Effects of the End Point Adjudication Process on the Results of the Perindopril Protection Against Recurrent Stroke Study (PROGRESS)

Toshiharu Ninomiya, MD, PhD; Geoff Donnan, MD; Neil Anderson, MD, PhD; Chris Bladin, MD; Brian Chambers, MD; Gary Gordon, MD; Norman Sharpe, MD; John Chalmers, MD, PhD; Mark Woodward, PhD; Bruce Neal, MD, PhD; for the PROGRESS Collaborative Group

Background and Purpose—End point adjudication committees (EPAC) are widely used in large-scale clinical trials to ensure the robustness of diagnosis for end points.

Methods—The Perindopril Protection Against Recurrent Stroke Study (PROGRESS) was a double-blind randomized trial of blood pressure lowering in 6105 participants with pre-existing cerebrovascular disease. Separate estimates of the effects of randomized treatment were determined using Cox regression models that were based on the unadjudicated events initially reported by the investigator and on the final events assigned by the EPAC.

Results—There were 992 strokes initially reported by the investigators and 894 (90%) retained these diagnoses after adjudication by the EPAC. The hazard ratios (95% CIs) for the effect of randomized treatment on stroke were 0.74 (0.64 to 0.85) based on the investigator diagnoses and 0.72 (0.62 to 0.83) based on the EPAC diagnoses (P homogeneity =0.7). For each stroke subtype reported, the corresponding numbers of diagnoses (investigators/EPAC) were ischemic (593/565), hemorrhagic (124/111), and unknown (124/93) with no impact of the EPAC review on the estimates of treatment effects (all P homogeneity >0.3). There was likewise no detectable effect of reclassification of diagnoses for the effect estimates calculated for myocardial infarction or the main causes of death (all P homogeneity >0.5).

Conclusion—The EPAC process had no discernible impact on the trial conclusions. Very large trials powered to detect effects on stroke subtypes might obtain real scientific gain from an EPAC, but in the case of PROGRESS, the value of the EPAC was in the reassurance it provided. (Stroke. 2009;40:2111-2115.)

Key Words: adjudication  clinical trial  outcomes

Cau sespecific mortality and major nonfatal events are the chief outcomes for many large-scale research projects. Much attention is given to the accurate classification of these events because large-scale, multicenter studies are considered to present particular challenges to the consistent diagnosis of outcomes.1 Standardized definitions of events and protocols for their assignment are used,2–6 but the possibility of misclassification remains. Accordingly, the design of most recent large-scale trials in vascular disease has included an end point adjudication committee (EPAC),7–9 charged with assuring the validity of the diagnoses for key end points such as death, stroke, and coronary heart disease. An EPAC may be particularly important for stroke studies because the diagnosis of stroke subtypes is complex.10 Likewise, where applications for new or expanded indications for a therapy are anticipated, an EPAC may be considered essential to meet regulatory requirements. There are, however, rather few data that have showed clear impact of EPACs on the main trial conclusions1,7 which is problematic because EPACs may be resource-intensive and expensive to operate.11 The goal of these analyses was to assess the impact of the EPAC on the conclusions of Perindopril Protection Against Recurrent Stroke Study (PROGRESS), a large double-blind randomized trial of blood pressure-lowering in patients with a history of cerebrovascular disease.12

Methods

Design of the PROGRESS

The design and main results of the PROGRESS study have been described in detail elsewhere.12,13 In brief, 6105 patients with a history of a cerebrovascular disease within the previous 5 years and no clear indication for, or contraindication to, treatment with an
angiotensin-converting enzyme inhibitor and a diuretic were recruited from 172 centers in 10 countries between 1995 and 1997. Eligible participants were subsequently randomly allocated to active treatment comprised of a regimen based on perindopril (4 mg daily) with the addition of indapamide (2.5 mg daily; or 2 mg daily in Japan) or matching placebo. The mean follow-up was 3.9 years.

