

# UCSF

## UC San Francisco Previously Published Works

### Title

Reassessing the Role of Surrogate End Points in Drug Development for Heart Failure

### Permalink

<https://escholarship.org/uc/item/1qp5m78m>

### Journal

Circulation, 138(10)

### ISSN

0009-7322

### Authors

Greene, Stephen J  
Mentz, Robert J  
Fiuzat, Mona  
[et al.](#)

### Publication Date

2018-09-04

### DOI

10.1161/circulationaha.118.034668

Peer reviewed



Published in final edited form as:

*Circulation*. 2018 September 04; 138(10): 1039–1053. doi:10.1161/CIRCULATIONAHA.118.034668.

## Reassessing the Role of Surrogate Endpoints in Drug Development for Heart Failure

Stephen J. Greene, MD<sup>1,2</sup>, Robert J. Mentz, MD<sup>1,2</sup>, Mona Fiuzat, PharmD<sup>1</sup>, Javed Butler, MD, MPH, MBA<sup>3</sup>, Scott D. Solomon, MD<sup>4</sup>, Andrew P. Ambrosy, MD<sup>1,2</sup>, Cyrus Mehta, PhD<sup>5</sup>, John R. Teerlink, MD<sup>6,7</sup>, Faiez Zannad, MD, PhD<sup>8</sup>, and Christopher M. O'Connor, MD<sup>9</sup>

<sup>1</sup>Duke Clinical Research Institute, Durham, NC, USA

<sup>2</sup>Division of Cardiology, Duke University School of Medicine, Durham, NC, USA

<sup>3</sup>Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA

<sup>4</sup>Division of Cardiology, Brigham and Women's Hospital, Boston, MA, USA

<sup>5</sup>Harvard School of Public Health, Boston, MA, USA

<sup>6</sup>Section of Cardiology, San Francisco Veterans Affairs Medical Center, San Francisco, CA, USA

<sup>7</sup>School of Medicine, University of California San Francisco, San Francisco, CA, USA

<sup>8</sup>Centre d'Investigation Clinique Plurithématique 1433, INSERM U1116, Université de Lorraine, CHRU de Nancy, Nancy, France

<sup>9</sup>Inova Heart and Vascular Institute, Falls Church, VA, USA

### Abstract

With few notable exceptions, drug development for heart failure (HF) has become progressively more challenging and there remain no definitively proven therapies for patients with acute HF or HF with preserved ejection fraction. Inspection of temporal trends suggests an increasing rate of disagreement between early phase and phase III trial endpoints. Preliminary results from phase II HF trials are frequently promising, but increasingly followed by disappointing phase III results. Given this potential disconnect, it is reasonable to carefully re-evaluate the purpose, design, and execution of phase II HF trials, with particular attention directed towards the surrogate endpoints commonly used by these studies. In this review, we offer a critical reappraisal of the role of phase II trials and surrogate endpoints, highlighting challenges in their use and interpretation, lessons learned from past experiences, and specific strengths and weaknesses of various surrogate outcomes. We conclude by proposing a series of approaches that should be considered for the goal of optimizing the efficiency of HF drug development. This review is based on discussions between scientists, clinical trialists, industry and government sponsors, and regulators that took place at the Cardiovascular Clinical Trialist Forum in Washington, DC on December 2, 2016.

**Address for Correspondence:** Stephen J Greene, MD, Duke Clinical Research Institute, 2400 Pratt Street, Room 311 Terrace Level, Durham, NC 27705; telephone: 919-684-8111; fax: 919-668-7078; stephen.greene@duke.edu.

**Twitter handle:** @SJGreene\_md, @DCRINews

## Keywords

heart failure; clinical trial; drug development; surrogate; endpoint

Over the past 30 years, outcomes for patients with chronic heart failure (HF) with reduced ejection fraction (HFrEF) have dramatically improved with development and implementation of multiple life-saving therapies. In the span of a decade, the field made relatively rapid advances with landmark trials generating foundational evidence for current guideline-directed medical therapy.<sup>1-3</sup> However, since those earlier times, the complexity and costs of drug development for HF have grown. Moreover, despite numerous specifically dedicated trials, patients with acute HF (AHF) and HF with preserved EF (HFpEF) remain without a single therapy definitively proven to improve outcomes.

In the context of efforts to improve HF drug development, available data from the past 2 decades suggest an increasing rate of disagreement between early phase and definitive phase III trials.<sup>4, 5</sup> Particularly in AHF and HFpEF, findings from phase II trials are frequently interpreted as promising, only to be followed by disappointing phase III results.<sup>6</sup> While reasons for a potential disconnect are complex and likely multifactorial, given the central nature of phase II trials and surrogate endpoints in the HF drug development paradigm, it is reasonable to re-evaluate their current and continued roles. In this review, we offer a critical reappraisal of the purpose of phase II trials and surrogate endpoints, highlighting challenges in their use and interpretation, lessons learned from past experiences, and specific strengths and weaknesses of various surrogate outcomes used in HFrEF, HFpEF, and AHF trials. We conclude by proposing a series of approaches that should be considered for optimizing the efficiency of HF drug development. This review is based on discussions between scientists, clinical trialists, industry and government sponsors, and regulators that took place at the Cardiovascular Clinical Trialist Forum in Washington, DC on December 2, 2016. Notably, the specific focus of this article is pharmacotherapy trials for HF and considerations outlined here should be distinguished from discussions related to device-based therapy.

## TRADITIONAL ROLES OF PHASE II AND SURROGATE ENDPOINTS IN DRUG DEVELOPMENT

Phase II studies traditionally form the foundation for “go/ no go” decisions on large time-consuming and costly phase III trials intended to provide definitive data for potential regulatory approval and labeling. Phase II may also influence many of the phase III study design features, including the ultimate target population, dose, and safety monitoring. Phase IIb studies, in particular, are similar to phase III in that they are frequently assigned a single or small number of primary efficacy endpoints. However, unlike phase III where HF trials are generally powered to evaluate mortality or hospitalization, primary efficacy endpoints for phase II trials are almost uniformly surrogate measures.<sup>4</sup> These endpoints facilitate limited sample size requirements and increase the feasibility of executing a smaller budget trial over a shorter duration, while also allowing investigators to see the effect of the agent on measures believed to have strong correlation with clinical outcomes.<sup>4</sup> Thus, under the traditional approach, the importance of the specific surrogate endpoint choice in early phase

trials cannot be overstated. Although far from the only factor, the ultimate decision to abandon or continue development of an investigational agent, as well as doses and outcomes to be tested, may be decided largely on its ability to meet these surrogate measures.

## GENERAL LIMITATIONS OF SURROGATE ENDPOINTS

Various types of measurements can be candidate surrogate endpoints, including laboratories, imaging studies, and physical exam findings. Moreover, while endpoints such as exercise tolerance and patient-reported outcomes (e.g. quality of life, dyspnea relief) are often termed “intermediate endpoints” since they measure their own independent clinical benefit, they are often viewed as potential surrogates for “hard” clinical outcomes such as death or hospitalization.<sup>7</sup>

Overall, the use of surrogates in cardiovascular medicine has been fraught with challenges and requires a cautious approach. A study by Patel et al. suggests that while rates of positive results in clinical endpoint trials have declined significantly over time, rates of positive surrogate and intermediate endpoint results remain consistently near 70% (**Figure 1**).<sup>5</sup> Lessons lie in numerous historical examples of premature false confidence in surrogate markers.<sup>8,9</sup> Specifically within the HF space, review of all phase II to phase IV HF trials from 2001–2012 shows that nearly 60% of all such trials utilized a surrogate or intermediate non-mortality primary endpoint.<sup>4</sup> These trials were more frequently positive than trials with endpoints including all-cause or cardiovascular mortality (74.4% vs 45.3%), consistent with a poor ability of phase II endpoints to predict late phase results.<sup>4</sup> Although factors such as an increasing number of trials over time and potential temporal shifts in goals of phase II work may be acknowledged as possible contributors, the available trial-level data suggest a substantial rate of disagreement between phase II and phase III/IV endpoints.<sup>4,5</sup>

Accordingly, the bar for declaring a measure a reliable surrogate endpoint should be set high. While virtually no surrogate is perfect, for a measure to serve as a reasonable and “reliable” marker for hard outcomes, several criteria are required, but not necessarily sufficient (**Table 1**).<sup>10,11</sup> Of these, perhaps the most difficult to achieve is for the correlation between therapeutically modulated levels of the marker and risk of outcome to be consistent across various classes of interventions and populations.<sup>10</sup> This latter point is particularly relevant to the use of surrogates in studies of novel therapies, where, by definition, the novelty of the intervention precludes prior validation of the biomarker in that setting. Indeed, investigational therapies are at risk of exerting deleterious effects that may neutralize or override any potential organ-specific benefits associated with change in the surrogate.

While multiple measures have been used as HF surrogate endpoints, each carries significant limitations and there remains no uniformly reliable surrogate efficacy endpoint for phase II drug trials in HF (**Table 2**). We again distinguish the present discussion as directed towards development of pharmacologic agents for HF and not necessarily applicable to medical devices, acknowledging that the role of surrogate endpoints in development and regulatory approval of devices may differ from drugs.<sup>12</sup> In the following sections, we consider the utility and challenges associated with frequently considered surrogate endpoints in early phase HF trials of pharmacologic agents.

## SURROGATE ENDPOINTS IN PHASE II DRUG TRIALS FOR HEART FAILURE

### Invasive Hemodynamics

Predating the contemporary HFrEF treatment paradigm centered on neurohormonal blockade, early treatment approaches focused on directly correcting hemodynamic abnormalities to reduce ventricular filling pressures and improve cardiac index. Based on encouraging hemodynamic data from early phase studies, flosequinan was approved for use in chronic HF in Europe in September 1992 and in the United States shortly thereafter.<sup>13–15</sup> A subsequent phase III outcome study of flosequinan was terminated early due to increased mortality risk, leading to withdrawal of the drug from all markets.<sup>16</sup> Similarly, oral milrinone sparked enthusiasm as a treatment for severe refractory HF based on favorable hemodynamic effects, only to be later shown to increase mortality and other adverse events, and is no longer available.<sup>17, 18</sup> The HF literature contains multiple other examples highlighting that short-term hemodynamic effects with an investigational agent cannot reliably predict effects on mortality and hospitalization in HFrEF.<sup>6</sup> This assertion stands in potential contrast to long-term hemodynamic improvements, which may be less driven by sudden changes in preload, afterload, or contractility, but are presumably more likely to reflect favorable reverse remodeling or ventricular recovery. For example, beta-blockers may sometimes cause acute hemodynamic worsening, but are cornerstone therapies for improving long-term outcomes.

