Mega-trials in heart failure

Davison, Beth A.; Takagi, Koji; Senger, Stefanie; Koch, Gary; Metra, Marco; Kimmoun, Antoine; Mebazaa, Alexandre; Voors, Adriaan A.; Nielsen, Olav W.; Chioncel, Ovidiu

Published in:
European Journal of Heart Failure

DOI:
10.1002/ejhf.1780

ESC European Society of Cardiology

European Journal of Heart Failure (2020)
doi:10.1002/ejhf.1780

RESEARCH ARTICL

Mega-trials in heart failure: effects of dilution in examination of new therapies

Beth A. Davison[1], Koji Takagi[2], Stefanie Senger[1], Gary Koch[3], Marco Metra[4], Antoine Kimmoun[5], Alexandre Mebazaa[2], Adriaan A. Voors[6], Olav W. Nielsen[7], Ovidiu Chioncel[8], Peter S. Pang[9], Barry H. Greenberg[10], Aldo P. Maggioni[11], Alain Cohen-Solal[12], Georg Ertl[13], Naoki Sato[14], John R. Teerlink[15], Gerasimos Filippatos[16], Piotr Ponikowski[17], Etienne Gayat[18], Christopher Edwards[1], and Gad Cotter[1]*

[1]Momentum Research, Inc., Durham, NC, USA; [2]Inserm 942-MASCOT, Department of Anaesthesiology and Critical Care Medicine, Assistance Publique – Hôpitaux de Paris, Saint Louis Lariboisière University Hospitals, University of Paris, Paris, France; [3]Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA; [4]Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, University of Brescia, Brescia, Italy; [5]Department of Intensive Medicine and Resuscitation Brabois, Regional University Hospitals of Nancy, University of Lorraine, Vandoeuvre les Nancy, France; [6]Department of Cardiology, University Medical Centre Groningen, Groningen, The Netherlands; [7]Department of Cardiology, Bispebjerg Hospital, University of Copenhagen, Copenhagen, Denmark; [8]Emergency Institute for Cardiovascular Diseases 'Prof. C.C. Iliescu', University of Medicine Carol Davila, Bucharest, Romania; [9]Department of Emergency Medicine, Indiana University School of Medicine, Indianapolis, IN, USA; [10]Division of Cardiology, University of California San Diego, San Diego, CA, USA; [11]Department of Internal Medicine, Associazione Nazionale Medici Cardiologi Ospedalien (ANMCO) Research Centre, Florence, Italy; [12]Cardiology Service, Lariboisière Hospital, Paris, France; [13]Julius-Maximilians University Würzburg, Wurzburg, Germany; [14]Cardiology and Intensive Care Unit, Nippon Medical School, Musashi-Kosugi Hospital, Kawasaki, Japan; [15]Section of Cardiology, San Francisco Veterans Affairs Medical Center and School of Medicine, University of California San Francisco, San Francisco, CA, USA; [16]Department of Cardiology, Heart Failure Unit, Athens University Hospital Attikon, Athens, Greece; [17]Department of Heart Diseases, Medical University, Military Hospital, Wroclaw, Poland; and [18]Faculty of Medicine Xavier Bichat, Paris Diderot University, Paris, France

Received 5 December 2019; revised 13 January 2020; accepted 7 February 2020

Aims Over the last 30 years, many medicine development programmes in acute and chronic heart failure (HF) with preserved ejection fraction (HFpEF) have failed, in contrast to those in HF with reduced ejection fraction (HFrEF). We explore how the neutral results in larger HF trials may be attributable to chance and/or the dilution of statistical power.

Methods and results Using simulations, we examined the probability that a positive finding in a Phase 2 trial would result in the study of a truly effective medicine in a Phase 3 trial. We assessed the similarity of clinical trial and registry patient populations. We conducted a meta-analysis of paired Phase 2 and 3 trials in HFrEF and acute HF examining the associations of trial phase and size with placebo event rates and treatment effects for HF events and death. We estimated loss in trial power attributable to dilution with increasing trial size. Appropriately powered Phase 3 trials should have yielded ~35% positive results. Patient populations in Phase 3 trials are similar to those in Phase 2 trials but both differ substantially from the populations of 'real-life' registries. We observed decreasing placebo event rates and smaller treatment effects with increasing trial size, especially for HF events (and less so for mortality). This was more pronounced in trials in acute HF patients.

Conclusions The selection of more positive Phase 2 trials for further development does not explain the failure of HFpEF and acute HF medicine development. Increasing sample size may lead to reduced event rates and smaller treatment effects, resulting in a high rate of neutral Phase 3 trials.

Keywords Acute heart failure • Chronic heart failure • Meta-analysis

*Corresponding author. Momentum Research, Inc., 807 East Main Street, Building 6, Suite 6-050, Durham, NC 27701, USA. Tel: +1 919 287 1824, Email: gadcotter@momentum-research.com

Introduction

Despite more than 30 years of attempts involving at least 15 development programmes, no new pharmacologic interventions have been shown to improve outcomes in trials in patients with either acute heart failure (AHF)[1–14] or heart failure (HF) with preserved ejection fraction (HFpEF).[15–23] In contrast, several new interventions have been shown to reduce morbidity and mortality in HF with reduced ejection fraction (HFrEF).[24–41] This discordance has occurred despite similar trial leadership, pharmaceutical industry and contract research organizations, and methods of transitioning from Phase 1 to Phase 2 and Phase 3.

Two main hypotheses have been put forward to explain the failure of new interventions. The first is 'random high bias', which means that ineffective interventions are carried forward to Phase 3 trials as a result of chance findings in Phase 2. In this case, these chance findings are then exposed in large Phase 3 trials, which, simply by their sample size, better represent a 'real-life' patient population, in which the medicines are found to be ineffective. A second argument posits that heterogeneity (in patients enrolled and subsequent treatment effects) inevitably results from a larger sample size in disease states in which the diagnosis is subjective. This argument would suggest that in some disease states (namely, HFrEF), in which the disease is clearly defined, large trials may increase our ability to detect smaller treatment effects, but this may not be the case in disease states such as AHF and HFpEF, in which the disease is less well defined. In this case, larger (and, hence, often less well-controlled) trials allow the enrolment of patients with different disease phenotypes or without the disease in question and in whom an intervention may not be effective.

Methods

Simulations to assess the role of chance findings in Phase 2 programmes

The probability that a favourable result in Phase 2 reflects a truly favourable treatment effect on a dichotomous outcome was examined through simulations. We simulated 10 000 trials, in which 300 patients were equally allocated to either an active or a control group (with a control event rate of 10%), for each of six scenarios assuming different distributions of true relative risks (RRs) for the interventions studied. The RRs assumed in all scenarios ranged from 0.5 (50% reduction in event rate with active treatment) to 1.15 (15% increase in event rate with active treatment) in 0.05 increments. The distribution of true RRs varied from pessimistic in Scenario 1 (most tested medicines are truly ineffective) to optimistic (many medicines are truly effective). Thus, for each simulated trial we know the true RR and we have an observed RR (resulting from chance variation around the true RR), and for each scenario the distribution of trials is shifted toward either fewer effective or more effective tested interventions. We also simulated a continuous surrogate outcome for each trial in which the mean difference between treatment groups was perfectly correlated with the true RR such that no effect on the dichotomous outcome (RR = 1) corresponded to no effect on the continuous variable, an RR of 0.5 corresponded to a mean difference of 1 standard deviation (SD) (both large beneficial treatment effects) and an RR of 1.15 corresponded to a mean difference of −0.3 SD (both somewhat harmful treatment effects). R was used for these simulations.[42] Details of the simulations are presented in the supplementary material online (Tables S1–S5 and Figures S1–S2). We considered an RR of 0.8 (a 20% reduction in the event rate with active treatment) to be clinically relevant.

