Issues in longitudinal studies of brain structure and function in small samples

Elsa G. Shapiro, Ph.D
Professor of Pediatrics and Neurology
University of Minnesota
Co- P.I. Lysosomal Disease
Network (RDCRN)

Lysosomal Disease Network

CENTER FOR neuroBEHAVIORAL DEVELOPMENT

Pediatric Clinical Neuroscience
UNIVERSITY OF MINNESOTA

Longitudinal research with small samples

1. Research designs using repeated measures in studies of neurodegenerative disease

1. Sensitive measures of outcome
   - -Neuroimaging
   - -Neuropsychology

Examples of longitudinal designs in rare diseases
(repeated measures designs)

- Natural history designs
- Classic randomized crossover design
- Randomized partial crossover design

Natural history designs

- Prospective observational/descriptive study measuring disease progression
- Hypothesis driven natural history study examining predictive or risk factors in disease or treatment outcome
- Historical control/case comparisons for a future treatment trial

Natural History studies

- Example 1: MPS longitudinal study of brain structure and function
  - Studying effects of disease, treatment, and risk factors on CNS over time
  - Specific localized brain findings are hypothesized for each MPS disorder
  - Additional goal: to develop sensitive MPS specific neuro-psychological and neuroimaging markers for future clinical trials
    - Slowly progressive diseases with small changes from year to year
  - Methods: Yearly focused exams for five years in at least 75 treated and untreated children with MPS I, II, and VI
    - Quantitative neuropsychological and neuroimaging measures will be applied
    - Stratified by disease and treatment

Natural History study as historical control

- Example 2: Sanfilippo natural history study: 1 year longitudinal study at two sites
  - Mild to moderately impaired children to provide a historical control for a subsequent treatment trial
  - Short study – rapidly progressive disease - baseline, 6 months, and one year
  - Measure cognitive decline, behavioral change, change in quantitative MRI, and biomarkers.
  - Both standard and theory-based disease-specific measures to measure disease progression

When treatment becomes available, this design will allow all eligible patients to enroll and they can be compared to this historical cohort.

Classic randomized crossover design

- Comparison of treatment and placebo or one treatment with another whose effect is known
  - A controlled trial where each study participant has both therapy (A) and no therapy/placebo (B), e.g, is randomized to A or B first, at the crossover point they then start B or A.
  - All participants serve as own controls reducing error variance thus a small sample can be used
  - Usually a blind study, but treatments in rare diseases rarely allow that
  - Disease progression during the trial controlled by the crossover
  - Problems: carryover effects in the second half of the trial
    - Usually a washout period with no treatment is used
    - Many of the treatments that are used in rare diseases have long term effects (e.g. ERT) or are permanent (HCT or gene therapy)
    - Poor cooperation with placebo/no treatment in fatal neurodegenerative diseases as half of the participants will need to wait for treatment.

Randomized partial crossover design

- Most commonly used in efficacy trials of ERT (enzyme replacement therapy)
- A controlled trial where each study participant is randomized to either therapy (A) or nontherapy/placebo (B) initially. After a period of time the B participants move to the therapy group (A)
- B subjects serve as own controls
- All subjects eventually receive treatment
- Poor cooperation with placebo in rapidly progressive fatal neurodegenerative diseases
- Averts problem of washout
- If N is very small, danger of non-comparability of the two groups.
- If placebo is used, can be blinded or not.

Randomized partial crossover design

- Example 3: Effects of intrathecal enzyme on cognitive decline in MPS I.

  Goal: To evaluate the efficacy of intrathecal enzyme to stabilize or reverse memory and intelligence deficits in MPS I patients with documented cognitive decline.