Independent of the exact measures themselves, use of invasive hemodynamic surrogate endpoints in HF drug development has practical concerns. Since the publication of the ESCAPE (Evaluation Study of Congestive Heart Failure and Pulmonary Artery Catheterization Effectiveness) trial, which showed no benefit with use of a pulmonary artery catheter (PAC) in AHF, clinicians may be reluctant to insert a PAC, particularly for patients without hypotension.<sup>19</sup> Indeed, current HF guidelines discourage use of a PAC in routine HF management, and recommend limiting to patients with cardiogenic shock or mechanical ventilation.<sup>20</sup> Thus, while guidelines make routine use of invasive hemodynamic endpoints impractical, difficulty rationalizing PAC use in trials is compounded by the common study requirement that patients be normotensive, or even relatively hypertensive. The phase II COMPOSE program offers an example.<sup>21</sup> Eligible patients were required to both have a “pre-existing requirement” for a PAC and a baseline systolic blood pressure >120 mmHg. The study was terminated early, largely due to patient recruitment futility.<sup>21</sup>

### Change in Natriuretic Peptide Level: Correlations with Patient Outcomes

**Heart Failure with Reduced Ejection Fraction**—Although natriuretic peptide (NP) measurements have been consistently associated with risk of adverse clinical events in the inpatient and outpatient settings, the ability of NPs to serve as a reliable surrogate for morbidity and mortality in early phase HF clinical trials is fraught with challenges. For example, the phase III ASTRONAUT (Aliskiren Trial on Acute Heart Failure Outcomes) trial found addition of aliskiren to standard HF therapy did not influence mortality or hospitalization endpoints, despite a statistically significant and sustained decrease in N-terminal pro-B-type natriuretic peptide (NT-proBNP) over 12-month follow-up (**Figure 2**).<sup>22</sup> A recent analysis across 16 chronic HF trials (excluding ASTRONAUT), most of which

were HF<sub>r</sub>EF, identified similar dissociations between therapeutic effects on NPs and all-cause mortality, although a correlation with HF hospitalization was seen.<sup>24</sup> Nonetheless, correlations with HF hospitalization were driven largely by trials of renin-angiotensin aldosterone system (RAAS) inhibitors, and thus may have less applicability to future drug development targeting other pathways.<sup>24</sup> In the context of multiple examples of concordance and discordance between change in NPs and mortality and hospitalization outcomes, such relationships may be most accurately described as therapy-specific.<sup>25, 26</sup> Although less data are available and rare examples of discordance exist, more consistent relationships may be seen when using treatment-related NP change as a surrogate for effects on patient-centered outcomes in chronic HF<sub>r</sub>EF, such as changes in symptoms, quality of life, or exercise capacity.<sup>22, 27, 28</sup>

Atrial fibrillation (AF) may further complicate the effective use of a NP endpoint. Data support a tendency for higher NP levels among HF patients with AF as compared to those without.<sup>28</sup> Thus, it is plausible that an investigational HF therapy may have differing ability to lower NP depending on the rhythm status of the patient, and that prevalence of AF in the study population could influence overall results. The ASTRONAUT investigators studied this hypothesis and found that the ability of aliskiren to lower NT-proBNP varied by baseline rhythm, with incremental NP lowering through 12 month follow-up only among patients without baseline AF (**Figure 3**).<sup>30</sup>

**Heart Failure with Preserved Ejection Fraction**—In addition to aforementioned issues, use of NP defined endpoints in HF<sub>p</sub>EF may carry additional challenges. Although NP levels carry prognostic significance comparable to HF<sub>r</sub>EF, absolute levels in HF<sub>p</sub>EF tend to be lower with up to one third of patients having normal levels despite significantly elevated filling pressures on invasive assessment.<sup>31</sup> In addition, while trial selection criteria based on elevated NP levels may facilitate recruitment of higher risk patients, it may signal inclusion of patients with advanced disease less likely to respond to study therapy. For example, in the I-PRESERVE (Irbesartan in Heart Failure with Preserved Ejection Fraction Trial) and TOPCAT (Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist) trials, patients with lower baseline NT-proBNP were more likely to respond to study therapy.<sup>32, 33</sup> Moreover, in the PARAMOUNT (Prospective Comparison of ARNI with ARB on Management of Heart Failure with Preserved Ejection Fraction) trial, despite a sustained robust reduction in NT-proBNP during follow-up, LCZ696 did not improve patient-centered outcomes, including quality of life or a composite including NYHA functional classification and patient global assessment.<sup>34</sup> These collective data suggest that conclusions on NP levels gained from HF<sub>r</sub>EF patients need not apply to HF<sub>p</sub>EF. Further research is needed to understand the interaction between HF<sub>p</sub>EF and NP levels before the biomarker can be used as a reliable surrogate endpoint in HF<sub>p</sub>EF drug development.

**Acute Heart Failure**—Several AHF trials have demonstrated that rapid and robust lowering of NP is possible with multiple therapies, but that these short-term changes may not reliably track with improved downstream clinical or patient-reported outcomes (PROs).<sup>35–37</sup> For instance, in the SURVIVE (Survival of Patients With Acute Heart Failure in Need

of Intravenous Inotropic Support) trial, AHF patients randomized to levosimendan had greater decreases in NP level at 24 hours and 5 days compared to those receiving dobutamine, but similar scores on global and dyspnea assessments at 24 hours and similarly high rates of 180-day mortality.<sup>35</sup>

### **Change in Natriuretic Peptide Level: Application within Clinical Trials**

The biological and analytical variability of NP measurements must be recognized in their use as endpoints. For example, among stable HF patients, intra-individual variability of NT-proBNP levels 1 hour and 1 week apart may approach 7% and 21%, respectively.<sup>38</sup> Likewise, the analytical variability (i.e., imprecision of the test) of the specific assay must be considered. To better distinguish altered patient status from inherent variability, prior work has described reference change values and suggested change in NT-proBNP >50–60% as clinically relevant.<sup>39, 40</sup>

Use of change in NT-proBNP as a surrogate endpoint is further complicated by the heterogeneous way in which the endpoint may be defined. These definitions may influence endpoint interpretation and highlight the need, at minimum, to ensure strict adherence to pre-specified statistical methods and for consideration of sensitivity analyses using alternative approaches.<sup>7</sup> For example, recent studies have been variable in their decision to evaluate absolute change in NP as a continuous variable versus the proportion of patients achieving a specified reduction from baseline (e.g., >30% relative reduction).<sup>41, 42</sup> Additionally, some trials take a ‘per-protocol’ or ‘completers’ approach with the primary analysis, whereby only patients who are alive with valid baseline and follow-up NT-proBNP measurements are used in the primary analysis.<sup>41</sup> Benefits of such a strategy include evaluation of complete data potentially most informative for dose-finding. Nevertheless, this approach is subject to downstream imbalances in clinical or safety events between study groups, as it is conceivable that death or withdrawal of ‘sicker’ patients in a particular study arm could bias the remaining population towards lower NT-proBNP levels. Alternatively, other studies have followed the ‘intention-to-treat’ principle, whereby ‘last observation carried forward’ or imputed values are inserted for patients not completing the study protocol.<sup>34</sup>

Lastly, although the randomization process should support similar patient characteristics between study arms, the generally modest size of phase II trials increases risk of random baseline NP imbalances. Study groups randomly starting with higher baseline NP levels may be more likely to experience a greater degree of subsequent decrease, independent of investigational therapy. For example, in both the ARTS-HF (Mineralocorticoid Receptor Antagonist Tolerability Study-Heart Failure) and SOCRATES-REDUCED (Soluble Guanylate Cyclase Stimulator in Heart Failure with Reduced Ejection Fraction Study) trials, control arm patients had substantially higher baseline NT-proBNP levels, potentially impacting the failure of both investigational agents to provide significant reduction relative to placebo and meet their respective primary endpoints.<sup>41, 42</sup>

## Change in Troponin Level

Across the spectrum of HF, troponin levels are elevated in a significant proportion of patients in the absence of overt clinical ischemia.<sup>43, 44</sup> Although the exact mechanism of this myocardial injury is unclear, data support troponin elevation as a prognostic marker and HF therapies that increase troponin (e.g., inotropes) have generally led to worse clinical outcomes.<sup>43–45</sup> However, unlike elevation in NPs, it is biologically plausible that troponin may be a potential mediator of poor outcomes, and thus may be a more suitable surrogate or therapeutic target. In addition, the biological variation profile of troponin relative to NPs may be better suited for serial measurement as a trial endpoint. Specifically, troponin measures generally have less intra-individual variation and reference change values suggest more modest changes are clinically relevant (i.e., ~30%).<sup>40</sup> Nonetheless, despite increased attention, HF therapies targeting troponin reduction have not reliably improved clinical outcomes and future work is needed to clarify the potential utility of troponin as an endpoint in HFrEF, HFpEF, and AHF. Most recently, the neutral top line results from the RELAX-AHF2 (Relaxin in Acute Heart Failure-2; NCT01870778) trial showed no long-term mortality benefit of serelaxin therapy in AHF, despite favorable short-term effects on troponin and several other markers of organ injury in an earlier study.<sup>36</sup> In contrast, among HFpEF patients enrolled in the PARAMOUNT trial, compared with valsartan, LCZ696 therapy resulted in further reduction in high-sensitivity troponin T and parallel improvements in NPs and left atrial size.<sup>46</sup> Of significant current interest, omecamtiv mecarbil has been consistently shown to produce low level troponin elevation and the definitive phase III outcome trial is ongoing.<sup>47, 48</sup>

## Cardiac Imaging

**Heart Failure with Reduced Ejection Fraction**—Imaging parameters have been regularly used as endpoints within early phase HFrEF trials with 2D echocardiography the most frequently utilized modality. Reverse remodeling has been a common focus, generally equated to an improvement in ejection fraction or ventricular volumes.<sup>49</sup> Rationale for these endpoints stems largely from a large meta-analysis across trials of patients with left ventricular dysfunction where short-term trial-level effects of a drug or device on ventricular remodeling were associated with long-term trial-level effects on mortality.<sup>49</sup>

However, practical limitations to echocardiographic endpoints exist. First, neither ejection fraction nor ventricular dimensions directly measure intrinsic myocardial function, as both are highly influenced by preload and afterload conditions. Multiple examples exist where short-term changes in ejection fraction are discordant from effects on survival.<sup>45</sup> Second, echocardiographic measurements may demonstrate moderate inter-reader variability and suboptimal reproducibility, particularly in the setting of relatively modest improvements expected in most trials.<sup>50</sup> Third, despite frequent use of core labs intended to standardize interpretation, quality of local image acquisition within multicenter trials may vary and issues may remain with calibration of serial measurements.

**Heart Failure with Preserved Ejection Fraction**—Aforementioned limitations notwithstanding, recent HFpEF trials have shown increasing focus on echocardiographic endpoints, including diastolic function and left atrial volume.<sup>34, 51</sup> However, despite



compelling biologic plausibility, the utility of these measures in predicting patient outcomes is unclear. For example, in the ALDO-DHF (Aldosterone Receptor Blockade in Diastolic Heart Failure) trial, randomization to spironolactone significantly improved E/e' ratio compared to placebo, but did not improve HF symptoms or quality of life and slightly reduced 6-minute walk distance.<sup>52</sup> Likewise, despite secondary analyses of the TOPCAT trial suggesting clinical benefits with spironolactone among patients from North and South America, no corresponding reduction in left atrial volume was seen.<sup>53</sup>

**Alternative Imaging Modalities Irrespective of Ejection Fraction**—More recently, there has been increasing consideration of cardiac magnetic resonance (CMR) or positron emission tomography (PET) within trial design.<sup>54, 55</sup> As compared with 2D echocardiography, CMR is a volumetric technique with high contrast and spatial resolution that favors improved accuracy and reproducibility for multiple measurements, including EF, volumes, and left ventricular mass.<sup>56, 57</sup> Because study power is dependent on effect size and inversely proportional to the variance of the measurement, use of CMR for endpoint assessment may facilitate a reduction in sample size.<sup>57, 58</sup> In addition, CMR also allows evaluation of myocardial constituents (i.e., myocyte, interstitium, microcirculation), as well as the degree of tissue viability and scar. Such detailed evaluation of the myocardium offers the potential of serial assessments to detect a specific therapeutic effect (e.g., change in myocardial fibrosis in response to an anti-fibrotic agent). Likewise, molecular imaging with PET could provide novel insight into drug dose and target receptor engagement, and the metabolic function of the heart.<sup>59</sup> Nonetheless, despite the promise of both CMR and PET as novel assessments of cardiac structure and function, to date, their use in drug development programs has been markedly limited. Thus, their effective use in reliably predicting a therapeutic response on clinical outcomes is unclear. Current hurdles to widespread use include limited availability at select centers, higher costs, longer exam times, and use of implantable cardioverter-defibrillators (for CMR).