Assessing the similarity of clinical trials and registries

We reviewed all Phase 2 and 3 trials of medicines for the treatment of AHF for the main baseline characteristics of the patients enrolled. We looked at trials that had no ejection fraction eligibility restriction and were conducted in the last 25 years.[1–6,43] Overall proportions are presented for categorical variables; overall means for continuous variables were calculated by weighting each trial's mean by the trial's sample size. These overall characteristics were then compared with the characteristics of patients enrolled in AHF registries.[44–52]

Assessing placebo event rates and treatment effects in chronic and acute heart failure trials

We searched both PubMed and ClinicalTrials.gov for clinical trials or randomized studies of medicines in humans with HF reported between January 1988 and December 2018, with clinical trials as a topic, excluding reviews (online supplementary *Figure S3*). Only articles published in English were considered. Trials in AHF secondary to myocardial infarction or surgery were excluded. Trials of a given medicine were included in the analyses if we found at least one pair of Phase 2 and Phase 3 trials both published before 1999, with a total sample size of at least 100 patients in the Phase 2 trial, and if both the Phase 2 and 3 trials reported placebo event rates and treatment effects (or reported the number of events from which placebo event rates and treatment effects could be estimated) for the outcomes of interest: all-cause mortality over 180 days, HF hospitalization over 180 days (in chronic HF trials), and worsening HF (WHF) (in trials in AHF). If multiple doses of a given medicine were tested, reported or estimated placebo event rates and treatment effects for the dose studied in Phases 2 and 3 were included. In studies in patients with chronic HF, only studies that referred to HFrEF were included.

For the estimation of 180-day mortality rates in placebo arms, only AHF studies with follow-up periods of at least 90 days were included (data were extrapolated to 180 days if necessary). If Kaplan–Meier event rates for all-cause mortality or for HF hospitalization were not reported for a 180-day observation period, 180-day placebo event rates were derived as follows, in hierarchical order and based on availability in each publication: (i) the Kaplan–Meier event rate reported at a given time-point was extrapolated to 180 days; (ii) an event rate at 180 days was estimated from a Kaplan–Meier plot; (iii) an event rate estimated from the reported number of events and the reported median or mean follow-up time was extrapolated to 180 days; and (iv) the reported crude rate at a given fixed follow-up time-point was extrapolated to 180 days.

With reference to placebo event rates for WHF, the reported observation periods ranged from 2 to 7 days after baseline, and Kaplan–Meier event rates or crude rates were used for analysis as reported. Estimated event counts (number of events) at the respective analysis time-point were derived from the estimated event rates as defined above.

The hazard ratio (HR) was used as the estimated treatment effect of active medicine vs. placebo and was assumed to be constant over the

Table 1 Results of the simulations under six scenarios of true relative risk (RR) distributions for equal allocations of 150 patients per group, a 10% control event rate, and a correlation of 0.5 between a continuous surrogate outcome and the true RR

	Scenario					
	1	2	3	4	5	6
True RR <0.8	6.0%	10.0%	20.0%	27.0%	36.0%	42.0%
Proceed to Phase 3: observed RR <0.8	28.5%	32.7%	36.9%	40.3%	45.1%	48.9%
Proceed to Phase 3: RR 95% UCL <1.0[a]	3.3%	4.3%	6.0%	7.1%	9.2%	10.4%
Proceed to Phase 3: one-sided $P < 0.025$ for continuous surrogate[b]	20.4%	31.3%	39.5%	47.0%	54.6%	60.5%
True RR <0.8 in those trials with observed RR <0.8	13.3%	19.8%	37.3%	45.4%	56.4%	61.1%
True RR <0.8 in those trials with RR 95% UCL <1.0	26.5%	35.9%	59.3%	66.8%	76.2%	79.6%
True RR <0.8 in those trials with one-sided $P < 0.025$ for continuous surrogate	29.4%	31.8%	50.5%	57.4%	65.8%	69.4%
Mean bias in those trials with observed RR <0.8 (observed RR−true RR)	−0.30	−0.27	−0.23	−0.21	−0.18	−0.17
Mean bias in those trials with RR 95% UCL <1.0 (observed RR−true RR)	−0.48	−0.44	−0.37	−0.35	−0.32	−0.31
Mean bias in those trials with one-sided $P < 0.025$ for continuous surrogate (observed RR−true RR)	0.00	0.01	0.02	0.03	0.03	0.03

[a]Upper 95% confidence limit for an RR of <1.0 for the clinical event (i.e. the event rate was statistically significantly reduced).
[b]One-sided $P < 0.025$ for test of continuous, surrogate outcome. The true standardized mean treatment difference (SMD) is a function of the true RR [SMD = 2 × (1-RR)]. An SMD of 0 and an RR of 1.0 represent no treatment effect.
UCL, upper confidence limit.

reported follow-up period. If an HR was not reported, it was estimated using the reported RR or the RR calculated based on the estimated event counts in the two groups.

Data were extracted by one analyst (KT) and verified by a second analyst (SS).

The potential modifying effects of trial phase and control group size on control event rates were examined using a random effects weighted least squares meta-regression with inverse variance weighting.[53] Non-linearity of the associations with trial size was assessed through the significance of a quadratic term in each of the regression models. The potential modifying effects of trial phase and size on the treatment effects on all-cause mortality, HF hospitalization and WHF were examined using similar meta-regressions. The effective sample size for trials with unequal allocation was estimated as the size of a trial with equal allocation that would have provided the same power.[54]

Power for diluted treatment effects

Power at the two-sided significance level of 0.05 was calculated using the POWER procedure in SAS Version 9.4 (SAS Institute, Inc., Cary, NC, USA). For the general case, HRs of 0.65 and 0.85 with placebo event rates of 10% and 20% at $n = 100$ were diluted as a linear function of sample size and power calculated for various dilutions of the HR. Dilution was computed as 1 minus the ratio of the distances of the HR from 1 with 6500 vs. 100 total patients, so that, for example, going from an HR of 0.5 with $n = 100$ to an HR of 0.75 with $n = 6500$ corresponds to a dilution of 50%.

Results

Random high bias or regression to the mean

In the 10 000 simulated trials for each of the six scenarios, the proportion of trials with a true RR of <0.8 (i.e. studying a truly effective drug), varied between 6.0% and 42.0% (*Table 1*). The probability that a 'positive' Phase 2 trial reflects a truly effective therapy (i.e. the true RR is <0.8 and would thus be detected in an appropriately powered Phase 3 trial) is highest when the criterion for Phase 2 'positivity' is an observed statistically significant RR (RR 95% upper confidence limit <1.0); the proportion of positive Phase 2 trials thus defined in which the therapy was truly effective ranged from 26.5% to 79.6% across the six scenarios. However, clearly Phase 2 studies do not have adequate power to detect meaningful effects on uncommon clinical events, and selecting medicines for further study on such a basis is not practical. Selection based on an observed RR of <0.8 may provide a lower 'positive predictive value'; the proportion of positive trials in which the therapy was truly effective ranged from 13.3% to 61.1% across the six scenarios. Like selection to move forward based on a statistically significant RR, selection based on an observed RR of <0.8 provides a biased, overly optimistic estimate of the RR. Continuous surrogate outcomes, which for smaller trials provide greater power, are often used to detect potential benefit; the observed RR in a Phase 2 trial with a statistically significant continuous surrogate outcome may be relatively unbiased. Similar results were obtained with 80 patients per group, with a control event rate of 20%, and with various correlations between the true RR and the surrogate outcome (online supplementary *Tables S2–S5*).

The proportion of all our simulated Phase 2 trials in which the true RR was <0.8 was 23.5% (*Table 1*, first row). If the decision to proceed to Phase 3 was based on an observed RR of <0.8, 38.9% of the selected trials would have a true RR of <0.8 (*Table 1*, fifth row) (i.e. 61.1% of these trials would have an RR of >0.8 and large trials would be likely to show a neutral result), and 35.0% of subsequent Phase 3 trials, each with 90% power, would be expected to be positive. Similar calculations result in expected proportions of positive Phase 3 trials of 45.6% based on continuing

because of a significant result for a perfectly correlated surrogate in Phase 2 (*Table 1*, seventh row) and 51.6% based on a statistically significant effect on the event rate (*Table 1*, sixth row). Regardless of the criterion for moving from Phase 2 to Phase 3, we would expect that even if few medicines tested were truly effective, the proportion of Phase 3 studies that should yield positive results if appropriately powered is not null. If the expected success rate were 35%, then 0/15 successful programmes have a probability of 0.00156.