Definitions for Adjudicated Outcomes
The primary outcome in the PROGRESS trial was “total stroke,” defined as an acute disturbance of focal neurological function with symptoms lasting more than 24 hours (or resulting in earlier death) and thought to be due to either intracranial hemorrhage or ischemia (International Classification of Diseases, 9th Revision codes 431, 432, 433, 434, 436, or 437). All stroke events were subclassified as ischemic (codes 433 or 434), hemorrhagic (code 431), or unknown type (codes 436 or 437). The Trial of Org 10172 in Acute Stroke Treatment criteria were used to further subclassify ischemic stroke as lacunar stroke, cardioembolic stroke, large artery stroke, or unknown ischemic type. The other outcomes for which adjudication was done were myocardial infarction and cause of death. The diagnosis of myocardial infarction (MI) was based on the combination of an appropriate clinical history supported by electrocardiographic changes and/or an elevation of cardiac enzymes or other biochemistry, hematology, radiology, and autopsy findings), which were translated as necessary before being provided to the EPAC. Using these additional data, the EPAC either confirmed or refuted the diagnosis initially assigned by the site investigators using the same standardized definitions laid out in the protocol. When the initial diagnosis made by the site investigator was refuted, either the event was determined not to meet the protocol criteria for a “serious adverse event” or an alternate diagnosis was provided and assigned the applicable International Classification of Diseases, 9th Revision code. As such, by the end of the end point reporting and adjudication process, every outcome had 2 diagnoses recorded, an initial diagnosis assigned by the site investigator and an adjudicated diagnosis assigned by the EPAC (Figure 1).

Outcome Diagnosis by the Investigator
All serious events recorded in the trial were first identified and a diagnosis assigned by one of the 172 site investigators using a standard “serious event form” according to the protocol definitions described previously (Figure 1). Investigators were a diverse group of physicians of different experience levels trained in multiple disciplines, mostly neurology or hypertension and in general, the participating sites were considered to be leading academic centers in the country. All outcomes were assigned initial International Classification of Diseases, 9th Revision codes on the basis of the site investigators’ diagnoses.

Outcome Diagnosis by the EPAC
For all deaths, and all other outcomes assigned a diagnosis of nonfatal stroke or nonfatal MI by the site investigators, additional information about the event was sought using a “stroke, myocardial infarction or death assessment form.” The supplementary data collected about the event included copies of investigation reports (eg, biochemistry, hematology, radiology, and autopsy findings), which were translated as necessary before being provided to the EPAC. Using these additional data, the EPAC either confirmed or refuted the diagnosis initially assigned by the site investigators using the same standardized definitions laid out in the protocol. When the initial diagnosis made by the site investigator was refuted, either the event was determined not to meet the protocol criteria for a “serious adverse event” or an alternate diagnosis was provided and assigned the applicable International Classification of Diseases, 9th Revision code. As such, by the end of the end point reporting and adjudication process, every outcome had 2 diagnoses recorded, an initial diagnosis assigned by the site investigator and an adjudicated diagnosis assigned by the EPAC (Figure 1).

Statistical Methods
For each outcome, the agreement of the diagnoses made by the site investigator and the EPAC was calculated. Percent agreement was calculated for each outcome by dividing the number of events assigned that diagnosis by the EPAC that was also initially assigned that same diagnosis by the site investigators. Kappa statistics were also calculated to summarize the overall coherence of the diagnoses for stroke subtypes and causes of death.

The impact of the EPAC adjudication process on the conclusions of the PROGRESS study was estimated by calculating the effects of study treatment on the outcomes of interest using univariate Cox regression models. The models were first fitted using the diagnoses initially reported by the site investigators and second using the diagnoses reported by the EPAC. All analyses were done according to the principle of intention to treat with relative risk reductions reported as percentage reductions ($[1 - hazard ratio] \times 100$). The homogeneity between the effects of treatment estimated using the site investigator diagnoses and the EPAC diagnoses was addressed through a test of the null hypothesis that the EPAC hazard ratio was equal to the investigator hazard ratio for each outcome. This test exploits the sequential nature of the diagnostic process; the EPAC diagnoses were made in the knowledge of the investigators’ diagnosis (Figure 1). In addition, we estimated the percentage of error saved by the EPAC process through a comparison of mean square errors (MSEs). Assuming that the effect estimate based on the EPAC diagnoses is unbiased, the percentage of error saved by the EPAC process was estimated as $\frac{\text{MSE}(e) - \text{MSE}(i)}{\text{MSE}(i)} \times 100$, where MSE(e) is the variance of the log hazard ratio based on the EPAC diagnoses and MSE(i) is the variance of the log hazard ratio based on the investigator diagnoses plus the squared difference between the log hazard ratio estimate based on the investigator and the EPAC diagnoses. The median value and interquartile range of the percentage of error saved by the EPAC process (which had a skewed distribution) was estimated from 1000 bootstrap samples. The SAS software package (SAS Institute, Inc, Cary, NC) was used to perform all statistical analyses. A probability value <0.05 was considered to indicate a result unlikely to have arisen by chance alone.

