benefit or a detrimental outcome associated with the new drug) but also for those with positive outcomes, where overall benefits are shown. In all the cases, the results of an undifferentiated aggregate analysis may cover up the existence of subgroups defined by biomarkers, which are associated with exceptionally beneficial or adverse responses, or no effect. Carrying out such an investigation as a potential rescue effort for a drug that has failed in clinical trials may be rather uncontroversial. However, persuading companies to do the same for a drug that has already shown overall efficacy in the aggregate will be much more difficult from a business standpoint, yet precisely the same public health and ethical imperatives apply. If such an approach identifies a subsegment of patients who should not receive the drug because of harm or lack of efficacy, the sponsor faces a potential diminution of what would under conventional paradigms have been the anticipated market for the product.

For example, if we hypothesize for the moment that the responder—genotype frequency in the dal-OUTCOMES trial had been somewhat larger, say 30%, the overall trial may have shown a modest, but statistically significant aggregate beneficial effect, sufficient for the regulatory authorities to approve the drug for all patients. Under these circumstances, there would be a strong disincentive for the sponsor to carry out any additional studies, for example, using the kind of genome screen applied by Tardif et al. In the first place, there currently exists a perhaps understandable reluctance in clinical drug development to meddle with positive trial data, as additional insights could come to light which could lead the regulatory authorities to reconsider their decision to approve. Simply stated, companies do not want to generate information that may complicate an already cumbersome regulatory approval process. Furthermore, stratification itself engenders costs. The prospective validation of the results of stratification engenders more and much higher costs. And with the added knowledge provided by the stratification, the company would be faced with a reduction (by 70% in our thought experiment) in eligible patients, as well as the added mandate to carry out genetic testing in all patients to determine eligibility. Although such a targeted use of the drug—given the data shown—would dramatically improve outcomes in eligible patients and reduce wasteful use in the 70% of people who cannot benefit and may see harm, from a business perspective this will probably be seen as highly undesirable, even if one accounts for the high penetration of the eligible market segment, improved drug adherence, longer duration of treatment, and avoidance of adverse event–related liabilities that are often cited as countering financial arguments in favor of personalized healthcare. Thus, without incentives or a regulatory mandate, few sponsors would carry out such costly post hoc screens for stratification parameters, particularly in the case of positive trial results, let alone the additional prospective clinical validation studies that may become necessary.
And Yet What Is at Stake If They Do Not?
In the present case, based on the data presented by Tardif et al., it could indeed be argued that it would be unethical not to carry out the requisite prospective follow-up studies on dalcetrapib. A new opportunity to offer highly effective cardioprotection to an easily identified subpopulation may be forfeited. By extension, the use of hypothesis-free whole genome screens in all new drug development trials, including those with successful outcomes in the aggregate analysis, offers a means of looking for clinically relevant subpopulations for each new medicine developed. Costly trial-and-error approaches to prescribing can be reduced; adverse reactions can be better predicted and avoided. This promise of targeted medicines, reduced suffering and improved cost-effectiveness does not rely on the previous elucidation of underlying physiological mechanisms. Instead, it marries the immense power of biomarkers and computational biostatistics in service of public health—a new paradigm indeed.

These ethical and clinical considerations mandate that we begin the conversation about how to ensure that carrying out biomarker screens in all clinical trials examining new drugs becomes standard procedure. There are precedents for requiring companies to act in service of the public good, even when it violates narrow business interests, such as the now compulsory disclosure of trial details and publication of all study outcomes. To require, as default, the kind of post hoc screening reported upon here, and, in particular, the conduct of additional prospective trials is a challenge of a different magnitude, but the potential benefit is also enormous. Although the cost of a genome screen is trivial compared with the overall expense of a drug’s clinical development program, the consequences of finding a biomarker–response association in subgroups that would have to be validated by additional prospective trials, and in particular the possibly resulting change in the drug’s anticipated addressable market, is certainly anything but trivial. It will, therefore, be important to establish not only relevant regulatory guidance mandating the requisite studies and establishing a level playing field for all applicants but also to find the appropriate mechanisms to provide incentives that counterbalance what by sponsors will a priori be viewed as additional new hurdles to obtaining market approval. Only then it can be ensured that subsets of probably drug responders, if they exist, will reliably be identified and will eventually gain access to medical benefits, although others that are likely not to benefit from a new medicine, or even be harmed, will be protected from receiving it.

How might this be accomplished? For 1, retrospective screens should always be conducted on phase II trials. If substantive marker-associated effects exist, they may well be apparent, and phase III trials could be designed accordingly. Given the limited size and statistical power of phase II trials, biomarker–outcome effects may, however, not be discovered; thus, retrospective screens should be mandatory for phase III studies as well. The customary process of carrying out 2 phase III studies could be adopted to include the hypothesis-free screen in 1, with the other being modified in case any findings are encountered. A different approach, if the results of a retrospective screen are sufficiently compelling, that would provide clear incentive to the sponsor may be to allow the drug to enter the market under a fast track or similar mechanism, for treatment of the putative beneficiaries under strict pharmacovigilance, and with an obligation for the sponsor to carry out appropriate postapproval studies. Yet another approach may be to provide public funding, such as may be available under the president’s new Precision Medicine Initiative, for prospective validation studies to be performed by the public sector, if the sponsor company decides not to pursue such studies. This might come with an obligation for the sponsor to share the ownership of the new drug with the public in some manner. Entering this obligation could become part of filing a New Drug Application for new investigational drugs.

Outlook
The brave new world of drug development research that includes the hypothesis-free retrospective investigations that Tardif et al.’s work heralds may be daunting, but it also offers a glimpse of hope that we may be at the brink of fundamentally improving medical care. Improving the failure-littered track record of clinical studies is certainly a prospect that sponsors should embrace. However, advocating for this approach as a universally applied principle of drug development offers an even more compelling vista for society by allowing new medicines that may be highly effective in addressing unmet medical needs to reach the specific patient subgroups who will benefit, instead of being prescribed in a shotgun manner to all comers or being abandoned because of lack of differentiated targeting.

Disclosures
None.

References

Key Words: Editorials ▪ biomarkers ▪ clinical trial ▪ drug approval ▪ genome-wide association study ▪ pharmacogenetics
Chasing—and Catching—the Wild Goose: Hypothesis-Free Post-Hoc Stratification Studies as a New Paradigm for Drug Development

Klaus Lindpaintner

Circ Cardiovasc Genet. 2015;8:253-255
doi: 10.1161/CIRCGENETICS.115.001037

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/8/2/253

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/