Title
With great power comes great… reliability.

Journal
European journal of heart failure, 22(9)

ISSN
1388-9842

Authors
Ziaeian, Boback
Butler, Javed
Fonarow, Gregg C

Publication Date
2020-09-01

DOI
10.1002/ejhf.1816

Peer reviewed

ESC European Society of Cardiology

European Journal of Heart Failure (2020)
doi:10.1002/ejhf.1816

EDITORIAL COMMENT

With great power comes great … reliability

Boback Ziaeian[1,2], Javed Butler[3], and Gregg C. Fonarow[2]*

[1]Division of Cardiology, David Geffen School of Medicine at University of California, Los Angeles, CA, USA; [2]Division of Cardiology, Veteran Affairs Greater Los Angeles Healthcare System, Los Angeles, CA, USA; and [3]Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA

This article refers to 'Mega-trials in heart failure: effects of dilution in examination of new therapies' by B.A. Davison *et al.*, published in this issue on pages xxx.

The sheer number of medical ailments that afflict humans far surpasses our ability to effectively alter the natural course of most diseases. Yet, modern evidence-based medicine has catapulted remarkable gains for both the prevention and chronic management of cardiovascular diseases. The recognition of smoking, hypertension, diabetes, and dyslipidaemia as strong modifiable atherosclerotic risk factors has slowly decreased rates of cardiovascular diseases and in turn have contributed to the overall improvement in population health. Nowhere is this more evident than in the management of chronic heart failure with reduced ejection fraction (HFrEF), where we now have over a dozen therapeutic options to improve quality of life and reduce morbidity and mortality[1] (*Figure 1*).

While a number of medical therapies – even after potentially promising signals in Phase II studies – failed to demonstrate benefits in large-scale trials in HFrEF, the benefits demonstrated in successful trials for chronic HFrEF have been reproduced in subsequent trials using other agents in the same pharmacologic class. These evidence-based medical therapies provide sequential incremental clinical benefits without noted heterogeneity across a multitude of subgroups of patients based on demographics, comorbidities, or severity of disease, enrolled in the landmark trials.[1,2] Treatment benefits stratified by severity estimated using validated risk scores report relative risk reductions of similar magnitude for treatment compared to placebo without important heterogeneity.[3,4] Although clinical trials may enrol a selective HFrEF patient population than those in routine clinical practice, similar relative risk reductions are observed in clinical effectiveness studies of registry populations with guideline-directed medical therapies, despite higher baseline risk and comorbidity burdens.[5]

In contrast to the success in demonstrating benefits with multiple medications for HFrEF, there has been consistent failure to identify therapies which improve outcomes for patients hospitalized with acute heart failure (AHF) or heart failure with preserved ejection fraction (HFpEF). Several theories have been put forth to explain these trends. For AHF, some of the early failures were attributed to initiating study medications too late after initial presentation, allowing patients with lower blood pressure to be enrolled, not identifying optimal doses of study medications, or using unreliable survey instruments to measure changes in symptoms. Subsequent trials attempting to target more defined patient populations and administer treatment soon after presentation were also unsuccessful. Questions have been raised as to whether AHF represents a distinct entity that requires acute intervention beyond intravenous diuretics and whether any short-term infusion of medication could meaningful improve clinical outcomes.[6] In this respect, it is interesting to note that initiating chronic guideline-directed medical therapy sacubitril/valsartan in AHF patients with HFrEF did result in improved post-discharge outcomes. For HFpEF trials, the lack of success has been attributed to the incomplete understanding of the pathophysiology of the disease state, marked heterogeneity among patients, inadequate phenotyping, unreliable surrogate markers, trials which may have enrolled patients that did not actually have HFpEF, as well as having not yet tested therapies with sufficient efficacy to improve outcomes.

In this issue of the Journal, Davison *et al.*[7] theorize that larger and larger trial sizes have led to the 'dilution of statistical power.' Furthermore, the lack of success of medical treatments for HFpEF and AHF may be secondary to these trial design decisions. They then use study simulations to examine the probability that a positive Phase II finding would be detected in confirmatory Phase III trials with their concern being that true causal effects of interventions are not being detected in confirmatory trials.