Placebo event rates and treatment effects in acute and chronic heart failure trials

We examined the associations of trial size with placebo event rates and treatment effects in trials in patients with chronic HFrEF and AHF. We examined all-cause mortality in 31 HFrEF and 15 AHF trials (online supplementary *Table S6*) and recurrent HF events in 17 HFrEF and 12 AHF trials (online supplementary *Table S7*). Recurrent HF events were defined as HF admissions in chronic HF and as WHF events in AHF.

With respect to all-cause mortality at 180 days (*Figure 1*), a slight, non-significant decline ($P = 0.3704$) in placebo event rate with increasing sample size was observed in HFrEF trials (a 1.4-fold decline in event rates between 50 and 3250 control patients), whereas a steeper decline with increasing sample size ($P = 0.0579$) was observed in AHF trials (a two-fold decline). In HFrEF trials, average placebo event rates were 10% [95% confidence interval (CI) 7–14%] in Phase 2 studies and 7% (95% CI 5–10%) in Phase 3 trials. In AHF, average placebo event rates were 21% (95% CI 12–30%) in Phase 2 and 14% (95% CI 11–16%) in Phase 3 trials. Medicine treatment effects did not vary with trial size in HFrEF trials ($P = 0.6832$), but declined somewhat, although not significantly ($P = 0.4508$), with increasing trial size in AHF trials. In HFrEF trials, the average treatment effect did not differ between Phases 2 and 3 [HR 0.92 (95% CI 0.64–1.34) vs. HR 0.94 (95% CI 0.80–1.10)]. Among the AHF trials, the average treatment effects were HR 0.77 (95% CI 0.49–1.22) in Phase 2 and HR 0.94 (95% CI 0.74–1.20) in Phase 3 trials.

With respect to recurrent HF events (*Figure 2*), we observed a significant reduction in placebo event rates with increasing trial size in both chronic HFrEF, with lower rates of HF admission in larger trials ($P = 0.0465$), and AHF studies, with lower rates of WHF in larger trials ($P = 0.0182$). The decline in placebo event rates was steeper in AHF trials (a 3.6-fold decline in event rates between 50 and 3250 control patients) than in chronic HFrEF trials (a 2.7-fold decline). Among the HFrEF trials, average placebo event rates were 10% (95% CI 5–16%) in Phase 2 trials and 7% (95% CI 5–10%) in Phase 3 trials. In AHF trials, average control event rates were 20% (95% CI 8–35%) in Phase 2 and 12% (95% CI 6–19%) in Phase 3. Apparent trends of smaller treatment effects in larger trials were observed in both chronic HFrEF trials and AHF trials, representing dilutions of 87% and 78%, respectively (*Figure 2B, D*), although neither reached statistical significance ($P = 0.0935$ and $P = 0.1795$, respectively). Average treatment effects in HFrEF trials were HR 0.64 (95% CI 0.42–0.99) in Phase 2 and HR 0.81 (95% CI 0.67–0.98) in Phase 3 ($P = 0.3205$). In AHF trials average treatment effects were HR 0.59 (95% CI 0.32–1.00) in Phase 2 and 0.79 (95% CI 0.60–1.03) in Phase 3.

Larger trials compared to real life

In online supplementary *Table S8*, we describe the baseline characteristics of patients enrolled in Phase 2 and 3 trials[1–6,43] compared to registries in AHF[44–52] in which there was no eligibility restriction with reference to left ventricular ejection fraction or other requirement for 'documented cardiac dysfunction'. As the table shows, some characteristics of patients in Phase 2 and Phase 3 trials showed lesser or greater degrees of similarity with those of registry patients (highlighted in green); however, there is no specific pattern showing that findings in Phase 3 trials are more similar to registry results than those in Phase 2 trials.

Do large trials have progressively more power regardless of size?

In *Figure 3*, we show the power for HRs of 0.65 (i.e. a 35% reduction in event incidence) and 0.80 (i.e. a 20% reduction in event incidence) for placebo event rates of 10% and 20%, for differing assumed dilutions of treatment effect with increasing trial size. A dilution of 50% for an HR of 0.65 would mean that we would expect the HR to be 0.65 at $n = 100$, 0.70 at $n = 2000$, and 0.82 at $n = 6500$. Thus, rather than a simple increase in power with increasing sample size under the usual assumption that the treatment effect is fixed (no dilution), a dilution of treatment effect and reduced placebo event rates with increasing sample size attenuate very substantially the power gained by increasing patient numbers.

Discussion

The results of our simulations suggest that chance alone (i.e. random high bias in which a new medicine is taken to Phase 3 because of a randomly large treatment effect in Phase 2) is unlikely to explain the complete failure to develop new interventions in AHF and HFpEF over the last decades. Success rates in a reasonable development programme for well-selected Phase 2 medicines should be at least 35%, which means that, given the success rates and effect sizes seen in Phase 2 trials, at least 35% of appropriately powered Phase 3 trials should have been positive. This is consistent with estimated probabilities of success of 32.3% in moving a cardiovascular clinical trial from Phase 3 to approval based on actual results in large clinical trial repositories.[57,58] The observation of no successful AHF or HFpEF development programme is thus highly improbable. Equally, the results of our analysis of Phase 2 and 3 trials in chronic HFrEF and AHF showing a larger effect of trial size on placebo event rates than treatment effect is not supportive of the random high bias theory. If random high bias is the reason why Phase 3 results are less positive, then placebo event rates in Phases 2 and 3 should be equal as it is unlikely that a medicine will be chosen for further exploration in Phase 3 just because the placebo

