SUMMARY OF DATA ANALYSIS PLAN

Parkinson's Progression Markers Initiative Statistics Core

Christopher S. Coffey
Department of Biostatistics
University of Iowa

OUTLINE

In this presentation, we will:

- Summarize the planned analyses
- Provide the justification for the sample size
- Discuss steps that can be taken by investigators to address future questions of interest

PLANNED ANALYSES

Planned Analysis #1: Comparison of Baseline Characteristics Among Health Subjects and PD Subjects.

- Continuous variables assessed using t-test
- Dichotomous variables assessed using chi-square test
- Appropriate assumptions will be assessed for each comparison and any necessary adjustments (i.e., transformations) will be made prior to analysis

PLANNED ANALYSES

Planned Analysis #2: Comparison of Short-Term Change in Progression Endpoints.

- Examine short-term change during first six months for each progression endpoint using mixed model (continuous endpoints) or logistic regression (dichotomous endpoints)
- Initial model will include all baseline characteristics, indicator for whether healthy control of PD patient, and all possible two-way interactions
- Will utilize backwards selection to build a model for each progression endpoint

PLANNED ANALYSES

Planned Analysis #3: Examination of Whether Short-Term Change in Progression Endpoints is Predictive of Change in Long-Term Endpoints

- Consider only progression endpoints that show differences between healthy subjects and PD patients
- Primary focus on long-term change in UPDRS score – additional long-term endpoints may be considered as well
- Ten-fold cross-validation procedure will be used to test predictive validity of each model
- If successful, final model will provide subset of short-term progression endpoints predictive of change in long-term endpoints – suggest biomarkers for future studies of interventions in PD patient populations

PLANNED ANALYSES

Planned Analysis #4: Examination of PD Subsets

- Each of first three sets of analyses will be repeated comparing subsets of PD patients
- If successful, final model will determine whether some short-term progression endpoints are more predictive of long-term endpoints for some subsets of PD patients and less predictive for other subsets

SAMPLE SIZE JUSTIFICATION

Because of exploratory nature of the planned analyses, it is very difficult to provide a formal sample size justification for the entire model building process.

However, we examined the ability of proposed sample size (400 PD patients/200 healthy controls) to detect meaningful effects of interest.

SAMPLE SIZE JUSTIFICATION

Total Sample Size	Detectable Correlation	Detectable Difference in Prevalence	Detectable Difference in Means (Standardized)
300	0.16	17%	0.33
400	0.14	14%	0.28
450	0.14	15%	0.28
600	0.11	13%	0.24

Last two rows correspond to first set of comparisons (PD patients vs. healthy controls)

First two rows correspond to second set of comparisons (among PD subsets)

ADDITIONAL ANALYSES

In addition to the planned analyses summarized above, the PPMI trial will involve the creation of a rich database.

It is hoped that the data from this trial will also allow assessing a number of additional questions.

Investigators are encouraged to bring possible future analyses to the table.

ADDITIONAL ANALYSES

There are two scenarios for future analyses:

1) Investigators can request the data needed to address the question and conduct their own analyses.

2) Investigators can propose a research question and work with the statistics core at Iowa to conduct analyses.