### Functional Capacity

Functional capacity is an important “intermediate” endpoint frequently evaluated in early phase HF trials. The most commonly utilized measures in routine practice and clinical research are the 6 minute walk distance (6MWD) and cardiopulmonary exercise testing (CPET) (e.g. peak oxygen consumption). Both measures have demonstrated consistent ability to predict downstream survival in HFrEF and HFpEF, and thus engendered great interest as surrogate endpoints in drug development.<sup>60–62</sup> Most recently, the NEAT-HFpEF (Nitrate’s Effect on Activity Tolerance in Heart Failure with Preserved Ejection Fraction) trial introduced a novel wearable accelerometer to measure daily patient activity level.<sup>63</sup> Although the utility of this novel measure in future HF programs remains to be seen, an ancillary study from NEAT-HFpEF failed to show significant correlations between therapeutically mediated changes in accelerometer units and changes in traditional functional assessments, including 6MWD.<sup>64</sup>

Despite evidence of prognostic value, use of serial exercise measurements as effective surrogates for mortality in clinical trials has been problematic, with multiple instances of HFrEF therapies with proven survival benefits offering little to no effect on exercise

tolerance.<sup>65, 66</sup> Likewise, inotropes may increase mortality despite improving functional capacity.<sup>67</sup> Reasons for such discordance are unclear, but a variety of issues could potentially obscure efficacy signals, including placebo effect (with improved exercise capacity with increased familiarization with the exercise protocol), uneven patient dropout between study arms, and uneven intensification of background therapy to maintain symptomatic stability between exercise assessments.<sup>68</sup> Moreover, with 6MWD in particular, concerns exist over the fidelity and reproducibility of serial measurements, as well as variation in clinician “coaching” and patient effort from one visit to the next and between trial sites. These issues may be less problematic with CPET where the respiratory equivalence ratio can objectively characterize patient effort and core labs can better standardize interpretation. Nonetheless, CPET testing may only be available at select centers and the burden of serial intensive exercise examinations on patients may present a challenge for trial enrollment.

### Dyspnea Relief and Patient Reported Outcomes

Although increasingly recognized as “intermediate” endpoints with their own important intrinsic value, PROs among HF patients have been frequently viewed in the context of surrogates for traditional mortality and hospitalization outcomes. Among PROs, dyspnea is the most common primary reason for HF hospitalization and is a frequently studied outcome in AHF trials. In contrast with other surrogates, dyspnea endpoints have extended beyond early phase trials to use as a primary endpoint in phase III AHF studies.<sup>69–72</sup> Although no longer necessarily the case, dyspnea relief had historically been a regulatory approval endpoint, for which there was precedent.<sup>69</sup> Regardless, use of dyspnea as an effective endpoint in HF trials has been consistently challenging. There remains no consensus on how to best measure dyspnea, and multiple grading instruments are available and inconsistently used. Indeed, by meeting only 1 of 2 co-primary endpoints, the RELAX-AHF (Relaxin in Acute Heart Failure) trial highlights how an investigational agent may improve dyspnea as measured by one scale, but not another.<sup>72</sup> Moreover, the URGENT (Ularitide Global Evaluation in Acute Decompensated Heart Failure) dyspnea study suggested that >75% of AHF patients have rapid dyspnea improvement within 6 hours of presentation with standard therapy.<sup>73</sup> Thus, dyspnea is often not severe by time of study enrollment, likely hindering ability of an investigational agent to show incremental improvement. In light of these ascertainment challenges, data linking dyspnea relief to improved longer-term quality of life, mortality, and rehospitalization outcomes are inconsistent.<sup>74</sup> Collectively, these data support dyspnea relief as an important patient-centered outcome, but not as a reliable substitute for clinical events, such as mortality or HF hospitalization. Similar conclusions can be applied to other PROs, including those derived from quality of life scales.<sup>65, 67, 75</sup>

### FUTURE DIRECTIONS

Given the challenges outlined, use of surrogate endpoints and phase II trials in HF drug development warrants reassessment. In an effort to maximize chances of successful therapeutic development and improve trial efficiency and costs, we propose the following strategies (**Figure 4**). While the effectiveness of these considerations requires prospective

validation, we believe them to be practical and to potentially offer key advantages over current approaches.

### Tailor Phase II to Key Drug-Specific Questions

Recognizing that there remains no fully reliable surrogate endpoint for clinical events, we believe the principal purpose of contemporary phase II trials should *not* be to predict phase III results. Likewise, it stands to reason that these studies should not be held to phase III standards, and thus not be simplistically determined ‘positive’ or ‘neutral/negative’ by virtue of effects on a primary endpoint. Rather, we propose that early phase studies be designed with intent to better understand the mechanism and biologic impact of therapy, including confirmation of ‘on target’ effects/ receptor engagement, dose finding, and pharmacokinetics/ pharmacodynamics. The ongoing development program for omecamtiv mecarbil may serve as an illustrative example. In this case, a pre-clinical and phase I program identified a clear mechanism of action (i.e., increasing the transition rate to the strongly bound myosin-actin cross-bridge state) with a biologically plausible and physiologic marker of drug response (i.e., systolic ejection time).<sup>76</sup> Subsequent phase II studies further validated drug effect by confirming dose-dependent effects on systolic ejection time and characterizing pharmacokinetics.<sup>47, 77</sup> Indeed, the primary endpoint of the largest phase II study of omecamtiv mecarbil in chronic HF (COSMIC-HF [Chronic Oral Study of Myosin activation to Increase Contractility in Heart Failure]) was the plasma concentration of the agent at 12 weeks, and not a generic surrogate endpoint (e.g., change in NT-proBNP level, functional status, etc).<sup>47</sup> Rather, investigators linked plasma concentration of the drug to secondary study endpoints specific to the drug mechanism (i.e., systolic ejection time). Although the clinical efficacy and safety of omecamtiv mecarbil remains to be definitively determined, a similar developmental approach centered on rigorously confirming biologic mechanism and ‘on target’ cardiac effects may prove more informative than evaluating isolated effects on a generic HF surrogate.<sup>78</sup>

### Clinical Composite Endpoints and Multi-dimensional Endpoint Assessment

Consistent with the above mentioned use of phase II studies to define the biologic impact of therapy, simultaneous evaluation of multiple endpoint domains may be considered. Under this approach, there would be understanding that limited power may make achievement of statistical significance of any particular endpoint difficult.<sup>79</sup> Rather, the premise would be that congruent signals of benefit on multiple measures, such as symptoms, functional status, biomarkers, quality of life, and clinical outcomes, may best represent the merits of continued study in phase III. Alternatively, instead of independent assessment of multiple endpoints, composite outcomes may be used. Assignment of weights (i.e., ranking) to individual components of a composite has been proposed, such as with use of a hierarchical clinical composite endpoint based on symptomatic, functional status, or quality of life improvement in the absence a worsening HF event (such as death or HF hospitalization).<sup>68</sup> Similarly, biomarkers may be incorporated within a global rank composite, as utilized in the recently published phase II FIGHT (Functional Impact of GLP-1 for Heart Failure Treatment) study.<sup>79, 80</sup> Others have proposed an average Z-score endpoint where components of the composite are unweighted.<sup>81</sup> Under this approach, the treatment effect on each component is converted to a Z score and Z scores across outcomes are averaged.<sup>81, 82</sup>

Regardless, while a multi-domain endpoint approach may offer greater insight on the biologic impact of therapy compared with studies using single or small numbers of single-domain primary endpoints, limitations should be acknowledged. We note that a multi-domain approach may still fail to predict phase III results, with positive results for serelaxin in the phase IIb Pre-RELAX-AHF study coupled with neutral results in the phase III RELAX-AHF2 trial being a notable example.<sup>83</sup> Moreover, composite endpoint results (e.g., a unit-less global rank score) and the corresponding effect sizes can be difficult to interpret clinically and subject to heterogeneity of effect across component outcomes. For example, a short follow-up duration diminishes the contribution of important but rare events (e.g., mortality and hospitalization) to a hierarchical composite endpoint and increases likelihood overall results are dominated by other components (e.g., traditional surrogates such as changes in dyspnea or NT-proBNP).<sup>84</sup> To better interpret composite endpoint results, Brown and Ezekowitz have supported use of the probability index (PI), a bounded measure between 0 and 1 with a value of 0.5 suggesting no difference.<sup>81</sup> The PI represents the probability of a randomly selected patient from the treatment arm having a response superior to a randomly selected patient from the control arm.<sup>81</sup> An advantage of the PI is that it readily describes the treatment effect size for component endpoints and overall composites using a common measure (rather than a combination of hazard ratios, differences between means, etc), thus facilitating assessment of heterogeneity and clinical significance.<sup>81</sup> While decisions for subsequent phase III investigation based on a series of singular measures or a composite outcome may inherently retain some degree of subjectivity, use of the PI and pre-specified procedures for considering endpoint heterogeneity and clinically meaningful benefits may favor more effective use of the multi-domain endpoint approach in phase II.

### Merging Phase II and Phase III with Adaptive Trial Design

While traditional trial design involves a fixed pre-specified sample size and final efficacy assessment after all subjects have been enrolled, an adaptive strategy monitors accruing data at set time points and includes pre-specified criteria for modifying design features as the study moves forward. Recognizing that phase II oftentimes offers questionable value and that only an adequate phase III trial can reliably prove efficacy and safety, drug development programs may consider streamlining the process by specifying conditions whereby phase II studies transform directly into phase III. A study may start with a phase II design, but include interim assessments where early intimation of efficacy, harm, or futility on phase III endpoints are made using pre-specified statistical boundaries (**Figure 5**).<sup>85</sup> Although a trial could begin with a traditional phase II sample size and surrogate endpoint, interval signals for “hard” efficacy and safety endpoints would be tracked and study leadership would reserve the pre-specified right to convert towards phase III by increasing trial enrollment and changing the primary efficacy or safety endpoint.<sup>85</sup> Thus, the trial would have potential to increase study power, via preset criteria based on interim data, such that therapeutic efficacy and safety was definitively determined. Similarly, a phase II study might start with a primary intent of dose-finding, and then continue directly to phase III with discontinuation of ineffective dosing arms.