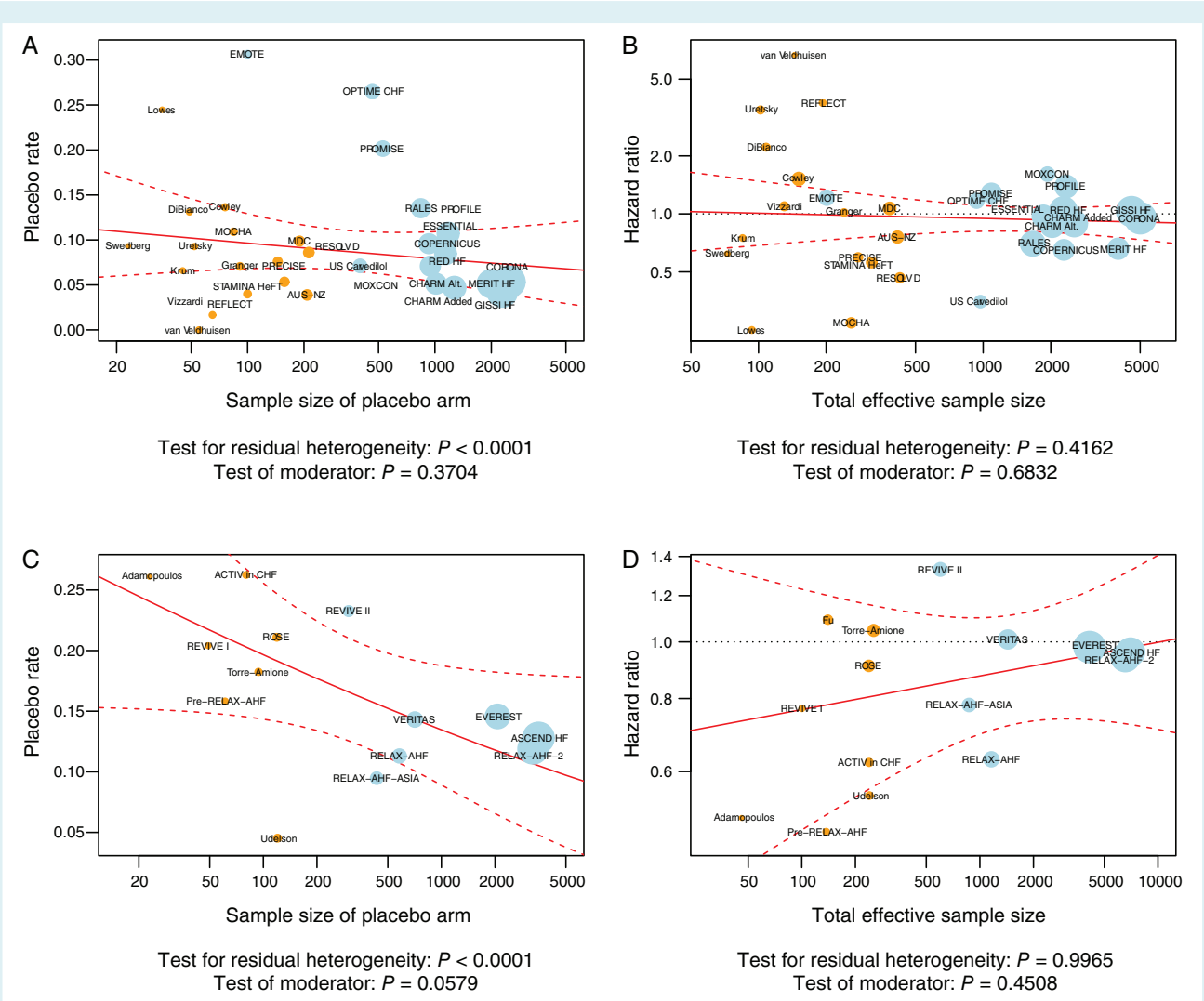

Figure 1 Association of sample size with placebo event rates and treatment effect for all-cause mortality at 180 days in trials in (*A*, *B*) chronic heart failure and (*C*, *D*) acute heart failure. Phase 2 studies are indicated in orange and Phase 3 studies in blue. In (C), one trial[55] lies outside the presented plotting area.

event rate in Phase 2 was high. Accordingly, the transition from Phase 2 to Phase 3, in contemporary trials in HF, is consistently associated with effect dilution.

Indeed, our analysis of Phase 2 and 3 acute and chronic HFrEF trials suggests that as trials become larger, placebo event rates decrease precipitously. This affects more disease-specific events such as HF recurrence than mortality and is more pronounced in AHF trials than in chronic HFrEF trials. The treatment effects of experimental medicines seem to be less affected by increasing trial size as we observed only non-significant trends of decreasing treatment effect on HF recurrence with increasing trial size, and no effect of trial size on treatment effects on mortality. What is the explanation for this phenomenon? Firstly, it is possible that in some indications in which the disease state is not easily and objectively defined, increasing trial size may shift the patient population toward patients who have risk factors for some cardiovascular adverse events but not for true HF. If this were to occur, it would explain the lower event rates, especially with respect to disease-related HF recurrent events, and it is also possible that such patients will be less responsive. An extreme example of such an occurrence was reported in the TOPCAT trial in which event rates outside the Americas were found to be substantially lower than those in the Americas.[59] It is possible that the globalization of clinical trials and the shift of recruitment to countries in which health care and general GDP per capita are lower[60] are resulting in the creation of incentives that may lead some to enrol patients who have milder cases of HF or none in pursuit of financial gain. In any case, the combination of a reduced event rate and reduced treatment effect leads to substantially reduced power and fewer positive trials. Secondly, it is possible that larger trials that are less well monitored under the new 'risk-based monitoring' paradigm[61] have lower event reporting rates (i.e. some sites fail to report some

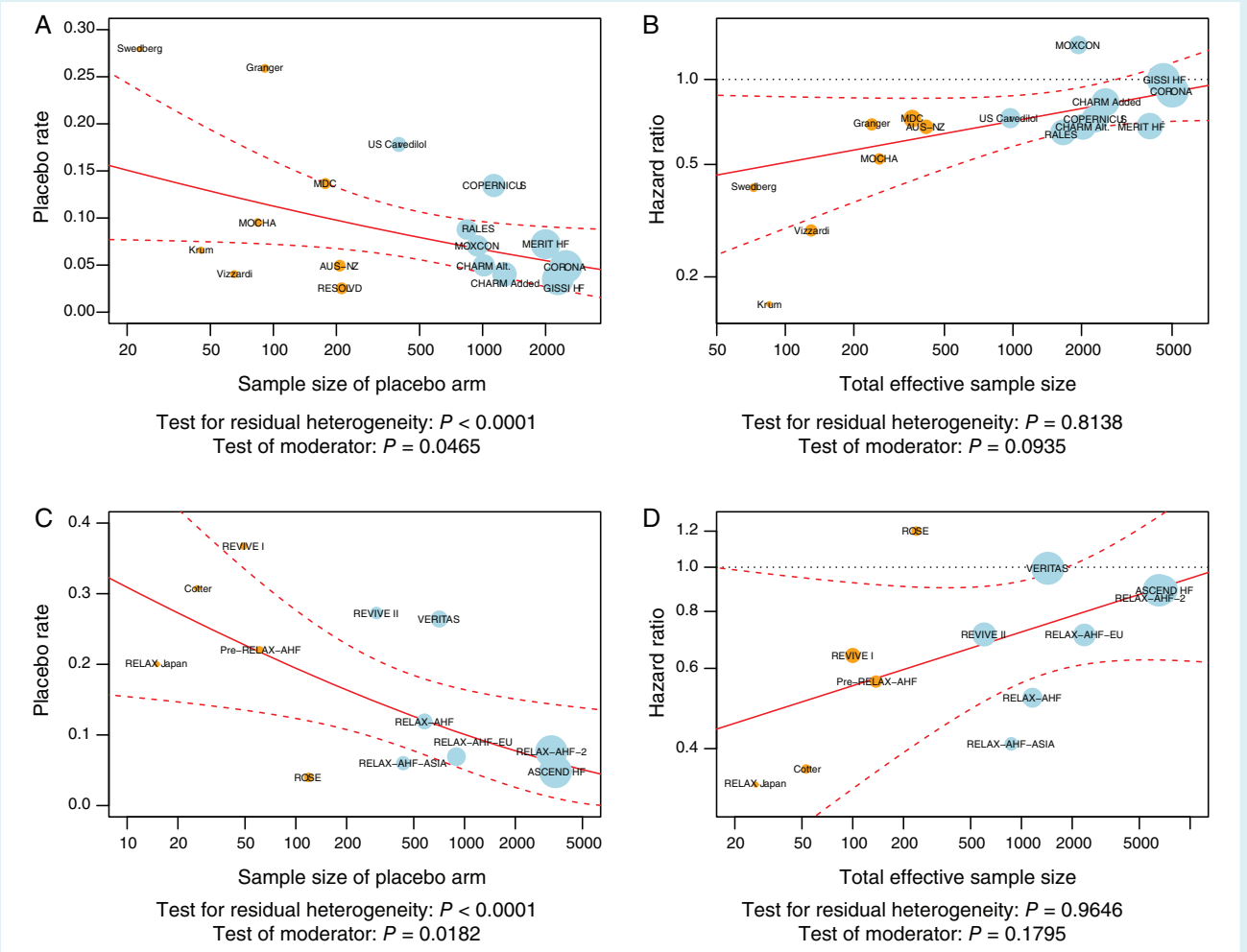

Figure 2 Association of sample size with placebo event rates and treatment effect for heart failure recurrence in trials in (*A*, *B*) chronic heart failure and (*C*, *D*) acute heart failure. Phase 2 studies are indicated in orange and Phase 3 studies in blue[56]. In (*B*) one trial (RESOLVD) lies outside the presented plotting area.

events). Although this may explain lower placebo event rates, it will not explain lower treatment effects on HF events as lower reporting is unlikely to preferentially occur in placebo arms.

In support of the mega (larger) trials, some experts have urged the performance of large trials because they consider such contexts to better represent 'real life'. However, the current analysis of trials performed in AHF suggests that larger trials are not more likely to better represent real-life contexts. In fact, the present analysis shows significant drifts away from rates observed in registries because registries have shown much higher event rates than the placebo event rates observed in larger Phase 3 trials. The current data suggest that placebo event rates are more similar to 'real-life' event rates in Phase 2 than in Phase 3 trials. For instance, WHF is an endpoint specific to AHF that occurs only in patients with real AHF; patients with severe stable chronic HF who are not acute (do not have AHF) will not have WHF. Our observation that in AHF trials the incidence of WHF is inversely correlated with trial size (i.e. occurs less in both active and placebo patients in larger trials) suggests that the larger trials simply included patients whose disease was 'less acute' or who had other disease that is in line with the higher event rates observed in registries for WHF but not in large AHF trials.

