



Published in final edited form as:

Clin Trials. 2012 February ; 9(1): 113–124. doi:10.1177/1740774511432726.

Midcourse correction to a clinical trial when the event rate is underestimated: the Look AHEAD (Action for Health in Diabetes) Study

Frederick L Brancati^a, Mary Evans^b, Curt D Furberg^c, Nancy Geller^d, Steven Haffner^e, Steven E Kahn^f, Peter G Kaufmann^d, Cora E Lewis^g, David M Nathan^h, Bertram Pittⁱ, and Monika M Safford^g on behalf of the Look AHEAD Study Group

^aThe Johns Hopkins University, Baltimore, MD, USA

^bNational Institute of Diabetes, Digestive, and Kidney Diseases, Bethesda, MD, USA

^cWake Forest University, Winston-Salem, NC, USA

^dNational Heart, Lung, and Blood Institute, Bethesda, MD, USA

^eSan Antonio, TX, USA

^fVA Puget Sound Health Care System, University of Washington, Seattle, WA, USA

^gThe University of Alabama Birmingham, Birmingham, AL, USA

^hMassachusetts General Hospital, Boston, MA, USA

ⁱThe University of Michigan, Ann Arbor, MI, USA

Abstract

The Look AHEAD (Action for Health in Diabetes) Study is a long-term clinical trial that aims to determine the cardiovascular disease (CVD) benefits of an intensive lifestyle intervention (ILI) in obese adults with type 2 diabetes. The study was designed to have 90% statistical power to detect an 18% reduction in the CVD event rate in the ILI Group compared to the Diabetes Support and Education (DSE) Group over 10.5 years of follow-up.

The original power calculations were based on an expected CVD rate of 3.125% per year in the DSE group; however, a much lower-than-expected rate in the first 2 years of follow-up prompted the Data and Safety Monitoring Board (DSMB) to recommend that the Steering Committee undertake a formal blinded evaluation of these design considerations. The Steering Committee created an Endpoint Working Group (EPWG) that consisted of individuals masked to study data to examine relevant issues.

The EPWG considered two primary options: (1) expanding the definition of the primary endpoint and (2) extending follow-up of participants. Ultimately, the EPWG recommended that the Look AHEAD Steering Committee approve both strategies. The DSMB accepted these modifications, rather than recommending that the trial continue with inadequate statistical power.

Trialists sometimes need to modify endpoints after launch. This decision should be well justified and should be made by individuals who are fully masked to interim results that could introduce

© The Author(s), 2011

Author for correspondence: Frederick L Brancati, Johns Hopkins University, 2024 East Monument St, 2-619, Baltimore, MD 21205, USA. fbrancat@jhmi.edu.

Reprints and permission: <http://www.sagepub.co.uk/journalsPermissions.nav>

bias. This article describes this process in the Look AHEAD study and places it in the context of recent articles on endpoint modification and recent trials that reported endpoint modification.

Introduction

Weight loss is commonly recommended to overweight and obese adults, especially those with type 2 diabetes, in order to reduce cardiovascular risk; however, the effect of weight loss on cardiovascular disease (CVD) outcomes has never been tested definitively [1]. The results of epidemiologic studies have been inconsistent, and randomized controlled trials of weight loss have generally focused on short-term changes in intermediate endpoints like blood pressure and serum lipids [2,3]. The Look AHEAD (Action for Health in Diabetes) Study [4] (Clinicaltrials.gov Identifier: NCT00017953) was designed to be the definitive test of the long-term health benefits of a lifestyle intervention aimed at weight loss in adults with type 2 diabetes.

The *primary hypothesis* was that the lifestyle intervention would reduce the incidence of a composite endpoint of incident CVD defined as cardiovascular death (including fatal myocardial infarction and stroke), nonfatal myocardial infarction, or nonfatal stroke. The *secondary hypothesis* was that the lifestyle intervention would reduce the incidence of a composite endpoint of all-cause death or incident CVD-related secondary outcomes defined in aggregate as myocardial infarction, stroke, coronary artery bypass graft (CABG) surgery or percutaneous coronary (PC) intervention, hospitalization for congestive heart failure (CHF), carotid endarterectomy, or surgical bypass or percutaneous intervention for peripheral arterial disease.

The sample size was based on the aim of detecting an 18% difference in the primary endpoint in the intensive lifestyle intervention (ILI) compared to the Diabetes Support and Education (DSE) control group. (During the design of the trial, targets of 15%–20% for interventions were identified as conveying significant public health benefit. The Steering Committee selected an 18% intervention effect on which to base power projections because this appeared feasible and required a cohort sufficiently large to meet other objectives of the trial.) We made the following assumptions: (1) The rate of incident CVD in the DSE group would be 3.125% per year (corresponding to the projected rate of incident CVD in a population of overweight and obese adults with type 2 diabetes eligible for participation in Look AHEAD); (2) the cohort would be recruited uniformly over 2.5 years of follow-up; (3) in all, 2% of participants would be lost to endpoint ascertainment annually; and (4) participants lost to follow-up would be similar to their counterparts in regard to both treatment and endpoint risk. Based on these assumptions, we determined that 5000 participants followed up for a maximum of 11.5 years would yield 92% power (with two-sided $\alpha = 0.05$) to detect an 18% relative difference in the composite primary endpoint (i.e., an absolute event rate of 3.125 per 100 person-years in the DSE group versus an absolute event rate of 2.562 per 100 person-years in the ILI group).

The 3.125% event rate was a best estimate based on reported CVD event rates among individuals similar to those who were to be recruited for the study. Specifically, we used longitudinal data from diabetic participants in the Atherosclerosis Risk in Communities (ARIC) study [5] and the Cardiovascular Health Study (CHS) [6]. We assumed that 75% of Look AHEAD participants would have no history of CVD and that their age distribution would represent an equal mixture of the ARIC (45–64 at baseline) and CHS (65–75 at baseline) samples. The overall event rate for the combined ARIC + CHS diabetic population was estimated to be 3.72%. This event rate was adjusted 5% upward to account for silent myocardial infarctions (which were not part of the ARIC/CHS event estimates) and then

20% downward to account for both a healthy volunteer effect and the overall decline in CVD event rates in the United States since the 1980s. This produced the anticipated 3.125% event rate used for power calculations ($3.72\% \times 1.05 \times 0.80 = 3.125\%$).