  - Two groups randomly assigned to no treatment or IT therapy- non-blinded
  - If interim analysis (one year) reveals benefits, all will go on treatment; if no significant differences, continue on the trial for additional year.
  - After the initial12 month trial period, both groups receive treatment. The no-treatment group moving to treatment is essentially serving as their own control
  - Problem: with few participants may result in baseline lack of comparability

CNS outcome measures

- Primary outcomes:
  - Changes in IQ (WASI) and memory score (Hopkins Verbal Learning Test) with baseline value as a covariate.
    - Memory tests need alternative forms due to practice effects (Hopkins Verbal Learning Test has 6 alternative forms)
    - Size of effect needs to be moderate to large in order to detect differences in a small sample; not subtle changes
    - To determine if it is a nonspecific improvement: Comparison of IQ and memory tests with a contrast test not expected to show change; the Test of Variables of Attention
- Secondary exploratory outcomes
  - Quantitative MRI – not enough data yet to develop specific hypotheses

Measuring CNS outcome in these three types of repeated measures designs

- Require repeated administration of the same tests
- Require sensitive measurements that can detect change over a relatively short period of time
- Baseline assessment is necessary to control for the participant's stage of disease

Best ways to measure CNS outcome

- Neuroimaging as a surrogate marker of brain structure and function
- Neuropsychological tests as a semi-direct marker of function

Goals

- To find sensitive measures of CNS disease progression
- To find sensitive measures of CNS change after treatment
- To employ measures that are feasible, practical, easily and reliably administered and that reflect functional change in the patient or be a surrogate marker of functional change.
- To employ measures that can be used in multi-center trials

Qualitative Clinical Neuroimaging

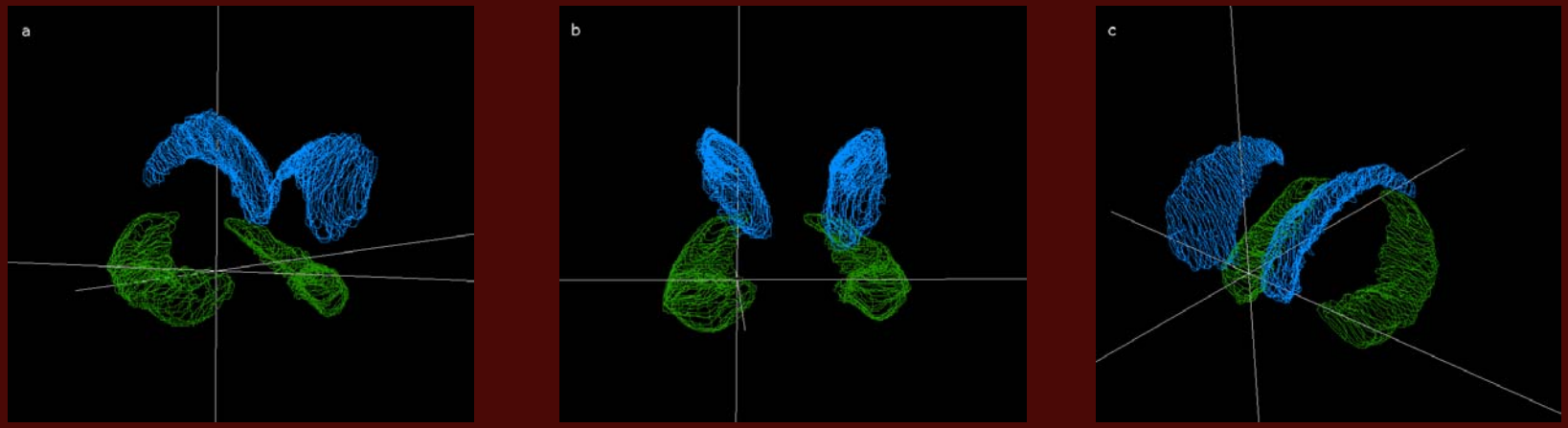
- Clinical diagnostic MRI protocols: designed to reach a diagnosis; not to understand the disease process.
- Clinical scans are not quantitative; don't have good reliability or correlation with function
- Previous longitudinal research has relied on administration of repeated clinical MRIs but often focused on irrelevant /stable aspects of the disease (e.g. enlarged perivascular or Virchow Robin spaces in MPS disorders).
- Qualitative approaches need to be replaced by quantitative reproducible methods.

