Pilot trials in Clinical Research

Of What Value Are They?

Joseph Loscalzo, MD, PhD

Pilot trials are exploratory studies limited in size and scope that give insight into the actions, efficacy, and safety of a drug or device but cannot provide definitive support for specific mechanistic or therapeutic claims. Defined in this way, pilot trials have been used to guide clinical and translational research for many years. As studies defined as pilot trials are regularly published in Circulation, I feel it is important for our readers to appreciate their appropriate use and key limitations for clinical and translational cardiovascular research.

Two broad classes of pilot trials can be identified: those designed a priori by the investigators and those redefined a posteriori. The first article published in Circulation identified by its title as a pilot study appeared in 1976, and since that time 146 such studies have graced our pages. Since the current editorial office’s tenure began in 2004, 41 pilot trials have been published in the journal, and many of these were not initially submitted as pilot studies: After reviewing the data carefully and recognizing their limitations, the editors requested that the authors define the trials as pilots to alert the reader to the preliminary nature of the results. It is likely that many trials published before 1976 in Circulation would have been redefined as pilot studies owing to similar limitations and rigorous attention paid to the statistical validity of the conclusions drawn in them.

Clinical pilot studies designed as such a priori are generally performed at the earliest phase of the development of a drug or device when little or no experience has been acquired with its use in humans. For the purpose of this editorial, pilot studies are defined as those clinical trials used to acquire specific essential information about a drug or device before beginning the pivotal trial (ie, the trial that will be used to make specific claims about efficacy and safety). Such information may address the underlying mechanism of a drug or the parameters needed to adjust the operating characteristics of a device in situ. In contrast to pivotal trials, pilot trials are not typically designed to test a critical hypothesis required for drug or device approval; rather, the data obtained from the pilot study are used to optimize the design of subsequent pivotal trials. Pilot trials defined in this way, then, include individual investigator–driven small-sized studies as well as formal industry-sponsored Phase I and early Phase II trials.

The objectives of a clinical pilot trial typically include assessing feasibility (eg, preliminary device performance), exploring eligibility criteria and their practical application for the pivotal randomized controlled trial, ascertaining potential harm (preliminary safety evaluations), studying drug mechanism, validating a method for determining an outcome measure, using a defined drug mechanism to validate a surrogate outcome measure, and evaluating the logistics of pivotal trial performance. The advantages of performing a clinical pilot trial follow from these objectives. Pilot trials can be used to predict the feasibility and operational acceptability of a protocol design planned for a pivotal trial and can achieve this end with comparatively few patients. Thus, the results of a pilot trial can help to guide the effective use of limited (financial and nonfinancial) resources essential for a successfully performed pivotal trial. Two other advantages include their use in identifying unpredicted harm early in the course of drug or device development and assessing the utility of a surrogate end point in the pivotal trial.

Notwithstanding these rational objectives and advantages of pilot clinical trials, their use is also associated with clear disadvantages. The feasibility and acceptability assessments may be misleading if only a limited number of highly motivated centers are included that are not representative of all of the centers in the pivotal trial. Complex pilot trials can be expensive relative to the information they provide. Owing to their size that is typically comparatively small and to the average frequency of clinical end points expected for most pivotal trials in the current era, they are unlikely to provide reliable estimates of sample size requirements for the definitive trial. Similarly, pilot trials are rarely powered adequately to detect harm with respect to clinically important end points, and, by their very design, they are underpowered to provide reliable estimates of benefit.

A major problem, then, with clinical pilot trials is that their results are often overinterpreted, misleading and misleading investigators and interested readers to consider potential benefit or potential harm when the statistical power to do so is woefully inadequate. The statistical basis for this conclusion can be illustrated by first considering the implications of small-sized studies in which an outcome of interest does not occur. With small sample sizes, the likelihood of observing even comparatively common occurrences is low. Yet even when no events of interest are observed, it may be necessary to estimate the true underlying event rate—or, at the very least, the upper limit of that event rate. For the purposes of this illustration, consider a pilot study of n independent patients and x adverse events, for which the probability of the adverse event is r, with 0< r<1. To generate a confidence interval for r in this small-sized sample, the probability of...
observing events can be best described by a binomial distribution in which the upper limit of the exact 1-sided \(100\times(1-\alpha)\%\) confidence interval for the unknown event rate, \(r\), is the value, \(r_{up}\), which yields \(\alpha\), the type I error rate, according to the following:

\[
\sum_{k=0}^{x} \frac{n!}{k!(n-k)!}\left[r_{up}(1-r_{up})^{n-k}\right] = \alpha
\]

The special case of zero events of interest (\(x=0\)) observed in pilot studies of comparatively few patients reduces to:

\[
(1-r_{up})^n = \alpha
\]

This equation can be rearranged to an expression for the upper confidence limit, \(r_{up}\) as:

\[
r_{up} = 1 - \alpha^{1/n}
\]

Solving this equation for a range of values of \(n\) patients from 1 to 15 and \(\alpha=0.05\) yields a range for the upper confidence limit from 0.95 to 0.18, respectively. Thus, when no adverse events occur (ie, \(x=0\)) in such relatively small-sized trials, the great uncertainty in \(r\) sorely limits any meaningful interpretation of the results. The practical implications of this straightforward analysis are clear-cut: one simply cannot infer that a therapy is harmless when no adverse effects are observed in a small-sized pilot trial.

