ESC European Society of Cardiology

European Journal of Heart Failure (2018) **20**, 1645–1648
doi:10.1002/ejhf.1360

VIEWPOINT

Future large-scale clinical trials in cardiovascular medicine: challenges and uncertainties

Milton Packer[1]* and Bertram Pitt[2]

[1]Baylor Heart and Vascular Institute, Baylor University Medical Center, Dallas, TX, USA; and [2]University of Michigan School of Medicine, Ann Arbor, MI, USA

Large-scale clinical trials have been the cornerstone of evidence-based medicine in cardiology for nearly 40 years. The results of these trials have led to justified confidence in most of our current treatments and have also identified interventions that have been ineffective or harmful. When planned and executed meticulously and analysed to minimize bias, large-scale trials have provided strong foundational support for the use of both drugs and devices for millions of people. However, recent changes in the practice of medicine and the business of clinical trials have threatened many of the operational principles that have successfully guided the conduct of these important studies. These changes may lead us to rethink our ability to rely on them for decision-making in cardiology, and particularly, with respect to the management of heart failure.

Risk–reward relationship for conducting large-scale trials has shifted

Demonstration of the efficacy and safety of cardiovascular interventions has generally required the enrolment of thousands of at-risk patients who are randomly assigned to one of several treatment arms and followed compulsively for long periods of time. This approach generates the large number of clinically relevant events required to demonstrate the modest treatment effect sizes typically seen in largely undifferentiated groups of patients with cardiovascular disease, including those with acute or chronic heart failure.

This traditional approach to trials in cardiology differs strikingly from the principles govern the conduct of trials in oncology.[1,2] Currently, most oncology trials evaluate treatments that are designed to disrupt a highly specific molecular target, whose presence in individual patients can be identified prior to enrolment. Therefore, although the eventual number of patients who might be treated in clinical practice may be small (< 10 000–100 000 patients), the expected magnitude of the treatment effect is often large (60–80% reduction in risk). Furthermore, oncology trials often are designed around biomarkers as their primary endpoints, typically those that reflect the abnormal proliferation of malignant cells. As a result, oncology trials frequently focus on 'disease or biomarker progression' rather than mortality, and because biomarker progression is a common occurrence and the effect sizes are often dramatic, the trials are typically small (< 500 patients). The trials recruit rapidly because patients with the genetically-defined cancer are already under the care of specialists who serve as investigators. If the disease has a highly predictable adverse outcome, a dramatic result can be persuasive, even in a small trial that is carried out on a surrogate endpoint, and often in the absence of a control group.

These features stand in marked contrast with the conditions that prevail when novel drugs are developed for patients with cardiovascular disease. An understanding of the genetics of cardiac disorders remains in its infancy, and specific molecular targets have not been identified as a valid basis for identifying subgroups with predictable responses to distinctive treatments. Interventions in cardiology typically focus on one of many pathophysiological (rather than genetic) mechanisms that contribute to but do not necessarily play a crucial role in a disease process. Interference with only one of many contributory mechanisms can be expected to have only a modest benefit (10–25% risk reduction), and there are few reliable surrogate endpoints. Hence, trials must track the occurrence of large numbers of fatal and non-fatal events of unquestioned clinical relevance. The most commonly measured non-fatal outcome in heart failure trials is hospitalization for worsening heart failure, but this endpoint depends on physician judgment, institutional policy and available healthcare resources. As event rates have declined over several decades, the size of large-scale outcome trials has necessarily soared, despite their

*Corresponding author. Baylor Heart and Vascular Institute, Baylor University Medical Center, 621 N. Hall Street, Dallas, TX 75226, USA. Tel: +1 212 3059260, Email: milton.packer@baylorhealth.edu

increasing reliance on composite endpoints. The typical large-scale cardiovascular outcome trial seeks to enrol 5000–20 000 patients and often takes 7–10 years to complete.

If a trial succeeds on its primary endpoint and the novel intervention is approved for clinical use, companies that develop drugs for cancer can price them at exorbitant levels (often exceeding $ 300 000 per year).[3] Although the number of patients who are eligible for treatment is small, the new drug is rapidly adopted by cancer specialists, and the resulting revenues rapidly exceed the costs of performing the small pivotal clinical trial that demonstrated efficacy. In contrast, a new cardiovascular drug — because of its modest effect size and potential use by millions of people — typically faces significant price constraints and onerous pre-authorization requirements, even if the treatment costs only a small fraction of that of a new oncology drug. Furthermore, most patients with common cardiovascular disorders (such as heart failure) are cared for by primary care physicians, who generally have not instituted the administrative procedures needed to tackle pre-authorization requirements. Therefore, the adoption of a new cardiovascular drug by generalists is slow; a sponsor can expect the revenues of a newly-approved cardiovascular drug to recoup the costs of large-scale trial only after many years, if ever.