Lower-than-expected event rates

As Look AHEAD reached the 2-year mark, the Data and Safety Monitoring Board (DSMB), charged with monitoring study progress, noted that the actual event rate in the DSE group was much lower than expected and informed the Steering Committee through the National Institute of Diabetes, Digestive and Kidney Diseases (NIDDK) Project Office. The Steering Committee therefore reconsidered the power projections in light of this lower-than-expected event rate. For example, at a hypothetical 2% event rate in the DSE group over 11.5 years of follow-up (considerably lower than the anticipated 3.125% rate), the trial would have 80% power to detect an 18% effect on the primary endpoint – less than the originally planned 90% power but still acceptable. However, as Look AHEAD reached the 3-year mark, the DSMB observed that the event rate in the DSE group was only 0.7% per year. In response to continuing DSMB concern, the Steering Committee created an Endpoint Working Group (EPWG) to carefully consider the ramifications of the lower-than-expected event rate and to recommend alternative approaches to preserve the study's integrity and maximize its scientific value.

EPWG

The EPWG included five Look AHEAD clinical investigators well versed in clinical trials, three National Institutes of Health (NIH) scientists, and two clinical trials experts not affiliated with Look AHEAD (Table 1). Two members also served on the study's Adjudication Committee, responsible for ascertaining endpoints. All EPWG members were masked to study results, including the three NIH scientific officers (Evans, Kaufman, and Geller). The EPWG was charged with evaluating all relevant aspects of the study bearing on the lower-than-expected event rate, including the duration of follow-up and definition of the primary endpoint. To avoid bias and to maximize masking of the study investigators, the EPWG only examined outcome data from the DSE control group and shared only summary data when making its recommendations to the Steering Committee. Neither the EPWG nor the Steering Committee was ever privy to the overall event rate (DSE and ILI combined), so data on the event rate in the DSE group did not effectively unblind them. The EPWG convened regularly to deliberate and to review updated event rates. A timeline of events and key study dates are provided in Figure 1.

Why Look AHEAD event rates were lower than expected

The EPWG identified three possible reasons for the unexpectedly low CVD event rates:

1. Secular trends have resulted in lower CVD incidence and CVD mortality in the United States: The downward trend in CVD appears to result from (a) improved control of dyslipidemia and high blood pressure, which have reduced the risk of clinically significant atherosclerosis and (b) improved care of chronic and acute coronary syndromes that have reduced CVD mortality [7,8].
2. Study participants were even healthier than expected: During trial design, the Coordinating Center and Steering Committee relied heavily on data from cohort studies like ARIC [5] and CHS [6]. The pattern of enrollment in cohort studies, however, is markedly different from trials: cohort studies like these enroll 50% or more of eligible individuals in the community, whereas trials forego representativeness in favor of motivation to participate fully in study interventions.

3. The Graded Exercise Test (GXT) excluded participants most likely to develop CVD events: The study included a GXT as an inclusion criterion based on concerns about safety and liability related to initiating an exercise program in adults with type 2 diabetes at high risk for CVD. The GXT effectively excluded some higher-risk participants (e.g., those with prevalent symptomatic CVD) who demonstrated electrocardiographic or blood pressure abnormalities during the test. The prospect of a GXT may have also discouraged some individuals with exercise-induced symptoms from attempting to enroll in the trial.

The Steering Committee weighed all three of these hypothetical concerns during the design phase and deflated the projected event rate by 20%. In retrospect, however, the original event rate projections were simply not conservative enough.

Options considered by the EPWG

1. Watchful waiting: Because of the healthy volunteer effect, incident CVD events may be delayed, yielding low event rates early in the study. To determine whether the lower-than-expected event rates were accurate and firm, the EPWG recommended watchful waiting beginning April 2004. After 3 years of watchful waiting, the event rate remained low. To be certain that the persistently low event rate was not related to delays in endpoint ascertainment and reporting, the EPWG recommended a 'Data Sweep' in late 2007. The Data Sweep confirmed that the cumulative event rate in the DSE group had remained nearly linear since 2003.
2. Extend study duration: Absent a substantial increase in event rates, extending the duration of the study would result in increased events and potentially boost power. However, with a stable event rate that was about one-third of the projected rate and trending linear, a two- to threefold extension of the study's duration would be required to obtain the required number of events. A more feasible and acceptable extension (e.g., 1–2 years) would provide too few events to achieve the required power.
3. Broaden definition of the primary endpoint: The third option considered by the EPWG was to broaden the definition of the primary endpoint. Such consideration was not taken lightly, since adherence to primary endpoints as specified in the original protocol is generally considered to be a cornerstone of clinical trial methodology. However, the EPWG considered the following:
 - a. During the design phase, the Steering Committee had debated the use of a broader endpoint definition, since obesity was thought to have adverse health effects well beyond CVD (e.g., certain cancers) but settled on traditional CVD endpoints primarily for comparability with other studies. After Look AHEAD was initiated, the evidence for other adverse health effects of obesity had grown [9,10].
 - b. After Look AHEAD was initiated, there was a shift in the consensus among CVD trialists regarding standard study endpoints, prompted by advances in diagnostic testing and downward secular trends in myocardial infarction and CVD death [11].
 - c. An influential paper provided a roadmap for modifying endpoints midstream in a long-term clinical trial when made by trialists blinded to treatment effect [12]. The framework addressed the possibility of extending duration and/or redefining study endpoints.