These discrepancies can be quantified by the analysis presented in *Figure 3*. If indeed placebo event rates and treatment effects do not drop with increasing trial size as in chronic HFrEF trials, this would correspond to the top power curves in *Figure 3* for no treatment effect dilution, whereby power increases monotonically with trial size. Even in the graphs on the right, which represent a halving of the placebo event rate, it is apparent that some power remains and trials may return positive results from time to time, something that has been observed in chronic HFrEF trials. However, in disease states such as AHF, in which both placebo event rates and treatment effects fall substantially and concurrently in larger trials, the increase in trial size not only does not compensate but actually precipitates an increasingly larger drop in power as the trial size increases, which potentially explains why

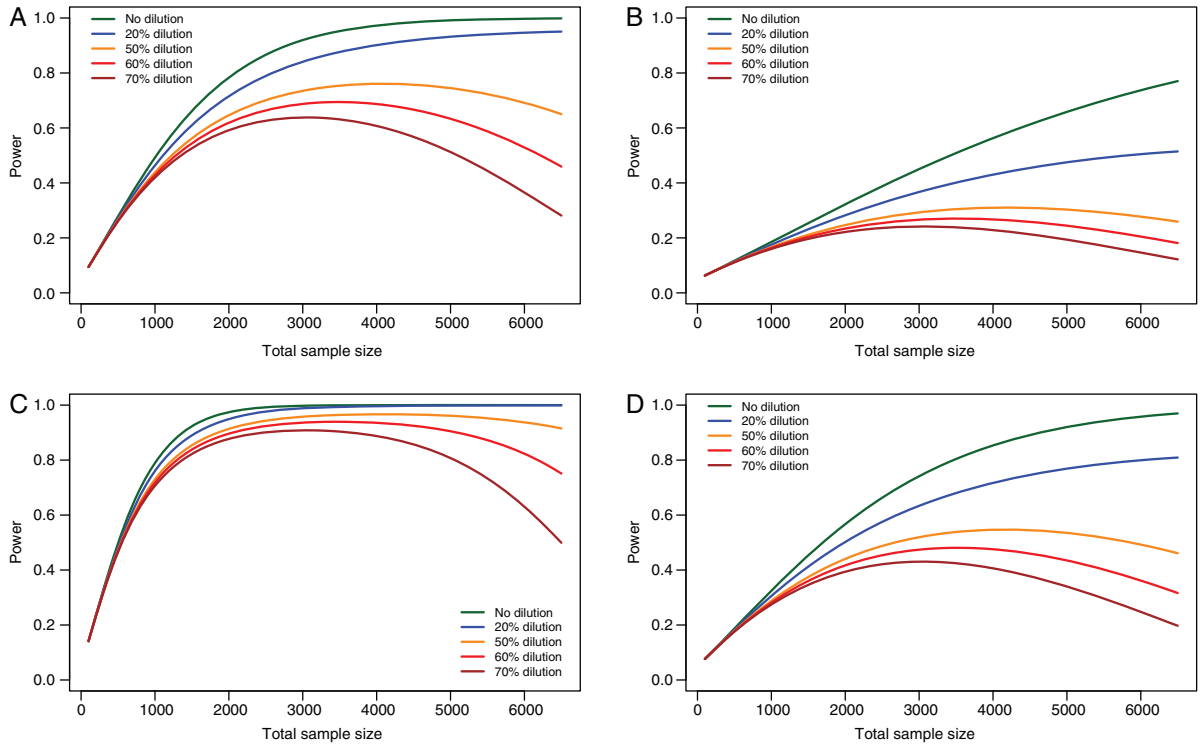

Figure 3 Power and effect dilution: power at a two-sided significance level of 0.05 to detect effect sizes with hazard ratios of 0.65 (*A*, *C*) or 0.8 (*B*, *D*) with placebo event rates of 10% (*A*, *B*) and 20% (*C*, *D*) at $n = 100$ at various treatment effect dilutions with increasing sample size.

no large AHF trial has ever shown a positive effect, something our simulations suggest is highly unlikely to be due to chance alone.

Conclusions and recommendations

Our study demonstrated that neutral Phase 3 trials in AHF were unlikely to relate to play of chance. Random high bias or regression to the mean may explain some of the failures in the development of new interventions that have led to the failure of about 65% of new medicines between Phases 2 and 3, as has been observed over all cardiovascular indications. We contend that chance cannot fully explain the complete failure to develop new interventions in diseases such as HFpEF and AHF. 'Mega-trials' may be vulnerable to both reductions in placebo event rates and dilutions of the treatment effect, most pronouncedly with disease-specific events such as recurrent HF. It is possible that this is explained by inappropriate heterogeneity of the target population, which will reduce both placebo event rates and treatment effects and have an adverse effect on the power to detect a real benefit of a new intervention. In HF, the effect seems to be larger in diseases that are less objectively defined, such as AHF and HFpEF, in which no Phase 3 trial has ever replicated the results of smaller Phase 2 trials. Therefore, although simple large trials may be appropriate for disease states with objective diagnostic criteria, trial sizes should be limited to approximately 2000–3000 patients in disease states in which the diagnosis is more subjective and less easy to ascertain, such as HFpEF, AHF and cardiogenic shock. The smaller trial can lead to substantial cost savings, which should be invested in strategies to ensure the tighter monitoring and control of the patients enrolled in the trial, such as in site selection, which should be strongly emphasized and conducted meticulously at the trial's outset, improved onsite and remote monitoring, implementation of statistical monitoring of blinded data to enable the assessment of whether the patients enrolled represent the intended population and whether patient characteristics and outcomes are equally distributed across sites and countries. Importantly, when issues are detected, study management should take decisive steps to correct them.

Limitations

The rates of false negative and positive Phase 2 trials estimated in the simulations are derived from the assumed distributions of effective medicines. In addition, the meta-analysis was based on a literature review, and some trials may have not been published and hence not included in this analysis.

Supplementary Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Distributions of trials according to true relative risk used for simulations.
Table S2. N/group = 150, 10% placebo event rate, correlation between continuous and dichotomous outcome 0.5.
Table S3. N/group = 80, 20% placebo event rate, correlation between continuous and dichotomous outcome 0.5.
Table S4. N/group = 80, 20% placebo event rate, correlation between continuous and dichotomous outcome 0.2.
Table S5. N/group = 150, 10% placebo event rate, correlation between continuous and dichotomous outcome 0.2.
Table S6. Overview of chronic and acute heart failure trials included in meta-analysis for all-cause mortality.
Table S7. Overview of chronic and acute heart failure trials included in meta-analysis for heart failure recurrence.
Table S8. Comparison of Phase 3, Phase 2 studies and registries in patients with acute heart failure with no restriction on ejection fraction.
Figure S1. Distribution of true relative risks assumed in the six scenarios.
Figure S2. Distributions of the observed relative risks in the simulated trials for the six scenarios, overlaid with the distribution of assumed true relative risks.
Figure S3. Identification of trials for inclusion in the meta-analyses.

