Pragmatic and mechanistic trials

The United States and many other societies are rapidly moving towards consolidating their health-care systems and allocating their limited resources. They will determine how best to spend these resources by using aggregate information to balance health benefits and risks against the costs of proposed therapeutic strategies. Cardiovascular disease therapies have a special place in this system because the specialty has highly quantified the cost of many of its procedures and treatments; thus, many of the precedents about how health-care systems make decisions using aggregate information about which therapies will be allowed will arise within the cardiovascular arena. One practical translation of this change is the recognition that the era of individual physician autonomy is ending and will be replaced by a more systematic approach to therapeutics with limited options. At the same time, the quantum improvement in the understanding of molecular mechanisms has allowed a plethora of potential therapeutic approaches to be considered.

This situation creates a dilemma: new discoveries have dramatically increased the opportunity to improve the outcome of patients with disease, while at the same time the level of evidence required to introduce a new therapy into a cost-conscious health-care system has escalated. Health-care providers and sponsors of research evaluating new and old therapies must continue to seek the most efficient methods of determining which therapies are most effective so that they will be available to patients and physicians. A critical component of therapeutic development is the type of clinical trial design.

Angiotensin-converting enzyme (ACE) inhibitors provide an excellent opportunity to examine different approaches to the development of knowledge about the efficacy of various therapies. A series of large ACE inhibitor trials have demonstrated a reduction in mortality\(^1\)\(^{-4}\), while smaller studies and substudies of the larger trials have shown an improvement in quality of life\(^5\)\(^{-7}\). Simultaneous with the success of ACE inhibitors as a class, several therapeutic approaches have been implemented in clinical practice based on a combination of pathophysiological reasoning and small supportive studies and have later been found to be detrimental to patient survival. Type 1 antiarrhythmic agents\(^8\), calcium antagonists\(^9\), flosequinan, and phosphodiesterase inhibitors\(^10\), are recent examples of this problem. In perhaps the most historically interesting example of the broad use of a therapy without documentation of benefit, digitalis has been used for centuries in the absence of reliable information about its mortality effects, a shortcoming that has finally been rectified by the unveiling of the first mortality trial evaluating digitalis for the treatment of congestive heart failure\(^11\).

Until the 1980s, the template for clinical trials in the United States was drawn predominantly from a conceptual model emanating from laboratory-based research. The dominant concepts were that the mechanisms of disease must be understood to devise rational therapies and that the most important experiments controlled for as many confounding variables as possible so that the experimental unit could be examined in isolation. Clinical studies based upon this philosophy, termed 'mechanistic' trials, have as their primary goal understanding how a therapy works so that the concept can be brought into the clinical arena and applied in practice.

The Oxford group led in the development of the contrasting philosophy of the 'practical' or 'pragmatic' trial\(^12\). This approach to clinical trials focused on the concept that, regardless of proposed mechanisms, the demonstration of benefit to the patient in a setting that reflects actual clinical practice is the most important goal. A number of considerations about the design, execution, and interpretation of trials have arisen from these contrasting approaches.

For the purposes of discussion it may be useful to consider these two types of trials, pragmatic and mechanistic, as contrasting although, in practice, there is obviously considerable overlap. Table 1 lists the characteristics of each type of trial.

The most obvious major difference is sample size. In order to demonstrate that a therapy produces a particular physiological or biochemical effect, a small sample is generally required. When the goal is to demonstrate that a therapy reduces mortality, however, a much larger sample is needed. A dramatic

Correspondence: Robert M. Califf, MD, Director, Duke Clinical Research Institute, 2024 West Main Street, Durham, NC 27705, U.S.A.
Table 1  Comparison of mechanistic and pragmatic trials

<table>
<thead>
<tr>
<th></th>
<th>Mechanistic trials</th>
<th>Pragmatic trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>Small</td>
<td>Large</td>
</tr>
<tr>
<td>Data collection</td>
<td>Voluminous</td>
<td>Minimal</td>
</tr>
<tr>
<td>Entry criteria</td>
<td>Strict</td>
<td>Minimal</td>
</tr>
<tr>
<td>Ancillary therapy</td>
<td>Strictly controlled</td>
<td>Reflects clinical practice</td>
</tr>
<tr>
<td>Multiple randomization</td>
<td>Forbidden</td>
<td>Allowed</td>
</tr>
<tr>
<td>Cost</td>
<td>Expensive per patient</td>
<td>Inexpensive per patient; costly related to size</td>
</tr>
<tr>
<td>Informed consent</td>
<td>Lengthy</td>
<td>Brief</td>
</tr>
</tbody>
</table>