Quantitative methods

- Easily available and allow measurement of small changes in brain structure and function.
- Standard protocols using fixed magnet strength, fixed and standard sequences to obtain disease specific change, and standard/reliable methods of analysis and quantification
- Comparability across centers using quantitative MRI is an obstacle

Examples of Quantitative Methods

- Volumetric changes of brain: examples white matter, gray matter, ventricle size, specific structures in the brain, whole brain.



Hippocampus and caudate volumetric images in 3D

- Diffusion Tensor Imaging –integrity of white matter pathways in the brain- technique most difficult in multicenter studies

- Magnetic Resonance Spectroscopy – applicable to a subset of disorders to obtain changes in neurochemical status

Neuropsychological testing

- Clinical neuropsychological tests: designed to reach a diagnosis or make specific clinical recommendations; not to understand the disease process.
- Many tests are not specific to a function: test performance results from more than one function or is confounded by other factors. If tests reflect compound psychological functions, they will not correlate with MRI or biomarkers.
  - A reductionist approach to test selection will yield more accurate data but has less practical utility.

  Results are often confounded by physical factors (sensory capacity, motor capacity, health, fatigue) as well as emotional/behavioral factors. Need to account for these.
  - Examples of corneal clouding and carpal tunnel syndrome in MPS disorders
- **These qualitative approaches need to be replaced by quantitative reproducible methods that bear a relationship to specific brain function.**

Selection of quantitative neuropsychological tests

- Use easily available tests and those specific to brain structure and function. Tests with minimal “noise.”
- Standard protocols should include the minimal number of tests to ensure compliance of both center and patient
- Comparability across centers is a goal in rare diseases-
  - Quality control in multi-center research is difficult especially in young children
  - Difficult to find experienced examiners with familiarity with the disease in question especially with skill in testing very young impaired children.
  - Centers with expertise need to be identified.