Conflict of interest: B.A.D. reports the receipt of personal fees from Novartis Pharma and grants from Amgen, Celyad, Cirius Therapeutics, Laguna Pharmaceuticals, Sanofi, Roche Diagnostics, Trevena, Novartis Pharmaceutical, the National Institutes of Health (NIH) and Ventrix. K.T. has received speaker's honoraria from Otsuka, Sumitomo Dainippon, AstraZeneca and Bayer, and consultancy fees from Terumo. S.S. reports the receipt of grants from Amgen, Celyad, Cirius Therapeutics, Laguna Pharmaceuticals, Sanofi, Roche Diagnostics, Trevena, Novartis Pharmaceutical, the NIH and Ventrix. G.K. is the principal investigator in a biostatistical agreement between Momentum Research and the University of North Carolina at Chapel Hill, which provided the structure for his activity for this article. He is also the principal investigator of many such biostatistical agreements with other biopharmaceutical sponsors, including AbbVie, Amgen, Arena, AstraZeneca, Eli Lilly, Forest Research Institute (Allergan), GlaxoSmithKline, Merck, Novartis, Otsuka, Pfizer and Sanofi, although his activities for those sponsors are not related to the content of this article. Information concerning all biostatistical agreements for which G.K. is the principal investigator is publicly available through the University of North Carolina at Chapel Hill. M.M. reports the receipt of personal fees from consulting honoraria as a committee member or for speeches from Abbott Vascular, Bayer, Edwards Therapeutics, Fresenius, Novartis, Servier, Vifor Pharma and Windtree Therapeutics. A.K. reports the receipt of fees from Baxter, MSD, Aspen and Gilead. A.M. reports the receipt of personal fees from Novartis, Orion, Roche, Servier and Sanofi, and grants and personal fees from Adrenomed and Abbott. A.A.V. reports the receipt of personal fees from Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Cytokinetics, GSK, Servier, Merck and Novartis, and grants and personal fees from Roche Diagnostics. O.W.N. reports the receipt of personal fees from Novartis. O.C. reports the receipt of grants from Servier and Vifor Pharma, and grants and other fees from Novartis. P.S.P. reports the receipt of grants, personal fees and non-financial support from BMS, Baxter, Roche, Trevena and Novartis, and grants from OrthoDiagnostics, the NIH, the National Heart, Lung and Blood Institute, the Agency for Healthcare Research and Quality and the American Heart Association. B.H.G. reports the receipt of personal fees from Novartis. A.P.M. reports the receipt of personal fees from Bayer, Novartis and Fresenius. A.C.-S. reports the receipt of personal fees from Novartis, Servier, MSD, AstraZeneca, Vifor Pharma, Amgen, CVRX, Menarini, Roche Diagnostics and BMS. G.E. reports the receipt of grants from Bayer, Novartis and Vifor Pharma. N.S. reports the receipt of personal fees from Novartis Otsuka, Ono and Terumo, and grants from Daichi-Sankyo and Roche Diagnostics Japan. J.R.T. reports the receipt of grants, personal fees and non-financial support from Novartis, Amgen and Trevena. G.F. reports activity as a steering committee member for Novartis, Medtronic, BI, Bayer and Vifor Pharma. P.P. reports the receipt of personal fees and other fees from Novartis, grants and other fees from Vifor Pharma, personal fees from Amgen, grants, personal fees and other fees from Bayer, BMS, Boehringer Ingelheim, Cardiorentis and Cibiem, grants from Singulex, and other fees from Fresenius. E.G. reports the receipt of personal fees from Magnisense and Sphingotec. C.E. reports the receipt of grants from Amgen, Celyad, Cirius Therapeutics, Laguna Pharmaceuticals, Sanofi, Roche Diagnostics, Trevena, Novartis Pharmaceutical, the NIH and Ventrix. G.C. reports the receipt of personal fees from Novartis Pharma, and grants from Amgen, Celyad, Cirius Therapeutics, Laguna Pharmaceuticals, Sanofi, Roche Diagnostics, Trevena, Novartis Pharmaceutical, the NIH and Ventrix.

References

1. O'Connor CM, Starling RC, Hernandez AF, Armstrong PW, Dickstein K, Hasselblad V, Heizer GM, Komajda M, Massie BM, McMurray JJV, 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, Corbalán 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, Méndez GF, Metra M, Mittal S, Oh BH, Pereira NL, Ponikowski P, Tang WHW, 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.
2. Voors AA, Dittrich HC, Massie BM, DeLucca P, Mansoor GA, Metra M, Cotter G, Weatherley BD, Ponikowski P, Teerlink JR, Cleland JG, O'Connor CM, Givertz MM. Effects of the adenosine A1 receptor antagonist rolofylline on renal function in patients with acute heart failure and renal dysfunction: results from PROTECT (placebo-controlled randomized study of the selective adenosine A1 receptor antagonist rolofylline for patients hospitalized with acute decompensated heart failure and volume overload to assess treatment effect on congestion and renal function). *J Am Coll Cardiol* 2011;**57**:1899–1907.
3. 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.
4. 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.