The first hypothesis of larger sample sizes diluting perceivable treatment effects will baffle members of the American Statistical Association. Statistical power is increased with larger sample sizes and small effects not meeting the minimal clinically important difference are more easily detected. How can it be justified to turn the concept of statistical power on its head? Davison *et al.* seem to suggest that the inclusion of heterogeneous patient populations may have questionable HFpEF or AHF diagnoses. Their simulation, however, does not answer or address this hypothesis. Whether

The opinions expressed in this article are not necessarily those of the Editors of the *European Journal of Heart Failure* or of the European Society of Cardiology. doi: 10.1002/ejhf.1780

*Corresponding author. Ahmanson–UCLA Cardiomyopathy Center, Ronald Reagan UCLA Medical Center, 10833 LeConte Avenue, Room A2-237 CHS, Los Angeles, CA 90095-1679, USA. Tel: +1 310 206-9112, Fax: +1 310 206-9111, Email: gfonarow@mednet.ucla.edu

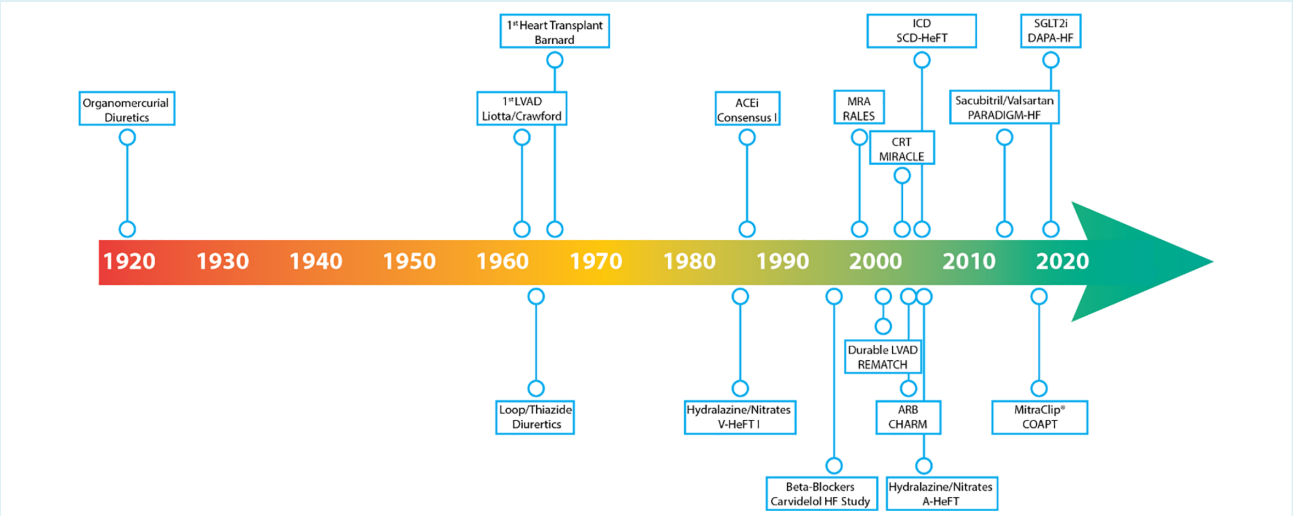

Figure 1 Timeline of medical and surgical interventions for heart failure with reduced ejection fraction. Organomercurial diuretics were used up until the 1960s when toxicities were appreciated. Furosemide was approved in 1964. The first left ventricular assist device (LVAD) was implanted by Dr. Domingo Liotta and Dr. Stanley Crawford in 1963. The first heart transplant was performed by Dr. Christiaan Barnard in 1967. The Vasodilator-Heart Failure I (V-HeFT) trial was published in 1986 using hydralazine–isosorbide dinitrate. The CONSENSUS I trial using the angiotensin-converting enzyme inhibitor (ACEi) enalapril was published 1987. The Carvedilol Prospective Randomized Cumulative Survival trial was published in 1996. The Randomized Aldactone Evaluation Study (RALES) was published in 1999. The Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH) trial was published in 2001. The Multicenter InSync Randomized Clinical Evaluation (MIRACLE) trial was published in 2002. The Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM-Alternative) trial was published in 2003. The African-American Heart Failure Trial (A-HeFT) was published in 2004. The Sudden Cardiac Death in Heart Failure Trial (SCD-HeFT) was published in 2005. The Prospective Comparison of ARNI (angiotensin receptor–neprilysin inhibitor) with ACEI (angiotensin-converting enzyme inhibitor) to Determine Impact on Global Mortality and Morbidity in Heart Failure (PARADIGM-HF) trial was published in 2014. The Cardiovascular Outcomes Assessment of the MitraClip Percutaneous Therapy for Heart Failure Patients with Functional Mitral Regurgitation (COAPT) trial was published in 2018. The Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure (DAPA-HF) trial was published in 2019. CRT, cardiac resynchronization therapy; ICD, implantable cardioverter-defibrillator; MRA, mineralocorticoid receptor antagonist; SGLT2i, sodium–glucose cotransporter 2 inhibitor.