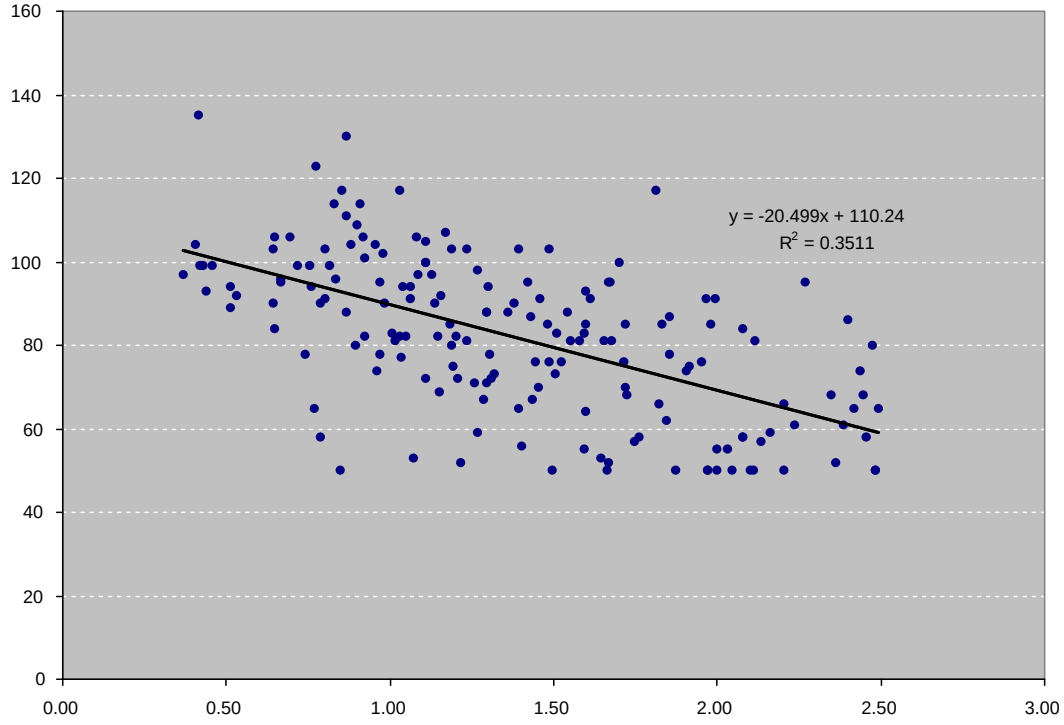
Technical problems in testing

- Practice effects
  - Practice effects vary with function- use of tests with multiple forms
  - Practice effects minimal with very young children due to rapidity of development
- Fatigue effects-
  - performance declines due to boredom or fatigue with repeated testing
  - Testing that lasts more than a few hours
- Age changes in test format
  - Different subtests used for different ages
  - Few tests available across a large age range and difficulty level
- Re-norming of tests which occurs every few years is an obstacle.
  - Cannot change tests during the course of a study.
  - In retrospective studies one may deal with multiple revisions of the same test which differ significantly from each other

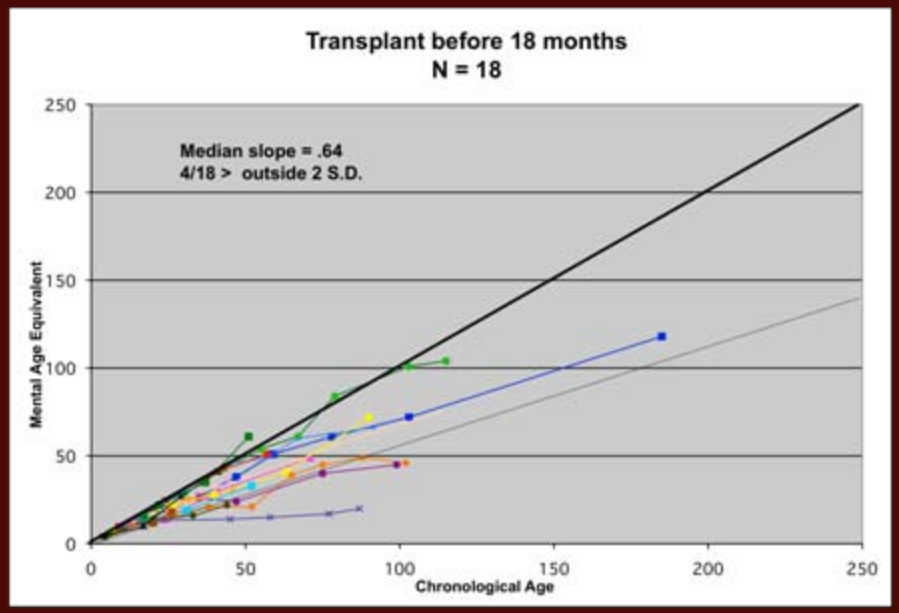
What metric to use?

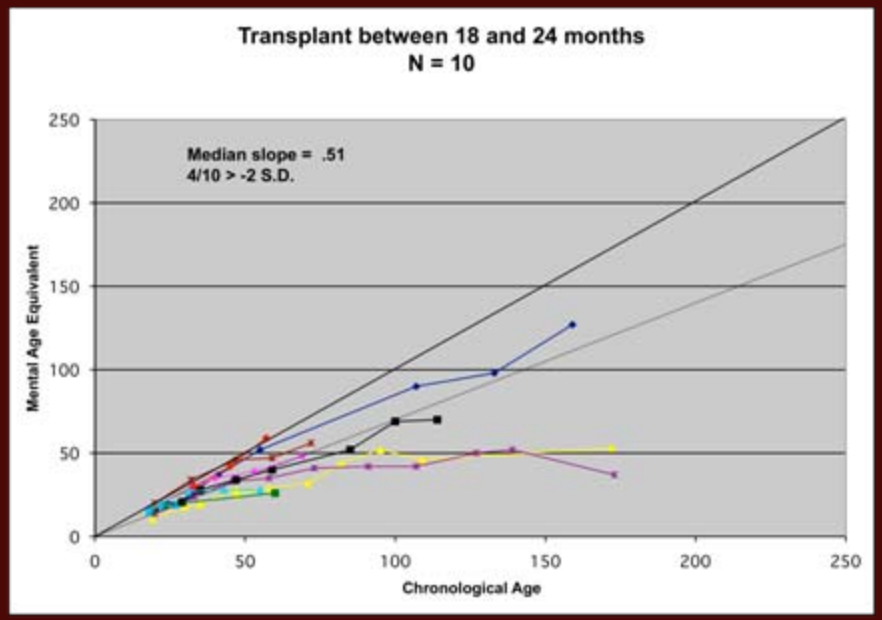
- Use of developmental growth curves calculating the slope of development/change vs. static age-normed measurements (IQ as example)
  - age-normed measures: cannot be used with severely impaired individuals- norms aren't low enough
  - age-normed measures: cannot determine if score reflects growth, plateau in development, or loss of skills
  - Use of raw scores/age equivalents to track change over time can alleviate these problems