Results

Diagnoses Made by Site Investigators and the EPAC
There were 992 diagnoses of stroke reported by the site investigators and 894 (90.1%) were confirmed by the EPAC. Of the 98 strokes not confirmed by the EPAC, 33 (3.3%) were assigned alternative diagnoses (transient ischemic attack, 10; seizure, 5; dementia, 5; subdural hematoma, 3; and various others, 10) and for the remainder (65), the event was deemed not to satisfy the protocol definition of a serious adverse
event. Among the 894 EPAC-confirmed strokes, there were a substantial number of reclassifications of the stroke subtypes by the EPAC (Table 1), although the kappa statistic of 0.53 indicated overall good agreement between the diagnoses made by the investigators and the EPAC. Agreement was 56% for diagnoses of ischemic stroke, 94% for hemorrhagic stroke, and 64% for stroke of unknown type. For ischemic stroke subtypes, agreement was lowest for large artery stroke (35%) and highest for ischemic stroke of unknown type (72%; Table 1).

There were 281 MIs reported by the site investigators and 234 (83.3%) of these were confirmed by the EPAC. Of the 47 MIs not confirmed by the EPAC, 14 (5.0%) were assigned alternate diagnoses and for the remainder (33 [11.7%]), the event was considered not to meet the criteria defining a serious adverse event. All deaths reported by the site investigators were confirmed by the EPAC with 94% agreement for cardiovascular causes of death, 93% for cancer deaths, and 69% for the other nonvascular deaths (Table 2) and the kappa statistic of 0.79 indicated excellent agreement.

Impact of End Point Adjudication on Estimates of Effect of Randomized Treatment
The hazard ratios (95% CIs) for the effect of randomized treatment on stroke were 0.74 (0.64 to 0.85) based on the investigator diagnoses and 0.72 (0.62 to 0.83) based on the EPAC diagnoses (P homogeneity = 0.7; Figure 2). Estimates of treatment effect based on investigator diagnoses and EPAC diagnoses were also highly comparable for MI and the main causes of death. The estimated median percentage of error saved by adjudication was 2% (interquartile range, −2% to 10%) for the primary outcome stroke and 12% (1% to 24%) for MI (Figure 2). The median errors avoided by adjudication were estimated to be larger for some stroke subtypes (15% [4% to 31%] for cardioembolic stroke, 32% [7% to 51%] for large artery stroke, and 45% [30% to 59%] for unknown ischemic stroke), although in no case was there evidence of heterogeneity in the relative risk reductions calculated on the basis of the diagnoses initially made by the site investigators and the diagnoses assigned by the EPAC (all P homogeneity >0.2; Figure 2).

Discussion
These analyses show that the end point adjudication process used in the PROGRESS trial had no substantive impact on the effect estimates obtained from the trial. Although it is likely that the use of an EPAC in PROGRESS was viewed positively by most users of the study results,1 the findings reported here raise important questions about the science underpinning the use of EPACs in clinical trials such as PROGRESS. Adjudicating the 1898 serious events reported by the investigators in the PROGRESS trial consumed very considerable resources and, although a formal estimate was not possible, certainly cost many hundreds of thousands of dollars to achieve. With the costs of conducting clinical research spiraling, it may be that EPACs could reasonably be excluded from many future studies.