Table 2  Expected effects of trial size on trial results

<table>
<thead>
<tr>
<th>Deaths</th>
<th>Approx. number of patients randomized if risk = 10</th>
<th>Approx. probability of failing to achieve $P&lt;0.01$ significance if true risk reduction = 1/4</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–50</td>
<td>&lt;500</td>
<td>0.9</td>
<td>Utterly inadequate</td>
</tr>
<tr>
<td>50–100</td>
<td>1000</td>
<td>0.7–0.9</td>
<td>Probably inadequate</td>
</tr>
<tr>
<td>150–350</td>
<td>3000</td>
<td>0.3–0.7</td>
<td>Possibly adequate, possibly not</td>
</tr>
<tr>
<td>350–650</td>
<td>6000</td>
<td>0.1–0.3</td>
<td>Probably adequate</td>
</tr>
<tr>
<td>&gt;650</td>
<td>10 000</td>
<td>&lt;0.1</td>
<td>Definitely adequate</td>
</tr>
</tbody>
</table>

Reprinted with permission13.

demonstration of this issue arose from the efforts of the ISIS-I group to determine the effects of intravenous beta-blockade on mortality in the setting of acute myocardial infarction. As shown in Table 2, a sample size of 10 000 patients was needed despite the fact that the largest previous trial had fewer than 500 patients13. Indeed, since therapies rarely exert an effect of more than a 25% reduction in mortality, similar sample sizes will be needed to fully evaluate the mortality implications of most major cardiovascular therapies. A second key issue in sample size requirements is the importance of unexpected adverse outcomes (such as intracranial haemorrhage with alteplase and sudden death with 'antiarrhythmic agents'). Although these events can occur at a low frequency that can be detected only with large numbers of patients, they may be the most important factor in determining the criteria for patient selection in the implementation of a therapeutic strategy, even when the majority of patients benefit from the therapy.

In mechanistic trials, extreme care must be taken to record in detail all observations pertaining to the mechanisms being examined. In contrast, a general principle of pragmatic trials is that the rate at which patients are enrolled and the accuracy of the data are each inversely proportional to the amount of data collected. Given the large number of patients required, multiple health-care providers and institutions must be involved, which reduces the probability of uniform definition and recording of information. Each data item in a pragmatic trial must be carefully examined as to its value and the cost of collecting it. In addition the pragmatic trial, because of the power to eliminate the effects of random error, can be satisfied with data collection that reflects the vagaries and inaccuracies of everyday clinical practice14.

As in laboratory experiments, mechanistic trials require a substantial number of inclusion and exclusion criteria in order to eliminate factors that might confound the mechanism being evaluated. In contrast, the pragmatic trial welcomes variability in the sample population. Once a therapy is deemed beneficial, practitioners will have to make decisions about whether to use it in patients with relative contraindications and different levels of risk; consequently the pragmatic trial makes an effort to enroll such patients to be able to inform the provider and the patient about the patient-specific risk/benefit situation as accurately as possible.

For similar reasons to the concerns about inclusion and exclusion criteria, ancillary therapy is most often explicitly controlled in mechanistic trials, which increases the likelihood that the experimental issue will be the only source of variability in the study. The pragmatic trial includes ancillary therapies reflective of standard clinical practice. Although this practice enables a therapy to be judged in the context of current clinical practice, pragmatic trial results must be analysed with special diligence and concern to eliminate the possibility that differences in post-randomization treatment — rather than the experimental agent — could have produced the outcome15.
Traditionally, randomizing the same patient into multiple clinical trials has been considered to be undesirable, and perhaps unethical. The reasons for this position are not entirely clear, particularly when only one of the randomizations involves a truly experimental agent with unknown risk. In a mechanistic trial an effort is made to completely control additional therapies so that the experimental and control groups are exposed to the same non-experimental factors. Some designers of pragmatic trials have argued that randomizing the allocation of commonly available therapies in the context of an ongoing trial, when their relative efficacy is uncertain, can only improve the knowledge base about these therapies and thus the clinical relevance of the results of the trials.