- d. All things considered, the DSMB deemed it scientifically more informative to evaluate the effect of the Look AHEAD intervention on a broader endpoint rather than to terminate the trial for futility, especially since the intervention had produced and sustained a substantial contrast between study groups in body weight and physical activity.

Additional endpoints considered by EPWG to expand the original composite endpoint

As the EPWG considered expanding the primary composite endpoint to include additional endpoints, it posed the following five questions: (1) Does obesity consistently predict the occurrence of the endpoint in longitudinal epidemiologic studies? (2) Is the endpoint of sufficient clinical importance to serve in a composite endpoint alongside myocardial infarction, stroke, and CVD death? (3) Is the endpoint related to atherosclerotic CVD? As originally written, the protocol enshrined atherosclerotic CVD as the study's main focus. Insofar as 'new' primary endpoints fall under this general category, they fit more naturally with the original conceptual framework. (4) How susceptible is the endpoint to ascertainment bias? It is crucial to avoid including in the primary endpoint any event that might be susceptible to ascertainment bias; for example, chronic stable angina might be differentially detected in the Lifestyle Intervention Group, because of more frequent exercise (which might induce symptoms) and more frequent contact with study staff (which might trigger safety measures leading to medical evaluation and treatment). (5) How acceptable is the endpoint to the study's stakeholders and scientific and clinical audience? Midstream changes in primary endpoints are apt to be viewed with some suspicion by Look AHEAD's stakeholders and audience: the less traditional or familiar the endpoint, the more suspicious they might be.

Criteria for an 'ideal' additional endpoint

The ideal additional endpoint(s) would therefore fit the following five criteria: (a) related to obesity, (b) high clinical importance, (c) related to atherosclerotic CVD, (d) low risk of ascertainment bias, and (e) acceptable to stakeholders and audience. With these criteria in mind, the EPWG reviewed the following nine endpoints:

1. *All-Cause Mortality* – Already a secondary endpoint. Since cardiovascular mortality was already a primary endpoint, moving 'all-cause' up from secondary to primary would amount to adding non-CVD deaths (e.g., cancer death and accidental death).
2. *Hospitalized Angina* – Angina of sufficient concern to warrant hospitalization (see Figure 2).
3. *Urgent Revascularization* – Angina of sufficient concern to warrant revascularization during the hospitalization (i.e., surgery and angioplasty).
4. Hospitalized CHF.
5. *Incident Chronic Kidney Disease (CKD)* – Based on serum creatinine at annual follow-up visits.
6. *Incident Obesity-Related Cancer* – Defined as cancer of the prostate, corpus uteri or endometrium, cervix, ovary, colon/rectum, biliary tract, esophagus, liver, pancreas, kidney, and non-Hodgkin's lymphoma, all of which appear to occur more frequently in obese adults.
7. *Incident Left Ventricular Hypertrophy (LVH)* – Based on study electrocardiograms (ECGs).

8. *Deep Venous Thrombosis/Pulmonary Embolism* – As abstracted from hospital records.
9. *Fractures* – Hip, upper leg, pelvis, knee, lower leg, ankle, foot (but not toe), coccyx, spine, lower arm, wrist, hand (not finger), elbow, upper arm, or shoulder.

The EPWG members selected these nine endpoints based on their experience in other trials of diabetes treatment and CVD prevention and on their knowledge of the epidemiology of type 2 diabetes and obesity.

Deliberations regarding possible additional endpoints

The EPWG evaluated each of these possible additional primary endpoints in detail. These deliberations are summarized in Table 2, with particular attention to how each endpoint might meet the five criteria listed above.

This deliberation led the EPWG to narrow the range of potentially acceptable endpoints to four: all-cause mortality, hospitalized angina, urgent revascularization, and hospitalized CHF. The rationale for excluding other endpoints was as follows: (1) Incident CKD, as defined by using serum creatinine to estimate glomerular filtration rate, was considered to have insufficient clinical importance; (2) cancer was too incongruent with our original endpoints and the evidence to determine which cancers are ‘obesity related’ was still unsettled; (3) LVH is primarily asymptomatic and suffers from disagreement about definition; (4) deep venous thrombosis/pulmonary embolism has a wide range of gravitas, is less related to atherosclerosis than are other vascular endpoints, and has not been widely used in trials or epidemiologic studies as part of a composite vascular endpoint; and (5) fractures have a wide range of severity and are inversely associated with body mass index.

Thus, the EPWG carefully considered the four remaining endpoints: all-cause mortality, hospitalized angina, urgent revascularization, and hospitalized CHF. These were reduced to three when the EPWG determined that virtually all of the urgent revascularizations in Look AHEAD occurred in participants who otherwise met criteria for hospitalized angina. The deliberations regarding all-cause mortality, hospitalized angina, and hospitalized CHF are summarized below.

All-Cause Mortality

The argument in favor was that all-cause mortality is the bottom line for patients and physicians and would provide a way to capture effects on important non-CVD events, like cancer or liver disease. The argument against was that (1) these non-CVD events are best treated as secondary endpoints, since the primary hypothesis focuses on CVD per se, and (2) all-cause mortality may introduce ‘noise’ in the form of accidental deaths and nonobesity-related cancers (e.g., brain and lung).

Hospitalized Angina

The argument in favor was that hospitalized angina would capture the ‘aborted’ myocardial infarctions related to secular improvements in acute cardiac care; would be consistent in tone, therefore, with recent thinking on CVD endpoints (see Luepker *et al.* [11]); and is fully congruent with the original hypothesis. The argument against was that (1) it might be difficult to distinguish ‘urgent’ cases from ‘chronic’ cases, the latter of which would be susceptible to ascertainment bias in an unblinded study, and (2) it might be difficult to agree upon a specific definition. However, the current Look AHEAD definition of hospitalized angina (see Figure 2) mitigated both concerns: the definition clearly excluded chronic stable angina and the definition had already been smoothly implemented by the Adjudication

Committee for several years without generating significant disagreements among committee members.