Adaptive design may offer several advantages over conventional approaches. By starting with a conventional phase II sample size, study sponsors would still receive the initial

clinical data they generally require before comfortably dedicating resources towards a large phase III trial. However, no patient data or study participants would be “wasted”, and such a “head start” on phase III enrollment may be particularly valuable in this era of difficult and slow phase III trial recruitment. Moreover, such design may reduce costs to the sponsor by consolidating site training and start-up costs, requiring fewer overall patient participants, and decreasing time sponsors dedicate to investigational agents that eventually prove ineffective. While an adaptive trial design does include limitations, such as data variability inherent to interim analyses, the requirement for carefully defined pre-specified criteria to guide trial modification, the potential for bias due to possible alterations in investigator behavior following an adaptive change, and the possible inability of a small sponsor to provide upfront support for a large seamless phase II-phase III program, this design strategy represents an attractive alternative that may be increasingly considered in HF. Effective examples of adaptive design have been seen in other areas of cardiovascular disease, as well as oncology.<sup>85, 86</sup>

### **Proceed Directly to Phase III Without Phase II**

In select circumstances, the most appropriate path for an investigational drug may be to skip phase II and proceed directly to phase III. While such course of action is non-traditional, the recent landmark PARADIGM-HF (Prospective comparison of ARNI with ACEI to Determine Impact on Global Mortality and morbidity in Heart Failure) trial and the development of sacubitril/valsartan have drawn attention to this possibility.<sup>87</sup> Omapatrilat, a dual acting angiotensin-converting enzyme inhibitor (ACEI)/neprilysin inhibitor was developed more than a decade ago and tested in the 5,770 patient phase III OVERTURE (Omapatrilat Versus Enalapril Randomized Trial of Utility in Reducing Events) trial where it was superior to enalapril for the endpoint of CV death or hospitalization and in analyses including recurrent HF hospitalizations.<sup>87</sup> Despite these promising findings, development of omapatrilat was terminated due to excess risk of angioedema. Nonetheless, these collective data provided both the rationale for replacement of the ACEI component with an angiotensin II receptor blocker to circumvent angioedema, as well as the foundational efficacy data supporting dual neprilysin and RAAS inhibition. A similar strategy of bypassing phase II may be reasonable if a drug compound shows an impressive HF efficacy signal in a large non-HF clinical trial. For example, findings from the EMPA-REG OUTCOME (Empagliflozin Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients—Removing Excess Glucose) trial among 7,020 patients with type 2 diabetes found treatment with empagliflozin associated with a 35% relative risk reduction for HF hospitalization, with findings consistent among subsets with and without prevalent HF.<sup>89</sup> Accordingly, the decision was made to initiate dedicated phase III trials among patients with established HF, irrespective of diabetes status (NCT03057977, NCT03057951).

Although conditions surrounding development of sacubitril/valsartan and empagliflozin may be rare, both examples highlight situations where it is unclear if a traditional modestly-sized phase II program with the specific agent in a HF population could have offered information superior to the larger scale, but only closely-related, data already available. Thus, these examples highlight potential opportunity for bypassing phase II when ample large-scale

randomized clinical data exist with a) a very closely-related drug in a HF population, or b) the same drug in a closely-related or overlapping trial population.

## CONCLUSIONS

Over the past two decades, the frequency of positive phase III trial results with investigational HF drugs has declined while rates of positive phase II results using surrogate endpoints have remained high. Despite many cases of compelling biologic plausibility, past experience strongly suggests that there remains no reliable and proven surrogate endpoint for phase III trial outcomes. Accordingly, a reappraisal of the design, execution, and purpose of phase II trials is warranted. It may be most appropriate to shift traditional thinking regarding phase II from intending to predict phase III results to a tailored approach specific to the study agent and focused on key questions relevant to the drug-patient interaction and drug mechanism. The overarching goal of such a strategy would be to provide investigators with information useful in optimizing the design and execution of a potential phase III program, rather than to predict the final result. Similarly, innovations in adaptive trial design have potential to streamline the drug development process and decrease the time and cost required for definitive efficacy and safety data. All relevant stakeholders, including patients, healthcare providers, academicians, regulators, and industry sponsors, have strong incentives to optimize drug development in efforts to improve HF patient outcomes. As part of this collective mission, added scrutiny towards the design, execution, and purpose of phase II trials and the use of surrogate endpoints carries the promise of significant potential benefit.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

None

### SOURCES OF FUNDING

None

### DISCLOSURES

Dr. Greene is supported by National Institutes of Health grant 5T32HL069749–14 and a Heart Failure Society of America/ Emergency Medicine Foundation Acute Heart Failure Young Investigator Award funded by Novartis, and has received research support from Amgen and Novartis. Dr. Mentz receives research support from Amgen, AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, Gilead, Novartis, Otsuka, and ResMed; and honoraria from Thoratec. Dr. Butler has received research support from the National Institutes of Health, PCORI and the European Union; and serves as a consultant for Amgen, Array, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, CVRx, G3 Pharmaceutical, Innolife, Janssen, Luitpold, Medtronic, Merck, Novartis, Relypsa, StealthPeptide, SC Pharma, Vifor, and ZS Pharma. Dr. Solomon has received grant funding, consultant fees, and travel support from Novartis. Dr. Ambrosy is supported by National Institutes of Health grant 5T32HL069749. Dr. Teerlink receives research/consulting fees from Amgen, Madeleine, Mast Therapeutics, Novartis, Relypsa, and Trevena. Dr. Zannad has received grant funding from Novartis, BG Medicine, and Roche Diagnostics; served on a board for Boston Scientific; and served as a consultant for Novartis, Takeda, AstraZeneca, Boehringer-Ingelheim, GE Healthcare, Relypsa, Servier, Boston Scientific, Bayer, Johnson & Johnson, and ResMed. Dr. O'Connor reports consulting fees from Novella and Amgen; ownership/ partnership/ principal in Biscardia, LLC; and research support from Otsuka, Roche Diagnostics, BG Medicine, Critical Diagnostics, Astellas, Gilead, GE Healthcare, and ResMed. All other authors report no conflicts.

## REFERENCES

1. SOLVD Investigators, Yusuf S, Pitt B, Davis CE, Hood WB, Cohn JN. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med.* 1991;325:293–302. [PubMed: 2057034]
2. CONSENSUS Trial Study Group. Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). *N Engl J Med.* 1987;316:1429–1435. [PubMed: 2883575]
3. Packer M, Bristow MR, Cohn JN, Colucci WS, Fowler MB, Gilbert EM, Shusterman NH. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. U.S. Carvedilol Heart Failure Study Group. *N Engl J Med.* 1996;334:1349–1355. [PubMed: 8614419]
4. Samman Tahhan A, Vaduganathan M, Kelkar A, Georgiopoulou VV, Kalogeropoulos AP, Greene SJ, Fonarow GC, Gheorghiade M, Butler J. Trends in Heart Failure Clinical Trials From 2001–2012. *J Card Fail.* 2016;22:171–179. [PubMed: 26106806]
5. Patel RB, Vaduganathan M, Samman-Tahhan A, Kalogeropoulos AP, Georgiopoulou VV, Fonarow GC, Gheorghiade M, Butler J. Trends in Utilization of Surrogate Endpoints in Contemporary Cardiovascular Clinical Trials. *Am J Cardiol.* 2016;117:1845–1850. [PubMed: 27085935]
6. Vaduganathan M, Greene SJ, Ambrosy AP, Gheorghiade M, Butler J. The disconnect between phase II and phase III trials of drugs for heart failure. *Nat Rev Cardiol.* 2013;10:85–97. [PubMed: 23296069]
7. Ferreira JP, Duarte K, Graves TL, Zile MR, Abraham WT, Weaver FA, Lindenfeld J, Zannad F. Natriuretic Peptides, 6-Min Walk Test, and Quality-of-Life Questionnaires as Clinically Meaningful Endpoints in HF Trials. *J Am Coll Cardiol.* 2016;68:2690–2707. [PubMed: 27978953]
8. AIM-HIGH Investigators Boden WE, Probstfield JL, Anderson T, Chaitman BR, Desvignes-Nickens P, Koprowicz K, McBride R, Teo K, Weintraub W Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. *N Engl J Med.* 2011;365:2255–2267. [PubMed: 22085343]
9. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, Arensberg D, Baker A, Friedman L, Greene HL, Huther ML, Richardson DW. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med.* 1991;324:781–788. [PubMed: 1900101]
10. Prasad V, Bonow RO. The cardiovascular biomarker conundrum: challenges and solutions. *JAMA.* 2011;306:2151–2152. [PubMed: 22089725]
11. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med.* 1989;8:431–440. [PubMed: 2727467]
12. Food and Drug Administration. Expedited Access for Premarket Approval and De Novo Medical Devices Intended for Unmet Medical Need for Life Threatening or Irreversibly Debilitating Diseases or Conditions; Guidance for Industry and Food and Drug Administration Staff. 2015 Available at: <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM393978.pdf>. Accessed September 2, 2017.
13. Packer M, Narahara KA, Elkayam U, Sullivan JM, Pearle DL, Massie BM and Creager MA. Double-blind, placebo-controlled study of the efficacy of flosequinan in patients with chronic heart failure. Principal Investigators of the REFLECT Study. *J Am Coll Cardiol.* 1993;22:65–72. [PubMed: 8509565]
14. Gottlieb SS, Kukin ML, Penn J, Fisher ML, Cines M, Medina N, Yushak M, Taylor M, Packer M. Sustained hemodynamic response to flosequinan in patients with heart failure receiving angiotensin-converting enzyme inhibitors. *J Am Coll Cardiol.* 1993;22:963–967. [PubMed: 8409070]
15. Vaduganathan M, Butler J, Gheorghiade M. Transforming Drug Development in Heart Failure: Navigating the Regulatory Crossroads. *Circ Heart Fail.* 2016;9:e003192 [PubMed: 27707750]
16. Packer M, Pitt B, Rouleau JL, Swedberg K, DeMets DL, Fisher L. Long-Term Effects of Flosequinan on the Morbidity and Mortality of Patients With Severe Chronic Heart Failure: Primary Results of the PROFILE Trial After 24 Years. *JACC Heart Fail.* 2017;5:399–407. [PubMed: 28501522]

