Extension Studies

Rutgers Biostatistics Day
(Feb 16, 2007)

Len Oppenheimer
Bruce Binkowitz
Amy Ko

Johnson and Johnson
Merck and Co.

Merck and Co.
Outline

- Overview
- Extension Study Design
- Hypotheses
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Overview

- Base study
- Extension study
- Advantages of extension studies
(Base) Study
(Standard Paradigm)

- Study Protocol
- Specific Hypotheses
- Recruit Patients
- Informed Consent
- Randomization
- LPO

- Data Base Lock
- Analyses
- Interpretation
  /Conclusions
- Disseminate Results
Extension Study

- Separate protocol (most cases)
- New hypotheses
- Same patient cohort
  - No new patients
  - Patients re-consented (most cases)
- Treatments can be re-assigned
  - Randomization
  - Systematic
Example: Extension Study
(Assessment of Reversibility)

Withdrawal of Hormone Therapy on BMD

Women between ages 45-59
Treated for 4 years; followed for 2 more years
Lumbar spine BMD - Change from Baseline

Mean Percent Change

Years

ALN 5 mg/ PBO
Hormone Therapy/Placebo
Placebo

Advantages of Extension Studies

- Advantages over additional or new studies
  - Results available sooner
    - no new study start up activities
      - recruitment of sites, investigators, and patients
  - Cost-effective
  - Can efficiently address additional questions via extension study design
    - safety & efficacy
    - vs comparators
    - dose response
    - long-term effects
Advantages of Extension Studies

Advantages over postmarket AE data for safety

- Defined denominator
- Comparators
- Higher quality reporting /monitoring
- Blinded assessments
- More risk /benefit results
- Results available sooner
Outline

- Overview
- Extension Study Design
- Hypotheses
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Extension Study Design

Examples of extension study design
- Single Arm /Open Label
- Multiple-Arm Design

Extension study design issues
- Bias
- Informed Consent
- Allocation schemes
- Randomization
- Impact on base study
Single Arm /Open Label

Base Study
Placebo → Proposed dose
Low Dose → Proposed dose
High Dose → Proposed dose

Extension Study
Single Arm /Open Label

Advantages
- Simple logistics
- Facilitates enrollment (all patients get treated)
- Patient retention
- Maximum number treated (rare AEs)
Disadvantages

- Increased potential for bias
  - patient/investigator unblinded
- Interpretation difficult
  - self selected cohort
  - dissimilar patient population – combined patients with heterogeneous base study experience
  - no comparator group – rely on historical or epidemiological data
Multiple-Arm Design

- Alternatives to single-arm extension study
  - Do not have to allocate 50% - 50%
  - Allocate 75% - 25%
  - Allocate 90% - 10%
  - Allocate X% to an active control (if exists)
Example: Multiple-Arm Design
Finasteride 4-Year Extension Study

Year 1

Finasteride 50%
(N = 779)

Placebo 50%
(N = 774)

2-Week Placebo Run-in
Example: Multiple-Arm Design
Finasteride 4-Year Extension Study
Example: Multiple-Arm Design
Finasteride 4-Year Extension Study

<table>
<thead>
<tr>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finasteride 45%</td>
<td>(N = 547)</td>
<td>(N = 474)</td>
<td>(N = 379)</td>
<td>(N = 323)</td>
</tr>
<tr>
<td>Finasteride 50%</td>
<td>(N = 65)</td>
<td>(N = 46)</td>
<td>(N = 39)</td>
<td>(N = 32)</td>
</tr>
<tr>
<td>Placebo 5%</td>
<td>(N = 60)</td>
<td>(N = 46)</td>
<td>(N = 33)</td>
<td>(N = 23)</td>
</tr>
<tr>
<td>Placebo 50%</td>
<td>(N = 543)</td>
<td>(N = 429)</td>
<td>(N = 353)</td>
<td>(N = 290)</td>
</tr>
<tr>
<td>Placebo 5%</td>
<td>(N = 779)</td>
<td>(N = 774)</td>
<td>(N = 46)</td>
<td>(N = 23)</td>
</tr>
</tbody>
</table>
Example: Finasteride Ext. Study Results
(Patients on Same Therapy for 5 Years)

The difference in hair count mean change from baseline of 277 hairs was seen between men taking PROPECIA vs placebo at 5 years.