there exist meaningful subgroups of HFpEF or AHF patients that would benefit from existing drug treatments is unknown. Diagnostic criteria for teasing out different HFpEF phenotypes are yet to be rigorously defined or tested. Clinical trials are yet to divide HFpEF populations in distinct subpopulations. A significant proportion of cardiac amyloidosis patients may have mixed into prior HFpEF and HFrEF trials.[8] Some studies have suggested subgroups of HFpEF patients that may benefit from select therapies, for example women with HFpEF in the Prospective Comparison of ARNI (angiotensin receptor–neprilysin inhibitor) with ARB (angiotensin receptor blockers) Global Outcomes in Heart Failure with Preserved Ejection Fraction (PARAGON-HF) trial may benefit more from sacubitril/valsartan when compared to men.[9] Secondary analyses are hypothesis generating and prone to multiple testing errors; any such hypothesis will require appropriately designed prospective trials.

Phase II trials have a high failure rate with only 30% advancing to Phase III.[10] Surrogate markers (i.e. biomarkers, natriuretic peptides, cardiac function) utilized in Phase II trials do not necessarily translate directly to clinical outcomes. As more and more markers are checked in a Phase II trial to provide hope of effectiveness, statistical considerations become increasingly challenging. Depending on small samples using weak outcomes will result in misdirected resources on no or low-value therapies along with false hope for the potential of novel treatments to improve outcomes for patients with HFpEF or AHF.

Ultimately, a treatment that works … works! The first cardiovascular trials of angiotensin-converting enzyme inhibitors in HFrEF compared to usual care demonstrated marked improvements in mortality with a modest sample size.[11] Each guideline-directed medical therapy for HFrEF has continued to show the same incremental progress in clinical trials. Slowing and reversing HFpEF is a challenge that has remained elusive despite many randomized trials. Myocardial fibrosis is prevalent and underlies much of the observed diastolic dysfunction in HFpEF. The potential to reverse or remodel progressive fibrosis in these patients may not be infeasible and may be the ultimate reason for recurrent failure of Phase III trials, and not inclusion/exclusion criteria or large clinical trial designs. Whether sodium–glucose cotransporter 2 inhibitors will improve HFpEF outcomes, while promising, remains to be demonstrated. Science is not linear, and neither is drug discovery. We may not expect successes seen in HFrEF at the turn of the century to necessarily translate to efficacious therapies for patients with HFpEF or AHF.

The efficiency of our clinical trial enterprise remains of great concern. The cost of undertaking clinical trials has grown exponentially.[12] Three fundamental approaches that are under-utilized in modern trials to reduce costs and improve inference are (i) baseline adjustment for covariates 'causal' on the outcomes of interest, (ii) adaptive Bayesian estimation, and (iii) pragmatic trials.[13–15] Baseline covariate adjustment requires inclusion of select factors determined *a priori*. Statistical models adjust for known risk factors along with randomized treatment. This approach reduces the statistical noise around estimating treatment effects and reduces sample sizes. Ultimately, it provides more efficient estimation of the 'true treatment effect.' Adaptive Bayesian trials allow for flexible recruitment of patients during randomization. The traditional frequentist paradigm of trial design requires a power calculation for estimated event rates and treatment effect sizes. These are frequently gamed to meet funding feasibility limits. An adaptive Bayesian trial allows for stoppage early when treatment is clearly futile, early stoppage for beneficial treatments, and continued recruitment for potentially successful interventions. Statistical pre-specification of the Bayesian estimation procedures allows for interval treatment effect measurement without contributing to a false positive or negative error rate. Pragmatistic trials are designed to use existing electronic health records or registry systems to serve as the backbone of follow-up data and event adjudication. Patients may be randomized to a treatment at an initial encounter with no further contacts by research staff. While cost savings, concerns regarding accurate event capture persist along with the loss of statistical efficiency and potential biases towards the null that may be introduced without monitoring treatment adherence.

Despite many advocates, these improved statistical methods and study designs are largely excluded from most published clinical trial protocols. Reversing our approach to the evaluation of novel therapies with smaller, haphazard clinical trials is not the answer to scientific and clinical advancement. Patients, clinicians, and regulators depend on clear demonstrations of the effectiveness of treatments through well-designed randomized clinical trials. It is not about bigger being better; it is about powering randomized clinical trials for an expected clinically relevant effect size to yield findings we will trust.