17. Simonton CA, Chatterjee K, Cody RJ, Kubo SH, Leonard D, Daly P, Rutman H. Milrinone in congestive heart failure: acute and chronic hemodynamic and clinical evaluation. *J Am Coll Cardiol.* 1985;6:453–459. [PubMed: 4019930]
18. Packer M, Carver JR, Rodeheffer RJ, Ivanhoe RJ, DiBianco R, Zeldis SM, Hendrix GH, Bommer WJ, Elkayam U, Kukin ML, Mallis GI, Sollano JA, Shannon J, Tandon PK, DeMets DL. Effect of oral milrinone on mortality in severe chronic heart failure. *N Engl J Med.* 1991;325:1468–1475. [PubMed: 1944425]
19. Binanay C, Califf RM, Hasselblad V, O'Connor CM, Shah MR, Sopko G, Stevenson LW, Francis GS, Leier CV, Miller LW. Evaluation study of congestive heart failure and pulmonary artery catheterization effectiveness: the ESCAPE trial. *JAMA.* 2005;294:1625–1633. [PubMed: 16204662]
20. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation.* 2013;128:e240–327. [PubMed: 23741058]
21. Gheorghiadu M, Greene SJ, Filippatos G, Erdmann E, Ferrari R, Levy PD, Maggioni A, Nowack C, Mebazaa A. Cinaciguat, a soluble guanylate cyclase activator: results from the randomized, controlled, phase IIB COMPOSE programme in acute heart failure syndromes. *Eur J Heart Fail.* 2012;14:1056–1066. [PubMed: 22713287]
22. Gheorghiadu M, Bohm M, Greene SJ, Fonarow GC, Lewis EF, Zannad F, Solomon SD, Baschiera F, Botha J, Hua TA, Gimpelewicz CR, Jaumont X, Lesogor A, Maggioni AP. Effect of aliskiren on postdischarge mortality and heart failure readmissions among patients hospitalized for heart failure: the ASTRONAUT randomized trial. *JAMA.* 2013;309:1125–1135. [PubMed: 23478743]
23. Gheorghiadu M, Böhm M, Greene SJ, Fonarow GC, Lewis EF, Zannad F, Solomon SD, Baschiera F, Botha J, Hua TA, Gimpelewicz CR, Jaumont X, Lesogor A, Maggioni AP. Effect of Aliskiren on Post-discharge Mortality and Heart Failure Readmissions Among Patients Hospitalized for Heart Failure: Aliskiren Trial on Acute Heart Failure Outcomes Oral Presentation at: American College of Cardiology (ACC) 2013 Scientific Sessions. 3 11, 2013 San Francisco, CA.
24. Vaduganathan M, Claggett B, Packer M, McMurray JJV, Rouleau JL, Zile MR, Swedberg K, Solomon SD. Natriuretic Peptides as Biomarkers of Treatment Response in Clinical Trials of Heart Failure. *JACC Heart Fail.* 2018 [Epub ahead of print] doi: 10.1016/j.jchf.2018.02.007.
25. Hartmann F, Packer M, Coats AJ, Fowler MB, Krum H, Mohacsi P, Rouleau JL, Tendera M, Castaigne A, Trawinski J, Amann-Zalan I, Hoersch S, Katus HA. NT-proBNP in severe chronic heart failure: rationale, design and preliminary results of the COPERNICUS NT-proBNP substudy. *Eur J Heart Fail.* 2004;6:343–350. [PubMed: 14987586]
26. Packer M, Colucci W, Fisher L, Massie BM, Teerlink JR, Young J, Padley RJ, Thakkar R, Delgado-Herrera L, Salon J, Garratt C, Huang B, Sarapohja T. Effect of levosimendan on the short-term clinical course of patients with acutely decompensated heart failure. *JACC Heart Fail.* 2013;1:103–111. [PubMed: 24621834]
27. Bhardwaj A, Rehman SU, Mohammed AA, Gaggin HK, Barajas L, Barajas J, Moore SA, Sullivan D, Januzzi JL. Quality of life and chronic heart failure therapy guided by natriuretic peptides: results from the ProBNP Outpatient Tailored Chronic Heart Failure Therapy (PROTECT) study. *Am Heart J.* 2012;164:793–799 e1. [PubMed: 23137512]
28. Shah KB, Kop WJ, Christenson RH, Diercks DB, Kuo D, Henderson S, Hanson K, deFilippi CR. Post-discharge changes in NT-proBNP and quality of life after acute dyspnea hospitalization as predictors of one-year outcomes. *Clin Biochem.* 2010;43:1405–1410. [PubMed: 20850426]
29. Kristensen SL, Jhund PS, Mogensen UM, Rorth R, Abraham WT, Desai A, Dickstein K, Rouleau JL, Zile MR, Swedberg K, Packer M, Solomon SD, Kober L, McMurray JJV. Prognostic Value of N-Terminal Pro-B-Type Natriuretic Peptide Levels in Heart Failure Patients With and Without Atrial Fibrillation. *Circ Heart Fail.* 2017;10:e004409. [PubMed: 29018174]
30. Greene SJ, Fonarow GC, Solomon SD, Subacius HP, Ambrosy AP, Vaduganathan M, Maggioni AP, Bohm M, Lewis EF, Zannad F, Butler J, Gheorghiadu M. Influence of atrial fibrillation on post-discharge natriuretic peptide trajectory and clinical outcomes among patients hospitalized for



- heart failure: insights from the ASTRONAUT trial. *Eur J Heart Fail.* 2017;19:552–562. [PubMed: 27748006]
31. Anjan VY, Loftus TM, Burke MA, Akhter N, Fonarow GC, Gheorghiade M, Shah SJ. Prevalence, clinical phenotype, and outcomes associated with normal B-type natriuretic peptide levels in heart failure with preserved ejection fraction. *Am J Cardiol.* 2012;110:870–876. [PubMed: 22681864]
  32. Anand IS, Rector TS, Cleland JG, Kuskowski M, McKelvie RS, Persson H, McMurray JJ, Zile MR, Komajda M, Massie BM, Carson PE. Prognostic value of baseline plasma amino-terminal pro-brain natriuretic peptide and its interactions with irbesartan treatment effects in patients with heart failure and preserved ejection fraction: findings from the I-PRESERVE trial. *Circ Heart Fail.* 2011;4:569–577. [PubMed: 21715583]
  33. Anand IS, Claggett B, Liu J, Shah AM, Rector TS, Shah SJ, Desai AS, O’Meara E, Fleg JL, Pfeffer MA, Pitt B, Solomon SD. Interaction Between Spironolactone and Natriuretic Peptides in Patients With Heart Failure and Preserved Ejection Fraction: From the TOPCAT Trial. *JACC Heart Fail.* 2017;5:241–252. [PubMed: 28359411]
  34. Solomon SD, Zile M, Pieske B, Voors A, Shah A, Kraigher-Krainer E, Shi V, Bransford T, Takeuchi M, Gong J, Lefkowitz M, Packer M, McMurray JJ. The angiotensin receptor neprilysin inhibitor LCZ696 in heart failure with preserved ejection fraction: a phase 2 double-blind randomised controlled trial. *Lancet.* 2012;380:1387–1395. [PubMed: 22932717]
  35. Mebazaa A, Nieminen MS, Packer M, Cohen-Solal A, Kleber FX, Pocock SJ, Thakkar R, Padley RJ, Poder P, Kivikko M. Levosimendan vs dobutamine for patients with acute decompensated heart failure: the SURVIVE Randomized Trial. *JAMA.* 2007;297:1883–1891. [PubMed: 17473298]
  36. Metra M, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Ponikowski P, Unemori E, Voors AA, Adams KF, Jr, Dorobantu MI, Grinfeld L, Jondeau G, Marmor A, Masip J, Pang PS, Werdan K, Prescott MF, Edwards C, Teichman SL, Trapani A, Bush CA, Saini R, Schumacher C, Severin T, Teerlink JR. Effect of serelaxin on cardiac, renal, and hepatic biomarkers in the Relaxin in Acute Heart Failure (RELAX-AHF) development program: correlation with outcomes. *J Am Coll Cardiol.* 2013;61:196–206. [PubMed: 23273292]
  37. Packer M, O’Connor C, McMurray JJV, Wittes J, Abraham WT, Anker SD, Dickstein K, Filippatos G, Holcomb R, Krum H, Maggioni AP, Mebazaa A, Peacock WF, Petrie MC, Ponikowski P, Ruschitzka F, van Veldhuisen DJ, Kowarski LS, Schactman M, Holzmeister J. Effect of Ularitide on Cardiovascular Mortality in Acute Heart Failure. *N Engl J Med.* 2017;376:1956–1964. [PubMed: 28402745]
  38. O’Hanlon R, O’Shea P, Ledwidge M, O’Loughlin C, Lange S, Conlon C, Phelan D, Cunningham S, McDonald K. The biologic variability of B-type natriuretic peptide and N-terminal pro-B-type natriuretic peptide in stable heart failure patients. *J Card Fail.* 2007;13:50–5. [PubMed: 17339003]
  39. Ohta Y, Shimada T, Yoshitomi H, Inoue S, Murakami Y, Shimizu H, Nakamura K, Ohta T, Katoh H and Ishibashi Y. Drop in plasma brain natriuretic peptide levels after successful direct current cardioversion in chronic atrial fibrillation. *Can J Cardiol.* 2001;17:415–20. [PubMed: 11329541]
  40. Meijers WC, van der Velde AR, Muller Kobold AC, Dijck-Brouwer J, Wu AH, Jaffe A, de Boer RA. Variability of biomarkers in patients with chronic heart failure and healthy controls. *Eur J Heart Fail.* 2017;19:357–365. [PubMed: 27766733]
  41. Gheorghiade M, Greene SJ, Butler J, Filippatos G, Lam CS, Maggioni AP, Ponikowski P, Shah SJ, Solomon SD, Kraigher-Krainer E, Samano ET, Muller K, Roessig L, Pieske B. Effect of Vericiguat, a Soluble Guanylate Cyclase Stimulator, on Natriuretic Peptide Levels in Patients With Worsening Chronic Heart Failure and Reduced Ejection Fraction: The SOCRATES-REDUCED Randomized Trial. *JAMA* 2015;314:2251–2262. [PubMed: 26547357]
  42. Filippatos G, Anker SD, Bohm M, Gheorghiade M, Kober L, Krum H, Maggioni AP, Ponikowski P, Voors AA, Zannad F, Kim SY, Nowack C, Palombo G, Kolkhof P, Kimmekamp-Kirschbaum N, Pieper A, Pitt B. A randomized controlled study of finerenone vs. eplerenone in patients with worsening chronic heart failure and diabetes mellitus and/or chronic kidney disease. *Eur Heart J.* 2016;37:2105–2114. [PubMed: 27130705]
  43. Felker GM, Mentz RJ, Teerlink JR, Voors AA, Pang PS, Ponikowski P, Greenberg BH, Filippatos G, Davison BA, Cotter G, Prescott MF, Hua TA, Lopez-Pintado S, Severin T, Metra M. Serial high