The underpinning of mechanistic research is careful, controlled observation in the experimental setting. From this perspective, one would need measurements made in the experiment to be well documented, reproducible, and clear. This type of thinking has led to a policy of 100% auditing of all data items, comparing them with the hospital or clinic record, in many clinical trials submitted for regulatory approval. In the pragmatic trial, the inherent variability in clinical observation is thought to overwhelm the value of this sort of auditing, but systematic spot checking, definition of ‘hard’ endpoints such as mortality, and use of blinded committees to assess important non-fatal endpoints are used to validate the observations.

The per patient cost of a mechanistic trial is often extremely high because of the large number of measurements that must be made, the stringent requirement for control of the clinical environment, and the frequent use of laboratory or invasive procedures. In contrast, the per patient cost of pragmatic trials is often much lower, since routine clinical practice is less significantly affected, the time spent on paperwork is much less, and ancillary testing is minimal. However, because of the large sample sizes and the cost of coordination, the aggregate cost of pragmatic trials can be much larger than typical mechanistic trials.

As a substitute for large, pragmatic trials the combining of multiple mechanistic trials is an attractive approach, using the techniques of systematic overviews or meta-analyses. In cardiovascular disease, the consistency of results was encouraging when meta-analyses and subsequent large trials of thrombolytic therapy and beta-blockers were compared. Recently, however, impressive meta-analysis results with magnesium and nitrates in acute myocardial infarction were not confirmed by large trials. According to some, these results reflect differences in protocol design, patient population, or dosing, while others have argued that in the absence of very large numbers in a single trial statistical certainty cannot be ensured.

The need for both pragmatic and mechanistic trials is highlighted by the ACE inhibitor data. Without proposed mechanisms for a pharmacological effect, new areas of investigation involving neuro-hormonal activation (carvedilol) and angiotensin receptor blockade would not be ongoing. Alternatively, mechanistic trials have identified a number of possible reasons for ACE inhibitors to reduce mortality in heart failure, including reduction of preload and afterload, blunting of neurohormonal response, stabilization of atherosclerotic plaques, vascular remodelling, and tissue effects in the left ventricle. However, only the pragmatic trials have provided definitive evidence that the ACE inhibitors reduce mortality; none of these mechanisms alone or in concert has proven to be responsible, and new mechanisms will likely be uncovered as investigation continues.

In summary, the strengths and disadvantages of pragmatic and mechanistic trials are becoming more evident, but the most effective approach to the timing and integration of mechanistic and pragmatic trials in defining the role of therapies in practice remains highly debatable. The time to launch a large, pragmatic trial and the degree to which meta-analysis should be used to guide treatment remain a matter of opinion, with substantial differences among knowledgeable people. This uncertainty becomes more bearable with the realization that the ability to accumulate large numbers of observations from patients with complex diseases has become possible only with the recent advent of effective communication systems and computer technology. The opportunity to develop a better understanding of how to mix pathophysiology and clinical outcome to define effective treatment strategies will grow as the field of clinical trials emerges from its current stage of infancy. Because of the magnitude of the patient volume needed to demonstrate clinical effectiveness, the clinical and statistical community must organize more effectively to provide an efficient mechanism for the conduct of large trials at lower cost and to develop the academic discipline of clinical investigation. Failure to do so will have tangible effects on the public’s access to improved therapies. The requirement for rapid performance of large, simple trials must not be accomplished to the detriment of excellent pathophysiological research aimed at elucidating disease mechanisms. Instead, increased efficiency in the performance of large trials should allow greater allocation of support to mechanistic studies. An alliance
aimed at developing and maintaining systems for aggregated information about outcome and cost will be needed to enable rational decisions to be made based on reliable information as potential therapies multiply and resources become even more constrained.

R. M. CALIFF
L. H. WOODLIEF
Duke Clinical Research Institute, Durham, NC, U.S.A.

References


Eur Heart J. Vol. 18, March 1997