The reasons why the EPAC process had no clear impact on the estimates of effect in the PROGRESS trial are likely severalfold. First, and most importantly, the site investigators, despite their diversity, made mostly correct diagnoses of the main study outcome, stroke, using the prespecified diagnostic criteria provided. Second, the misclassification that did occur was not substantial and was nondifferential between intervention and control groups. Small amounts of random misclassification would be expected to produce a slight underestimation of the true effect.
mate of the treatment effect and the exclusion of some events by the EPAC process would be expected to produce slightly wider CIs.\textsuperscript{11,17} Third, the evaluation of this adjudication procedure is subject to the same issues of power that affect most pieces of research. A process such as that evaluated here is only likely to show clear effects if the chance of diagnostic error is substantial and there are a large number of events that are misclassified. In PROGRESS, the events most prone to misclassification were the stroke subtypes,\textsuperscript{10} but there were only relatively small numbers of each recorded.\textsuperscript{15} Finally, if the effect of randomized treatment is broadly consistent across the different types of outcomes being adjudicated, then misclassification (and its correction) would be expected to have little impact on the conclusions drawn. In regard to this last point, targeting adjudication efforts toward event types that interventions might differentially influence may increase the value of the process.

This report is not the first to question the true value of EPACs for clinical trials,\textsuperscript{1,11} although most previous evaluations of EPACs have focused on the positive impact that the adjudication process has had on the reliability of the final diagnoses.\textsuperscript{2–6,8} Certainly, there can be little doubt that an EPAC results in more reliable diagnoses being made, even if the impact on the conclusions pertaining to the effectiveness of randomized interventions may be fairly limited in most cases. The real potential for adjudication processes like those used in PROGRESS is probably for studies of other designs or seeking to achieve different goals. For example, a study wanting to establish the incidence of stroke and its subtypes would have its conclusions seriously adversely impacted by misclassification at the level seen in the PROGRESS trial and would obtain real scientific gain from an EPAC. Likewise, studies using softer end points, which are more difficult to diagnose, studies with fully or partially unblinded designs, and studies in which outcomes derive from less reliable sources or are not defined by prespecified diagnostic criteria (eg, mortality registers\textsuperscript{18–20}) may achieve important gains from an EPAC process.

Figure 2. Effects of randomized treatment calculated using diagnoses assigned by site investigators and diagnoses assigned by the EPAC. cardio, cardioembolic; Invest, Investigator; IQR, interquartile range. *The median percentage of error in the treatment effect saved by the adjudication process (assuming the estimate based on the EPAC diagnoses is true).
Summary
In conclusion, these data highlight the need for more careful consideration of the likely impact of an EPAC on the conclusions of clinical trials. The appointment of an EPAC has become a knee-jerk reaction in the design of large-scale clinical trials, but actually requires the same careful scientific consideration as other aspects of the trial design. With clinical research costs escalating, it is requisite on researchers to consider the likely impact of an EPAC on the conclusions of clinical trials. The appointment of an EPAC is no doubt of substantial importance. Although the reassurance that the EPAC provided to the users of the PROGRESS trial was no doubt of substantial importance, there are probably more efficient ways of achieving this goal.

Sources of Funding
PROGRESS was funded by grants from Servier, the Health Research Council of New Zealand, and the National Health and Medical Research Council (NHMRC) of Australia. The study was designed, conducted, analyzed, and interpreted by the investigators independent of all sponsors. The conduct of the present analyses has been supported by fellowships awarded to T.N. by the Banyu Life Science Foundation, Japan, and the Foundation for High Blood Pressure Research in Australia.

Disclosures
None.

References
Effects of the End Point Adjudication Process on the Results of the Perindopril Protection Against Recurrent Stroke Study (PROGRESS)
Toshiharu Ninomiya, Geoff Donnan, Neil Anderson, Chris Bladin, Brian Chambers, Gary Gordon, Norman Sharpe, John Chalmers, Mark Woodward and Bruce Neal
for the PROGRESS Collaborative Group

Stroke. 2009;40:2111-2115; originally published online April 9, 2009;
doi: 10.1161/STROKEAHA.108.539601

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://stroke.ahajournals.org/content/40/6/2111

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke 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 Stroke is online at:
http://stroke.ahajournals.org//subscriptions/