Establishing Treatments as Disease-Modifying in Neurodegenerative Disease

Challenges in Trial Design and Analysis

Michael P. McDermott, Ph.D.
University of Rochester
Outline

• Problems with single-period designs
• Two-period designs
  – Withdrawal vs. delayed start vs. “complete” two-period designs
  – Statistical issues
  – Sample size issues
• Other design considerations
What is Disease Modification?

• Parkinson’s disease (PD)
  – Slowing/halting neuronal degeneration

• Alzheimer’s disease (AD)
  – Slowing/halting the appearance of AD neuropathology

• Treatments exert their effects on the underlying pathology or pathophysiology of the disease rather than simply on symptoms
Single-Period Designs

• Problems with single period designs
  – Difficult to determine whether treatment effect is symptomatic, disease-modifying, or both
  – Absence of valid markers of underlying disease progression
  – Reliance on clinical measures of outcome
    • UPDRS (PD)
    • ADAS-Cog (AD)
    • Total Functional Capacity (HD)
DATATOP Trial

• The Parkinson Study Group. NEJM 1989; 321:1364-1371

• Randomized trial of deprenyl and vitamin E in 800 participants with early PD

• Primary outcome variable
  – Time to reach a level of functional disability sufficient to warrant initiation of levodopa therapy, in the judgment of the enrolling investigator
Figure 1. Cumulative Probability of Reaching the End Point (Kaplan-Meier Estimate), According to Treatment Group.

The hazard ratio for the comparison of group B to group A with respect to the risk of reaching the end point per unit of time is 0.43 ($P<10^{-12}$; 95 percent confidence limits, 0.33 and 0.55). See text for details.
DATATOP Trial

• Evaluation of disease modification was confounded by the apparent short-term symptomatic effects of deprenyl

• Mean changes in total UPDRS scores at Month 3
  – Deprenyl: +1.48 (n = 383)
  – No deprenyl: -1.63 (n = 365)
  – P-value < 0.0001
Single-Period Designs

- Problems with simple slope comparisons
  - Magnitude of the symptomatic effect may depend on factors that change over time
    - True disease state
    - Measured disease state (clinical rating scale)
    - Age
    - Drug exposure
  - May need to discard data from visits that occur before symptomatic effects are fully apparent
Withdrawal Design

• Leber (1996)
• Subjects randomized to active treatment followed by placebo (A/P) vs. placebo followed by placebo (P/P)
• Period 1: Estimation of total treatment effect ($\alpha_T = \alpha_D + \alpha_S$)
• Period 2: Estimation of symptomatic ($\alpha_S$) and disease-modifying ($\alpha_D$) components
Withdrawal Design

• The DATATOP trial had a withdrawal component
  – Two-month duration of washout period was criticized for being too brief

• Potential problems
  – Lack of blinding in Period 2
    • Differential bias in the evaluation of those declining more rapidly?
  – Difficulties in subject recruitment and retention
Delayed Start Design

- Leber (1996); Bodick et al. (1997)
- Subjects randomized to active treatment followed by active treatment (A/A) vs. placebo followed by active treatment (P/A)
- Period 1: Estimation of total treatment effect ($\alpha_T = \alpha_D + \alpha_S$)
- Period 2: Estimation of symptomatic ($\alpha_S$) and disease-modifying ($\alpha_D$) components
Delayed Start Design in the ADAGIO Trial

Delayed Start Design

• Potential problem
  – Lack of blinding in Period 2
    • Differential bias in the evaluation of those improving more rapidly?

• May improve subject recruitment/retention compared to the withdrawal design
Randomized Withdrawal/Start Designs

• Whitehouse et al. (1998)
• Randomized withdrawal design
  – Subjects randomized to three groups
    • P/P, A/P, A/A
• Randomized start design
  – Subjects randomized to three groups
    • P/P, P/A, A/A
• Blinding is improved, but at the cost of efficiency
Complete Two-Period Design

• Whitehouse et al. (1998); McDermott et al. (2002)
• Combination of withdrawal and delayed start designs
  – Subjects randomized to four groups
    • P/P, P/A, A/P, A/A
• Blinding is maintained throughout
• Efficient (under certain assumptions)
Delayed Start Design