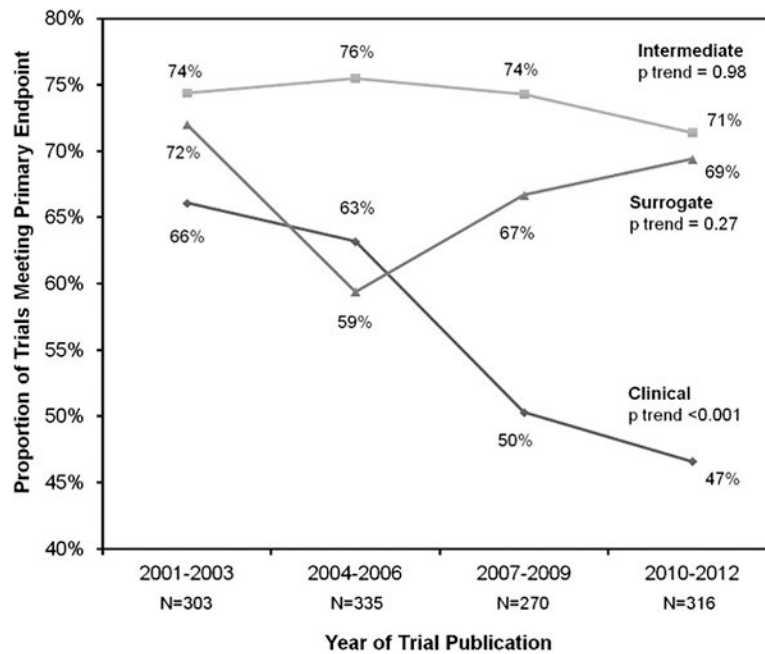
- sensitivity cardiac troponin T measurement in acute heart failure: insights from the RELAX-AHF study. *Eur J Heart Fail*. 2015;17:1262–1270. [PubMed: 26333655]
44. Pandey A, Golwala H, Sheng S, DeVore AD, Hernandez AF, Bhatt DL, Heidenreich PA, Yancy CW, de Lemos JA, Fonarow GC. Factors Associated With and Prognostic Implications of Cardiac Troponin Elevation in Decompensated Heart Failure With Preserved Ejection Fraction: Findings From the American Heart Association Get With The Guidelines-Heart Failure Program. *JAMA Cardiol*. 2017;2:136–145. [PubMed: 28030747]
  45. Francis GS, Bartos JA, Adaya S. Inotropes. *J Am Coll Cardiol*. 2014;63:2069–2078. [PubMed: 24530672]
  46. Jhund PS, Claggett BL, Voors AA, Zile MR, Packer M, Pieske BM, Kraigher-Krainer E, Shah AM, Prescott MF, Shi V, Lefkowitz M, McMurray JJ, Solomon SD. Elevation in high-sensitivity troponin T in heart failure and preserved ejection fraction and influence of treatment with the angiotensin receptor neprilysin inhibitor LCZ696. *Circ Heart Fail*. 2014;7:953–959. [PubMed: 25277997]
  47. Teerlink JR, Felker GM, McMurray JJ, Solomon SD, Adams KF, Jr, Cleland JG, Ezekowitz JA, Goudev A, Macdonald P, Metra M, Mitrovic V, Ponikowski P, Serpytis P, Spinar J, Tomcsanyi J, Vandekerckhove HJ, Voors AA, Monsalvo ML, Johnston J, Malik FI, Honarpour N. Chronic Oral Study of Myosin Activation to Increase Contractility in Heart Failure (COSMIC-HF): a phase 2, pharmacokinetic, randomised, placebo-controlled trial. *Lancet*. 2016;388:2895–2903. [PubMed: 27914656]
  48. Teerlink JR, Felker GM, McMurray JJV, Ponikowski P, Metra M, Filippatos GS, Ezekowitz JA, Dickstein K, Cleland JGF, Kim JB, Lei L, Knusel B, Wolff AA, Malik FI, Wasserman SM. Acute Treatment With Omecamtiv Mecarbil to Increase Contractility in Acute Heart Failure: The ATOMIC-AHF Study. *J Am Coll Cardiol*. 2016;67:1444–1455. [PubMed: 27012405]
  49. Kramer DG, Trikalinos TA, Kent DM, Antonopoulos GV, Konstam MA, Udelson JE. Quantitative evaluation of drug or device effects on ventricular remodeling as predictors of therapeutic effects on mortality in patients with heart failure and reduced ejection fraction: a meta-analytic approach. *J Am Coll Cardiol*. 2010;56:392–406. [PubMed: 20650361]
  50. Jenkins C, Bricknell K, Hanekom L, Marwick TH. Reproducibility and accuracy of echocardiographic measurements of left ventricular parameters using real-time three-dimensional echocardiography. *J Am Coll Cardiol*. 2004;44:878–886. [PubMed: 15312875]
  51. Pieske B, Butler J, Filippatos G, Lam C, Maggioni AP, Ponikowski P, Shah S, Solomon S, Kraigher-Krainer E, Samano ET, Scalise AV, Muller K, Roessig L, Gheorghide M. Rationale and design of the SOLuble guanylate Cyclase stimulator in heart failure Studies (SOCRATES). *Eur J Heart Fail*. 2014;16:1026–1038. [PubMed: 25056511]
  52. Edelmann F, Wachter R, Schmidt AG, Kraigher-Krainer E, Colantonio C, Kamke W, Duvinage A, Stahrenberg R, Durstewitz K, Loffler M, Dungen HD, Tschope C, Herrmann-Lingen C, Halle M, Hasenfuss G, Gelbrich G, Pieske B. Effect of spironolactone on diastolic function and exercise capacity in patients with heart failure with preserved ejection fraction: the Aldo-DHF randomized controlled trial. *JAMA*. 2013;309:781–791. [PubMed: 23443441]
  53. Shah AM, Claggett B, Sweitzer NK, Shah SJ, Deswal A, Anand IS, Fleg JL, Pitt B, Pfeffer MA, Solomon SD. Prognostic Importance of Changes in Cardiac Structure and Function in Heart Failure With Preserved Ejection Fraction and the Impact of Spironolactone. *Circ Heart Fail*. 2015;8:1052–1058. [PubMed: 26475142]
  54. Schelbert EB, Fonarow GC, Bonow RO, Butler J, Gheorghide M. Therapeutic targets in heart failure: refocusing on the myocardial interstitium. *J Am Coll Cardiol*. 2014;63:2188–2198. [PubMed: 24657693]
  55. Papadimitriou L, Smith-Jones PM, Sarwar CM, Marti CN, Yaddanapudi K, Skopicki HA, Gheorghide M, Parsey R, Butler J. Utility of positron emission tomography for drug development for heart failure. *Am Heart J*. 2016;175:142–152. [PubMed: 27179733]
  56. Marwick TH, Neubauer S, Petersen SE. Use of cardiac magnetic resonance and echocardiography in population-based studies: why, where, and when? *Circ Cardiovasc Imaging*. 2013;6:590–6. [PubMed: 23861451]

57. Strohm O, Schulz-Menger J, Pilz B, Osterziel KJ, Dietz R, Friedrich MG. Measurement of left ventricular dimensions and function in patients with dilated cardiomyopathy. *J Magn Reson Imaging*. 2001;13:367–371. [PubMed: 11241808]
58. Bellenger NG, Davies LC, Francis JM, Coats AJ, Pennell DJ. Reduction in sample size for studies of remodeling in heart failure by the use of cardiovascular magnetic resonance. *J Cardiovasc Magn Reson*. 2000;2:271–278. [PubMed: 11545126]
59. Gheorghide M, Larson CJ, Shah SJ, Greene SJ, Cleland JG, Colucci WS, Dunmon P, Epstein SE, Kim RJ, Parsey RV, Stockbridge N, Carr J, Dinh W, Krahn T, Kramer F, Wahlander K, Deckelbaum LI, Crandall D, Okada S, Senni M, Sikora S, Sabbah HN, Butler J. Developing New Treatments for Heart Failure: Focus on the Heart. *Circ Heart Fail*. 2016;9:e002727. [PubMed: 27166246]
60. Keteyian SJ, Patel M, Kraus WE, Brawner CA, McConnell TR, Pina IL, Leifer ES, Fleg JL, Blackburn G, Fonarow GC, Chase PJ, Piner L, Vest M, O'Connor CM, Ehrman JK, Walsh MN, Ewald G, Bensimhon D, Russell SD. Variables Measured During Cardiopulmonary Exercise Testing as Predictors of Mortality in Chronic Systolic Heart Failure. *J Am Coll Cardiol*. 2016;67:780–789. [PubMed: 26892413]
61. Shafiq A, Brawner CA, Aldred HA, Lewis B, Williams CT, Tita C, Schairer JR, Ehrman JK, Velez M, Selektor Y, Lanfear DE, Keteyian SJ. Prognostic value of cardiopulmonary exercise testing in heart failure with preserved ejection fraction. The Henry Ford Hospital CardioPulmonary EXercise Testing (FIT-CPX) project. *Am Heart J*. 2016;174:167–172. [PubMed: 26995385]
62. Bittner V, Weiner DH, Yusuf S, Rogers WJ, McIntyre KM, Bangdiwala SI, Kronenberg MW, Kostis JB, Kohn RM, Guillothe M, Greenberg B, Woods PA, Bourassa MG. Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. *JAMA*. 1993;270:1702–1707. [PubMed: 8411500]
63. Redfield MM, Anstrom KJ, Levine JA, Koeppe GA, Borlaug BA, Chen HH, LeWinter MM, Joseph SM, Shah SJ, Semigran MJ, Felker GM, Cole RT, Reeves GR, Tedford RJ, Tang WH, McNulty SE, Velazquez EJ, Shah MR, Braunwald E. Isosorbide Mononitrate in Heart Failure with Preserved Ejection Fraction. *N Engl J Med*. 2015;373:2314–2324. [PubMed: 26549714]
64. Snipelisky D, Kelly J, Levine JA, Koeppe GA, Anstrom KJ, McNulty SE, Zakeri R, Felker GM, Hernandez AF, Braunwald E, Redfield MM. Accelerometer-Measured Daily Activity in Heart Failure With Preserved Ejection Fraction: Clinical Correlates and Association With Standard Heart Failure Severity Indices. *Circ Heart Fail*. 2017;10:e003878. [PubMed: 28588021]
65. Packer M, Colucci WS, Sackner-Bernstein JD, Liang CS, Goldscher DA, Freeman I, Kukin ML, Kinhal V, Udelson JE, Klapholz M, Gottlieb SS, Pearle D, Cody RJ, Gregory JJ, Kantrowitz NE, LeJemtel TH, Young ST, Lukas MA, Shusterman NH. Double-blind, placebo-controlled study of the effects of carvedilol in patients with moderate to severe heart failure. The PRECISE Trial. Prospective Randomized Evaluation of Carvedilol on Symptoms and Exercise. *Circulation*. 1996;94:2793–2799. [PubMed: 8941104]
66. Gundersen T, Swedberg K, Amtorp O, Remes J, Nilsson B. Absence of effect on exercise capacity of 12-weeks treatment with ramipril in patients with moderate congestive heart failure. Ramipril Study Group. *Eur Heart J*. 1994;15:1659–1665. [PubMed: 7698136]
67. Massie BM, Berk MR, Brozena SC, Elkayam U, Plehn JF, Kukin ML, Packer M, Murphy BE, Neuberg GW, Steingart RM, Levine TB, DeHaan H. Can further benefit be achieved by adding flosequinan to patients with congestive heart failure who remain symptomatic on diuretic, digoxin, and an angiotensin converting enzyme inhibitor? Results of the flosequinan-ACE inhibitor trial (FACET). *Circulation*. 1993;88:492–501. [PubMed: 8339411]
68. Packer M. Development and Evolution of a Hierarchical Clinical Composite End Point for the Evaluation of Drugs and Devices for Acute and Chronic Heart Failure: A 20-Year Perspective. *Circulation*. 2016;134:1664–1678. [PubMed: 27881506]
69. Publication Committee for the VMAC Investigators. Intravenous nesiritide vs nitroglycerin for treatment of decompensated congestive heart failure: a randomized controlled trial. *JAMA*. 2002;287:1531–1540. [PubMed: 11911755]
70. McMurray JJ, Teerlink JR, Cotter G, Bourge RC, Cleland JG, Jondeau G, Krum H, Metra M, O'Connor CM, Parker JD, Torre-Amione G, van Veldhuisen DJ, Lewsey J, Frey A, Rainisio M, Kobrin I. Effects of tezosentan on symptoms and clinical outcomes in patients with acute heart

failure: the VERITAS randomized controlled trials. *JAMA*. 2007;298:2009–2019. [PubMed: 17986694]

