REVIEW

THE AMERICAN JOURNAL of MEDICINE®

Individual and Composite Study Endpoints: Separating the Wheat from the Chaff

Robert Goldberg, PhD,[a,b] Joel M. Gore, MD,[c] Bruce Barton, PhD,[a] Jerry Gurwitz, MD[b,c]

[a]Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester; [b]Meyers Primary Care Institute, Worcester, Mass; [c]Department of Medicine, University of Massachusetts Medical School, Worcester.

ABSTRACT

We provide an overview of the individual and combined clinical endpoints and patient-reported outcomes typically used in clinical trials and prospective epidemiological investigations. We discuss the strengths and limitations associated with the utilization of aggregated study endpoints and surrogate measures of important clinical endpoints and patient-centered outcomes. We hope that the points raised in this overview will lead to the collection of clinically rich, relevant, measurable, and cost-efficient study outcomes.
 • *The American Journal of Medicine (2014) 127, 379-384*

KEYWORDS: Clinical epidemiology; Study design; Study outcomes and endpoints

The information obtained and reported from observational clinical/epidemiological research studies and randomized clinical trials (RCTs) provides clinicians, health policymakers, and prevention-oriented practitioners with much needed information about the causes of disease, their prevention, and the most effective ways to manage and improve the prognosis of individuals diagnosed with disease as well as improve the general health status of communities.[1-4] The validity, reliability, and relevance of these studies is based, in part, on the size and descriptive characteristics of the study population, their representativeness to the broader universe of patients with the condition under study, accurate measurement of key exposure and potentially confounding factors, and careful and unbiased assessment of primary and secondary study endpoints.

While observational epidemiologic investigations and RCTs remain the cornerstone approaches in the development and evaluation of new therapies and lifestyle intervention approaches, especially the "gold standard" RCTs, these investigations are often scientifically and logistically complex, with considerable costs associated with their design, conduct, and analysis. Given the enormous costs, resources, and logistics involved in carrying out these investigations, researchers and funding agencies continue to explore novel approaches and strategies, such as pragmatic clinical trials and quasi-experimental designs that would result in more efficient and cost-effective approaches to the enhanced primary and secondary prevention of chronic and infectious diseases at both the individual and community levels.[5-9]

The selection of carefully considered, measured, and categorized study outcomes in these investigations is crucial to the successful assessment, meaning, and eventual incorporation of the study results into clinical practice and broader public policy initiatives. Randomized trials and appropriately designed and performed observational studies can provide complementary insights into a broad range of clinical problems, particularly in the current era of more widespread implementation of electronic health records at individual medical practices and in larger health care systems.[10] To provide meaningful results, these investigations must have sufficient statistical power to address clinically important questions, balance a variety of known and unknown potentially confounding factors between the respective comparison

Funding: National Institutes of Health (RO1 HL35434). Partial salary support is provided to JMG and RG by National Institutes of Health grant 1U01HL105268. The development of this publication was partially supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award number UL1TR000161 (BB).

Conflict of Interest: None.

Authorship: All authors had access to the data and a role in writing this manuscript.

Requests for reprints should be addressed to Robert J. Goldberg, PhD, Division of Epidemiology of Chronic Diseases and Vulnerable Populations, Department of Quantitative Health Sciences, University of Massachusetts Medical School, 368 Plantation Street, Worcester, MA 01605.

E-mail address: Robert.Goldberg@umassmed.edu

0002-9343/$ -see front matter
http://dx.doi.org/10.1016/j.amjmed.2014.01.011

groups (in RCTs), and provide estimates of treatment effects with narrow confidence intervals.

The purpose of this article is to describe the use of individual and combined clinical endpoints and patient-reported outcomes in clinical research studies and the strengths and limitations associated with the utilization of aggregated endpoints and surrogate measures of important clinical outcomes.

