JACC: HEART FAILURE VOL. ■, NO. ■, 2018

PUBLISHED BY ELSEVIER

EDITORIAL COMMENT

Will Biomarkers Succeed as a Surrogate Endpoint in Heart Failure Trials?*

James L. Januzzi, Jr, MD

During the past 20 years, it is estimated the number of therapeutic clinical studies has increased dramatically, with the number of trials recorded in international clinical trial registries increasing from 3,294 in 2004 to 23,384 in 2013 (1). This remarkable growth in the number of trials has come with a troubling development: more than ever, pivotal studies are often slow in execution, are frequently enormous in size, and may be troublingly expensive.

What solutions exist to this problem? A clear explanation is to do something to shorten the length of clinical trials, as it is estimated it takes 14 years from a therapy being identified to patient care. Surprisingly, however, despite all efforts to make studies more efficient, the amount of time to execute clinical trials has paradoxically increased: in an analysis of more than 17,000 studies performed between 2005 and 2015, the cycle time for phase II and phase III trials appears to have lengthened to 39 months (phase II) and 42.9 months (phase III) (2).

Perhaps most surprising in this statistic is how long phase II trials now take. This may be due to increased efforts to "de-risk" the phase III pivotal trial through vigorous scrutiny earlier in development programs: phase II studies are larger (e.g., median size of a phase II trial grew from 88 to 108 subjects between 2005 and 2015) and take longer (with a 23% increase in duration when scrutinizing trials performed between 2013 and 2015 compared with between 2010 and 2012) (2). Furthermore, complexity of phase II trial design has increased with gathering of numerous physiologic and intermediate endpoints such as biomarkers, in an effort to "predict the future" for a successful pivotal trial effort.

Phase III trials remain slow for another reason: dogmatic use of "harder" clinical endpoints such as mortality or hospitalization due to the fear of using endpoints that are not sufficient for regulatory approval or reimbursement. Even with strategies to enroll higher risk patients, use of cause-specific endpoints (e.g., cardiovascular death rather than all-cause mortality), time to first event analyses, and composite endpoints with clinical events equated with risk for mortality (e.g., hospitalization for heart failure or cardiovascular death), it still takes time to perform event-based pivotal studies.

To overcome the need for larger and slower studies based on lesser common clinical outcomes such as death or hospitalization, there has been great interest in using novel trial designs using alternative endpoints, such as the hierarchical combination of post-treatment change in symptoms, either alone or in combination with canonical endpoints such as mortality. In theory, a trial using such novel endpoints could more rapidly provide similar outcome information as a larger study that uses canonical endpoints such as death or hospitalization alone. Into this discussion comes whether other alternative measures of prognosis such as post-treatment change in a relevant biomarker may be introduced as an endpoint for trials, either alone or as a part of a hierarchical endpoint approach.

It may be intuitively obvious why use of a circulating biomarker to substitute for an endpoint

*Editorials published in *JACC: Heart Failure* reflect the views of the authors and do not necessarily represent the views of *JACC* or the American College of Cardiology.

From the Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts; Harvard Medical School, Boston, Massachusetts; and the Cardiometabolic Trials, Baim Institute for Clinical Research, Boston, Massachusetts. Dr. Januzzi is supported in part by the Hutter Family Professorship; has received grant support from Roche Diagnostics, Siemens, Cleveland Heart Labs, and Prevencio; consulting income from Roche Diagnostics, Abbott, Phillips, and Novartis; and has participated in clinical endpoint committees or data safety monitoring boards for Abbvie, Bayer, Pfizer, Novartis, Amgen, Janssen, and Boehringer Ingelheim.

ISSN 2213-1779/$36.00 https://doi.org/10.1016/j.jchf.2018.02.008

would be attractive in a heart failure (HF) trial. Markers such as B-type natriuretic peptide (BNP) or its amino terminal equivalent (N-terminal pro-B-type natriuretic peptide [NT-proBNP]) are easily and accurately measured, they are inexpensive, and there is a huge fund of knowledge suggesting they (and other biomarkers) may be strongly prognostic in those affected by HF. However, it remains discussed and appropriately debated whether a post-intervention change in a biomarker concentration is acceptable for consideration as a tool for regulatory approval of a novel HF therapy. This discussion has been going on for more than a decade, and we have not moved forward much, in contrast to colleagues in other medical specialties. One might ask, what are we waiting for?

The answer may be surprising to some. Although secular trends of many biomarkers have been strongly linked to outcomes such as worsening HF, hospitalization for HF, and cardiovascular (CV) death in patients randomized in clinical studies, there is much that remains unknown about such findings.

It is not entirely clear, for example, when one should measure a post-treatment NT-proBNP after therapy initiation to best determine prognosis, and it is similarly uncertain just how much change in the biomarker is absolutely necessary to predict a longer-term favorable effect of a therapy. Also, there is a risk a favorable trend in a biomarker might not reveal other risks from a therapy (3). Last, it is well established that some therapies may foster favorable changes in biomarker concentrations, but such changes may be counterintuitive at times, and may occur at a different rate than with other treatments. For example, in early studies of beta-blocker therapy, concentrations of BNP and NT-proBNP were noted to rise in some patients despite favorable outcomes. Had natriuretic peptides been used as surrogate endpoints in these pilot trials, larger studies might not have been performed. Although it is now well understood that beta-blocker therapy chronically lowers BNP and NT-proBNP in parallel with its benefits, these experiences illustrate importance of a good understanding of biomarker-drug interface, to avoid type I or II error.

Thus, remarkably, though numerous studies suggest post-treatment biomarker changes may be leveraged for "predicting the future," we still lack understanding of optimal timing for post-treatment biomarker measurement along with clarity regarding the biomarker-therapy interface (which markers to measure for which therapy).

In this issue of the *JACC: Heart Failure*, Vaduganathan et al. (4) provide useful information to help reduce uncertainty regarding how biomarker changes in clinical trials may fit in. In their study, the investigators conducted a trial-level analysis of 16 phase III chronic HF trials completed between 1987 and 2013 studying 18 therapeutic comparisons in 48,844 patients. They calculated weighted Pearson correlation coefficients between average control versus therapy changes in NPs and the longer-term treatment effects on clinical endpoints. Biomarker concentrations were measured at a median of 4 months after randomization. The authors found changes in NP concentrations were surprisingly not correlated to the effect of treatment on all-cause mortality, but appeared modestly correlated to occurrence of HF hospitalization. Correlations between NP change and HF hospitalization were strongest in studies of drugs inhibiting the renin-angiotensin-aldosterone system, where change in biomarker concentration was nearly perfectly correlated with hospitalization ($r = 0.97$; $p = 0.0002$). Somewhat surprisingly, these results remained robust even if the post-treatment sample was early (e.g., 1 to 3 months), intermediate (3 to 6 months), or later (>6 months) after randomization.

Three main questions remain about their results:

- First, timing of the post-treatment sample was not standardized across the trials analyzed; though their analysis placing studies into early, intermediate, and later time for resampling offers reassurance, it would be of value to know if it is absolutely necessary to wait more than 6 months to remeasure a biomarker to inform benefit of a therapy.
- Second, how much NP lowering (or how much of a difference between treatment and placebo) needed to determine benefit was not clarified; the nature of the analysis performed by the authors does not inform whether a "target" value or a percentage change would be best to predict benefit. Thus, it is possible a nonlinear association between biomarker change and mortality was present.
- Third, despite well-established associations between change in NP concentrations with mortality, the startling result suggesting no link between NP change and all-cause mortality requires careful scrutiny; though one might argue CV death may be more likely linked to NP change, the majority of deaths in most HF trials are CV in nature. Could this have been due to inclusion of beta-blocker studies, where an early rise in NP concentrations did not associate with improved mortality seen in these studies? If so, would a later NP measurement have affected the authors' result? Or could the

TABLE 1 Information Required to Confidently Use a Biomarker as a Surrogate Endpoint

- A strong correlation between the impact of an intervention on the biomarker and the impact of that intervention on a clinically meaningful endpoint.
- The outcome measure of interest predicted by the biomarker should be remediable through therapeutic intervention.
- Preferably, a biomarker should reflect both benefit as well as risk related to drug exposure.
- Sampling strategies for measurement of the surrogate biomarker, relative to risk for outcome measure must be established, with time course between change in biomarker and substituted outcome well understood.

Reprinted with permission from Ibrahim et al. (5).

result be due to inclusion of trials focused on implantable cardioverter-defibrillators, where the main outcome prevented was arrhythmic death, something BNP or NT-proBNP are less strongly predictive of?

The results from Vaduganathan et al. (4) affirm previous suggestions regarding required knowledge before biomarkers may be best leveraged as an endpoint in clinical trials (Table 1). Studies to establish answers to these requirements should be executed at this point if biomarkers can be judged as acceptable endpoints for clinical trials. Without such understanding—which is still lacking for nearly all cardiac biomarkers—we cannot move forward as our colleagues have in oncology trials, where efficiently run biomarker-supported pivotal trials are now the norm, rapidly delivering approvable and reimbursable therapies to patients in need.

ADDRESS FOR CORRESPONDENCE: Dr. James L. Januzzi, Jr., Cardiology Division, Massachusetts General Hospital, 32 Fruit Street, Yawkey 5984, Boston, Massachusetts 02114. E-mail: jjanuzzi@partners.org.

REFERENCES

1. Viergever RF, Li K. Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. BMJ Open 2015;5:e008932.

2. Martin L, Hutchens M, Hawkins C. Trial watch: clinical trial cycle times continue to increase despite industry efforts. Nat Rev Drug Discov 2017;16:157.

3. Mebazaa A, Nieminen MS, Packer M, et al. Levosimendan vs dobutamine for patients with acute decompensated heart failure: The survive randomized trial. JAMA 2007;297:1883–91.

4. Vaduganathan M, Claggett B, Packer M, et al. Natriuretic peptides as biomarkers of treatment response in clinical trials of heart failure. J Am Coll Cardiol HF 2018;6:XXX–XXX.

5. Ibrahim NE, Gaggin HK, Konstam MA, Januzzi JL Jr. Established and emerging roles of biomarkers in heart failure clinical trials. Circ Heart Fail 2016;9:e002528.

KEY WORDS biomarker, heart failure, outcomes