Mean Change in UPDRS Score

Period 1

Period 2

P/A

A/A
Statistical Model

• Assume normally distributed responses
  – End of Period 1: \( Y_1 \)
  – End of Period 2: \( Y_2 \)
  – Period 2 – Period 1 difference: \( Y_2 – Y_1 \)

• Variances assumed to be the same across treatment groups, but may vary with time
## Statistical Model for Mean Responses

<table>
<thead>
<tr>
<th>Group</th>
<th>$E[Y_1]$</th>
<th>$E[Y_2]$</th>
<th>$E[Y_2 - Y_1]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>P/P</td>
<td>$\mu_1$</td>
<td>$\mu_2$</td>
<td>$\mu_{21} = \mu_2 - \mu_1$</td>
</tr>
<tr>
<td>A/P</td>
<td>$\mu_1 + \alpha_D + \alpha_S$</td>
<td>$\mu_2 + \alpha_D$</td>
<td>$\mu_{21} - \alpha_S$</td>
</tr>
<tr>
<td>P/A</td>
<td>$\mu_1$</td>
<td>$\mu_2 + \alpha_T'$</td>
<td>$\mu_{21} + \alpha_T'$</td>
</tr>
<tr>
<td>A/A</td>
<td>$\mu_1 + \alpha_D + \alpha_S$</td>
<td>$\mu_2 + \alpha_D + \alpha_T'$</td>
<td>$\mu_{21} + \alpha_T' - \alpha_S$</td>
</tr>
</tbody>
</table>
Model Assumptions

• Disease-modifying effect acquired during Period 1 remains with the subject through the end of Period 2

• Symptomatic effect acquired during Period 1 disappears by the end of Period 2
  – Withdrawal design
  – Can be difficult to verify
Model Assumptions

• Incremental effect acquired during Period 2 is the same for the P/A and A/A groups
  – Magnitude of the incremental effect ($\alpha'_{\tau}$) is independent of whether or not the subject received active treatment during Period 1
  – Strong assumption that may compromise the ability of a delayed start design to provide a conclusive distinction between symptomatic and disease-modifying effects
Model Assumptions

• Incremental effect acquired during Period 2 is the same for the P/A and A/A groups
  – Disease-modifying effect ($\alpha_D$) is underestimated if the incremental effect ($\alpha'_T$) is greater for the P/A group
  – Disease-modifying effect ($\alpha_D$) is overestimated if the incremental effect ($\alpha'_T$) is greater for the A/A group

• This assumption can be tested in the complete two-period design
### Statistical Model for Mean Responses

<table>
<thead>
<tr>
<th>Group</th>
<th>$E[Y_1]$</th>
<th>$E[Y_2]$</th>
<th>$E[Y_2 - Y_1]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>P/P</td>
<td>$\mu_1$</td>
<td>$\mu_2$</td>
<td>$\mu_{21} = \mu_2 - \mu_1$</td>
</tr>
<tr>
<td>A/P</td>
<td>$\mu_1 + \alpha_D + \alpha_S$</td>
<td>$\mu_2 + \alpha_D$</td>
<td>$\mu_{21} - \alpha_S$</td>
</tr>
<tr>
<td>P/A</td>
<td>$\mu_1$</td>
<td>$\mu_2 + \alpha_T'$</td>
<td>$\mu_{21} + \alpha_T'$</td>
</tr>
<tr>
<td>A/A</td>
<td>$\mu_1 + \alpha_D + \alpha_S$</td>
<td>$\mu_2 + \alpha_D + \alpha_T'$</td>
<td>$\mu_{21} + \alpha_T' - \alpha_S$</td>
</tr>
</tbody>
</table>
Parameter Estimation

- Disease-modifying effect ($\alpha_D$)
  - Difference between means in the A/P and P/P groups at the end of Period 2 (Withdrawal)
  - Difference between means in the A/A and P/A groups at the end of Period 2 (Delayed Start)
  - Average of the above two estimators (Complete Design)
Parameter Estimation

• Symptomatic effect \((\alpha_s)\)
  
  – Difference between mean Period 2 – Period 1 differences in the P/P and A/P groups (Withdrawal)
  
  – Difference between mean Period 2 – Period 1 differences in the P/A and A/A groups (Delayed Start)