5. Cotter G, Dittrich HC, Weatherley BD, Bloomfield DM, O'Connor CM, Metra M, Massie BM; Protect Steering Committee, Investigators and Coordinators. The PROTECT pilot study: a randomized, placebo-controlled, dose-finding study of the adenosine A1 receptor antagonist rolofylline in patients with acute heart failure and renal impairment. *J Card Fail* 2008;**14**:631–640.
6. Teerlink JR, Voors AA, Ponikowski P, Pang PS, Greenberg BH, Filippatos G, Felker GM, Davison BA, Cotter G, Gimpelewicz C, Boer-Martins L, Wernsing M, Hua TA, Severin T, Metra M. Serelaxin in addition to standard therapy in acute heart failure: rationale and design of the RELAX-AHF-2 study. *Eur J Heart Fail* 2017;**19**:800–809.
7. Konstam MA, Gheorghiade M, Burnett JC Jr, Grinfeld L, Maggioni AP, Swedberg K, Udelson JE, Zannad F, Cook T, Ouyang J, Zimmer C, Orlandi C; Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan (EVEREST) Investigators. Effects of oral tolvaptan in patients hospitalized for worsening heart failure: the EVEREST Outcome Trial. *JAMA* 2007;**297**:1319–1331.
8. Mebazaa A, Nieminen MS, Packer M, Cohen-Solal A, Kleber FX, Pocock SJ, Thakkar R, Padley RJ, Põder P, Kivikko M; SURVIVE Investigators. Levosimendan vs. dobutamine for patients with acute decompensated heart failure: the SURVIVE Randomized Trial. *JAMA* 2007;**297**:1883–1891.
9. Maggioni AP, Greene SJ, Fonarow GC, Bohm M, Zannad F, Solomon SD, Lewis EF, Baschiera F, Hua TA, Gimpelewicz CR, Lesogor A, Gheorghiade M; ASTRONAUT Investigators and Coordinators. Effect of aliskiren on post-discharge outcomes among diabetic and non-diabetic patients hospitalized for heart failure: insights from the ASTRONAUT trial. *Eur Heart J* 2013;**34**:3117–3127.
10. 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; VERITAS Investigators. Effects of tezosentan on symptoms and clinical outcomes in patients with acute heart failure: the VERITAS randomized controlled trials. *JAMA* 2007;**298**:2009–2019.
11. 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; REVIVE Heart Failure Study Group. Effect of levosimendan on the short-term clinical course of patients with acutely decompensated heart failure. *JACC Heart Fail* 2013;**1**:103–111.
12. Shah SJ, Blair JE, Filippatos GS, Macarie C, Ruzyllo W, Korewicki J, Bubenek-Turconi SI, Ceracchi M, Bianchetti M, Carminati P, Kremastinos D, Grzybowski J, Valentini G, Sabbah HN, Gheorghiade M; HORIZON-HF Investigators. Effects of istaroxime on diastolic stiffness in acute heart failure syndromes: results from the hemodynamic, echocardiographic, and neurohormonal effects of istaroxime, a novel intravenous inotropic and lusitropic agent: a randomized controlled trial in patients hospitalized with heart failure (HORIZON-HF) trial. *Am Heart J* 2009;**157**:1035–1041.
13. Coletta AP, Cleland JG. Clinical trials update: highlights of the scientific sessions of the XXIII Congress of the European Society of Cardiology – WARIS II, ESCAMI, PAFAC, RITZ-1 and TIME. *Eur J Heart Fail* 2001;**3**:747–750.
14. Felker GM, Leimberger JD, Califf RM, Cuffe MS, Massie BM, Adams KF Jr, Gheorghiade M, O'Connor CM. Risk stratification after hospitalization for decompensated heart failure. *J Card Fail* 2004;**10**:460–466.
15. Pfeffer MA, Swedberg K, Granger CB, Held P, McMurray JJV, Michelson EL, Olofsson B, Östergren J, Yusuf S. Effects of candesartan on mortality and morbidity in patients with chronic heart failure: the CHARM-Overall programme. *Lancet* 2003;**362**:759–766.
16. Massie BM, Carson PE, McMurray J, Komajda M, McKelvie R, Zile MR, Anderson S, Donovan M, Iverson E, Staiger C, Ptaszynska A; I-PRESERVE Investigators. Irbesartan in patients with heart failure and preserved ejection fraction. *N Engl J Med* 2008;**359**:2456–2467.
17. Pitt B, Pfeffer MA, Assmann SF, Boineau R, Anand IS, Claggett B, Clausell N, Desai AS, Diaz R, Fleg JL, Gordeev I, Harty B, Heitner JF, Kenwood CT, Lewis EF, O'Meara E, Probstfield JL, Shaburishvili T, Shah SJ, Solomon SD, Sweitzer NK, Yang S, McKinlay S; TOPCAT Investigators. Spironolactone for heart failure with preserved ejection fraction. *N Engl J Med* 2014;**370**:1383–1392.
18. Deswal A, Richardson P, Bozkurt B, Mann DL. Results of the randomized aldosterone antagonism in heart failure with preserved ejection fraction trial (RAAM-PEF). *J Card Fail* 2011;**17**:634–642.
19. Komajda M, Isnard R, Cohen-Solal A, Metra M, Pieske B, Ponikowski P, Voors AA, Dominjon F, Henon-Goburdhun C, Pannaux M, Böhm M; prEserveD left ventricular ejectIon fraction chronic heart Failure with ivabradine studY (EDIFY) Investigators. Effect of ivabradine in patients with heart failure with preserved ejection fraction: the EDIFY randomized placebo-controlled trial. *Eur J Heart Fail* 2017;**19**:1495–1503.
20. Flather MD, Shibata MC, Coats AJ, van Veldhuisen D, Parkhomenko A, Borbola J, Cohen-Solal A, Dumitrascu D, Ferrari R, Lechat P, Soler-Soler J, Tavazzi L, Spinarova L, Toman J, Böhm M, Anker SD, Thompson SG, Poole-Wilson PA; SENIORS Investigators. Randomized trial to determine the effect of nebivolol on mortality and cardiovascular hospital admission in elderly patients with heart failure (SENIORS). *Eur Heart J* 2005;**26**:215–225.
21. Ahmed A, Rich MW, Fleg JL, Zile MR, Young JB, Kitzman DW, Love TE, Aronow WS, Adams KF Jr, Gheorghiade M. Effects of digoxin on morbidity and mortality in diastolic heart failure: the ancillary digitalis investigation group trial. *Circulation* 2006;**114**:397–403.
22. Edelmann F, Wachter R, Schmidt AG, Kraigher-Krainer E, Colantonio C, Kamke W, Duvinage A, Stahrenberg R, Durstewitz K, Löffler M, Düngen HD, Tschöpe C, Herrmann-Lingen C, Halle M, Hasenfuss G, Gelbrich G, Pieske B; Aldo-DHF Investigators. 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.
23. Havranek EP, Thomas I, Smith WB, Ponce GA, Bilsker M, Munger MA, Wolf RA. Dose-related beneficial long-term hemodynamic and clinical efficacy of irbesartan in heart failure. *J Am Coll Cardiol* 1999;**33**:1174–1181.
24. McMurray JJ, Krum H, Abraham WT, Dickstein K, Køber LV, Desai AS, Solomon SD, Greenlaw N, Ali MA, Chiang Y, Shao Q, Tarnesby G, Massie BM; ATMOSPHERE Committees Investigators. Aliskiren, enalapril, or aliskiren and enalapril in heart failure. *N Engl J Med* 2016;**374**:1521–1532.
25. Packer M, Fowler MB, Roecker EB, Coats AJ, Katus HA, Krum H, Mohacsi P, Rouleau JL, Tendera M, Staiger C, Holcslaw TL, Amann-Zalan I, DeMets D; Carvedilol Prospective Randomized Cumulative Survival (COPERNICUS) Study Group. Effect of carvedilol on the morbidity of patients with severe chronic heart failure: results of the carvedilol prospective randomized cumulative survival (COPERNICUS) study. *Circulation* 2002;**106**:2194–2199.
26. Poole-Wilson PA, Swedberg K, Cleland JGF, di Lenarda A, Hanrath P, Komajda M, Lubsen J, Lutiger B, Metra M, Remme WJ, Torp-Pedersen C, Scherhag A, Skene A. Comparison of carvedilol and metoprolol on clinical outcomes in patients with chronic heart failure in the Carvedilol or Metoprolol European Trial (COMET): randomised controlled trial. *Lancet* 2003;**362**:7–13.
27. Kjekshus J, Apetrei E, Barrios V, Böhm M, Cleland JG, Cornel JH, Dunselman P, Fonseca C, Goudev A, Grande P, Gullestad L, Hjalmarson A, Hradec J, Jánosi A, Kamenský G, Komajda M, Korewicki J, Kuusi T, Mach F, Mareev V, McMurray J, Ranjith N, Schaufelberger M, Vanhaecke J, van Veldhuisen D, Waagstein F, Wedel H, Wikstrand J; CORONA Group. Rosuvastatin in older patients with systolic heart failure. *N Engl J Med* 2007;**357**:2248–2261.
28. McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, Rouleau JL, Shi VC, Solomon SD, Swedberg K, Zile MR; PARADIGM-HF Investigators and Committees. Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N Engl J Med* 2014;**371**:993–1004.
29. Zannad F, McMurray JJV, Krum H, van Veldhuisen DJ, Swedberg K, Shi H, Vincent J, Pocock SJ, Pitt B. Eplerenone in patients with systolic heart failure and mild symptoms. *N Engl J Med* 2011;**364**:11–21.
30. Pitt B, Remme W, Zannad F, Neaton J, Martinez F, Roniker B, Bittman R, Hurley S, Kleiman J, Gatlin M; Eplerenone Post-Acute Myocardial Infarction Heart Failure Efficacy and Survival Study Investigators. Eplerenone, a selective aldosterone blocker, in patients with left ventricular dysfunction after myocardial infarction. *N Engl J Med* 2003;**348**:1309–1321.
31. Homma S, Thompson JL, Sanford AR, Mann DL, Sacco RL, Levin B, Pullicino PM, Freudenberger RS, Teerlink JR, Graham S, Mohr JP, Massie BM, Labovitz AJ, di Tullio MR, Gabriel AP, Lip GY, Estol CJ, Lok DJ, Ponikowski P, Anker SD; WARCEF Investigators. Benefit of warfarin compared with aspirin in patients with heart failure in sinus rhythm: a subgroup analysis of WARCEF, a randomized controlled trial. *Circ Heart Fail* 2013;**6**:988–997.
32. Swedberg K, Komajda M, Böhm M, Borer JS, Ford I, Dubost-Brama A, Lerebours G, Tavazzi L. Ivabradine and outcomes in chronic heart failure (SHIFT): a randomised placebo-controlled study. *Lancet* 2010;**376**:875–885.
33. Konstam MA, Neaton JD, Dickstein K, Drexler H, Komajda M, Martinez FA, Riegger GAJ, Malbecq W, Smith RD, Guptha S, Poole-Wilson PA. Effects of high-dose versus low-dose losartan on clinical outcomes in patients with heart failure (HEAAL study): a randomised, double-blind trial. *Lancet* 2009;**374**:1840–1848.
34. McMurray JJ, Östergren J, Swedberg K, Granger CB, Held P, Michelson EL, Olofsson B, Yusuf S, Pfeffer MA. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. *Lancet* 2003;**362**:767–771.
35. Granger CB, McMurray JJV, Yusuf S, Held P, Michelson EL, Olofsson B, Östergren J, Pfeffer MA, Swedberg K. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function intolerant to angiotensin-converting-enzyme inhibitors: the CHARM-Alternative trial. *Lancet* 2003;**362**:772–776.
36. Cleland JG, Findlay I, Jafri S, Sutton G, Falk R, Bulpitt C, Prentice C, Ford I, Trainer A, Poole-Wilson PA. The Warfarin/Aspirin Study in Heart Failure (WASH): a randomized trial comparing antithrombotic strategies for patients with heart failure. *Am Heart J* 2004;**148**:157–164.