Hospitalized CHF

The argument in favor was that CHF is common and important, it was already a component of the composite secondary endpoint, and it might be improved by weight loss along a variety of physiologic pathways (e.g., better exercise tolerance, reduced reliance on thiazolidinediones, and improved lung function). The argument against was that (1) CHF is a heterogeneous syndrome related not only to atherosclerosis but also to hypertension, renal disease, and other causes (e.g., valvular heart disease) generally not discernable from the records available to the study adjudicators, and (2) it is often difficult to distinguish from other causes of acute dyspnea, especially chronic obstructive pulmonary disease and pneumonia.

After deliberation, the EPWG unanimously favored hospitalized angina and unanimously rejected hospitalized CHF. A large majority was against all-cause mortality, but a minority favored it. Further discussion led to a consensus that the additional primary endpoint plus all-cause mortality should be an additional major secondary analysis in the main results of this study.

Event rates if primary endpoint definition were expanded

Having evaluated the possible options for expanding the primary endpoint on purely scientific grounds, the EPWG then turned to the practical matter of whether the potential additional endpoint occur frequently enough to augment the overall event rate. The coordinating center determined that adding hospitalized angina to the primary endpoint definition would approximately double the event rate in the DSE (control) group to 1.25%–1.35% per year.

Final EPWG recommendation for response to lower-than-expected event rate

After over 2 years of monitoring, research, and deliberation, the EPWG made the following recommendations to the Look AHEAD Steering Committee as a means to address the lower-than-expected event rate:

1. Expand the primary endpoint to include hospitalized angina
2. Extend the duration of the study by 24 months
3. Specify two additional secondary endpoints
 - a. Original primary composite endpoint
 - b. Expanded primary composite endpoint plus all-cause mortality

Net effect of these recommendations on statistical power

Either (a) expanding the primary endpoint to include hospitalized angina without extending study duration or (b) extending study duration by 2 years without expanding the primary endpoint would increase statistical power to detect an effect of 18% from roughly 50% to only 70%. Adopting both recommendations would push statistical power to roughly 75%, assuming no change in the event rate in the second half of the study. If the underlying event rate increased in the second half of the study, then adoption of both recommendations would increase power to above 80% – that is, into the conventional range for randomized controlled trials. EPWG members recommended applying the same effect size (18%

difference in risk of the composite primary endpoint) to the expanded endpoint, because the causal pathways to hospitalized angina were so similar to the pathways leading to myocardial infarction and CVD death. The change in the definition required an adjustment in the formal statistical monitoring of the trial to preserve the prior level of spending while transitioning to rules based on the expanded endpoint.

Acceptability of recommendations to the scientific audience outside Look AHEAD

Even with the strongest scientific rationale, the EPWG considered how these recommendations might be viewed by the scientific audience outside Look AHEAD. In the end, the EPWG consensus was that the recommended modifications to the study protocol would be well accepted for the following four reasons: (1) they are responsive to generally recognized secular trends in CVD, (2) they are concordant with the study's original conceptual framework and primary hypothesis, (3) they were developed and proposed by a group that was aware only of event rates in the comparison group and otherwise fully blinded to treatment effects, and (4) they were proposed 4 years before the originally planned date of study close-out (December 2012).

Final decision by the Look AHEAD Steering Committee

The EPWG presented its recommendations to the Steering Committee on 8 April 2008. The Steering Committee unanimously supported the recommendation. Throughout the decision-making process, the DSMB chose to remain silent on the specifics of the protocol change, because it believed that its unblinded status could otherwise introduce bias. Beyond asking the Steering Committee to review the low event rate in the DSE group, the DSMB never indicated (a) whether it would have ended the trial for futility had the protocol not been changed or (b) whether the specific changes allayed its original concern. Although neither the EPWG nor the Steering Committee were privy to DSMB deliberations, both groups recognized that the DSMB was in the awkward position of having prompted an inquiry into event rates and endpoint definition without having the latitude to comment on the investigators findings or response. From the perspective of the EPWG and the Steering Committee, the investigators had responded to the DSMB and saw only that the DSMB continued to allow the trial to continue.

Review of prior literature

Modifying endpoints after launch

Rigorous adherence to study protocol is an acknowledged cornerstone of trial methodology. In practice, however, it appears that many trialists modify their approach after the study goes into the field. In fact, Chan [13,14] and Mathieu [15] estimate that about one-third of properly registered trials undergo a modification of the primary outcome(s) following registration. That such changes often appear to favor the intervention [15] heightens the fear that failure to adhere to predetermined endpoints can inflate type I error rate.

Nonetheless, trialists do recognize that there may be appropriate reasons for modifying endpoints after launch. For example, in 2002, Wittes [12] argued for some 'agility' in study design, especially for long-term trials that may see secular trends in standard of care or relevant endpoints, or that may simply have been underpowered based on a priori calculations. Indeed, over the past decade, several large trials have changed endpoints after launch. Table 3 summarizes changes made in six major trials over the past 7 years. Most have done so apparently without sacrificing integrity, impact, or acceptability. The one possible exception is the PROActive trial [16], which drew criticism for relying heavily on a

secondary endpoint that was introduced only a few weeks before the trial was closed [17,18]. In two of these trials (FIELD [19] and PEACE [20]), the definition of the primary endpoint was expanded to increase the event rate in the face of unexpectedly low power – similar to the situation in Look AHEAD. In two trials, the primary endpoint was narrowed in light of new data that came to light after launch (NAVIGATOR [21,22] and EUROPA [23]).