CLINICAL SIGNIFICANCE

- Issues involved in the collection and measurement of high-quality and meaningful outcomes data in clinical and epidemiological research investigations are discussed.
- Collection of clinical endpoints and patient-reported outcomes, use of single versus multiple study endpoints, prespecification of primary and secondary study outcomes, effect of endpoint selection on study sample size, and strengths and limitations of using composite and surrogate endpoints are highlighted.

STUDY ENDPOINTS

There are a variety of endpoint/outcome measures that can be investigated in a clinical/epidemiological research project. These include patient-associated morbidity (eg, recurrent episodes of disease), mortality (total and cause-specific), quality of life (general and disease-specific), health services utilization, and changes in various lifestyle practices and physiologic parameters over the course of an observational longitudinal study or RCT.

In addition to the conventional "hard" (eg, morbidity, mortality) event-type endpoints typically examined in clinical research investigations, and more "soft" study endpoints such as hospitalizations, symptomatology, and changes in selected physiologic measures, patient-reported outcomes are being used more frequently in the study of chronic diseases and their precursor conditions. These outcomes represent the status of a patient's health condition and are elicited directly from the patient without any interpretation of the patient's responses by a health care provider.[11] These outcomes are being utilized on a more frequent basis in clinical research because they are of considerable importance to patients and their families; are able to be measured effectively with standardized instruments; are generally brief in nature; and can utilize different means of administration to assess these self-reported outcomes over the course of a longitudinal study or RCT.[12-15] For example, questionnaires assessing quality of life, social support, and cognitive status are being used more frequently in research studies, as are more novel approaches such as cell phone applications, to record daily pain levels in patients with various underlying illnesses, such as rheumatoid arthritis. Indeed, the CONsolidated Standards Of Reporting Trials (CONSORT) statement has been recently updated to include standards for reporting patient-centered outcomes in RCTs.[12,16]

USE OF SINGLE VERSUS MULTIPLE STUDY CLINICAL ENDPOINTS AND PATIENT-REPORTED OUTCOMES

The majority of observational studies and RCTs prespecify both a primary and, typically, several secondary study endpoints/outcomes, and estimate the sample size for the study based on clinically meaningful differences in the primary study outcome between intervened and nonintervened individuals in RCTs and between exposed and nonexposed individuals in longitudinal studies. For example, many clinical trials that have evaluated the use of novel therapies in patients hospitalized with acute myocardial infarction have examined differences in hospital case-fatality rates in patients receiving, as compared with those not receiving, a study drug of interest, and have estimated the sample size for these trials based on what were considered to be meaningful differences in the primary trial outcome of all-cause mortality.

Clinical and public health researchers need to carefully consider the pertinent endpoints they intend to monitor a priori, and the accuracy, time, and costs associated with their measurement and potentially independent validation/adjudication. Indeed, even the choice of more conventional "clinical" outcomes can be fraught with difficulty in measurement and interpretation.

For example, the multicenter Rapid Early Action for Coronary Treatment (REACT) trial was designed to examine the effects of a multipronged community intervention on patient's care-seeking behavior in adults presenting with signs and symptoms of acute coronary disease to more than 40 hospitals in 10 states throughout the US.[17] This trial recruited enough patients to detect meaningful differences in prehospital delay times in the 10 pair-matched intervention and reference communities.[17] This principal trial outcome was defined as the time interval from self-reported acute symptom onset to arrival at the emergency department, as recorded in hospital medical records. However, the quantification of patient's care-seeking behavior is fraught with potential difficulties and problems, including the extent and accuracy of patient recall and the systematic elicitation and recording of this information by health care professionals. Furthermore, in observational studies and clinical trials, to standardize classification of what may appear to be a relatively straightforward "hard" clinical endpoint, such as cardiovascular-related mortality, a study adjudication committee may be created to develop predefined criteria for this endpoint.

EFFECT OF ENDPOINT SELECTION ON STUDY SAMPLE SIZE

Once the primary study outcome has been agreed upon, the investigators will need to determine, especially for longitudinal studies that involve following patients over a