37. Pitt B, Segal R, Martinez FA, Meurers G, Cowley AJ, Thomas I, Deedwania PC, Ney DE, Snavely DB, Chang PI. Randomised trial of losartan versus captopril in patients over 65 with heart failure (Evaluation of Losartan in the Elderly Study, ELITE). *Lancet* 1997;**349**:747–752.
38. 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. US Carvedilol Heart Failure Study Group. *N Engl J Med* 1996;**334**:1349–1355.
39. Krum H, Sackner-Bernstein JD, Goldsmith RL, Kukin ML, Schwartz B, Penn J, Medina N, Yushak M, Horn E, Katz SD, Levin HR, Neuberg GW, DeLong G, Packer M. Double-blind, placebo-controlled study of the long-term efficacy of carvedilol in patients with severe chronic heart failure. *Circulation* 1995;**92**:1499–1506.
40. Krum H, Ashton E, Reid C, Kalff V, Rogers J, Amarena J, Singh B, Tonkin A. Double-blind, randomized, placebo-controlled study of high-dose HMG CoA reductase inhibitor therapy on ventricular remodeling, pro-inflammatory cytokines and neurohormonal parameters in patients with chronic systolic heart failure. *J Card Fail* 2007;**13**:1–7.
41. Solomon SD, Hee Shin S, Shah A, Skali H, Desai A, Kober L, Maggioni AP, Rouleau JL, Kelly RY, Hester A, McMurray JJV, Pfeffer MA; Aliskiren Study in Post-MI Patients to Reduce Remodeling (ASPIRE) Investigators. Effect of the direct renin inhibitor aliskiren on left ventricular remodelling following myocardial infarction with systolic dysfunction. *Eur Heart J* 2011;**32**:1227–1234.
42. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2018.
43. Packer M, O'Connor C, McMurray J, 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 D, Kowarski LS, Schactman M, Holzmeister J; TRUE-AHF Investigators. Effect of ularitide on cardiovascular mortality in acute heart failure. *N Engl J Med* 2017;**376**:1956–1964.
44. Kociol RD, Hammill BG, Fonarow GC, Klaskala W, Mills RM, Hernandez AF, Curtis LH. Generalizability and longitudinal outcomes of a national heart failure clinical registry: comparison of Acute Decompensated Heart Failure National Registry (ADHERE) and non-ADHERE Medicare beneficiaries. *Am Heart J* 2010;**160**:885–892.
45. Atherton JJ, Hayward CS, Wan Ahmad WA, Kwok B, Jorge J, Hernandez AF, Liang L, Kociol RD, Krum H; ADHERE International–Asia Pacific Scientific Advisory Committee. Patient characteristics from a regional multicenter database of acute decompensated heart failure in Asia Pacific (ADHERE International-Asia Pacific). *J Card Fail* 2012;**18**:82–88.
46. Sato N, Kajimoto K, Asai K, Mizuno M, Minami Y, Nagashima M, Murai K, Muanakata R, Yumino D, Meguro T, Kawana M, Nejima J, Satoh T, Mizuno K, Tanaka K, Kasanuki H, Takano T; ATTEND Investigators. Acute decompensated heart failure syndromes (ATTEND) registry. A prospective observational multicenter cohort study: rationale, design, and preliminary data. *Am Heart J* 2010;**159**:949–955.e1.
47. Cleland JG, Swedberg K, Follath F, Komajda M, Cohen-Solal A, Aguilar JC, Dietz R, Gavazzi A, Hobbs R, Korewicki J, Madeira HC, Moiseyev VS, Preda I, van Gilst W, Widimsky J, Freemantle N, Eastaugh J, Mason J; Study Group on Diagnosis of the Working Group on Heart Failure of the European Society of Cardiology. The EuroHeart Failure Survey programme – a survey on the quality of care among patients with heart failure in Europe Part 1: patient characteristics and diagnosis. *Eur Heart J* 2003;**24**:442–463.
48. Nieminen MS, Brutsaert D, Dickstein K, Drexler H, Follath F, Harjola VP, Hochadel M, Komajda M, Lassus J, Lopez-Sendon JL, Ponikowski P, Tavazzi L; EuroHeart Survey Investigators, Heart Failure Association, European Society of Cardiology. EuroHeart Failure Survey II (EHFS II): a survey on hospitalized acute heart failure patients: description of population. *Eur Heart J* 2006;**27**:2725–2736.
49. Senni M, Gavazzi A, Oliva F, Mortara A, Urso R, Pozzoli M, Metra M, Lucci D, Gonzini L, Cirrincione V, Montagna L, di Lenarda A, Maggioni AP, Tavazzi L; IN HF Outcome Investigators. In-hospital and 1-year outcomes of acute heart failure patients according to presentation (de novo vs worsening) and ejection fraction. Results from IN-HF Outcome Registry. *Int J Cardiol* 2014;**173**:163–169.
50. Lee SE, Lee HY, Cho HJ, Choe WS, Kim H, Choi JO, Jeon ES, Kim MS, Kim JJ, Hwang KK, Chae SC, Baek SH, Kang SM, Choi DJ, Yoo BS, Kim KH, Park HY, Cho MC, Oh BH. Clinical characteristics and outcome of acute heart failure in Korea: results from the Korean Acute Heart Failure Registry (KorAHF). *Korean Circ J* 2017;**47**:341–353.
51. Curtis LH, Greiner MA, Hammill BG, DiMartino LD, Shea AM, Hernandez AF, Fonarow GC. Representativeness of a national heart failure quality-of-care registry: comparison of OPTIMIZE-HF and non-OPTIMIZE-HF Medicare patients. *Circ Cardiovasc Qual Outcomes* 2009;**2**:377–384.
52. Trullas JC, Miró Ò, Formiga F, Martín-Sánchez FJ, Montero-Pérez-Barquero M, Jacob J, Quirós-López R, Herrero Puente P, Manzano L, Llorens P; Members of the RICA and EAHFE Registries. The utility of heart failure registries: a descriptive and comparative study of two heart failure registries. *Postgrad Med J* 2016;**92**:260–266.
53. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Soft* 2010;**36**:1–48.
54. Wittes J. Sample size calculations for randomized controlled trials. *Epidemiol Rev* 2002;**24**:39–53.
55. Fu S, Yi S, Zhu B, *et al.* Efficacy and safety of a modified dosage regimen of nesiritide in patients older than 75 years with acute heart failure. *Aging Clin Exp Res* 2012;**24**:524–529.
56. Effects of metoprolol CR in patients with ischemic and dilated cardiomyopathy: the randomized evaluation of strategies for left ventricular dysfunction pilot study. *Circulation* 2000;**101**:378–384.
57. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. *Biostatistics* 2019;**20**:273–286.
58. Wong CH, Siah KW, Lo AW. Corrigendum: estimation of clinical trial success rates and related parameters. *Biostatistics* 2019;**20**:366.
59. Pfeffer MA, Claggett B, Assmann SF, Boineau R, Anand IS, Clausell N, Desai AS, Diaz R, Fleg JL, Gordeev I, Heitner JF, Lewis EF, O'Meara E, Rouleau JL, Probstfield JL, Shaburishvili T, Shah SJ, Solomon SD, Sweitzer NK, McKinlay SM, Pitt B. Regional variation in patients and outcomes in the Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist (TOPCAT) trial. *Circulation* 2015;**131**:34–42.
60. Ferreira JP, Rossignol P, Dewan P, Lamiral Z, White WB, Pitt B, McMurray JJV, Zannad F. Income level and inequality as complement to geographical differences in cardiovascular trials. *Am Heart J* 2019;**218**:66–74.
61. Cotter G, Davison BA. Unmodifiable events, heart failure research, and 'risk-based monitoring' in large studies – the unholy triumvirate. *Eur J Heart Fail* 2018;**20**:1639–1644.