Wittes [12] and Evans [24] agree that individuals who are blinded to trial results should be the ones who decide about changing the primary endpoint; unblinded individuals might be tempted to modify endpoints so as to favor positive results. Evans [24] poses a series of questions to trialists who are considering a change: (1) What data triggered the review? (2) Have interim results been reviewed? (3) Who is making the decision? Ideally, endpoint reviews would be triggered by factors other than low event rates (e.g., secular trends in endpoint classification) and there would be no review of interim results. In Look AHEAD, internal data (lower-than-expected event rates) triggered the review and the DSMB and unblinded Coordinating Center investigators had reviewed interim results. However, the decision makers were investigators in the EPWG and Steering Committee and NIH project scientists, all of whom were fully blinded. From Evans' perspective, therefore, these answers seem to put Look AHEAD in a 'gray zone' with regard to propriety – hence we prepared this article to make our decision-making process fully transparent [24].

Modifying study duration after launch

In studies of CVD, extending the duration of follow-up is a logical remedy for decreased power related to lower-than-expected event rates. Because CVD is the leading cause of death in the general population, longer follow-up will inevitably lead to more events, especially in middle-aged and older populations with underlying CVD risk factors, like type 2 diabetes and obesity. Less has been written about decision making leading to a change in study duration.

Longer duration poses a special challenge in trials designed to test a behavioral intervention for two reasons: (1) It may be difficult to maintain the contrast between the intervention group and the comparison group in later years; (2) as the cohort ages, the accumulating comorbid conditions might interfere with adherence. The decision to increase Look AHEAD's duration by 2 years required reconsenting participants for extended duration, but the intervention was otherwise unmodified.

Recent trends in cardiovascular endpoint definitions

In the later part of the twentieth century, CVD researchers began to observe secular trends in the incidence, treatment, and natural history of coronary heart disease that had important ramifications for endpoint definition [25]. To create a new standard for CVD endpoints across studies, an international panel of experts representing the American Heart Association, the National Heart, Lung, and Blood Institute (NHLBI), and the Centers for Disease Control, the World Heart Federation, and the European Society of Cardiology released a scientific statement on 'Case Definitions for Acute Coronary Heart Disease in Epidemiology and Clinical Research Studies' [11]. The statement identifies a standard approach in using 'unstable angina pectoris' as part of a component definition of coronary heart disease along with nonfatal myocardial infarction (MI). According to this scheme, 'unstable angina' is defined as new or changing cardiac symptoms with positive ECG findings. The additional Look AHEAD primary endpoint of 'hospitalized angina' (see Figure 2) is designed to capture episodes of unstable angina as part of a composite definition of CVD. Indeed, over the past decade, most major cardiovascular trials have used composite endpoints, typically with three or four components [26]. This is certainly true for trials of diabetes treatment [27].

Lessons learned

The experience regarding endpoint modification in Look AHEAD teaches several lessons for long-term trials.

First, in a long-term trial, it would be prudent to prespecify plans for checking event rates during the course of the trial. The plans should include options that could be triggered if the observed rate is much lower than expected. The options might include (a) extending duration, (b) expanding the endpoint, or (c) stopping the trial for lack of statistical power. In Look AHEAD, the Coordinating Center conducted regular rate checks under the supervision of the DSMB, but there was no prespecified plan to react when the rates were low.

Second, even absent low event rates, it might be useful to prespecify a point at which the endpoint definition is reexamined in light of secular trends and prevailing practice.

Third, it might be helpful to use the number of events to drive duration, rather than setting duration a priori based on estimated event rates. Of course, this approach requires that the funding source makes a somewhat open-ended commitment in terms of duration, which may not be feasible in some circumstances.

Finally, it is good to be as conservative as feasible about projected event rates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

See Appendix (supplementary material).

Funding This study is supported by the Department of Health and Human Services through the following cooperative agreements from the National Institutes of Health: DK57136, DK57149, DK56990, DK57177, DK57171, DK57151, DK57182, DK57131, DK57002, DK57078, DK57154, DK57178, DK57219, DK57008, DK57135, and DK56992. The following federal agencies have contributed support: National Institute of Diabetes and Digestive and Kidney Diseases; National Heart, Lung, and Blood Institute; National Institute of Nursing Research; National Center on Minority Health and Health Disparities; Office of Research on Women's Health; the Centers for Disease Control and Prevention; and the Department of Veterans Affairs. This research was supported in part by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases. The Indian Health Service (IHS) provided personnel, medical oversight, and use of facilities. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the IHS or other funding sources.

Additional support was received from The Johns Hopkins Medical Institutions Bayview General Clinical Research Center (M01RR02719) and the Prevention & Control Core of the Baltimore Diabetes Research & Training Center (P60KD079637); the Massachusetts General Hospital Mallinckrodt General Clinical Research Center and the Massachusetts Institute of Technology General Clinical Research Center (M01RR01066); the University of Colorado Health Sciences Center General Clinical Research Center (M01RR00051) and Clinical Nutrition Research Unit (P30 DK48520); the University of Tennessee at Memphis General Clinical Research Center (M01RR0021140); the University of Pittsburgh General Clinical Research Center (GCRC) (M01RR000056), the Clinical Translational Research Center (CTRC) funded by the Clinical & Translational Science Award (UL1 RR 024153) and NIH grant (DK 046204); and the Frederic C. Bartter General Clinical Research Center (M01RR01346)

The following organizations have committed to make major contributions to Look AHEAD: FedEx Corporation; Health Management Resources; LifeScan, Inc., a Johnson & Johnson Company; OPTIFAST® of Nestle HealthCare Nutrition, Inc.; Hoffmann-La Roche, Inc.; Abbott Nutrition; and Slim-Fast Brand of Unilever North America.

