Overview

- **Biomarkers**
  - must have a scientific basis
  - a change in the marker must reflect a change in disease progression
  - be measurable and reproducible

Non-cognitive biomarkers are often used because cognitive measures often do not have a tight link to disease severity modifications.
• The goal of biomedical imaging is to understand biophysical processes. Visualization can aid understanding.

• As medical knowledge progresses, the effects we study become ever more subtle.

• Statistically, we then need a larger number of independent measurements to get accurate and precise results for subtle effects. → acquire data at multiple sites…
Overviews

- Which biomarkers are predictive?
  - do they track with disease progression and treatment effect?
- Can we measure these biomarkers in a multi-site trial, i.e., what is the size of the variability of:
  - site effects?
  - subject effects?
  - choice of processing algorithms and input parameters?
• Non-neuroimaging example:

With imaging, we can visualize the effect of the drug, which can help decide drug efficacy. Later, it can monitor the effects of refinement in the drug formulation and dosing.
Overview – Example I

Treatment, 100mg/kg 5-FU

<table>
<thead>
<tr>
<th>ADC</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor Volume</td>
<td>0.60 cm³</td>
<td>0.70 cm³</td>
<td>0.95 cm³</td>
<td>0.86 cm³</td>
<td>0.71 cm³</td>
<td>0.76 cm³</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ADC</th>
<th>Day 7</th>
<th>Day 8</th>
<th>Day 9</th>
<th>Day 10</th>
<th>Day 11</th>
<th>Histology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor Volume</td>
<td>1.13 cm³</td>
<td>1.36 cm³</td>
<td>1.60 cm³</td>
<td>1.79 cm³</td>
<td>2.08 cm³</td>
<td></td>
</tr>
</tbody>
</table>
What if the disease effect on the biomarker is small?

“Several studies by Jack et al. (1998, 2000, 2004) of older sporadic AD cases (mean ages 74–79 years in the different studies) have reported annualized rates of hippocampal atrophy of 3–4% with age-matched controls (mean ages 77–80 years) having rates of atrophy of 1.4–1.7% per year and similarly aged MCI subjects have intermediate rates at 1.8–3.7% per year.”


Alzheimer’s Disease

- Clinical symptoms of AD are due to the loss of neurons and loss of viable connections between neurons.
- The medial temporal lobe (MTL) has the highest density of histopathological markers.

Deficits in verbal memory correlate with atrophy in left hippocampal volume and deficits in non-verbal memory correlate with atrophy in the right hippocampal volume.

Alzheimer’s Disease

Fig. 3.  Postulated sequence of spread of neurofibrillary pathology in AD, showing the medial aspect of the cerebral cortex. The depth of the red color is in proportion to the density of tangles (based on refs. 24 and 28). Several of the red areas showed atrophy in the study by Scialli et al. (6).

Courtesy of National Academy of Sciences, U. S. A. Used with permission. Copyright (c) 2002 National Academy of Sciences, U.S.A.

ADNI – Biomarkers for AD

• Alzheimer’s Disease Neuroimaging Initiative

A longitudinal multisite study of elderly people with either mild cognitive impairment (MCI, N=400), Alzheimer’s Disease (AD, N=200) or normal cognition (N=200).

Data was collected at 55 sites.

Half of the subjects were imaged using FDG positron emission tomography (PET). All were imaged using MRI on a 1.5T scanner with a structural imaging protocol.
• Healthy controls sampled at: 0, 6, 12, 24, and 36 months
• MCI subjects sampled at: 0, 6, 12, 18, 24, and 36 months
• AD subjects sampled at: 0, 6, 12, and 24

In addition, urine, serum, and CSF biomarkers were acquired in addition to neuropsychiatric evaluations.
ADNI

• ADNI Goals:
  1) Identify best biomarkers for early diagnosis
  2) Identify best biomarkers for following disease progression
  3) Develop surrogate endpoints for clinical trials
  4) Establish methods for dealing with multisite data

• ADNI Imaging Goals:
1) Link all data at each time point and share data with public
2) Develop technical standards for imaging in longitudinal studies
3) Optimize acquisition and analysis
4) Validate imaging and biomarker data with psychometric and clinical assessments
5) Improve clinical trial methods

While humans can make sense of images with minor artifacts, this is not usually true of automated processing pipelines.

Therefore:

1. use larger fields-of view and many slices
2. no parallel imaging
3. no partial k-space imaging
4. correct for chemical shift artifacts
5. correct for intensity inhomogeneity
Multi-site Trials

- Variance due to differences in sites!

Same subject, same slice, different sites ‘best’ scan

Courtesy of Jessica Turner. Used with permission.