*Changes in hair count were measured within a 1-inch diameter circle at the anterior leading edge of the thinning vertex area.

Pooled data from 2 vertex studies (mean baseline hair count=876),
P<0.001 PROPECIA vs baseline at each time point.
P<0.001 placebo vs baseline at each time point.
Generic Extension Study Design

- **Treated**
  - Long-term maintenance of effect
  - Rebound (estrogen) or reversibility of effect

- **Placebo**
  - Catch up,
  - Disease modification,
  - Additional safety info

- **“Natural” disease progression**

---

**0 year** → **1 year** → **2 years**

- Base Study → Extension Study
Example: Finasteride Ext. Study Results
(All Treatment Sequences over 5 Years)
Extension Study Design Issues

- Bias
- Informed Consent
- Allocation schemes
- Randomization
- Impact on base study
Extension Study Design Issues
Avoiding Bias

- Avoiding /minimizing bias - critical design feature for all studies/base study/ext. studies
  - Blinding
  - Randomization
  - Appropriate comparators

- Even more critical for extension studies since
  - More discontinuations /withdrawals /lost to follow-up
  - Extent of discontinuations /withdrawals /lost to follow-up may differ by treatment group
When to Obtain Informed Consent for Extension

- Start of the “base study”
  - Fewer drop-outs since do not need to re-consent

- Start of the “extension study”
  - Easier to enroll the “base study”
  - Greater flexibility to make changes
  - More appropriate with multiple extensions
Extension Study Design Issues
Allocation Schemes for Extension

- Randomized allocation
- Systematic allocation
- Combination
- Example

High dose → High dose
Low dose → Low dose
Placebo → Active Comparator
Base Study → Extension Study
When to Randomize for Extension Treatments

Start of “base study” vs. start of “extension”

- Active Comparator
  - 1: Active Comparator
  - 2: Placebo

- Experimental
  - 3: Experimental
  - 4: Placebo

- Placebo
  - 5: Experimental
  - 6: Active Comparator

Base Study ➔

Extension Study
Extension Study Design
Impacting Base Study

- Base Study: Includes Placebo group
- Extension Study: Active treatment only
- Dropout of base study allowed to enter ext.
  - Impact: Earlier and higher frequency base study drop-outs
Example: Extension Study Design
(4S Study - No Follow-up Visits)

- 4S (Simvastatin Scandinavian Survival Study)
- Simvastatin vs. Placebo (Secondary prevention cardiovascular disease)
- Primary endpoint: total mortality
- Secondary endpoints: major events
  - cardiovascular mortality
  - cardiovascular events
  - cancer events, etc.
- Study termination: event based
Example: 4S Study Results

- Base study results (median follow-up 5.4 years)
- Extension Study Design (5 additional yrs.)
  - No assigned treatment (standard of care)
  - No follow-up visits
  - Objectives: Examine the long-term effect in terms of risk benefit
- Data Collection
  - Death Registries over 10 years
  - Cancer Registries over 10 years
## Example: 4S Study Results

<table>
<thead>
<tr>
<th></th>
<th>Double-blind trial</th>
<th>5-year extension</th>
<th>10-year follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All causes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of deaths</td>
<td>256 182</td>
<td>212 232</td>
<td>468 414</td>
</tr>
<tr>
<td>K-M estimate</td>
<td>12.4% 8.7%</td>
<td>10.8% 11.4%</td>
<td>21.3% 19.9%</td>
</tr>
<tr>
<td><strong>All cardiovascular</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of deaths</td>
<td>207 136</td>
<td>128 155</td>
<td>335 291</td>
</tr>
<tr>
<td>K-M estimate</td>
<td>10.2% 6.7%</td>
<td>6.6% 7.7%</td>
<td>15.6% 14.6%</td>
</tr>
<tr>
<td><strong>Cancer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of deaths</td>
<td>35 33</td>
<td>65 52</td>
<td>100 85</td>
</tr>
<tr>
<td>K-M estimate</td>
<td>1.7% 1.6%</td>
<td>3.4% 2.7%</td>
<td>5.1% 4.3%</td>
</tr>
</tbody>
</table>