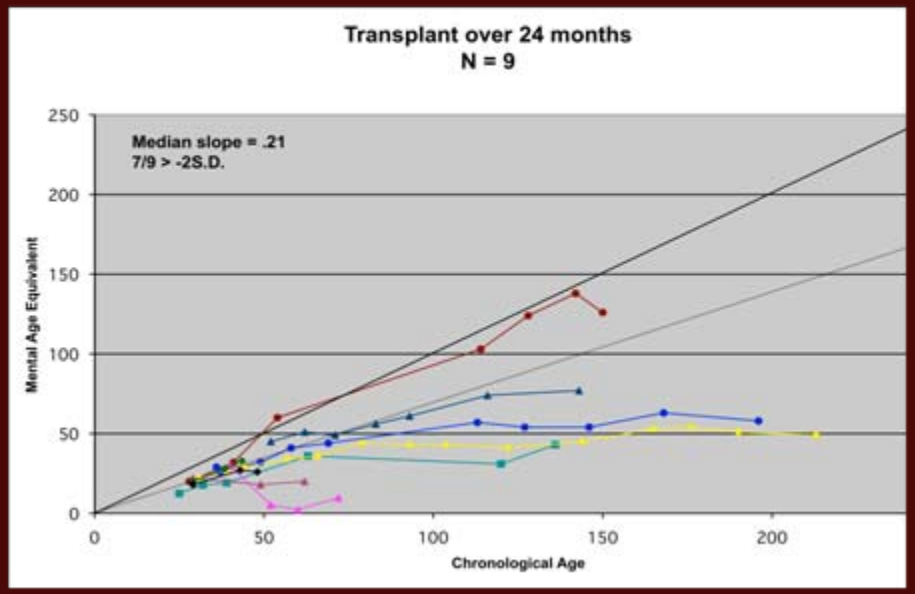
Age and Mental Development in Severe MPS I (Hurler disease) in untreated children or prior to treatment in 137 children and 174 observations