Source: Jessica Turner (FBIRN), Univ. California, Irvine
Figure 2.
Poor SNR with single-channel birdcage coils in first version of protocol. As indicated in Table 1, the protocol using a single-channel birdcage coil differs from the phased array protocol. Left: When 1.5 T images are acquired using the phased array protocol with a birdcage coil, poor SNR results. Right: Making the parameter adjustments listed in Table 1 resolves the problem without increasing chemical shift.
Figure 6.
Intensity in-homogeneity correction. Phased array coil acquisition at 1.5 T before (left) and after (right) intensity nonuniformity correction. Images have been reformatted from the sagittal into the axial plane to illustrate the intensity in-homogeneity anteriorly prior to correction.
Figure 5.
Effect of gradwarp. Spherical phantom with rectilinear grid inclusion before (left) and after (right) gradwarp correction.
Multi-Center Imaging Study

Non-imaging measures

Image Acquisition

Central Database

Other databases

Collaborators

Investigators

Data sharing

Processing & applications

Source: Dan Marcus (WUSTL)

Courtesy of Daniel Marcus, Ph.D. Used with permission.
Overall goal of the Morphometry BIRN:

To develop the ability to process and analyze, as a single data set, MRI data acquired across multiple sites, using tools developed at multiple sites. In addition, to allow data to be shared with the larger community.
mBIRN Use Case – Shape Analysis in AD

(45 subjects: 21 non-demented controls, 18 very mild Alzheimer's Disease and 6 semantic dementia).
mBIRN Use Case – Shape Analysis

LD = Linear Discriminant

(45 subjects: 21 non-demented controls, 18 very mild Alzheimer's Disease and 6 semantic dementia).
mBIRN Use Case – Shape Analysis

Image removed due to copyright restrictions.
• **2006.05.02** (A. Kolasny)
The 101 subject lddmm processing has been completed. This required 244,824 cpu/hrs of processing and 40TB of storage.
SDSC and NCSA TeraGrid sites, BIRN SDSC cluster and JHU CIS cluster were used for processing the 40,804 lddmm jobs. We will now begin the statistical analysis and visualization processing.

• **2006.03.14** (A. Kolasny)
Added an additional 10TB of storage to the JHU CIS storage repository. This used to complete the right hippocampus processing. Experimenting with sshfs and unionfs to assist in cluster processing.

• **2006.02.10** (A. Kolasny)
Completed lddmm processing for the 101 left hippocampus data sets. This computation required a total of 13 cpu/years of computing. We utilized the JHU CIS cluster, SDSC BIRN Cluster and TeraGrid for the processing. The processing required 20TB of storage which is being stored on the BIRN rack and the TeraGrid /gpfswan storage repository. Started processing the right hippocampus data.
• Changes in cortical thickness accompany normal aging, AD, Huntington’s disease, amyotrophic lateral sclerosis, multiple sclerosis, …

Can cortical thickness be used as a reliable biomarker?

Fig. 3. Histogram of thickness values in cortical regions of the subject shown in Fig. 2. More than 99% of the surface is between 1- and 4.5-mm thick.

Courtesy of National Academy of Sciences, U. S. A. Used with permission.
Copyright © 2000 National Academy of Sciences, U.S.A.

mBIRN Use Case II – Cortical Thickness

Fig. 4. Average cortical thickness across 30 subjects, with primary auditory (A1), somatosensory (S1), and visual (V1) cortices indicated by the white arrows.

Fig. 5. Map of the standard deviations of the thickness measurements across 30 subjects. Noncortical regions have been excluded on the medial aspect of the surface.

Courtesy of National Academy of Sciences, U. S. A. Used with permission. Copyright © 2000 National Academy of Sciences, U.S.A.

mBIRN Use Case II – Cortical Thickness

- BIRN cortical thickness reliability study
- 15 healthy older subjects scanned a total of four times. The time between scans was two weeks and three different scanners were used:

Scan 1: Siemens Sonata 1.5T
Scan 2: Siemens Sonata 1.5T (rescan)
Scan 3: GE Signa 1.5T (cross platform)
Scan 4: Siemens Trio 3.0T (cross field strength)

(plus pulse sequence, multiple scans, smoothing)
mBIRN Use Case II – Results


Fig. 2. Maps of thickness measurement variability for four test-retest comparisons (the left hemisphere only; the right hemisphere is similar). Subcortical regions and corpus callosum area are masked out since thickness is not defined there. The measurement variability is less than 0.12 mm for most the cortex for the within-scanner test-retest comparison and less than 0.2 mm for across platform comparisons. Left column: lateral view; right column: medial view.
mBIRN Use Case II – Results

- Within-scanner variability < 0.03 mm
- Cross-scanner variability = 0.15 mm
- Cross-field strength variability = 0.17 mm
- Measurements across field strength biased to the higher field strength (thicker)
- No effect from using the average of multiple runs, however using the 1st run as an initial guess for the processing resulted in a statistically significant reduction in variability.
Function BIRN – Multi-site functional MRI

• The goal of the function BIRN is similar to that of mBIRN, but with a focus on functional MR imaging and schizophrenia used as the target population.

