BIOMARKERS AND OUTCOME MEASURES IN NEUROLOGY CLINICAL TRIALS

Jeremy M. Shefner
Kemper and Ethel Marley Professor and Chair of Neurology
Barrow Neurological Institute
Disclosures

• Personal compensation received from Cytokineti
  ctics, Biogen, MT Pharma, Neuraltus, Brainstorm, Pharnext

• Research funding received from Cytokinetics,
  Biogen, Synapse, Neuraltus, Biotie, Amylyx, ALS
  Association, MDA, NINDS
What is a Biomarker?

- generally refers to a measurable indicator of some biological state or condition. (Wikipedia)
- a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention (NIH)
- Definition is so broad that a biomarker can be any one of above
- Functional and clinically relevant endpoints can also be a biomarkers
<table>
<thead>
<tr>
<th>Biomarker Category</th>
<th>Utility</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Engagement</td>
<td>• The drug interacts with its intended molecular target <em>in vivo</em></td>
<td>• PET receptor occupancy studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Measurement of molecular complexes in vivo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Binding in a surrogate compartment (e.g., lymphocytes)</td>
</tr>
<tr>
<td>Pharmacokinetic</td>
<td>• The drug reaches its desired molecular site of action</td>
<td>• Pharmacokinetics in CSF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• CNS uptake studies</td>
</tr>
<tr>
<td>Pharmacodynamic</td>
<td>• The intended molecular effect produces the desired biological effect.</td>
<td>• Effect on Molecular Target:</td>
</tr>
<tr>
<td></td>
<td>• Useful for determining therapeutic dose range;</td>
<td>• Effect on Presumed Downstream Marker</td>
</tr>
<tr>
<td></td>
<td>• potential candidate for becoming a surrogate</td>
<td>• Plasma proteomics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Plasma metabolomics</td>
</tr>
<tr>
<td>Diagnosis/Stratification</td>
<td>• The targeted disease state is present, and/or the desired patient</td>
<td>• Genetics</td>
</tr>
<tr>
<td></td>
<td>population can be stratified to optimize risk benefit ratio and</td>
<td>• Blood-based makers</td>
</tr>
<tr>
<td></td>
<td>probability of success</td>
<td>• CSF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Imaging</td>
</tr>
<tr>
<td>Disease Outcome</td>
<td>• Assessment of effect on Clinical or Pathological Disease measures</td>
<td>• Clinical Outcome Measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Imaging</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Anatomical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Functional</td>
</tr>
<tr>
<td>Safety</td>
<td>• Presence and/or severity of potential target organ toxicity is</td>
<td>• Biochemical (common/special labs)</td>
</tr>
<tr>
<td></td>
<td>measurable</td>
<td>Electrophysiological (QTc)</td>
</tr>
</tbody>
</table>
Integration of Biomarker Strategies into Drug Development Decision Making

<table>
<thead>
<tr>
<th>Decision Points</th>
<th>PDCR</th>
<th>FIH</th>
<th>Ph.2</th>
<th>Ph.3</th>
<th>Filing</th>
<th>Launch</th>
<th>LCM</th>
<th>Post-launch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead Opt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Mechanism of Action**
- **Pharmacokinetics**
- **Pharmacodynamics**
- **Disease Dx/Stratification**
- **Disease Progression**
- **Safety**

*Courtesy of Jesse Cedarbaum*
**P<sub>D</sub> Markers:** measure of compounds ability to interact with its intended target leading to a biological effect.