Slope of development for three age groups after HCT for MPS IH







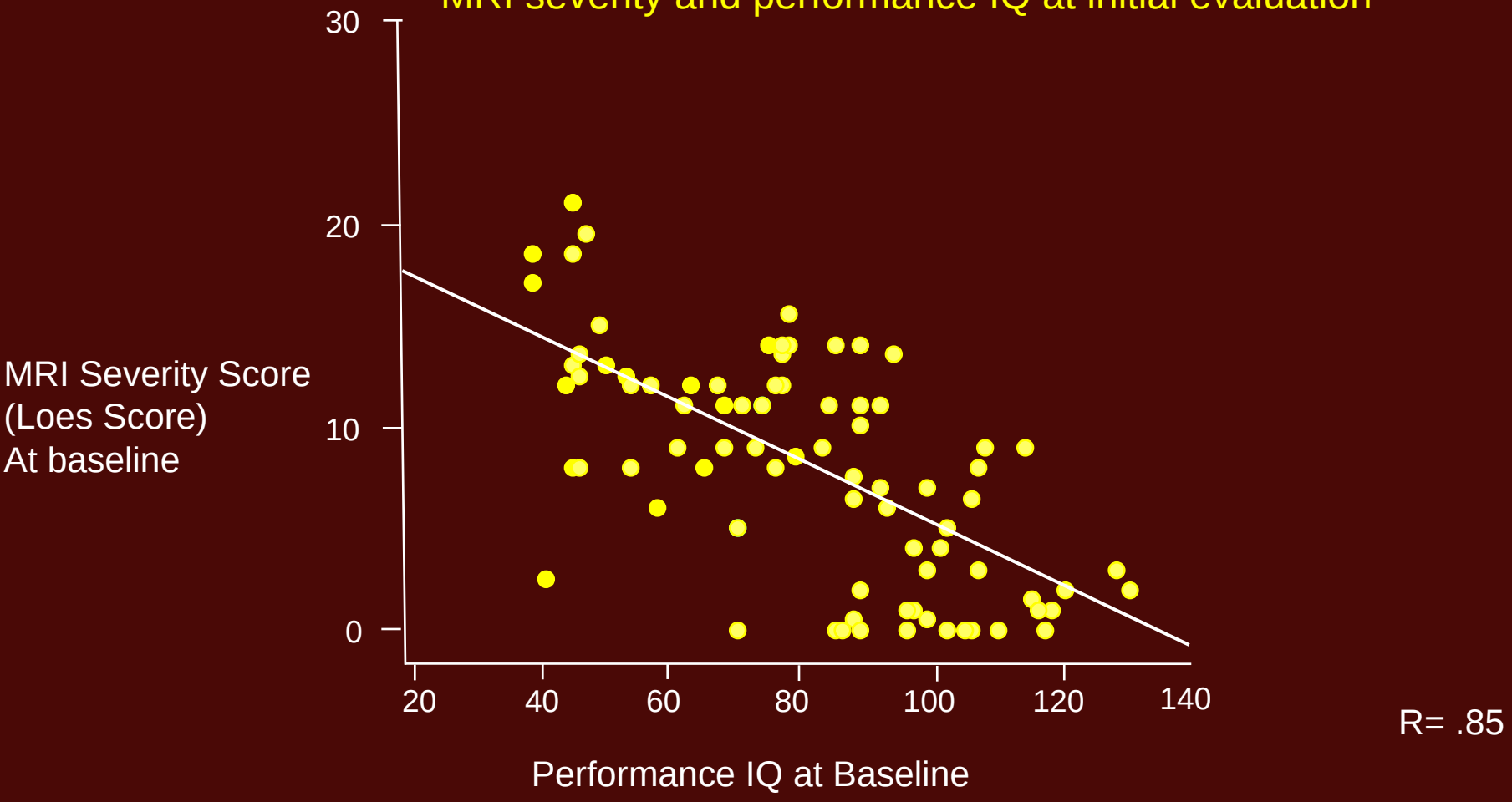
Age at HCT and outcome

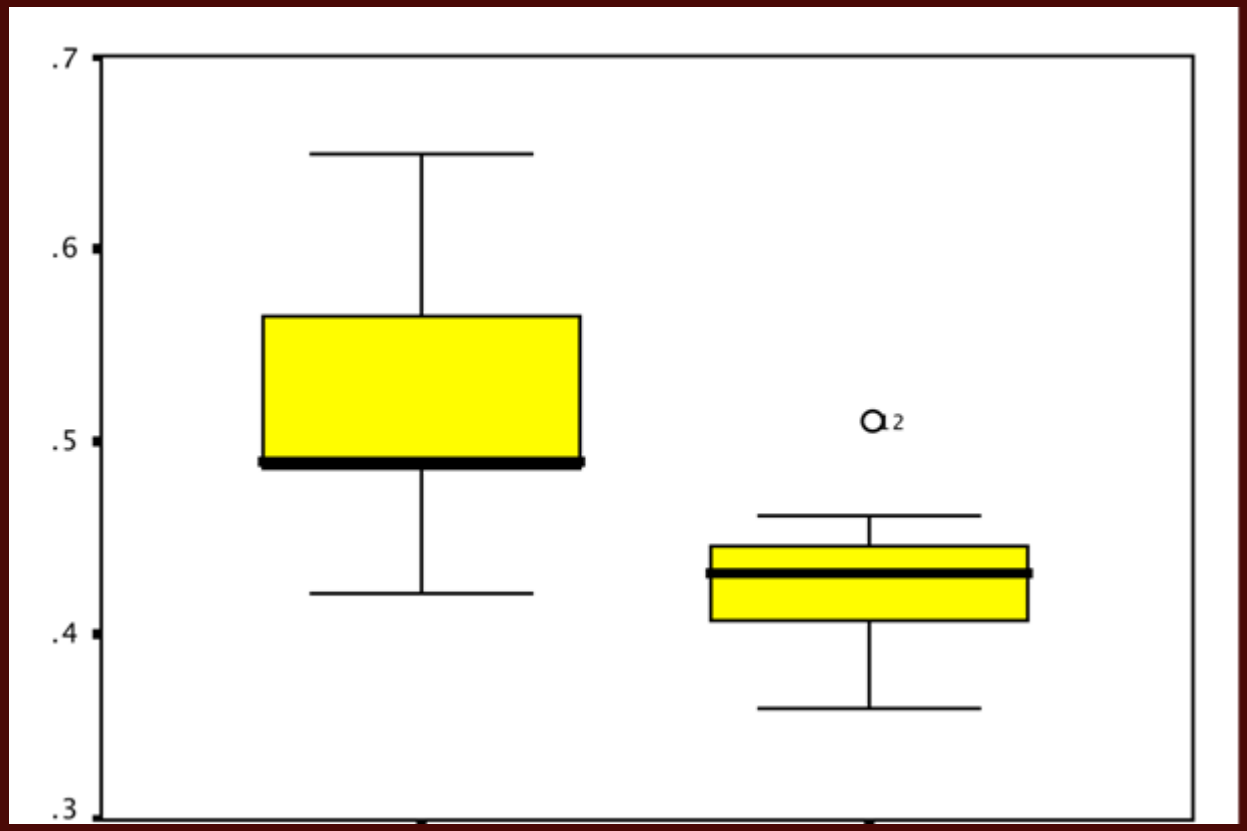
- The overall correlation of age at transplant with slope of development was .46, $p < .001$.
- Comparing slopes of mental development for HCT after 24 months = 0.24 with HCT before 24 months = 0.62.
- Difference is significant $p < .004$.
- The earlier the transplant the better the overall cognitive outcome.

Validation of results

- Correlation of two methods of measuring outcome- <u>imaging and neuropsych test results</u>
- Important that the relationship between the two methods be based on theory
- If there is a relationship, strengthens the rationale for their use in measuring outcome