  – Average of the above two estimators (Complete Design)
Subject Allocation

• Allocation to minimize the variances of the estimates of $\alpha_D$ and $\alpha_S$
  – Equal allocation to the P/P and A/P groups
  – Equal allocation to the P/A and A/A groups
  – Arbitrary allocation to these two components

• May want to allocate fewer subjects to the P/P and A/P groups (withdrawal design component) for purposes of recruitment and retention
Issues in Implementation

• Lengths of follow-up periods
  – Period 1 needs to be long enough for the disease-modifying effects of treatment to become apparent
  – Period 2 needs to be long enough to allow the symptomatic effects of treatment to completely disappear (withdrawal) and/or become fully apparent (delayed start)
    • Problems with long duration or slow-onset of the symptomatic effect
Issues in Implementation

• Lengths of follow-up periods
  – Above assumptions are very difficult to test
  – Concerns about retention and the need for therapy are greater as the required length of follow-up increases
    • Dopaminergic therapy (PD)
    • Acetylcholinesterase inhibitor therapy (AD)
    • Periods beyond 6-9 months stretch the limits of feasibility
Issues in Statistical Analysis

• Three main steps in the analysis
  – Estimation of the total treatment effect at the end of Period 1
    • Not essential if demonstration of this effect is not a goal of the trial
  – Estimation of disease-modifying and symptomatic effects at the end of Period 2
    • Comparison of end-of-trial means
Issues in Statistical Analysis

• Three main steps in the analysis
  – Comparison of the pattern of mean responses near the end of Period 2
    • Need to establish constant (or non-decreasing) separation between the group means over time
    • Need to define “near the end”, i.e., which time points to use in the comparisons
    • Non-inferiority comparison
      – How should the non-inferiority margin be chosen?
      – Large sample size requirement
Issues in Statistical Analysis

• Subject withdrawal
  – Need for treatment of emerging symptoms
    • Move participants directly into Period 2 if treatment is needed (delayed start design)?
    • Preservation of time scale
  – Other reasons for withdrawal
  – Clearly one of the major limitations of two-period designs
Issues in Statistical Analysis

• Subject withdrawal
  – Analysis strategies
    • Last-observation-carried-forward (LOCF)
    • Complete cases
    • Cases that make it to Period 2
      – Statistical adjustment for imbalances in baseline covariates, propensity score
    • Likelihood-based (MAR assumption)
    • Multiple imputation (MAR assumption)
    • Sensitivity analyses
Sample Size Determination

• Minimally important effect size for disease modification?
  – Extrapolation of benefits over very long treatment periods
  – What can be anticipated to accrue over relatively short (6-9 month) treatment periods?
Sample Size Determination

• Non-inferiority comparison at the end of Period 2
  – Number and timing of observations in Period 2 may be an important design consideration

• Need to account for subject withdrawal
  – Particularly a problem in Period 2

• Need to account for misdiagnosis
  – Particularly a problem if those with early disease are of interest
Other Design Considerations

• Eligibility criteria
  – Best to treat as early as possible with a disease-modifying treatment in neurodegenerative disease
    • Recruitment and misdiagnosis issues
  – Need to minimize co-morbid conditions
    • Retention
• Allowance of concomitant medications
• Primary outcome variable
Alternative Approaches

• Use of biomarker/surrogate marker
  – Ideal solution to the problem?
  – Validity as a marker of the effect of treatment on the underlying disease process
  – Validity of a biomarker as a surrogate outcome
    • Must yield the same inference regarding the treatment as the meaningful clinical outcome
Regulatory Opinion

• FDA
  – Conference held in April 2008 to discuss relevant issues in PD trials
  – Remains open to the idea of approval of an agent for disease modification on the basis of clinical outcomes only

• EMA
  – Will only consider strong biomarker evidence along with evidence of treatment effects on meaningful clinical outcomes
Summary

• Two-period designs may be able to distinguish between symptomatic and disease-modifying effects when there are no available direct measures of underlying disease progression

• These designs are challenging in terms of:
  – Assumptions
  – Feasibility
  – Execution
ELLDOPA Trial


• Randomized trial of levodopa (150, 300, and 600 mg/day) vs. placebo in 361 participants with early PD

• Primary outcome variable
  – Change from baseline to Week 42 (after 2-week washout of study medication) in the total UPDRS score