One can similarly calculate the upper confidence limit for non-zero adverse events occurring within a small-sized pilot trial by expanding equation (1). For example, doing so for \(x=1, 2, \) or \(5\) adverse events for a trial of size \(n=15\) patients and \(\alpha=0.05\) yields values for the upper confidence limit, \(r_{up}\), of 0.28, 0.36, and 0.58, respectively. The implications of this calculation are also straightforward: owing to this wide range of the upper bound for the confidence interval, there will be considerable overlap with the confidence interval for the control sample, and, as a result, even frequent adverse events in a small-sized pilot trial do not necessarily predict the true event rate in an adequately sized pivotal trial.

An analysis of the constraints on the lower confidence bound for small-sized trials yields equally informative results with regard to beneficial outcomes. The lower limit of the exact 1-sided confidence interval for the unknown beneficial event rate is the value, \(r_{lo}\), that yields \(\alpha\) according to the following:

\[
\sum_{k=x}^{n} \frac{n!}{k!(n-k)!}\left[r_{lo}(1-r_{lo})^{n-k}\right] = \alpha
\]

This equation can be expanded to determine the lower confidence limit for non-zero beneficial outcomes occurring within a small-sized pilot trial. For example, doing so for \(x=5, 8, \) or \(9\) beneficial outcomes for a trial of size \(n=15\) patients and \(\alpha=0.05\) yields values for the lower confidence limit, \(r_{lo}\), of 0.14, 0.30, and 0.36, respectively. This wide range of values for the lower confidence limit indicates that there will be considerable overlap with the confidence interval for the control sample in a small-sized pilot trial, and, thus, even reasonably frequent beneficial events do not necessarily predict the true event rate in an adequately sized pivotal trial.

Thus, one can neither infer that a therapy leads to harm nor offers benefit when events occur in a small-sized pilot trial because doing so would generate undue optimism or unnecessary pessimism, respectively, in expectations for the pivotal trial. That pilot trials offer little insight into benefit or harm of new therapies is indirectly supported by the fact that ineffective drugs and devices continue to be tested in pivotal trials, and harmful agents continue to gain traction throughout the later stages of the clinical trial process until sufficient numbers of infrequent events are observed in large follow-up studies to define the true adverse event rates.

In view of these problems, it is reasonable to consider alternative approaches to the design and interpretation of clinical pilot studies. First and foremost, pilot trials must be approached rigorously and with the same level of scrutiny as pivotal trials, including public registration. The investigators must identify feasible end points, and the trial must be sized to provide meaningful information about those end points. Clinical end points, beneficial or harmful, must be reported with caution and must include an estimate of the confidence intervals for those end points. Importantly, improved methodology is required to define optimal pilot trial design, analysis, and interpretation of results. The zero event rate problem can, for example, be addressed from a Bayesian statistical perspective, in which a prior distribution that defines the initial uncertainty about the false-positive rate of a small-sized study is specified, and then updated as new evidence is acquired from the ongoing trial. Bayesian approaches, however, cannot be used to reduce sample size in Phase III trials because doing so could lead to an underestimation of risk, especially for infrequent harmful events. Other statistical investigators have suggested that small-sized pilot trials should be designed to test whether any individual subject experiences an effect of benefit or harm, rather than the group on average. Whereas this approach forces the investigator to focus on benefit or harm in each patient, it offers little additional insight into the implications of the trial outcomes for the design of the pivotal trial. Including pilot trial results in the course of an evolving adaptive trial design, although posing its own challenges to data interpretation and therapeutic implications, represents another timely approach to handling pilot trial data efficiently and with appropriate caution. Adaptive designs for Phase III trials can also be developed on the basis of signals from pilot studies; for example, subgroups that may be at great risk of harm might be monitored more frequently than other patients in Phase III trials and the trial design adapted accordingly.

Pilot trial design and analysis is an area of clinical research that warrants further study, as a means to ensure both effective use of limited resources and appropriate interpretation of results. Pilot trial outcomes need to be interpreted so as to avoid inappropriate enthusiasm for potential benefit and inappropriate concern for potential harm when the data do not justify drawing these inferences. A more rigorous statistically rational approach to pilot trials can prevent the exclusion of potentially useful drugs or devices and prevent the inclusion of potentially harmful drugs or devices from further study.

For the readers of *Circulation*, it is important to know that pilot studies are defined as such owing to the uncertainty.
about the generalizability of the results they report. These studies are interesting and quite often novel, but assessment of their therapeutic implications must await adequately sized definitive pivotal trials.

Acknowledgments
The author wishes to thank Drs Elliott Antman, Martin Larson, and Joseph Vita for their helpful comments.

Disclosures
None.

References

Key Words: Editorials / clinical trials as topic / pilot studies / trials
Pilot Trials in Clinical Research: Of What Value Are They?
Joseph Loscalzo

Circulation. 2009;119:1694-1696
doi: 10.1161/CIRCULATIONAHA.109.861625
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/119/13/1694

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/