Translational Science in Drug Development: Surrogate Endpoints, Biomarkers, and More

Duke-Margolis Center for Health Policy | Virtual Meeting
May 24-25, 2022
Welcome and Overview | Day 2

Mark McClellan
Director, Duke-Margolis Center for Health Policy
Statement of Independence

The Robert J. Margolis, MD, Center for Health Policy is part of Duke University, and as such it honors the tradition of academic independence on the part of its faculty and scholars. Neither Duke nor the Margolis Center take partisan positions, but the individual members are free to speak their minds and express their opinions regarding important issues.

For more details on relevant institutional policies, please refer to the Duke Faculty Handbook, including the Code of Conduct and other policies and procedures. In addition, regarding positions on legislation and advocacy, Duke University policies are available at http://publicaffairs.duke.edu/government.
Disclaimer

Funding for this workshop was made possible in part by a cooperative agreement from the U.S. Food and Drug Administration. The views expressed in written workshop materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services nor does mention of trade names, commercial practices, or organizations imply endorsements by the U.S. Government.
Meeting Agenda (Day 2)

12:00 pm Welcome and Overview

12:10 pm Session 3: Clinical Validation and Regulatory Acceptance of Biomarkers as Surrogate Endpoints

1:50 pm Break

2:05 pm Session 4: Beyond Surrogate Endpoints: Other Ways Translational Science Can Support Drug Development

3:30 pm Session 5: Opportunities and Challenges for Incorporation of Translational Science in Clinical Development Programs

4:15 pm Closing Remarks

4:25 pm Adjournment
Session 3: Clinical Validation and Regulatory Acceptance of Biomarkers as Surrogate Endpoints

12:10 pm – 1:50 pm EST
Steve Ryder
Chief Medical Officer
Rallybio Inc.
Translational Science in Drug Development: Surrogate Endpoints, Biomarkers, and More
May 24th and 25th 2022
Session 3: Clinical Validation and Regulatory Acceptance of Biomarkers as Surrogate Endpoints

Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

Steve Ryder
Chief Medical Officer
Rallybio

Contributors: David Thompson, Tino Melian, Kenji Fujita and Colleagues
Disclosure

I am a full-time employee and hold an equity interest in Rallybio
At the time the presented work was done, I and all contributors were
full-time employees of and held equity interests in Alexion
Pharmaceuticals
Overview

- Rare/ultra-rare diseases are generally poorly understood and poorly researched.
- This extends to both the preclinical and clinical areas.
- Almost always there is no precedent for designing studies in the treatment of rare/ultra-rare disease. Irreversible disease morbidity/mortality may constrain design and analytical approaches.
- Assessment tools are often unavailable and almost never validated in the rare/ultra-rare disease under study.
- One approach to improve the availability of assessment tools is to thoroughly review assessment tools in alternative disease areas with relevant morbidity/functional disability and pre-apply them to natural history cohorts.
- This importation and logical application of assessment tools was successfully used in the development of asfotase alfa (Strensiq®) in the treatment of patients with juvenile-onset hypophosphatasia (HPP).
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

Hypophosphatasia (HPP)

Genotypes vs Phenotypes/Onset

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Phenotypes/Onset (Perinatal/Infantile, Juvenile, Adult)</th>
</tr>
</thead>
</table>

Health

Hypophosphatasia

Ca, P, TNSALP, ANKH, ENPP1, ATP, HAP, PPi

Osteoblast
Chondrocyte
Ameloblast
Odontoblast
Cementoblast

5/22/2022  Bowden and Foster, Drug Design, Development and Therapy (2018)
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

Strensiq®

A

TNALP dimer

IgG₁ Fc

D₁₀

TNALP-D₁₀

HAP

B

Percent survival

WT

Alpl⁻/⁻ ERT

Alpl⁻/⁻
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

Strensiq® (asfotase alfa) [hypophosphatasia; HPP]