- **P<sub>D</sub> type:**
  - Biochemical:
    - Enzyme substrate
    - mRNA/ Protein
  - Imaging:
    - PET
    - MRI
    - CT
  - Physiology:
    - Axonal excitability
    - MUNE

- **P<sub>D</sub> use:**
  - Test biological hypothesis in human
  - Combine with P<sub>K</sub>
  - Select dose:
    - Efficacious range
    - Safe range
CSF SOD1 as a PD Biomarker for ALS

• SOD1 Antisense Oligonucleotides (ASO) lower SOD1 and prolong survival in animal models
• SOD1 natural history data suggests we will be able to determine benefit
• ASOs safe in prior IONIS/Biogen Phase I in SOD1 ALS
Antisense Oligos Decrease CSF in SOD1 G93A Rats

- **Brain SOD1 Protein** (% Saline Control)
- **CSF SOD1 Protein** (% Saline Control)

Bar graph showing the decrease in SOD1 protein in CSF and brain after oligo treatment compared to saline control. The graph includes error bars for each group.

Scatter plot showing the correlation between brain SOD1 percent decrease and CSF SOD1 percent decrease with an R² value of 0.9377.

Tim Miller
SOD1 in CSF Varies Little Over Time

Winer et al. 2013, JAMA Neurology
Regulatory T Cell and their function are reduced in ALS

Henkle et al., 2013
pNFH levels correlate to patient survival

Steinacker et al., (2015): 253 ALS Subjects

Oeckl et al., (2016): Correlation to survival

Level of pNFH in the blood or CSF is a prognostic for patient survival and rate of disease progression
pNFH or NFL levels are relatively stable over time.

Level of pNFH or NFL in blood or CSF could be used to monitor drug effects.

Turner and colleagues (2016)

Bowser lab
[\textsuperscript{11}C]PBR-28 identifies activated microglia in ALS

Increased binding to activated microglia in Motor cortex and other areas of interest for ALS.

Potential use as PD marker in trials that target microglial activation (RNS60, ibudilast)

• 257 patients, 3 doses vs placebo for 24 weeks
• Primary endpoint: new GdE lesions
  – Clear dose response; lesions reduced by 69% at highest dose
• Secondary endpoint: relapse rate
  – No dose response; overall, relapse rate declined by 32% (p=0.27)

Kappos et al
Lancet 2008
RRMS: Gd+ lesions
A marker of disease activity

Kappos et al
Lancet 2008
Placebo-Controlled Phase 3 Study of Oral BG-12 for Relapsing Multiple Sclerosis

Ralf Gold, M.D., Ludwig Kappos, M.D., Douglas L. Arnold, M.D., Amit Bar-Or, M.D., Gavin Giovannoni, M.D., Krzysztof Selmaj, M.D., Carlo Tornatore, M.D., Marianne T. Sweetser, M.D., Ph.D., Minhua Yang, M.S., Sarah I. Sheikh, M.D., and Katherine T. Dawson, M.D.,
for the DEFINE Study Investigators*

Hazard ratio vs placebo (95% CI)
BG-12 bid: 0.51 (0.40 – 0.66); p < 0.001
BG-12 tid: 0.50 (0.39 – 0.65); p < 0.001

Estimated proportion with relapse at 2 years
BG-12 bid: 27%
BG-12 tid: 26%
placebo: 46%

No. at risk
Placebo 408 356 321 282 243 224 205 190 115
BG-12 bid 410 353 324 303 286 267 255 243 154
BG-12 tid 416 346 322 301 286 270 251 244 166
Functional markers serve as intermediate stage endpoints

- Strength
- Pulmonary function
- 6 minute walk
- Timed up and go
- Many others
Methods of assessment can be very important

- Strength is a functional marker that may be important in studying many diseases
- However, how it is measured affects its utility
  - Single muscle group
    - Vital capacity
    - Handgrip
  - Global Assessment
    - MRC manual muscle testing
      - Any number of muscle can be tested on a 0-5 ordinal scale
    - Quantitative muscle testing
      - TQNE
      - HHD
Uneven Steps Between MRC Grades

MMT scale compared with actual dynametric force measurement of the biceps brachii

(modified from van der Ploeg: J Neurol, 1984)
Quantitative Muscle Testing: Standardized Training and Validation