Outline

- Overview
- Extension Study Design
- Hypotheses / Objectives
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Hypotheses / Objectives (Extension Study)

Safety:
- Long-term safety (multiple extensions)
  - Higher incidence AE’s – comparator group required
  - Low incidence (rare) AE’s – comparator less necessary

Efficacy:
- Resolution / reversibility of effect-rebound
- Maintenance of effect (non-inferiority)
- Dose response (titration) vs. active comparators
- Disease modification
- Natural history of disease

(See study design)
Outline

- Overview
- Extension Study Design
- Hypotheses
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Analysis Issues

- Missing data
- Selection Assessment
- Disclosure of Results
- Comparison to Baseline and Choice of Metric
- Unequal Treatment Allocation
Analysis Issue

- **Major issue:** Missing Data
  - Reduced Power
  - Selected Cohort: potential bias

- **Selection factors**
  - Patient self-selection
    - treatment of symptoms
    - symptomatic AEs
  - Protocol mandated withdrawal
Analysis Issue

- When to stop?
  - Overall Sample Size
  - Within a treatment group
Example: Finasteride (Propecia)

N = Number of patients entering the extension.
Example: Finasteride (Propecia)

Pooled data from 2 vertex studies (mean baseline hair count=876). 
P<0.001 PROPECIA vs baseline at each time point. 
P<0.001 placebo vs baseline at each time point.

Reference: www.propecia.com
Selection Assessment
(Cohort Comparison)

- Cohorts
  - Randomized population
  - Completers: Base study
  - Those who enter extension(s)

- Comparisons of cohorts based on:
  - Baseline characteristics
  - Base study results
    - efficacy
    - safety
Selection Assessment
(Cohort Comparison)

- Often patients continuing into extensions have
  - better efficacy
  - better safety

- Selection bias exacerbates missing data issue

- More missing data - less likely to be MCAR
Selection Assessment
(Cohort Comparison)

■ Objective vs. subjective endpoints
  – BMD vs. patient / investigator assessment
  – Less subject to self-selection
  – More subject to protocol mandated selection criteria

■ Include sensitivity analysis to ensure proper interpretation

■ Hypotheses → Objectives → Estimation
Disclosure of Results

- Potential bias to extension study conduct due to disclosing base study result
  - Importance of comparator group

- Who /when?
  - Patient
  - Investigator
  - Sponsor
Disclosure of Results

- What to disclose?
  - High level only
  - Group results vs. individual patient results
  - Details
    - Secondary or tertiary variables
    - sub-groups
    - full data
Comparisons to Baseline

Which “Baseline”?  
- Base study baseline  
- Extension study baseline

If efficacy responders continue longer  
- change from base study baseline > change from extension study baseline  
- Randomized comparator group may alleviate / eliminate bias
Choice of Metric

Baseline = 100

End of base study, start of extension

Baseline

Observed value = 80

Change from baseline = -20 (-20%)

End of Extension

Observed value = 60

Change from start of extension = -20 (-25%)

Total Change from Base study baseline = -40 (-40%)

- Incremental changes (-20 + -20) add up to the total change, but the percent changes (-20% + -25%) not equal to -40%
- Potential for confusion for non-statistical readers/reviewers
Unequal Treatment Allocation

- Assume there is no true underlying AE rate difference between two treatment groups
- 2 scenarios: equal vs. unequal allocation
  - Equal (100 per group)
  - Unequal (150/50 per group)
- Assume AE relatively rare (≈ 5% or less)
Unequal Treatment Allocation

- A spurious AE disparity is more likely to occur in the unequal allocation scenario than in the equal
- For equal allocation, need a ratio of about 1:1
- For unequal (in this example) need a ratio of 3:1
- BUT, 3:0 can more easily be obtained in the unequal scenario than the equal, leading to a potentially spurious finding
- Up to 4X more likely
Unequal Treatment Allocation

Probability of observing 3:0 for 150:50 vs 100:100 when the underlying probability is the same for both groups

![Graph showing probability vs. Common AE probability for N=150/50 and N=100/100]
Outline

- Overview
- Extension Study Design
- Hypotheses
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Reporting/Presenting Results

Scale break between base study and extension -- discontinuity

The difference in hair count mean change from baseline of 277 hairs was seen between men taking PROPECIA vs placebo at 5 years.