Adrenoleukodystrophy
MRI severity and performance IQ at initial evaluation



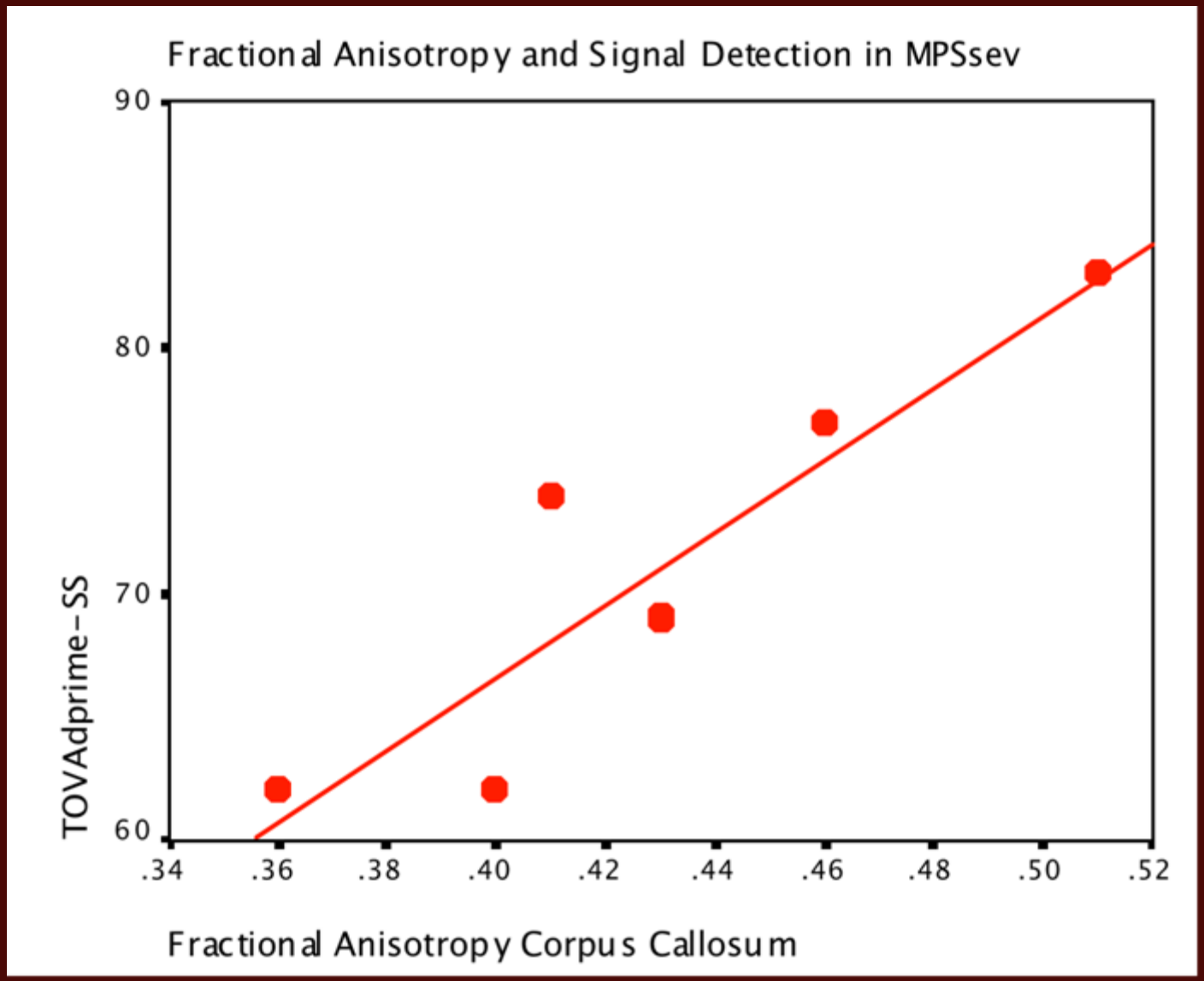


Callosal Fractional Anisotropy

Mann Whitney U
P< .04

MPS I attenuated
ERT – no HCT

MPS I severe (Hurler)
HCT – no ERT


Fractional Anisotropy and Signal Detection in MPSsev
90
80
70
60
TOVAdprime-SS
.34
.36
.38
.40
.42
.44
.46
.48
.50
.52
Fractional Anisotropy Corpus Callosum

Conclusions

- Research design for small sample longitudinal studies or clinical trials will depend on disease, type of treatment, and purpose of the study
- Outcome measures (either MRI or neuropsych) for such studies need to be
  - Quantitative
  - Sensitive to change from either disease or treatment
  - Reproducible
  - Theoretically sound

Thanks to…..

Our patients and research participants

Lysosomal Disease Network group at the University of Minnesota
- Chester Whitley, Ph.D., M.D.
- Igor Nestrasil, M.D.
- Alia Ahmed, M.D.
- Kendra Bjoraker, Ph.D
- Richard Ziegler, Ph.D.
- Julie Eisengart, Ph.D.
- Kate Delaney
- Brenda Diethelm Okita
- David Erickson

Blood and Marrow Transplant group
- Paul Orchard, M.D.
- Jakub Tolar, M.D.
- Teresa Kivisto, R.N.

Our research uses the resources of
- University of Minnesota Center for Magnetic Resonance Research
- University of Minnesota Center for Neurobehavioral Development
- University of Minnesota Supercomputer Center

Support from NIH U54 NS065768-01, the National MPS Society,

Genzyme, Shire HGT, and Biomarin