References

1. [accessed 6 December 2010] The practical guide: Identification, evaluation, and treatment of overweight and obesity in adults. 2000. NIH Publication No. 00-4084. http://www.nhlbi.nih.gov/guidelines/obesity/prctgd_c.pdf
2. Andersen RE, Wadden TA, Bartlett SJ, et al. Effects of lifestyle activity vs. structured aerobic exercise in obese women: A randomized trial. *JAMA*. 1999; 281(4):335–40. [PubMed: 9929086]
3. Effects of weight loss and sodium reduction intervention on blood pressure and hypertension incidence in overweight people with high-normal blood pressure. The Trials of Hypertension Prevention, phase II. The Trials of Hypertension Prevention Collaborative Research Group. *Arch Intern Med*. 1997; 157(6):657–67. [PubMed: 9080920]
4. Ryan DH, Espeland MA, Foster GD, et al. Look AHEAD (Action for Health in Diabetes): Design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes. *Control Clin Trials*. 2003; 24(5):610–28. [PubMed: 14500058]
5. Rosamond WD, Folsom AR, Chambless LE, Wang CH, the ARIC Investigators. Coronary heart disease trends in four United States communities. The Atherosclerosis Risk in Communities (ARIC) study 1987-1996. *Int J Epidemiol*. 2001; 30(Suppl. 1):S17–S22. [PubMed: 11759846]
6. Siscovick DS, Fried L, Mittelmark M, et al. Exercise intensity and subclinical cardiovascular disease in the elderly. The Cardiovascular Health Study. *Am J Epidemiol*. 1997; 145(11):977–86. [PubMed: 9169906]
7. Ford ES, Ajani UA, Croft JB, et al. Explaining the decrease in U.S. deaths from coronary disease, 1980–2000. *N Engl J Med*. 2007; 356:2388–98. [PubMed: 17554120]
8. Paynter NP, Sharrett AR, Louis TA, et al. Paired comparison of observed and expected coronary heart disease rates over 12 years from the Atherosclerosis Risk in Communities Study. *Ann Epidemiol*. 2010; 20(9):683–90. [PubMed: 20696408]
9. Berrington de Gonzalez A, Hartge P, Cerhan JR, et al. Body-mass index and mortality among 1.46 million white adults. *N Engl J Med*. 2010; 363:2211–19. [PubMed: 21121834]
10. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med*. 2003; 348(17):1625–38. [PubMed: 12711737]
11. Luepker RV, Apple FS, Christenson RH, et al. Case definitions for acute coronary heart disease in epidemiology and clinical research studies: A statement from the AHA Council on Epidemiology and Prevention; AHA Statistics Committee; World Heart Federation Council on Epidemiology and Prevention; the European Society of Cardiology Working Group on Epidemiology and Prevention; Centers for Disease Control and Prevention; and the National Heart, Lung, and Blood Institute. *Circulation*. 2003; 108(20):2543–9. [PubMed: 14610011]
12. Wittes J. On changing a long-term clinical trial midstream. *Stat Med*. 2002; 21(19):2789–95. [PubMed: 12325094]
13. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *JAMA*. 2004; 291(20):2457–65. [PubMed: 15161896]
14. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*. 2004; 171(7):735–40. [PubMed: 15451835]
15. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA*. 2009; 302(9):977–84. [PubMed: 19724045]
16. Dormandy JA, Charbonnel B, Eckland DJ, et al. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAZone Clinical Trial in macroVascular Events): A randomised controlled trial. *Lancet*. 2005; 366(9493):1279–89. [PubMed: 16214598]
17. Freemantle N. How well does the evidence on pioglitazone back up researchers' claims for a reduction in macrovascular events? *BMJ*. 2005; 331(7520):836–8. [PubMed: 16210286]

18. Skyler JS. PROactive: A sad tale of inappropriate analysis and unjustified interpretation. *Clin Diabetes*. 2006; 24(2):63–5.
19. Keech A, Simes RJ, Barter P, et al. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): Randomised controlled trial. *Lancet*. 2005; 366(9500):1849–61. [PubMed: 16310551]
20. Braunwald E, Domanski MJ, Fowler SE, et al. Angiotensin-converting-enzyme inhibition in stable coronary artery disease. *N Engl J Med*. 2004; 351(20):2058–68. [PubMed: 15531767]
21. Califf RM, Boolell M, Haffner SM, et al. Prevention of diabetes and cardiovascular disease in patients with impaired glucose tolerance: Rationale and design of the Nateglinide and Valsartan in Impaired Glucose Tolerance Outcomes Research (NAVIGATOR) Trial. *Am Heart J*. 2008; 156(4):623–32. [PubMed: 18946890]
22. NAVIGATOR Study Group. Holman RR, Haffner SM, et al. Effect of nateglinide on the incidence of diabetes and cardiovascular events. *N Engl J Med*. 2010; 362(16):1463–76. [PubMed: 20228402]
23. Fox KM, the EUROPA study investigators. On reduction of cardiac events with perindopril in stable coronary artery disease investigators. Efficacy of perindopril in reduction of cardiovascular events among patients with stable coronary artery disease: Randomised, double-blind, placebo-controlled, multicentre trial (the EUROPA study). *Lancet*. 2003; 362(9386):782–8. [PubMed: 13678872]
24. Evans S. When and how can endpoints be changed after initiation of a randomized clinical trial? *PLoS Clin Trials*. 2007; 2(4):e18. [PubMed: 17443237]
25. Newby LK, Alpert JS, Ohman EM, Thygesen K, Califf RM. Changing the diagnosis of acute myocardial infarction: Implications for practice and clinical investigations. *Am Heart J*. 2002; 144(6):957–80. [PubMed: 12486420]
26. Lim E, Brown A, Helmy A, Mussa S, Altman DG. Composite outcomes in cardiovascular research: A survey of randomized trials. *Ann Intern Med*. 2008; 149(9):612–17. [PubMed: 18981486]
27. McCallum RW, Fisher M. Review: Comparing cardiovascular outcomes in diabetes studies. *Br J Diabetes Vasc Dis*. 2006; 6:111–18.
28. Baigent C, Landray MJ, Reith C, et al. The effects of lowering LDL cholesterol with simvastatin plus ezetimibe in patients with chronic kidney disease (Study of Heart and Renal Protection): a randomised placebo-controlled trial. *Lancet*. 2011; 377:2181–92. [PubMed: 21663949]
29. LaRosa JC, Grundy SM, Waters DD, et al. Intensive lipid lowering with atorvastatin in patients with stable coronary artery disease. *N Engl J Med*. 2005; 352:1425–32. [PubMed: 15755765]