*Changes in hair count were measured within a 1-inch diameter circle at the anterior leading edge of the thinning vertex area.

Pooled data from 2 vertex studies (mean baseline hair count=876).
P<0.001 PROPECIA vs baseline at each time point.
P<0.001 placebo vs baseline at each time point.
Reporting/Presenting Results

- Effects over time (efficacy or safety)
  - follow same cohort
  - non-responders drop-out

- Events per 100 patient-years
  - Constant hazard rate
  - Need to confirm assumption
  - Are 100 patients followed for 5 years equivalent to 1000 patients followed for 6 months?
Outline

- Overview
- Extension Study Design
- Hypotheses
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Interpretation

- Generalizability
- Inferences to what population
  - potential for selection bias
  - for those patients completing the base study and entering the extension
- Reduced sample size
Outline

- Overview
- Extension Study Design
- Hypotheses
- Analysis Issues
- Reporting/Presenting Results
- Interpretation
- Summary
Summary

■ Value of extension studies
  – Answer important questions re. safety/efficacy
  – Sooner than follow-up studies
  – More effective than some alternative (post marketing AE reports)

■ Study planning /design
  – Address selection bias
  – Maintain blind
  – Re-randomize
  – Comparator treatments
Summary

- Limitations on hypotheses
  - focus on safety
  - objectives or estimation

- Analyses
  - plan to address potential confounding and bias
  - sensitivity analyses

- Presentation of results
Summary

- Interpretation
  - Identify inference population
  - Cautious interpretation of p-values
  - Clearly identify issues impacting conclusions
Thank You!
Backup Slides
“Administrative” Interim Analysis

- Interim analysis usually done for IIA/IIB studies
  - Not for early stopping, but for “administrative purposes”
  - Establish POC
  - Initiate Phase III planning
  - Initiate manufacturing /marketing activities

- Study would complete even with positive interim analysis results
Interim Analysis or Extension Study?

- Six-month study with six-month double blind extension
  or
- Twelve-month study with six-month interim analysis

Answer: Should be based on primary question that needs to be addressed.
Example: Extension Study
(Fosamax® Regulatory Filings)

- Fosamax: N-containing bisphosphonate
- Treatment for osteoporosis
- Slows bone resorption
- Increases bone mineral density (BMD)
- Reduces fracture risk
- Regulatory requirements at that time
  - EU: 2 years efficacy /safety
  - FDA: 3 years efficacy /safety
Example (continued): Fosamax Regulatory Filings

- Two year base study with one year double blind extension
  - Patients remained on same therapy for Year 3
  - Did not unblind 2-year data until last patient completed 3 years to minimize potential for bias
  - Filed 2-year data with EU
  - Subsequently filed 3-year data with FDA, and EU was provided 3-year data as an update.
Example: Extension Study Design
Impacting Base Study

Discontinuation Rates: Protocols 024 and 096

<table>
<thead>
<tr>
<th></th>
<th>“Fast forwarded”</th>
<th>Not “Fast forwarded”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>024</td>
<td>096</td>
</tr>
<tr>
<td>Etori PBO</td>
<td>62.2</td>
<td>32.2</td>
</tr>
<tr>
<td>Etori 90 mg</td>
<td>28.8</td>
<td>25.7</td>
</tr>
<tr>
<td>Etori Nap</td>
<td>43.5</td>
<td>21.2</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of</td>
<td>54.5</td>
<td>26.6</td>
</tr>
<tr>
<td>Efficacy</td>
<td>21.7</td>
<td>17.6</td>
</tr>
<tr>
<td></td>
<td>36.5</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12.1</td>
</tr>
</tbody>
</table>
Example: Extension Study Design Impacting Base Study

