ADNI Biostatistics conference call, 15 January 2008
Present on call: Laurel, Danielle, Hao, Qian, Thomas Blaettler from BMS, John,

Danielle gave a brief overview of the available data:
Clinical data are available (bulk download). Imaging data summaries:
Cross-sectional: Paul Thompson, Gene Alexander SPM and VBM,
Longitudinal: Norbert Schuff hippocampal volumes, Nick Fox boundary shift integral,
PET: Bill Jagust's glucose metabolism, Norman Foster also.
Charlie DeCarli says he has a large block of data on white matter hyperintensities but not submitted yet, will go in soon.
The biomarker data are completed for baseline but not submitted yet.
The push is to complete as much longitudinal imaging data summaries as possible so that we can present longitudinal data analysis at the Chicago meeting in April. A list has been circulated of people with 0, 6 and 12 month data and priority will be for those, especially MCI and AD folks.

A report is being organized by ADCS folks that will be available on LONI web site that will show status of numbers of people who got clinical eval posted, had imaging, had raw images uploaded, processed, and summaries uploaded at each time point for follow-up. This will be a real-time summary so it will update whenever data are uploaded.
Cliff Jack's lab is worried about why some scans are missing; ADCS has circulated a list of people who withdrew and this was very helpful. Sometimes site says they have uploaded but LONI has no record. Tracking is a challenge and people are working on it.

We at UCD are working on several prototype analyses to post for other groups. One issue we want to show is the use of cross-validation, both the leave-k-out and the 40-60 split. The second issue to show is the correlation between change in potential marker and change in clinical endpoint (incident AD or MCI; change in key neuropsych tests, e.g. ADAS-COG and MMSE).

The UCD and UCSD folks are also working on 3 papers involving baseline data: the baseline clinical data paper (led by Ron Peterson), baseline biomarker data (led by John Trojanowski and Les Shaw), and a PET paper looking at FTD cases that show up amongst the AD (led by Norman Foster). We have data for the first one and have discussed the other two and should get data for the second soon.

Two issues about how to analyze data have come up in the baseline clinical paper that will apply to other papers as well. First, we would like to have a consensus on who is included in the study. There were 821 people randomized to study arms, but 2 or 3 were excluded before baseline. We proposed to use just the baseline people and exclude those who had screening only. Second, there are a few discrepancies between screening diagnosis, which determined study arm and thus what follow-up measurements and timing were used, and baseline diagnosis. We are not sure which should be the primary diagnosis for analyses that involve diagnostic category. One could make a case for using screening diagnosis, on the analogy with clinical trials and randomization to a specific arm, but one could also make a case for using the presumably more accurate diagnosis at baseline. The numbers affected are very small, and thus it is unlikely to make much difference which we use. We will continue discussing this with the clinical core, and Qian is going to get the exact numbers to compare.

Another concern raised by Laurel and Danielle in the biomarker paper involves reporting of sensitivity, specificity, and ROC curves for comparison of the normal and AD groups in this study. These two groups were chosen to have very pure diagnostic criteria; e.g. no strokes, mixed dementias, etc. Thus diagnosis is made artificially easy, compared to a clinical population that has a broad spread of performance and many mixed or complicated cases. We have raised concerns that high sensitivity and specificity may mislead labs or primary care doctors to think that our markers have been demonstrated to work for diagnosis in a clinical setting, and possibly even for early warning signs of when to start patients on drugs. The ADNI study was not designed to assess performance in a clinic or population-based setting and cannot inform us about the potential use of biomarkers in such settings. The discussion among biostat folks on the call was strongly in agreement with this point of view and argued for great caution in reporting and presenting between-group comparisons. All of us would prefer not to see results reported in terms of ROC curves but agreed that if this is required by the journal, we should push for inclusion of the strongest possible language of caution on interpretation and an editorial if possible.

Leaders are working on the renewal and what we would like to do that is new, and why. Each core leader has been asked to provide a one-page summary. Discussions in the imaging and biomarker groups so far have focused on new or refined technology. Our group agreed in this call that we would like to emphasize design concerns, especially the need for a translation to clinical trials in the next phase. The industry groups contributed $20 million in hopes that new methods would help to speed clinical trials by development of better surrogate markers and tools. We think the most attractive proposal will feature assessment of performance in actual settings as a marker for drug efficacy (or, perhaps, for diagnosis.) We and the industry statisticians will try to develop this idea further, and have them present it at the April meeting.

Next call is January 29, also at 10 AM Pacific time.

Next call