Funding

Boback Ziaeian's research is supported by American Heart Association (AHA) SDG 17SDG33630113 and the National Institutes of Health (NIH)/National Center for Advancing Translational Science (NCATS) UCLA CTSI Grant Number KL2TR001882. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Conflict of interest: B.Z.: none. J.B. consultant for Abbott, Amgen, Applied Therapeutics, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, CVRx, Janssen, LivaNova, Luitpold, Medtronic, Merck, Novartis, Relypsa, and Vifor. G.C.F.: research: NIH; consulting: Abbott, Amgen, AstraZeneca, Bayer, CHF Solutions, Janssen, Medtronic, Merck, and Novartis.

References

1. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Colvin MM, Drazner MH, Filippatos GS, Fonarow GC, Givertz MM, Hollenberg SM, Lindenfeld J, Masoudi FA, Mcbride PE, Peterson PN, Stevenson LW, Westlake C. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *J Am Coll Cardiol* 2017;**70**:776–803.
2. Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, Palensky J, Wittes J. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *N Engl J Med* 1999;**341**:709–717.
3. Collier TJ, Pocock SJ, McMurray JJ, Zannad F, Krum H, Van Veldhuisen DJ, Swedberg K, Shi H, Vincent J, Pitt B. The impact of eplerenone at different levels of risk in patients with systolic heart failure and mild symptoms: insight from a novel risk score for prognosis derived from the EMPHASIS-HF trial. *Eur Heart J* 2013;**34**:2823–2829.
4. Simpson J, Jhund PS, Silva Cardoso J, Martinez F, Mosterd A, Ramires F, Rizkala AR, Senni M, Squire I, Gong J, Lefkowitz MP, Shi VC, Desai AS, Rouleau JL, Swedberg K, Zile MR, McMurray JJ, Packer M, Solomon SD; PARADIGM-HF Investigators and Committees. Comparing LCZ696 with enalapril according to baseline risk using the MAGGIC and EMPHASIS-HF risk scores: an analysis of mortality and morbidity in PARADIGM-HF. *J Am Coll Cardiol* 2015;**66**:2059–2071.
5. DeVore AD, Mi X, Thomas L, Sharma PP, Albert NM, Butler J, Hernandez AF, Patterson JH, Spertus JA, Williams FB, Duffy CI, McCague K, Fonarow GC. Characteristics and treatments of patients enrolled in the CHAMP-HF registry compared with patients enrolled in the PARADIGM-HF trial. *J Am Heart Assoc* 2018;**7**:e009237.
6. Packer M. Why are physicians so confused about acute heart failure? *N Engl J Med* 2019;**381**:776–777.
7. Davison B, Tokagi K, Senger S, Koch G, Metra M, Kimmoun A, Mebazaa A, Voors AA, Nielsen OW, Chioncel O, Pang PS, Greenberg BH, Maggioni AP, Cohen-Solal A, Ertl G, Sato N, Teerlink JR, Filippatos G, Ponikowski P, Gayat E, Edwards C, Cotter G. Mega-trials in heart failure: effects of dilution in examination of new therapies. *Eur J Heart Fail* 2020 Mar 30. https://doi.org/10.1002/ejhf.1780 [Epub ahead of print].
8. Manolis AS, Manolis AA, Manolis TA, Melita H. Cardiac amyloidosis: an underdiagnosed/underappreciated disease. *Eur J Intern Med* 2019;**67**:1–13.
9. JJ MM, Jackson AM, Lam CS, Redfield MM, Anand IS, Ge J, Lefkowitz MP, Maggioni AP, Martinez F, Packer M, Pfeffer MA, Pieske B, Rizkala AR, Sabarwal SV, Shah AM, Shah SJ, Shi VC, van Veldhuisen DJ, Zannad F, Zile MR, Cikes M, Goncalvesova E, Katova T, Kosztin A, Lelonek M, Sweitzer N, Vardeny O, Claggett B, Jhund PS, Solomon SD. Effects of sacubitril-valsartan versus valsartan in women compared with men with heart failure and preserved ejection fraction: insights from PARAGON-HF. *Circulation* 2020;**141**:338–351.
10. van Norman GA. Phase II trials in drug development and adaptive trial design. *JACC Basic Transl Sci* 2019;**4**:428–437.
11. The 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.
12. Lauer MS, Gordon D, Wei G, Pearson G. Efficient design of clinical trials and epidemiological research: is it possible? *Nat Rev Cardiol* 2017;**14**:493–501.
13. Steyerberg EW, Bossuyt PM, Lee KL. Clinical trials in acute myocardial infarction: should we adjust for baseline characteristics? *Am Heart J* 2000;**139**:745–751.
14. Ford I, Norrie J. Pragmatic trials. *N Engl J Med* 2016;**375**:454–463.
15. Bhatt DL, Mehta C. Adaptive designs for clinical trials. *N Engl J Med* 2016;**375**:65–74.