Discontinuation Rates: Protocols 025 and 097

<table>
<thead>
<tr>
<th></th>
<th>“Fast forwarded”</th>
<th>Not “Fast forward”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>025</td>
<td>097</td>
</tr>
<tr>
<td></td>
<td>Etori</td>
<td>Rofecoxib</td>
</tr>
<tr>
<td></td>
<td>PBO</td>
<td>90 mg</td>
</tr>
<tr>
<td>Overall</td>
<td>32.2</td>
<td>16.7</td>
</tr>
<tr>
<td>Lack of Efficacy</td>
<td>25.2</td>
<td>12.5</td>
</tr>
</tbody>
</table>
Example: Dose Escalation Design (Etoricoxib 8-Week Extension)

617 Patients Randomized

- Placebo: N=60
- 5 mg Etoricoxib: N=117
- 10 mg Etoricoxib: N=114
- 30 mg Etoricoxib: N=102
- 60 mg Etoricoxib: N=112
- 90 mg Etoricoxib: N=112

50% of patients in each group:

- Diclofenac: N=102
- 30 mg Etoricoxib: N=198
- 60 mg Etoricoxib: N=102
- 90 mg Etoricoxib: N=148

Gottesdiener K, et al, Rheumatology 2002;41:1052-1061
## Generic Example:
Cohort Comparison

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; Year</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entered 2&lt;sup&gt;nd&lt;/sup&gt; Year Mean (N=602)</td>
<td>869.3</td>
<td>962.3</td>
<td>93.0</td>
</tr>
<tr>
<td>All patients Mean (N=679)</td>
<td>874.8</td>
<td>962.3</td>
<td>87.5</td>
</tr>
</tbody>
</table>
Generic Example: Cohort Comparison

Results at the end of the 1\textsuperscript{st} year

<table>
<thead>
<tr>
<th>Entered 2\textsuperscript{nd} Year</th>
<th>Baseline</th>
<th>1\textsuperscript{st} Year</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (N=602)</td>
<td>869.3</td>
<td>962.3</td>
<td>93.0</td>
</tr>
<tr>
<td>All patients</td>
<td>Mean (N=679)</td>
<td>874.8</td>
<td>962.3</td>
</tr>
<tr>
<td>Chose to stop after 1 yr.</td>
<td>Mean (N=77)</td>
<td>917.8</td>
<td>962.3</td>
</tr>
</tbody>
</table>
Extension Example Background
Fracture Intervention Trial (FIT)

Background for the base study (FIT):

- Randomized Clinical Trial of Fosamax vs. Placebo
- 6459 women with low hip bone mineral density (BMD)
- 3 to 4.5 years
- Primary endpoint: fracture
- Secondary endpoint: BMD, biochemical markers
FIT Long-Term Extension Study (FLEX)

Research question:
- In patients who completed FIT on Fosamax®:
  - What is the effect of ALN on bone mineral density over an additional 5 years of ALN therapy compared to PBO?

Patient eligibility
- Received at least 3 yrs of Fosamax during FIT
- Total hip BMD at FLEX baseline
  - T-score >−3.5 and > FIT baseline
- Willing to be randomized (blinded) for 5 more years to:
  - Placebo (40%), Fosamax 5 mg/day (30%), Fosamax 10 mg/day (30%)
FLEX Results

- Follow-up: 91% completed the 5 year extension
- Fosamax® preserved BMD compared to PBO
- %change from baseline hip BMD 2.8%
- %change from baseline spine BMD 3.8%

Black, et. al., ASBMR 2004
Additional Issues - FLEX

- 10 of 11 original FIT clinics agreed to continue
- Assay drift of biochemical markers (concurrent assays were used to eliminate potential drift)
- Storage of samples over time
- Changing technology over time (DXA machine)
Extra Example of an Extension Study

Hair Weight Results  J-Am-Acad-Dermatol. 2002 Apr; 46(4): 517-23

In the longest reported controlled clinical study of male pattern hair loss patients ever conducted ...

PROPECIA® (finasteride) grew thicker and/or faster-growing natural hair

In a 192-week study, a 46% difference (P<0.001) was observed in the mean percent change from baseline in hair weight between men treated with PROPECIA and men treated with placebo at Week 192.²