- Standardized positions
- Video and hands on training
- Requirement for demonstration of adequate training
- Test-retest reliability criterion
Decline in individual muscle groups

Biogen Empower Study
Proportion of zero force per muscle

Liu et al., 2017
Sensitivity of time to first zero muscle compared to survival

Liu et al., 2017
## HHD0 vs other measures

<table>
<thead>
<tr>
<th>Hazard Ratio of Treated versus Control</th>
<th>Cumulative Proportion of Zero Score Events at Month 12</th>
<th>Sample Size Required* per Treatment Group for 90% Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control*</td>
<td>Treatment</td>
</tr>
<tr>
<td>0.5</td>
<td>64%</td>
<td>40%</td>
</tr>
<tr>
<td>0.4</td>
<td>64%</td>
<td>33%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hazard Ratio</th>
<th>Cumulative Proportion of Deaths at Month 12</th>
<th>Sample Size Required* per Treatment Group for 90% Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control*</td>
<td>Treatment</td>
</tr>
<tr>
<td>0.6</td>
<td>17%</td>
<td>11%</td>
</tr>
<tr>
<td>0.5</td>
<td>17%</td>
<td>9%</td>
</tr>
<tr>
<td>0.4</td>
<td>17%</td>
<td>7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difference at Month 12</th>
<th>Mean Change from Baseline over 12 Months</th>
<th>Sample Size Required* per Treatment Group for 90% Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control*</td>
<td>Treatment</td>
</tr>
<tr>
<td>1.5</td>
<td>-11</td>
<td>-9.5</td>
</tr>
<tr>
<td>2.0</td>
<td>-11</td>
<td>-9</td>
</tr>
<tr>
<td>2.2</td>
<td>-11</td>
<td>-8.8</td>
</tr>
</tbody>
</table>

Liu et al., 2017
Electrical Impedance Myography (EIM)

- Pioneered by Seward Rutkove
- Technique based on the application of high-frequency electrical current to localized areas of muscle with measurement of resulting voltages.
  - Painless
  - Non-invasive
  - Can apply to virtually any superficial muscle
    - Tongue, paraspinals, proximal muscles all possible
- Sensitive to alterations in muscle composition, structure, atrophy
A. Healthy Muscle

Applied current

Surface voltages result as current flows through resistance and capacitance in the tissue. The capacitance also causes a phase shift.

B. Diseased Muscle

Applied current

Increased tissue resistance in diseased muscle results in a larger voltage, and reduced capacitance results in less phase shift.
EIM has been studied in several NM diseases

- ALSA-funded Longitudinal Study in ALS
- Ongoing SBIR
- Neuralstem study of stem cells in ALS
- SMA
- Animal models
- A variety of muscle diseases
EIM vs other measures

Coefficient of Variation: 0.62

From: Rutkove et al., 2012

Shefner et al., 2011

Coefficient of Variation: 0.81

Coefficient of Variation: 0.93

MUNE
Coefficient of Variation: 0.72

From: Rutkove et al., 2012
Shefner et al., 2011

Coefficient of Variation: 0.62
Clinically Relevant Endpoints

• Clinically relevant endpoints are required for phase 3 trials
  • May be subjective (I feel better) or objective (I can walk across the room better)
  • Survival
  • Time to event
• However:
  • Clinical Relevance is often a fuzzy target
    • Is vital capacity clinically relevant?
    • Is strength clinically relevant?
  • Clinical relevance does not necessarily imply relevance to potential therapeutic mechanism
  • Issues of variability may limit utility
    • Disease related
    • Measure related
Functional Scales

• Functional Scales are considered clinically relevant
  • They directly ask patients about functional capacity, or assess these functions by observation
  • However, size of effect that is important is not always clear
  • The scale properties are critical and often undefined
    • Interval Scaling
    • Continuous vs discrete
Functional Scales