Courtesy of Jessica Turner. Used with permission.
Functional MRI – Multisite Issues

• Differences in:
  - site hardware and software
  - data processing
  - subject’s cortical structures
  - subject’s activation magnitude on that day
  - brain networks elicited for each task
Phase I – Travelling Human Phantoms

Subjects traveled around the country to be scanned at all FBIRN sites

Unique dataset: Subject x site interactions can be measured for the first time

Courtesy of Jessica Turner. Used with permission.
Variance Components Analysis

### ROI – Top 10% of Activated Voxels

<table>
<thead>
<tr>
<th>Variance Source</th>
<th>Auditory</th>
<th>Hand</th>
<th>Visual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>18.8</td>
<td>18.3</td>
<td>21.8</td>
</tr>
<tr>
<td>Site</td>
<td>43.0</td>
<td>21.0</td>
<td>43.8</td>
</tr>
<tr>
<td>Day</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Run</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Subject X Site</td>
<td>3.6</td>
<td>14.6</td>
<td>10.5</td>
</tr>
<tr>
<td>Subject X Site+</td>
<td>20.7</td>
<td>35.2</td>
<td>20.0</td>
</tr>
<tr>
<td>Residual</td>
<td>1.5</td>
<td>4.2</td>
<td>1.5</td>
</tr>
</tbody>
</table>

VCA is a method to identify the individual contributions to the overall variance from the various possible sources.

Courtesy of Jessica Turner. Used with permission.
fBIRN Calibration Methods

- To remove “spikes” that occur in image collection
  - In the traveling subjects dataset, these occurred in .01% of the images
  - Looking for them visually is not feasible

- Automated spike detection and removal allows millions of images to be processed correctly

Courtesy of Jessica Turner. Used with permission.
Scaling by the breathhold response increases group effect size.

Differences between young and old subjects in an fMRI task, before and after correcting for BOLD differences in a separate breathhold task.

Courtesy of Jessica Turner. Used with permission.
fBIRN Calibration Methods

Smoothing to a common level reduces inter-site effects

Site: MGH  Minn.  Iowa  N. Mex.

UNSMOOTHED

SMOOTHED

Courtesy of Jessica Turner. Used with permission.
Impact of fBIRN Calibration Methods

ANOVA Observed Effect Size

- Group by Site
- Group

Cohen's $f$

None | Smooth | Smooth, BH Calibrate | Smooth, BH Calibrate, BH Screen

Remaining challenge

Courtesy of Jessica Turner. Used with permission.
Biomarkers - Conclusion

• The establishment of reliable biomarkers necessitates understanding and overcoming sources of variability due to subject and site.

• Access to a common set of acquisition protocols, processing and analysis tools, and data sharing infrastructure increase the chances of success.
Discover

Analyze

Collaborate

Access

Tools accessed from portal; Eclipse Workflow Client

Global search portlet

Identity and access

Registration portlets

DB dashboards and backup

MBAT/WOMBAT; query interfaces

Annotation interfaces

Video/3D viewer

Registration workflows: Clinical/behavioral

MR scanner

Microarray

Microscopy/atlas

data (2D, 3D…)

ontologies

Find commonalities

Adjust to AIDB,MADB

Subject IDs generator

Project IDs generator

Subject ID mapper

Atlas and Query tools: Concept-based query, view-based query

Atlas Interoperability server and API

Data integration services: terms mapping

Data Quality Assurance, specific for data classes

Data serve and download (e.g. for large images)

Image QA on grid (+curation dashboard)

FAQs, tutorials, API registry, collaborative programming tools.

SRB URIs collection management

Provenance management

XCEDE 2.0

Improved data security

DB schemas and GUI (portal dashboard), versioning and regression tests (HID)

Oracle upgrade

Replication services

SRB upgrade

Database backup

Distributed MCAT

Better compute grid access

Performance improvements, better error handling, utilization of authentication resources, integration of application services with ROCKS, BIRN authentication, portal via application launcher…

 Courtesy of Jeffrey Grethe. Used with permission.