As a result of the interplay of these socioeconomic forces, the development of novel cardiovascular drugs has slowed markedly in recent years, and, in parallel, the number of large-scale cardiovascular outcome trials has dwindled.[1] Given the high likelihood of failure and (when successful) the long lead-time to recoup the costs of a substantial financial investment (often more than $ 300 million), the risk–reward relationship for carrying out these trials is no longer favourable.

Investigator base for large-scale trials has been transformed

Twenty to 30 years ago, it was common for large-scale cardiovascular outcome trials to be carried out by highly motivated clinical investigators, who largely worked at academic medical centres and were allowed (and even encouraged) to devote a portion of their professional time and energy to participation in clinical research. North American and Western European centres played a major role in trials, and the infrastructure to support the recruitment of patients (e.g. experienced study coordinators) was well-established. However, in recent years, healthcare systems have placed enormous pressures on physicians to generate clinical revenues, and thus, the time available for clinical research activities has evaporated. In any fixed time frame, the revenues that a clinical investigator can generate from clinical trials are no longer competitive with the revenues that they can generate by seeing large numbers of patients at a rapid pace or by performing interventional or imaging procedures. Furthermore, many physicians no longer feel an obligation to contribute to medical progress or may fear surrendering control of their patients. At the same time, some patients may harbour suspicions that their participation in a long-term trial is likely to provide greater financial rewards to the clinical investigator than direct health benefits to themselves.

As a result, recruitment efforts in large-scale clinical trials have shifted to Eastern Europe, Asia and Latin America, where enthusiastic investigators can still be found and where the patient–physician relationship is likely to be more trustful. In some geographical regions, participation in a clinical trial may provide the only opportunity for patients to obtain good medical care. The cost of establishing an investigative site in these geographical regions is comparatively low, and the revenues generated from clinical research often provide a critical supplement to the relatively meager salaries of investigators. In North America and Western Europe, private practices have superseded academic medical centres as investigative sites, because they are able to structure the recruitment and care of patients far more efficiently than their counterparts at medical schools. Regardless of geography, many (perhaps most) investigative sites now view clinical trials primarily as a business opportunity, rather than a research enterprise. This shift in thinking means that investigators are primarily motivated to maximize trial revenues rather than to answer a specific research question.

This shift in the investigator base has affected trials regardless of the sponsoring organization. Even trials sponsored by the National Heart, Lung and Blood Institute (NHLBI) have moved to unfamiliar geographical regions or investigative sites. Because these sites are often lightly supervised, recent NHLBI-sponsored cardiovascular trials (such as TOPCAT and SPRINT) have been characterized by a significant heterogeneity in the types of patients who have been enrolled or in the ways that study procedures have been carried out. This heterogeneity has led to a meaningful degree of uncertainty about the results of these trials, leading to controversy about their implications for clinical practice.[4,5]

Contract research organizations supervise the quality of work

The ability of trials to effectively answer a research question depends substantially on the quality of work carried out at investigative sites, and in the past, assurances about quality were possible because the work could be closely monitored by auditors who were able to confirm the validity of recorded data. Although research carried out by the NHLBI was not typically closely monitored, trials conducted by large pharmaceutical companies were meticulously supervised, often by teams of full-time highly trained professionals with a primary allegiance to the integrity of the trial.

However, during the past 20–30 years, these experienced in-house monitoring teams have largely disappeared. Large pharmaceutical companies have eliminated much of this internal infrastructure in their efforts to streamline costs. Small start-up companies never had the opportunity or resources to establish these internal quality-assurance teams in the first place. As a result, drug sponsors — regardless of size — now rely heavily on contract research organizations (CRO) to identify and supervise investigative sites.

Not surprisingly, these CROs emerged as a business enterprise in their own right. They were designed not only to deliver a

product, but also to make a profit. Capitalizing on the price sensitivity of pharmaceutical sponsors (large or small), it was tempting for CROs to submit a low bid for services but then to deliver even lower-quality services. Recently, to further reduce expenses, some CROs have shifted to 'risk-based monitoring', a strategy where only sites with the highest possibility of mischief would be closely monitored. However, it has been difficult to validate the worthiness of the algorithms used to identify risk, and these algorithms are not necessarily updated if the study personnel at a previously reliable site change. The inevitable result of cost-cutting is that certain operational deficiencies are unlikely to be detected. To complicate matters further, since CROs are under significant financial pressure to meet recruitment targets, they may have little motivation to pursue potentially suspicious investigative activities and are understandably reluctant to take punitive actions at a questionable site that is recruiting well. The CROs understood that the likelihood that the sponsor would ever become aware of or be able to prove a lapse in study quality was remote.