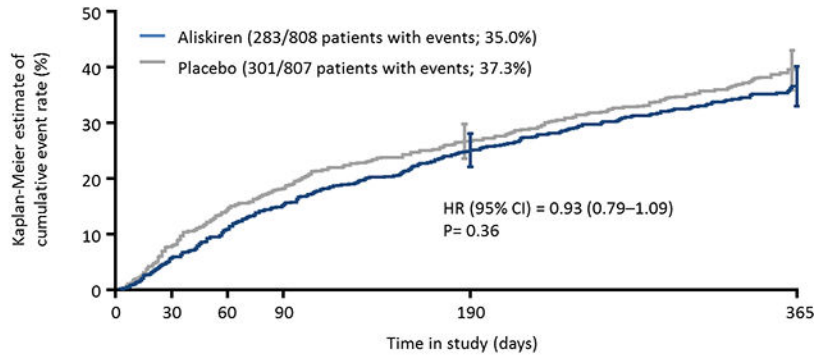
71. O'Connor CM, Starling RC, Hernandez AF, Armstrong PW, Dickstein K, Hasselblad V, Heizer GM, Komajda M, Massie BM, McMurray JJ, Nieminen MS, Reist CJ, Rouleau JL, Swedberg K, Adams KF, Jr., Anker SD, Atar D, Battler A, Botero R, Bohidar NR, Butler J, Clausell N, Corbalan R, Costanzo MR, Dahlstrom U, Deckelbaum LI, Diaz R, Dunlap ME, Ezekowitz JA, Feldman D, Felker GM, Fonarow GC, Gennevois D, Gottlieb SS, Hill JA, Hollander JE, Howlett JG, Hudson MP, Kociol RD, Krum H, Laucevicius A, Levy WC, Mendez GF, Metra M, Mittal S, Oh BH, Pereira NL, Ponikowski P, Tang WH, Tanomsup S, Teerlink JR, Triposkiadis F, Troughton RW, Voors AA, Whellan DJ, Zannad F, Califf RM. Effect of nesiritide in patients with acute decompensated heart failure. *N Engl J Med*. 2011;365:32–43. [PubMed: 21732835]
72. Teerlink JR, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Ponikowski P, Unemori E, Voors AA, Adams KF, Jr., Dorobantu MI, Grinfeld LR, Jondeau G, Marmor A, Masip J, Pang PS, Werdan K, Teichman SL, Trapani A, Bush CA, Saini R, Schumacher C, Severin TM, Metra M Serelaxin, recombinant human relaxin-2, for treatment of acute heart failure (RELAX-AHF): a randomised, placebo-controlled trial. *Lancet*. 2013;381:29–39. [PubMed: 23141816]
73. Mebazaa A, Pang PS, Tavares M, Collins SP, Storrow AB, Laribi S, Andre S, Mark Courtney D, Hasa J, Spinar J, Masip J, Frank Peacock W, Sliwa K, Gayat E, Filippatos G, Cleland JG, Gheorghide M. The impact of early standard therapy on dyspnoea in patients with acute heart failure: the URGENT-dyspnoea study. *Eur Heart J*. 2010;31:832–841. [PubMed: 19906690]
74. Ambrosy AP, Khan H, Udelson JE, Mentz RJ, Chioncel O, Greene SJ, Vaduganathan M, Subacuis HP, Konstam MA, Swedberg K, Zannad F, Maggioni AP, Gheorghide M, Butler J. Changes in Dyspnea Status During Hospitalization and Postdischarge Health-Related Quality of Life in Patients Hospitalized for Heart Failure: Findings From the EVEREST Trial. *Circ Heart Fail*. 2016;9:e002458. [PubMed: 27140204]
75. Lewis EF, Kim HY, Claggett B, Spertus J, Heitner JF, Assmann SF, Kenwood CT, Solomon SD, Desai AS, Fang JC, McKinlay SA, Pitt BA, Pfeffer MA. Impact of Spironolactone on Longitudinal Changes in Health-Related Quality of Life in the Treatment of Preserved Cardiac Function Heart Failure With an Aldosterone Antagonist Trial. *Circ Heart Fail*. 2016;9:e001937. [PubMed: 26962133]
76. Teerlink JR, Clarke CP, Saikali KG, Lee JH, Chen MM, Escandon RD, Elliott L, Bee R, Habibzadeh MR, Goldman JH, Schiller NB, Malik FI, Wolff AA. Dose-dependent augmentation of cardiac systolic function with the selective cardiac myosin activator, omecamtiv mecarbil: a first-in-man study. *Lancet*. 2011;378:667–75. [PubMed: 21856480]
77. Cleland JG, Teerlink JR, Senior R, Nifontov EM, Mc Murray JJ, Lang CC, Tsyrlin VA, Greenberg BH, Mayet J, Francis DP, Shaburishvili T, Monaghan M, Saltzberg M, Neyses L, Wasserman SM, Lee JH, Saikali KG, Clarke CP, Goldman JH, Wolff AA, Malik FI. The effects of the cardiac myosin activator, omecamtiv mecarbil, on cardiac function in systolic heart failure: a double-blind, placebo-controlled, crossover, dose-ranging phase 2 trial. *Lancet*. 2011;378:676–683. [PubMed: 21856481]
78. Gheorghide M, Pang PS, O'Connor CM, Prasad K, McMurray J, Teerlink JR, Fiuzat M, Sabbah H, Komajda M. Clinical development of pharmacologic agents for acute heart failure syndromes: a proposal for a mechanistic translational phase. *Am Heart J*. 2011;161:224–232. [PubMed: 21315202]
79. Felker GM, Maisel AS. A global rank end point for clinical trials in acute heart failure. *Circulation Heart failure*. 2010;3:643–646. [PubMed: 20841546]
80. Margulies KB, Hernandez AF, Redfield MM, Givertz MM, Oliveira GH, Cole R, Mann DL, Whellan DJ, Kiernan MS, Felker GM, McNulty SE, Anstrom KJ, Shah MR, Braunwald E, Cappola TP. Effects of Liraglutide on Clinical Stability Among Patients With Advanced Heart Failure and Reduced Ejection Fraction: A Randomized Clinical Trial. *JAMA*. 2016;316:500–508. [PubMed: 27483064]
81. Brown PM, Ezekowitz JA. Composite End Points in Clinical Trials of Heart Failure Therapy: How Do We Measure the Effect Size? *Circ Heart Fail*. 2017;10.
82. Pang PS, Butler J, Collins SP, Cotter G, Davison BA, Ezekowitz JA, Filippatos G, Levy PD, Metra M, Ponikowski P, Teerlink JR, Voors AA, Bharucha D, Goin K, Soergel DG, Felker GM. Biased

- ligand of the angiotensin II type 1 receptor in patients with acute heart failure: a randomized, double-blind, placebo-controlled, phase IIB, dose ranging trial (BLAST-AHF). *Eur Heart J*. 2017;38:2364–2373. [PubMed: 28459958]
83. Teerlink JR, Metra M, Felker GM, Ponikowski P, Voors AA, Weatherley BD, Marmor A, Katz A, Grzybowski J, Unemori E, Teichman SL, Cotter G. Relaxin for the treatment of patients with acute heart failure (Pre-RELAX-AHF): a multicentre, randomised, placebo-controlled, parallel-group, dose-finding phase IIb study. *Lancet*. 2009;373:1429–1439. [PubMed: 19329178]
84. Brown PM, Ezekowitz JA. Letter by Brown and Ezekowitz Regarding Article, “Development and Evolution of a Hierarchical Clinical Composite End Point for the Evaluation of Drugs and Devices for Acute and Chronic Heart Failure: A 20-Year Perspective”. *Circulation*. 2017;135:e889–e891.
85. Bhatt DL, Mehta C. Adaptive Designs for Clinical Trials. *N Engl J Med*. 2016;375:65–74. [PubMed: 27406349]
86. Berry DA, Herbst RS, Rubin EH. Reports from the 2010 Clinical and Translational Cancer Research Think Tank meeting: design strategies for personalized therapy trials. *Clin Cancer Res*. 2012;18:638–644. [PubMed: 22298897]
87. McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, Rouleau JL, Shi VC, Solomon SD, Swedberg K, Zile MR. Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N Engl J Med*. 2014;371:993–1004. [PubMed: 25176015]
88. Packer M, Califf RM, Konstam MA, Krum H, McMurray JJ, Rouleau JL, Swedberg K. Comparison of omapatrilat and enalapril in patients with chronic heart failure: the Omapatrilat Versus Enalapril Randomized Trial of Utility in Reducing Events (OVERTURE). *Circulation*. 2002;106:920–926. [PubMed: 12186794]
89. Fitchett D, Zinman B, Wanner C, Lachin JM, Hantel S, Salsali A, Johansen OE, Woerle HJ, Broedl UC, Inzucchi SE. Heart failure outcomes with empagliflozin in patients with type 2 diabetes at high cardiovascular risk: results of the EMPA-REG OUTCOME trial. *Eur Heart J*. 2016;37:1526–1534. [PubMed: 26819227]



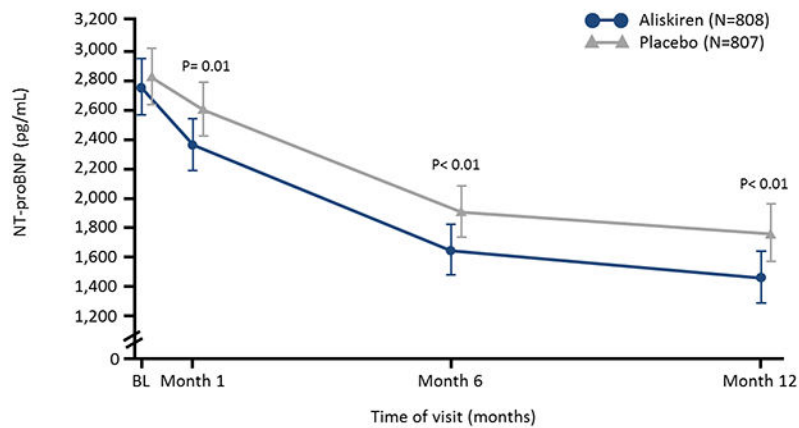
**Figure 1.** Proportion of cardiovascular trials with positive outcomes by the type of primary endpoint over time. Cardiovascular trials using intermediate or surrogate endpoints had similar proportions of positive outcomes over time, whereas trials including clinical endpoints had less positive outcomes over time. Analyses of temporal trends across trial time frames were performed using nonparametric tests for trend. Adapted with permission from Patel RB et al.<sup>5</sup>

**A. Cardiovascular Death or Heart Failure Re-hospitalization Within 12 months**



Patients at risk	0	30	60	90	180	365
Aliskiren	808	762	716	679	597	204
Placebo	807	743	690	655	578	196

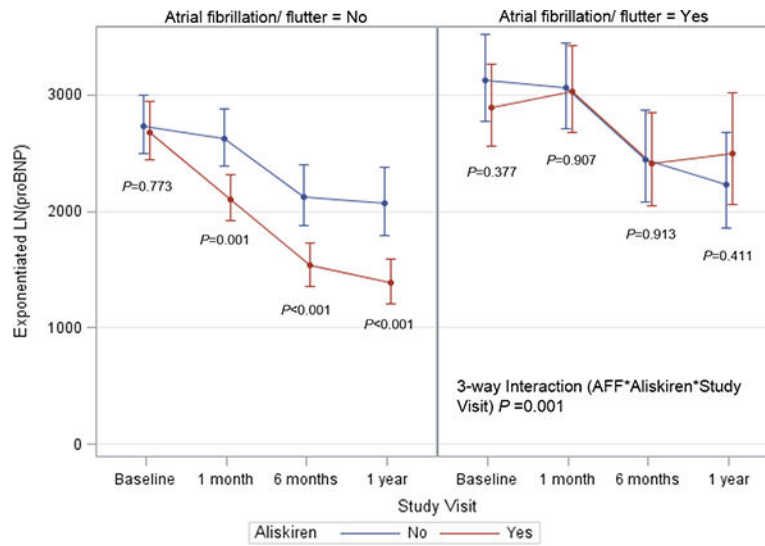
**B. Change in NT-proBNP With Time**



Number of subjects	BL	Month 1	Month 6	Month 12
Aliskiren	778	673	573	451
Placebo	776	678	562	430