2001	First approved protocol with original endpoints 02/16/01 First participant randomized 08/22/01
2002	
2003	DSMB asks Steering Comm to check on event rate 09/10/03 AHA recommends changes in CVD endpoint definition 09/03 EPWG established 10/02/03
2004	
2005	
2006	
2007	EPWG shares DSE rate data with Steering Comm 04/18/07
2008	Steering approves changes offered by EPWG 04/08/08 EPWG and SC cease to receive rate data 04/09/08 NIDDK Director approves changes 07/08
2009	
2010	
2011	
2012	Original study close out 12/31/12
2013	
2014	Revised study close out 12/31/14

↑
Watchful
Waiting
↓
↑
Sweep
↓

Figure 1. Timeline of events and key study dates. The Look AHEAD study. EPWG: Endpoint Working Group; AHA: American Heart Association; DSMB: Data and Safety Monitoring Board; NIDDK: National Institute of Diabetes, Digestive and Kidney Diseases; AHEAD: Action for Health in Diabetes.

Hospitalized angina is defined as a “Yes” to question 1 below, *plus* a ‘yes’ to either 2 or 3 and at least one ‘yes’ to questions 4 through 8.

1. Is this a case of angina pectoris (including unstable angina) requiring and/or occurring during hospitalization? (Angina is defined as either chest pain, tightness, or shortness of breath produced by myocardial ischemia that does not result in infarction--usually caused by coronary insufficiency).

If YES, then mark all that apply:

2. Physician diagnosis of angina and receiving medical treatment for angina on this admission (e.g., nitrate, beta-blocker, or calcium-channel blocker).

3. Physician diagnosis of angina and receiving medical treatment for angina on this admission plus current medical record documenting a history of coronary heart disease by previous catheterization or revascularization procedure.

4. CABG surgery or other revascularization procedure on this admission.

5. 70% or greater obstruction of any coronary artery on angiography on this admission.

6. Horizontal or down-sloping ST-segment depression or abnormal ST elevation ≥ 1 mm on exercise or pharmacological stress testing with pain on this admission or immediately preceding and leading to this admission.

7. Scintigraphic or echocardiographic stress test positive for ischemia on this admission or immediately preceding and leading to this admission.

8. Resting ECG shows horizontal or down-sloping ST depression or abnormal ST elevation ≥ 1 mm with pain that is not present on ECG without pain on this admission.

Figure 2.
Look AHEAD adjudication criteria for hospitalized angina.
AHEAD: Action for Health in Diabetes.

Table 1**Membership of Endpoint Working Group**

Frederick L Brancati, MD, MHS; Look AHEAD PI, Johns Hopkins University; Chair
Mary Evans, PhD; Look AHEAD Project Scientist, NIDDK
Curt Furberg, MD; Trialist, Wake Forest University
Nancy Geller PhD; Director, Office of Biostatistics Research, NHLBI
Steven Haffner, MD, MPH; Look AHEAD PI, University of Texas, San Antonio (2000-2009).
Steven E Kahn, MB, ChB; Look AHEAD PI, VA Puget Sound Health Care System and University of Washington
Peter G Kaufmann, PhD; Look AHEAD Project Officer, NHLBI
Cora Beth Lewis, MD, MPH; Look AHEAD PI and Chair of the Adjudication Committee, University of Alabama, Birmingham
David Nathan, MD; Look AHEAD PI, Massachusetts General Hospital
Bertram Pitt, MD; Cardiologist, Trialist, University of Michigan
Monika M Safford, MD, MPH; Look AHEAD Co-Investigator and Adjudication Committee member, University of Alabama, Birmingham

All members were and remain blinded to interim results.

PI, Principal Investigator.

Table 2

Summary of EPWG deliberations regarding full range of possible endpoints

Endpoint	Relationship to obesity	Clinical gravitas	Related to ASCVD?	Susceptibility to ascertainment bias in an unblinded trial	Acceptability to stakeholders and audience
All-cause mortality	Yes, BMI mortality	++++	Only to the extent that cancer and other non-CVD deaths are influenced by underlying CVD.	Minimal to none	Widely accepted as a secondary endpoint.
Hospitalized angina	Yes, BMI CVD	+++	Yes, strongly	Yes. LSI staff and/or PCPs react to exercise-induced symptoms.	Difficulty in adjudication and concerns reascertainment bias have discouraged wider use.
Urgent revascularization	Yes, BMI CVD	+++	Yes, strongly	Possibly. LSI staff and/or PCPs may react to exercise-induced symptoms	Concerns about ascertainment bias have traditionally discouraged experts in unblinded trials
Hospitalized CHF	Yes, BMI CVD CHF; BMI BP CHF; and BMI Hypo-ventilation RHF; and BMI hyperglycemia TZD CHF	+++	Yes, but also related to non-ASCVD causes. But COPD and pneumonia – sometimes mistaken for CHF – are unrelated to CVD	Possibly. LSI staff and/or PCPs react to exercise-induced symptoms	Difficulty in adjudication and complexity of underlying causes have traditionally discouraged experts. But pro-BNP now eases distinction from COPD and pneumonia
Incident CKD (stage 3/4)	Yes, BMI BP CKD; And BMI hyperglycemia CKD	++	Yes, insofar as CKD is a strong risk factor for CVD and vice versa and they share many upstream risk factors	Minimal, since periodic 'gold standard' assessments of CKD are conducted per study protocol. Some concerns about classification bias across ethnic groups	Endpoints based on serum creatinine (e.g., doubling) are widely used as primary endpoint in trials of CKD prevention. Not commonly used in CVD trials. Variability and threshold effects raise concerns
Obesity-related cancer	Yes, BMI several obesity-related cancers. But, no evidence that weight loss improves mediating factors	+ to +++	No	Minimal to none	Combining cancer with CVD endpoints would be unusual, since the causal pathways are different
Incident LVH	Yes, BMI BP LVH	++	Yes	Minimal to none	Commonly used as a surrogate but not commonly combined with clinical endpoints. A variety of conflicting ECG definitions raises concerns
DVT/PE	Yes, BMI and DVT DVT PE	+ to +++	Thrombosis, yes; atherosclerosis, no	Possibly. LSI staff and/or PCPs react to exercise-induced symptoms	Not commonly used. DVT occurrence often does not affect long-term prognosis
Fractures	Yes, but inverse: MI fewer fractures	+ to +++	No	Minimal to none	Inverse relation inappropriate for primary endpoint