With the outsourcing of many (and often most) elements of the clinical trial operations and execution, pharmaceutical sponsors effectively delegated large components of a trial's conduct to business entities that had no vested interest in the study outcome. For many companies, there was no viable alternative. Privately, sponsors prayed that their novel drug would be so effective that the trial would be able to detect a meaningful treatment effect, even if issues related to study quality created a substantial amount of background noise. Additionally, since lapses in study quality typically generate data that bias a result to the null, sponsors often compensated by making their trials even larger — and thus, more susceptible to the vagaries of recruitment and quality control.

Shift towards large-scale pragmatic trials carries major risks

Given these uncertainties, corporate and government sponsors have been reluctant to make the enormous investment that has traditionally been required to conduct large-scale cardiovascular outcome trials. Understandably, they have become keenly interested in any proposals that might allow the execution of these enormous trials at a fraction of their current cost. Suppose a trial could be carried out with minimal involvement by investigators and with minimal supervision. Theoretically, potential patients could respond to an online announcement; a website could provide all informational elements and informed consent and allow self-recruitment. Patients could be sent their assigned medication by mail or could collect it at central distribution facilities. Visits could be conducted electronically; adherence with the assigned medication (and the reporting of adverse effects) could be documented by self-report; and information about endpoint events could be obtained through electronic medical records or public databases. Such measures, taken collectively, constitute the primary elements of what is currently referred to as a 'pragmatic' randomized trial.

Technically, the 'pragmatic' trial has been distinguished from the 'explanatory' trial.[6–8] The traditional 'explanatory' trial seeks to control all potentially confounding factors in order to discern evidence for a treatment effect. In contrast, although treatment assignments are still randomized, the 'pragmatic' trial seeks to control as few operational issues as possible, thus allowing the recruitment of large numbers of patients with minimal infrastructure and cost. Although some might believe that a pragmatic trial determines whether a drug works in 'real life', the distinction between an 'explanatory' and 'pragmatic' trial lies not in its clinical applicability, but in its operational features. The pragmatic trial places most of the responsibility of study conduct on the trial participants, virtually eliminating the need for investigators, research coordinators or supervisory structure. In its most basic form, a trial can be carried out entirely based on the interaction between a study participant and a website.

What kind of information can we expect when these pragmatic trials are applied to questions in cardiovascular medicine? We can reasonably anticipate — if appropriately publicized — that recruitment will be rapid; potentially, tens of thousands of people could be randomized in months, rather than years. Furthermore, the costs of the trial could be reduced dramatically, perhaps down to 5% of the cost of a conventionally-executed phase III trial. Yet, every other aspect of the conduct of a pragmatic trial would increase uncertainty. Left unsupervised, participants might or might not be adherent with their study medication, compliant with the forms required at each study visit or responsive to phone calls or other electronic means of follow-up. Unless actively motivated, people might be less likely to maintain participation for long periods of time, and the ascertainment of study endpoints is likely to be less complete. The predictable consequence of a pragmatic trial would be an increase in background noise, thus driving the estimated effect size towards the null, assuming that the trial were analysed according to the intention-to-treat principle. Of course, a treatment effect might still be detected and reach conventional levels of statistical significance if the trial recruited an exceptionally large number of patients.

Are these pragmatic trials the wave of the future? The question is not whether these trials will be less costly, but whether such trials will be capable of generating persuasive evidence that can change clinical practice. Suppose a trial were to enrol 50 000 patients, who were randomized 1:1 to one of two treatment arms. Suppose that 20% of patients did not take the study medication; another 20% did not comply with study procedures; and another 20% did not provide data on endpoints. In the end, by intention-to-treat analysis, one group of patients might have a lower risk of a major adverse clinical outcome ($P < 0.001$). Yet, because of the numerous factors that nudge the effect size towards the null, the magnitude of the actually observed treatment effect may be quite small, even for an intervention that truly has a meaningful therapeutic benefit. A drug that would be shown to reduce mortality by 20–25% in a conventionally executed trial might produce only a 3–10% reduction in relative risk in a large-scale pragmatic trial. At its completion, the trial might achieve pre-planned levels of statistical significance, but will anyone care? At worst, the factors that bias the result to the null might entirely obliterate the ability

of the study to detect a true treatment difference. Certainly, if the pragmatic trial yields a false negative result or reports an effect size that appears tiny and unpersuasive by conventional standards, then even the paltry sums invested in it will have been wasted. The widely-held belief that a pragmatic trial provides 'real-world' evidence is seductively appealing, but no one has shown that such trials are more likely to yield the truth. They are certainly more likely to yield results that do not matter.