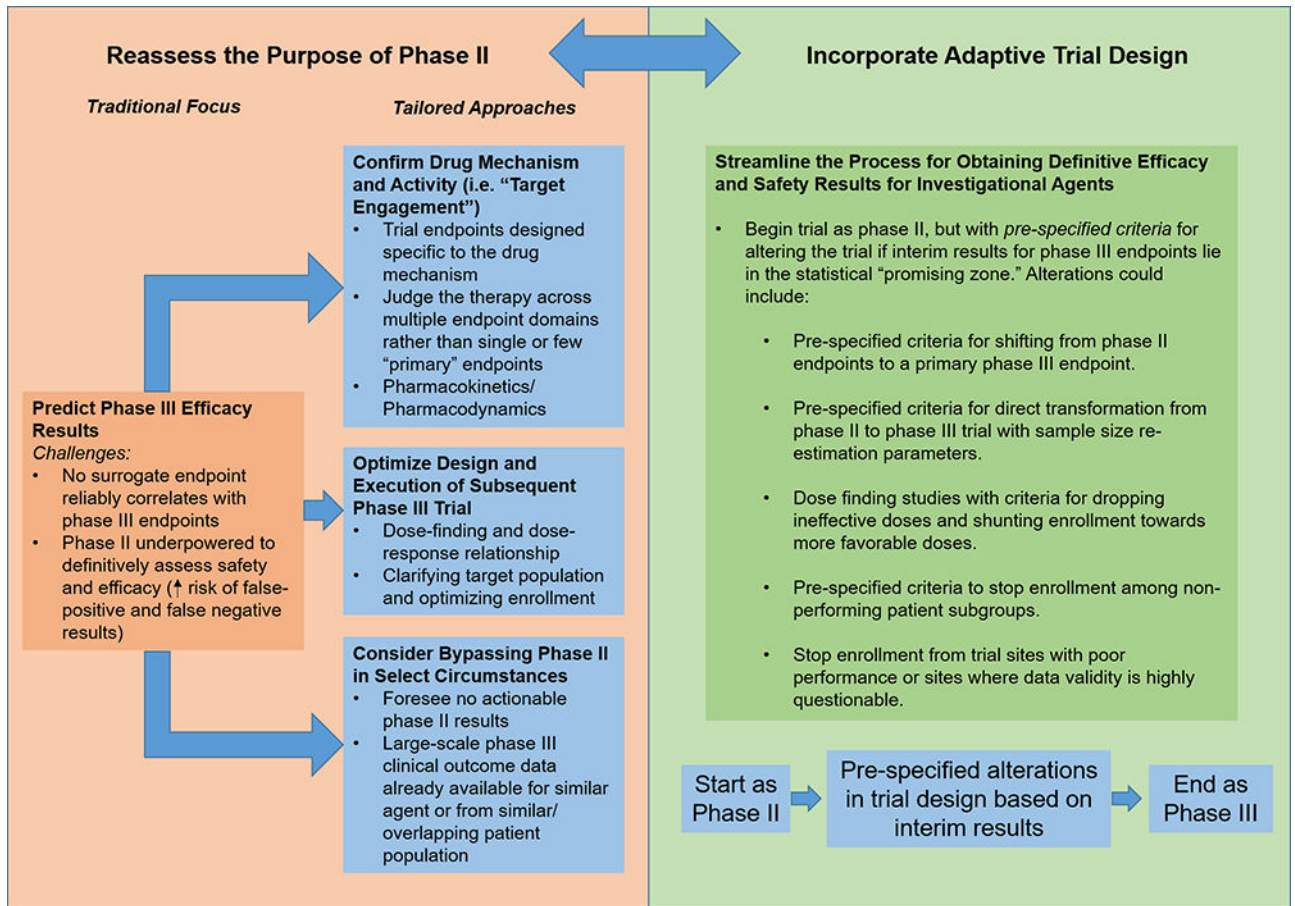
Data presented as geometric mean +/- 95% confidence interval  
 BL=baseline; NT-proBNP=N-terminal pro-B-type natriuretic peptide

**Figure 2.** Cumulative event rates for cardiovascular death or heart failure re-hospitalization within 12 months (A) and the change in NT-proBNP over follow-up (B) from the ASTRONAUT trial. Adapted from Gheorghide M et al. Late Breaking Clinical Trial Presentation at ACC 2013 Meeting, March 11, 2013. <sup>23</sup>



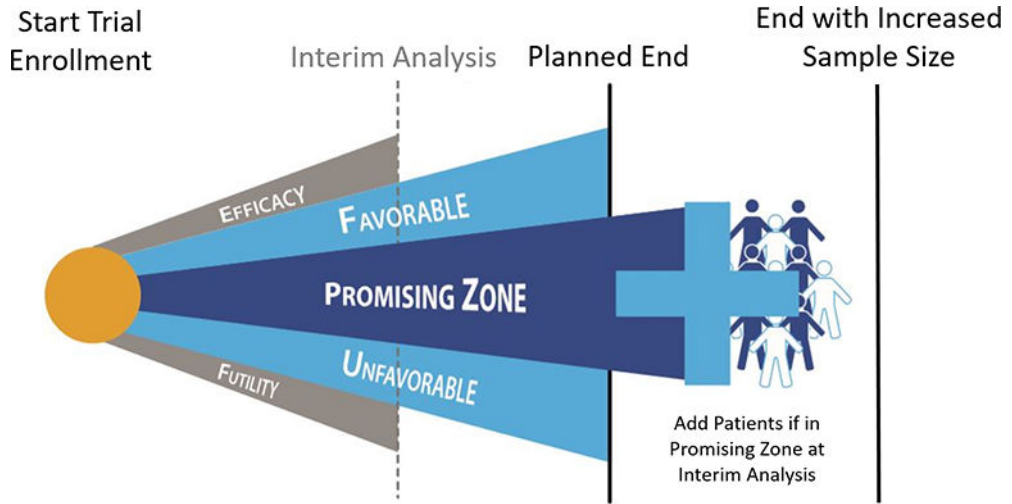
**Figure 3.** Influence of aliskiren on longitudinal NT-proBNP level in the ASTRONAUT trial by atrial fibrillation/flutter status. The y-axis represents the estimated NT-proBNP level in pg/mL, derived from exponentiation of the log-transformed NT-proBNP level at each time point. NTproBNP, N-terminal pro-B-type natriuretic peptide. Adapted with permission from Greene SJ et al.<sup>30</sup>





**Figure 4.**

Potential framework for re-appraising the future design, purpose, and execution of phase II studies in HF drug development. Recognizing that there are no reliable surrogate outcomes for phase III approval endpoints, we propose that the central goal of phase II should not be to predict phase III results but to optimize phase III execution and to clarify the drug mechanism and drug-patient interaction. Adaptive trial design may allow for a more efficient drug development process and may more directly allow information gained from phase II to shape the phase III trial.



**Figure 5.** Conceptual framework of adaptive trial design. Trial would start with a planned end. Interim analysis in the statistical “promising” zone would prompt sample size re-estimation with addition of further patients and extension of the trial beyond the planned end, increasing chances of trial finding definitive result. Interim analysis in the “favorable” or “unfavorable” zone would not prompt extension of the trial, as the initial planned end would have reasonable probability of arriving at a definitive result. Interim analysis in the “efficacy” or “futility” zones would prompt the trial being stopped early before the planned end.

**Table 1.**

## Requirements for a “Reliable” Surrogate Endpoint for Clinical Outcomes

• The measure must correlate with risk of the outcome in patients without a given medical intervention.
• The measure must continue to correlate with risk of the outcome after modulation with a given therapy.
• Therapeutically modulated levels of the measure must predict the net effect of the treatment on the clinical outcome.
• The correlation between therapeutically modulated levels of the measure and risk of outcome must be consistent across various classes of interventions and population subsets.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.****Challenges Specific to Common Surrogate Endpoints Used in Early Phase Heart Failure Trials**

<p><b><i>Invasive Hemodynamics</i></b></p> <hr/> <ul style="list-style-type: none"> <li>• Multiple examples of dissociation between hemodynamic improvement (e.g., improved cardiac index) and subsequent worse clinical outcomes (e.g., flosequinan, milrinone).</li> <li>• Cardiac index and ventricular filling pressures can change acutely with changes in loading conditions and contractility and may not necessarily represent intrinsic cardiac recovery.</li> <li>• Current guidelines discourage routine PAC placement. Common indication for PAC is cardiogenic shock but most HF trials require patients to be normotensive or hypertensive, thus making trial enrollment challenging.</li> </ul> <hr/>
<p><b><i>Change in Natriuretic Peptide (NP) Level</i></b></p> <hr/> <ul style="list-style-type: none"> <li>• Multiple examples of dissociation between treatment-related NP change and clinical outcomes (e.g., levosimendan, aliskiren).</li> <li>• Comorbid atrial fibrillation may influence NP levels and hinder the ability of study therapy to decrease NP levels in HFpEF. The ability of a study therapy to meet an NP-defined endpoint may be influenced by the prevalence of AF in the overall trial population.</li> <li>• A significant minority of HFpEF patients have normal NP levels. Other HFpEF patients may have only minimal baseline NP elevation and it is unclear to what degree this can be modified with therapy.</li> <li>• Biological and analytical variability must be recognized in interpretation of serial NP measurements.</li> <li>• Heterogeneity in how NP endpoints defined: <ul style="list-style-type: none"> <li>○ Intention-to-treat principle versus per-protocol/ 'completers' approach</li> <li>○ Continuous absolute change versus categorical endpoint defined by threshold level of relative reduction from baseline (i.e., &gt;30%)</li> </ul> </li> </ul> <hr/>
<p><b><i>Change in Troponin Level</i></b></p> <hr/> <ul style="list-style-type: none"> <li>• Limited understanding of the mechanisms of troponin elevation in HF</li> <li>• Biological and analytical variability must be recognized in interpretation of serial troponin measurements (although may be less of a challenge compared to NPs).</li> </ul> <hr/>
<p><b><i>Cardiac Imaging</i></b></p> <hr/> <ul style="list-style-type: none"> <li>• Ejection fraction and ventricular volumes can change acutely with changes in loading conditions and contractility and may not necessarily represent intrinsic cardiac recovery.</li> <li>• Relationship between changes in diastolic function with therapy and subsequent clinical outcomes is poorly established in HFpEF.</li> <li>• 2D echocardiography may present challenges with inter-reader variability, reproducibility, and image acquisition.</li> <li>• CMR and PET imaging are available only at select specialized centers and at relatively high cost.</li> <li>• Use of CMR limits enrollment of patients with implantable cardiac devices</li> </ul> <hr/>
<p><b><i>Functional Capacity (e.g, peak VO<sub>2</sub>, 6MWD)</i></b></p> <hr/> <ul style="list-style-type: none"> <li>• Multiple examples of dissociation between change in exercise capacity and clinical outcomes (e.g., flosequinan, ramipril, carvedilol).</li> <li>• Trial protocols subject to accentuated placebo effect whereby exercise capacity improves with familiarization with exercise protocol.</li> <li>• Potential for uneven intensification of background therapy between study arms during follow-up to maintain symptomatic stability.</li> <li>• 6MWD may suffer from modest reproducibility and variability in clinician "coaching" and patient effort from one visit to the next and between study sites.</li> <li>• CPET testing only available at select centers and burden of serial intensive exercise examinations on patients may be a challenge for patient recruitment.</li> </ul> <hr/>
<p><b><i>Quality of Life, Dyspnea Relief, and Other Patient-reported Outcomes</i></b></p> <hr/> <ul style="list-style-type: none"> <li>• Multiple examples of dissociation between change in quality of life and clinical outcomes (e.g., flosequinan, spironolactone).</li> <li>• Potential for uneven intensification of background therapy between study arms during follow-up to maintain symptomatic stability.</li> </ul>

- Multiple grading instruments available and often no clear consensus on which is the “gold standard.” Examples exist where a therapy (e.g., serelaxin) may improve a patient-reported outcome (e.g., dyspnea relief) as measured by one grading instrument but not another.
- Majority of acute HF patients experience rapid (e.g., within hours) and robust dyspnea relief with standard therapy. Dyspnea may no longer be severe by the time of patient enrollment and it may be difficult for investigational therapies to show incremental improvement in dyspnea relief over standard therapy.

6MD, 6-minute walk distance; CMR, cardiac magnetic resonance; HF, heart failure; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; NP, natriuretic peptide; PAC, pulmonary arterial catheter; PET, positron emission tomography

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript