How the ACCOMPLISH trial was stopped early for efficacy

Paul Gallo
Novartis Pharmaceuticals
Rutgers Biostatistics Day
April 3, 2009
Background

- Optimal therapy strategies for hypertension continue to evolve.
- Current guidelines recommend initial therapy with a combination of drugs, with diuretics included in the regimen.
- Classes of drugs work by differential mechanisms, so their benefits may extend beyond simply their BP-lowering effects.
- The mechanisms of calcium channel blockers and ACE inhibitors suggest that combinations might be particularly beneficial.
ACCOMPLISH trial

- The **ACCOMPLISH** trial investigated the hypothesis that an ACE inhibitor (*benazepril*) combined with the CCB *amlodipine* would result in better outcomes than the same ACE inhibitor combined with a diuretic.

- Compared initial combination therapies:
  - benazepril / amlodipine (*B/A*)
  - benazepril / hydrochlorothiazide (*B/H*)

- were compared using a dose titration scheme to achieve BP control, in patients at high risk for CV events in US and Nordic countries.
ACCOMPLISH design

- The primary endpoint was a CV mortality / morbidity composite (time-to-event).
- The trial was designed for 90% power to detect a 15% risk reduction (i.e., hazard ratio = 0.85) for B/A patients.
- The target sample size was 12000 patients, in an event-driven trial planned for 1642 patients to reach the primary endpoint.
- Endpoints underwent a process of central adjudication.
Monitoring - DSMB

- An independent *Data Safety Monitoring Board* (DSMB) periodically reviewed trial results, to ensure patient safety and implement an efficacy monitoring scheme.
  - Results remained confidential, outside of the DSMB and the *independent* statistical support staff supplying the results to them.
  - DSMB statistician: Lloyd Fisher

- Efficacy was governed by an *O’Brien-Fleming*-type spending function.
Example – 5 look O’Brien-Fleming scheme

- Stopping boundaries on \( z \)-score scale
Example – 5 look O’Brien-Fleming scheme

- Stopping boundaries on *risk reduction* scale
At 2 early meetings, there were few events, nothing striking; then:

<table>
<thead>
<tr>
<th>Trt X Events</th>
<th>Trt Y events</th>
<th>Hazard ratio</th>
<th>z-score</th>
<th>Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>182</td>
<td>144</td>
<td>1.27</td>
<td>2.05</td>
<td>4.99</td>
</tr>
<tr>
<td>295</td>
<td>246</td>
<td>1.20</td>
<td>2.07</td>
<td>3.73</td>
</tr>
<tr>
<td>402</td>
<td>318</td>
<td>1.27</td>
<td>3.18</td>
<td>3.20</td>
</tr>
</tbody>
</table>
Issues

- The steps in the adjudication process inevitably led to a lag in information accrual, and a backlog of pending and potential events.

- The DSMB expected that the results likely would be across the boundary already if the resolution of pending cases were known.
DSMB recommendation

- The DSMB recommended to look again soon – perhaps in a 2-4 month timeframe, rather than the typical 6 months.

- The trial team was instructed to expend all possible efforts to maximize the number of adjudicated cases by the time of their next meeting.

- Specific information was of course not conveyed, but speculation was inevitable.
Question - criteria

- What stopping criteria would govern this next look?

- By choosing the timing based on the current results, we’re breaking out of the formal spending function framework, and risking inflating the false positive rate.
DSMB statistician’s proposal

➢ Compute the boundary for this ‘alternately scheduled’ look so that:

- the probability of boundary crossing given the current data and conditional on the null hypothesis (i.e., equality) is equal to the same quantity computed for the next look in the originally planned scheme.

➢ The Independent Statistician and DSMB statistician would jointly derive this criterion.
Criteria issues

- Who should decide what the criterion should be?

- The trial team *could not be* part of the discussion at this point.

- The DSMB needs a definition to guide their actions, but they are not the party that will be responsible before health authorities.

- While the DSMB preference may carry some weight with authorities, it’s not *binding*. 
Implementing the proposal

- The “originally planned scheme” is not so clearly defined:
  - the next look would have been sometime in the summer of 2007
  - with *how many events***??*

- Some reasonable values based on trial history and recent trends were chosen, and the specific approach was documented in a memo dated prior to the analysis (but remaining confidential within the DSMB).
Simple alternate approach

- The next analysis might be viewed as far enough away from the previous one so that, using simplicity as a tie-breaker, just deriving the criterion as if this were the next scheduled look in the original scheme might be OK.

- e.g., Dave DeMets has noted that spending function schemes are hard to “break”.
Derivation of the criterion

- The current analysis had 720 events.
- A reasonable event total for the next “planned” analysis would be 900 events.
- The spending function would yield a boundary of \( z = 2.843 \).
  \[
P ( Z_{900} > 2.843 \mid \text{current data, } H_0 ) = .4995
\]
- If the analysis includes \( N \) events, choose \( Z^* \) so that
  \[
P ( Z_N > Z^* \mid \text{current data, } H_0 ) = .4995
\]
Characterization of O’Brien-Fleming

- In a “classical” (i.e., equally spaced) O’Brien-Fleming design, let’s say that at one of the analyses the test statistic was exactly equal to the boundary value.

- If, between that look and any later look the estimated difference in the new data was zero, then the test statistic would again be exactly equal to the boundary value.

- Or, the conditional boundary crossing chance under the null hypothesis is 50%.
Example – 5 look O’Brien-Fleming scheme

- Stopping boundaries on \( z \)-score scale

![Graph showing stopping boundaries on z-score scale](image-url)
Results

- The next DSMB meeting took place in April 2007, and the analysis included 130 new events, 850 in total.
  - The boundary computed according to the revised approach was \( z = 2.92 \)
  - The boundary computed ‘naively’ using the spending function was \( z = 2.94 \)
- The results of the analysis were: \( z = 2.93 \)
Recommendation

- There had been 65 new event patients in each group.
  - Thus, there was no *current* signal of differential outcome.
  - Plus, the overall signal of a difference for more serious outcomes had lessened.

- The DSMB announced that the trial should continue.
  - The next look should be ‘*back on schedule*’ in about 6 months.
The DSMB next met in October 2007. Analysis results:

2. 67 new events for Trt. X and 62 for Trt. Y
3. This was enough to push the result ($z = 2.92$) beyond the boundary ($z = 2.74$).

The DSMB recommended that the trial be terminated.
Paradox?

- Multiple testing decision rules can lead to apparent paradoxes.
- The analysis at which the event totals were 402 vs 318 did not meet the stopping criterion.
- When more data was added with the following breakdown: 132 vs 127, it *did* meet the criterion.
Trial shut-down

- Shortly, plans had been made for all patients to be brought in for a final visit during a period extending into January 2008.
- Database lock was anticipated to be around mid-year.
- The study Executive Committee decided that the information was too important to withhold that long, as the results might impact medical thinking and practice.
Announcement plans

- Though adjudications would not be complete, the EC proposed announcing the results at the *American College of Cardiology* conference in March 2008.
  - Given the inherent adjudication lag, and the length of time between the DSMB meeting and ACC, there would be many new events – perhaps 200.
  - Some risk that the results would tell a *different story* than the DSMB report.
  - But hopefully the data would be close enough to complete so that results would not differ materially from the eventual *final* results.
Database issues

- Data collection and cleaning efforts focused on the primary endpoint.
- The public presentation at ACC would similarly focus on the primary endpoint, and would emphasize that announcement of other data/endpoints, and full interpretation of the trial results, would await the final data.
Results shown at ACC

- Analyses run just *days* prior to ACC were the basis for the public presentation. These strengthened the signal of difference between the treatments:
  - B/H: 653 (11.4%), B/A: 530 (9.3%)
  - 119 new events for B/H, and 85 for B/A
  - \( z = 3.72 \), corresponding to a HR of 0.80.
Results shown at ACC

ACCOMPLISH: ACE Inhibitor Plus Calcium-Channel Blocker Best for Reducing Clinical Events in Hypertensive Patients

News Author: Michael O'Riordan
CME Author: Désirée Lie, MD, MSEd
Disclosures

Release Date: April 1, 2008; Valid for credit through April 1, 2009

Credits Available
Physicians - maximum of 0.25 AMA PRA Category 1 Credit(s)™ for physicians;
Family Physicians - up to 0.25 AAFP Prescribed credit(s) for physicians

From American College of Cardiology (ACC) 57th Annual Scientific Session
From Heartwire — a professional news service of WebMD

April 1, 2008 (Chicago) — New data from the Avoiding Cardiovascular Events in Combination Therapy in Patients Living with Systolic Hypertension (ACCOMPLISH) trial were presented today at the American College of Cardiology 57th Annual Scientific Session [1]. They showed that a single-tablet dual-mechanism therapy initiated in high-risk hypertensive patients significantly reduced the risk of morbidity and mortality by 20% compared with conventional therapy.

ACCOMPLISH, a major morbidity and mortality trial, compared the effects of two forms of antihypertensive combination therapies on major fatal and nonfatal cardiovascular events. It was stopped early because treatment with antihypertensive combination therapy — the angiotensin-converting enzyme (ACE) inhibitor benazepril plus the calcium-channel blocker amlodipine — was more effective than treatment with the ACE inhibitor plus diuretic.

Learning Objectives
Upon completion of this activity, participants will be able to:
1. Inform clinicians of the latest medical information comparing the combination of an angiotensin-converting enzyme inhibitor and calcium channel blocker with an angiotensin-converting enzyme inhibitor and another calcium channel blocker.
Final results

- Main trial manuscript was recently published:

The NEW ENGLAND JOURNAL of MEDICINE

Benazepril plus Amlodipine or Hydrochlorothiazide for Hypertension in High-Risk Patients

Kenneth Jamerson, M.D., Michael A. Weber, M.D., George L. Bakris, M.D., Björn Dahlof, M.D., Bertram Pitt, M.D., Victor Shi, M.D., Allen Hester, Ph.D., Jitendra Gupte, M.S., Marjorie Gatlin, M.D., and Eric J. Velazquez, M.D., for the ACCOMPLISH trial investigators*

ABSTRACT

BACKGROUND
The optimal combination drug therapy for hypertension is not established, although current U.S. guidelines recommend inclusion of a diuretic. We hypothesized that treatment with the combination of an angiotensin-converting-enzyme (ACE) inhibitor and a dihydropyridine calcium-channel blocker would be more effective in reducing the rate of cardiovascular events than treatment with an ACE inhibitor plus a thiazide diuretic.
Final results

➢ As reported in *NEJM*, December 2008:
  o B/H: 679 (11.8%), B/A: 552 (9.6%)
  o Hazard ratio = 0.80
  o There was a high degree of consistency across primary endpoint components, and key subgroups.