- Can be disease or attribute specific
- Scoring of individual items should have characteristics of an interval scale: i.e., a change of 1 unit should be the same anywhere on the scale
- Often comprised of well defined domains capable of assessing different aspects of function
Limitations of Functional Scales

- Often combine attributes so it is difficult to attribute a change to a specific function
- The minimum clinically significant change is undetermined
- Lack of interval scaling may mask small changes
- Variability of scoring may limit use or increase sample size
- Individual items are usually strikingly non-linear; averaging many items together can create appearance of linearity
Commonly Used Functional Scales

- Kurtzke EDSS
- ALS Functional Rating Scale- Revised (ALSFRS-R)
- Unified Parkinson’s Disease Rating Scale (UPDRS)
- Alzheimer’s Disease Assessment Scale-cognitive subscale (ADAS-cog)
- Modified Rankin Scale
Disability Scores do not linearly decline in MS

From: www.mult-sclerosis.org
ALSFRS-R

- Speech
- Salivation
- Swallowing
- Handwriting
- Cutting food, handling utensils
- Dressing and Hygiene
- Turning in bed and adjusting bed clothes
- Walking
- Climbing stairs
- Dyspnea
- Orthopnea
- Respiratory insufficiency

From: Cedarbaum et al, 1999
ALSFRS-R Sub-Domains

- Changes in sub-domain scores validated across two studies conducted a decade apart in time

Respiratory questions are 25% of the scale, but only 13% of the change over time

Cedarbaum et. al 1999
Edaravone Phase 3 Trial

Edaravone ALS 19 Study Group, 2017
Binary/Time to Event

• Advantages
  – Easy to understand
  – Power calculations are straightforward

• Disadvantages
  – Only subjects who reach endpoint are useful
  – Only 1 change of state is deemed important
Time to Event: Survival

• Useful only when events are likely to occur
  – Stroke
  – SAH
  – ALS

• Depending on disease state and target, may not be sensitive to experimental intervention
  – Nuedexta for Emotional Lability
    • Approved for ALS, but unlikely to impact survival
Survival as an outcome measure in ALS

From: Drachmann et al., 2000

From Lacomblez et al., 1996
Time to event is an example of a binary endpoint

- Time to event endpoints
  - Survival
  - Hospital readmission
  - Time to new vascular event
  - Time to initiation of NIV

- Other binary endpoints
  - Achieving functional independence
  - Achieving independent ambulation
Binary outcomes

A 7 point scale is often dichotomized (0-2 vs 3-6) for primary analysis

### Table 1. Modified Rankin Scale

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No symptoms at all</td>
</tr>
<tr>
<td>1</td>
<td>No significant disability despite symptoms; able to perform all usual duties and activities</td>
</tr>
<tr>
<td>2</td>
<td>Slight disability; unable to perform all previous activities, but able to take care of self without assistance</td>
</tr>
<tr>
<td>3</td>
<td>Moderate disability; requiring some help, but able to walk without assistance</td>
</tr>
<tr>
<td>4</td>
<td>Moderately severe disability; unable to walk without assistance and unable to attend to own bodily needs without assistance</td>
</tr>
<tr>
<td>5</td>
<td>Severe disability; bedridden, incontinent, and requiring constant nursing care and attention</td>
</tr>
<tr>
<td>6</td>
<td>Dead</td>
</tr>
</tbody>
</table>

Adapted from Saver, 2007
How does one choose which biomarker/endpoint is most appropriate?

• Development stage
• Qualities intrinsic to marker/endpoint
  – Relevance to clinically important endpoints
  – Variability
    • Measurement related
    • Disease related
  – If a binary endpoint, how many events expected?
Summary

• The choice of endpoint is critical in the design of clinical trials
• Endpoints should be reliable, meaningful, and sensitive to disease modification
• An appropriate choice of endpoint should increase the probability of correctly determining whether the goals of the study are met
• The currently available toolbox of measures is not adequate to meaningfully shorten trials or reduce sample size for most neurological diseases