BMI: body mass index; LSI: lifestyle intervention; BP: blood pressure; LVH: left ventricular hypertrophy; CKD: chronic kidney disease; ECG: electrocardiogram; CHF: congestive heart failure; PE: pulmonary embolism; CVD: cardiovascular disease; TZD: thiazolidinedione; DVT: deep venous thrombosis; PCP: primary care provider; ASCVD: atherosclerotic cardiovascular disease; RHF: right heart failure; COPD: chronic obstructive pulmonary disease; BNP: brain natriuretic peptide.

Table 3

Selected features of some CVD and diabetes trials in which the primary endpoints were modified after launch, 2003–2011

Trial/PI/ reference	Design/ intervention	Original primary endpoint(s)	How endpoints were modified/ when changes were made	Rationale for modification
SHARP, <i>Lancet</i> , 2011 [28]	Parallel group simvastatin plus ezetimibe vs. placebo	Composite of nonfatal myocardial infarction or any cardiac death, any stroke, or any arterial revascularization excluding dialysis access procedures	Narrowed the endpoint by changing 'any cardiac death' to coronary death and 'any stroke' to nonhemorrhagic stroke. Decision made in October 2009, about 4 years after randomization and about 2 years before close	Review of blinded data on clinical outcomes showed that about one-third of the original protocol-defined primary outcome was noncoronary cardiac deaths or hemorrhagic strokes, which had been found in other trials not to be prevented by statin therapy
NAVIGATOR, Califf <i>et al.</i> , <i>NEJM</i> , 2010 [22]	2 × 2 factorial design; nateglinide +/- Valsartan	Two coprimary endpoints: (1) new onset of type 2 diabetes and (2) extended CVD endpoint, including CVD death, nonfatal MI, nonfatal stroke, hospitalization for HF, arterial revascularization, and hospitalized unstable angina	Third coprimary endpoint added: (3) core CVD endpoint (extended CVD endpoint minus revascularization and unstable angina). Decision made in 2008, 6 years after launch and about 2 years before close in 2010	In a meta-analysis conducted after launch, the investigators discovered a pattern of risk in prior studies that suggested that Valsartan might have stronger effects on a narrower, harder CVD endpoint
TNT, LaRosa <i>et al.</i> , <i>NEJM</i> , 2006 [29]	Parallel group; 10 mg vs. 80 mg of atorvastatin daily; goal of LDL: 100 vs. 75 mg/dL.	Major cardiovascular event, defined as death from CHD, nonfatal nonprocedure-related MI, resuscitation after cardiac arrest, or fatal or nonfatal stroke	In February 2003, the Steering Committee added stroke (fatal or nonfatal) to the primary efficacy outcome. Decision was made 5 years after launch in July 1998 and about 2 years before close in 2005	After launch, there was accumulating evidence of the beneficial role of statins in reducing the risk of stroke
PROactive [16]	Matched parallel group; pioglitazone vs. placebo, in addition to other meds	All-cause mortality, nonfatal MI (including silent MI), stroke, acute coronary syndrome, cardiac intervention, leg amputation (above ankle) or revascularization	Added a new composite secondary outcome: all-cause mortality, nonfatal MI (excluding silent MI), and stroke. Decision made in May 2005, 4 years after launch and 2 weeks before database was locked	During an interim analysis, investigators recognized that they had failed to focus on this clinically important composite endpoint
FIELD, <i>Lancet</i> , 2005 [19]	Parallel group; once-daily micronized fenofibrate 200 mg vs. placebo	CHD death	Nonfatal MI added to primary endpoint definition. Decision made in December 2002, about 4 years after launch and about 3 years before close	To maintain the study's power in light of interim data on medication adherence, commencement of open-label lipid lowering treatment, and CVD event rates
PEACE, <i>NEJM</i> , 2004 [20]	Parallel group; trandolapril vs. placebo	CVD death or nonfatal MI	Expanded the primary endpoint to include revascularization. Decision made in October 1997, 11 months after launch, after 1584 patients had undergone randomization; 6 years before the trial ended in 2003	To maintain the study's power in the face of slower than expected recruitment. The revised sample size dropped from 14,100 to 8100
EUROPA [23]	Parallel group perindopril vs. placebo	Composite of total mortality, nonfatal MI, unstable angina, and cardiac arrest with successful resuscitation	Dropped non-CVD mortality. Dropped unstable angina without evidence of myocardial necrosis. To accrue sufficient number of events, the trial was also extended by 1 year. Decision made in January 2002, about 4 years after launch and about 1 year prior to close	(1) In light of advances in diagnostic assays, unstable angina without myocardial necrosis was no longer judged an appropriate endpoint given its subjective diagnosis and favorable prognosis; (2) non-CVD mortality turned out to be higher than expected, about 40%. ACE inhibition would be unlikely

Trial/PI/ reference	Design/ intervention	Original primary endpoint(s)	How endpoints were modified/ when changes were made	Rationale for modification
				to affect non-CVD mortality

CVD: cardiovascular disease; CHD: coronary heart disease; MI: myocardial infarction; HF: heart failure; LDL: low density lipoprotein; ACE: Angiotensin converting enzyme.