For these cogent reasons, pragmatic trials are unlikely to be utilized by the drug or device industry for the registration of novel interventions. However, pragmatic trial designs might have a role in answering important questions about already-approved medications, e.g. particularly in addressing uncertainties related to dose of an established drug or the comparative effectiveness of two treatments. The NHLBI is currently carrying out two large-scale pragmatic trials in patients with chronic heart failure. The TRANSFORM-HF trial[9] will enrol 6000 recently-hospitalized patients who will be randomized (open-label) to furosemide or torsemide, with total mortality as the primary endpoint. The SPIRRIT trial[10] (carried out in partnership with the Swedish Heart Foundation and the Swedish Ministry of Health and Social Affairs) will enrol 3200 patients with chronic heart failure and an ejection fraction $>40\%$, who will be randomized (open-label) to receive spironolactone or no spironolactone and be followed for the occurrence of cardiovascular death or hospitalization for heart failure. Ascertainment of study endpoints will be accomplished through national death index or healthcare registry. Neither trial will rely on physical visits by participants to a clinical research site. Conceivably, further technological advances (e.g. geofencing) may allow trialists to reliably obtain information on patient outcomes without any involvement of a physician or other healthcare provider.

Conclusions

If research in cardiovascular disease fails to identify critical causal abnormalities in highly select populations and if drug development continues to focus on the modulation of mechanisms that play only a small contributory role, then traditional large-scale clinical trials involving thousands of patients will be needed to demonstrate worthwhile treatment effects. Because of the need to minimize operational factors that create statistical noise, these trials will continue to be extraordinarily expensive, and current trends in society and in medicine will continue to present important operational challenges. If a traditionally-executed trial demonstrates a modest therapeutic effect but the price of the novel agent carries an innovative premium, the new drug may not be rapidly adopted, and the revenues that were used to justify the trial's expense may not materialize in a timely manner. Pragmatic trials are seductively less expensive, but may yield results that generate little enthusiasm or be difficult to interpret. Therefore, unless the science underlying cardiovascular drug development changes dramatically, we may not need to worry about how large-scale clinical trials in cardiovascular medicine will be performed. They will simply cease to exist.

Conflict of interest: M.P. has received consulting fees from Actavis, Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Cardiorentis, Celyad, Daiichi Sankyo, Gilead, Novartis, NovoNordisk, Relypsa, Sanofi, Synthetic Biologics, Takeda, Teva and ZS Pharma. B.P. has received consulting fees from AstraZeneca, Bayer, KBP Pharmaceuticals, Relypsa, Sanofi, Stealth Peptides, scPharmaceuticals, G3 Pharmaceuticals, Tricida and Sarfez Pharmaceuticals, and also reports ownership interest in Relypsa, KBP Pharmaceuticals, scPharmaceuticals and G3 Pharmaceuticals.

References

1. Packer M. The imminent demise of cardiovascular drug development. *JAMA Cardiol* 2017;**2**:1293–1294.
2. Packer M. Angiotensin neprilysin inhibition for patients with heart failure: what if sacubitril/valsartan were a treatment for cancer? *JAMA Cardiol* 2016;**1**:971–972.
3. Tran G, Zafar SY. Financial toxicity and implications for cancer care in the era of molecular and immune therapies. *Ann Transl Med* 2018;**6**:166.
4. 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.
5. Agarwal R. Implications of blood pressure measurement technique for implementation of Systolic Blood Pressure Intervention Trial (SPRINT). *J Am Heart Assoc* 2017;**6**:e004536.
6. Correia LC, Correia VC, Souza TM, Cerqueira AM Jr, Alexandre FK, Garcia G, Ferreira FR, Lopes FO. Inconsistency prevents the valuable synergism of explanatory and pragmatic trails. *J Evid Based Med* 2018;**11**:105–111.
7. Dal-Ré R, Janiaud P, Ioannidis JP. Real-world evidence: how pragmatic are randomized controlled trials labeled as pragmatic? *BMC Med* 2018;**16**:49.
8. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;**20**:637–648.
9. ToRsemide compArisoN With furoSemide FORManagement of Heart Failure (TRANSFORM-HF). ClinicalTrials.gov Identifier: NCT03296813. https://clinicaltrials.gov/ct2/show/NCT03296813?term=NCT03296813&rank=1 (29 October 2018).
10. Spironolactone Initiation Registry Randomized Interventional Trial in Heart Failure With Preserved Ejection Fraction (SPIRRIT). ClinicalTrials.gov Identifier: NCT02901184. https://clinicaltrials.gov/ct2/show/NCT02901184?term=NCT02901184&rank=1 (29 October 2018).