### Biology
- **TNSALP**
- Substrates (PPI, PLP)

### Pathology
- **Skeletal System**
  - Bone Mineralization (Bone Biopsy; DEXA)
  - Rickets Severity (RGI-C; RSS)
  - Growth

### Clinic
- **Physical Function**
  - Ambulation (6MWT)
  - Development Milestones (BSID-III; BOT-2)
  - Strength (HDD)

### QoL / Survival
- **Survival/Respiratory Status**
- **Activities of Daily Living/Pain** (CHAQ; PODCI; LEFS; BPI-SF)

6MWT = 6 minute walk test; BOT 2 = Bruininks-Oseretsky Test of Motor Proficiency; BPI SF = Brief Pain Inventory-Short Form; BSID III = Bayley Scales of Infant and Toddler Development; CHAQ = Child Health Assessment Questionnaire; DEXA = dual energy x ray absorptiometry; HDD = hand-held dynamometry; LEFS = Lower Extremity Functional Scale; PODCI = Pediatric Outcomes Data Collection Instrument; PLP = pyridoxal 5'-phosphate; PPI = inorganic pyrophosphate; RGI C = Radiographic Global Impression of Change; RSS = rickets severity scale
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

Strensiq® (asfotase alfa) [hypophosphatasia; HPP]

Perinatal/Infantile-onset

Biology
- TNSALP
- Substrates (PPI, PLP)

Pathology
- Skeletal System
  - Bone Mineralization (Bone Biopsy; DEXA)
  - Rickets Severity (RGI-C; RSS)
  - Growth

Clinic
- Physical Function
  - Ambulation (6MWT)
  - Development Milestones (BSID-III; BOT-2)
  - Strength (HDD)

QoL / Survival
- Survival/Respiratory Status
  - Activities of Daily Living/Pain (CHAQ; PODCI; LEFS; BPI-SF)

6MWT = 6 minute walk test; BOT 2 = Bruininks-Oseretsky Test of Motor Proficiency BPI SF = Brief Pain Inventory-Short Form; BSID III = Bayley Scales of Infant and Toddler Development; CHAQ = Child Health Assessment Questionnaire; DEXA = dual energy x ray absorptiometry; HDD = hand-held dynamometry; LEFS = Lower Extremity Functional Scale; PODCI = Pediatric Outcomes Data Collection Instrument; PLP = pyridoxal 5′ phosphate; PPI = inorganic pyrophosphate; RGI C = Radiographic Global Impression of Change; RSS = rickets severity scale
Ventilatory Support and Patient Outcomes: 20 Historical Control (top) and 14 Asfotase Alfa–Treated (bottom) Patients. Whyte et al JCEM 2016
QoL/Survival

Perinatal/Infantile-onset

STRENSIQ™ (asfotase alfa) injection; US Label; Figure 1: Overall Survival in STRENSIQ-Treated versus Historical Control Patients with Perinatal/Infantile-Onset HPP

x-axis truncated at Week 528; there were 6 historical control patients censored after truncation with censored times ranging from 5.98 to 10.30 weeks
Strensiq® (asfotase alfa) [hypophosphatasia; HPP]

Juvenile-onset

**Biology**
- TNSALP
- Substrates (PPI, PLP)

**Pathology**
- Skeletal System
  - Bone Mineralization (Bone Biopsy; DEXA)
  - Rickets Severity (RGI-C; RSS)
  - Growth

**Clinic**
- Physical Function
  - Ambulation (6MWT)
  - Development Milestones (BSID-III; BOT-2)
  - Strength (HDD)

**QoL / Survival**
- Survival/Respiratory Status
  - Activities of Daily Living/Pain (CHAQ; PODCI; LEFS; BPI-SF)

6MWT = 6 minute walk test; BOT 2 = Bruininks-Oseretsky Test of Motor Proficiency BPI SF = Brief Pain Inventory-Short Form; BSID III = Bayley Scales of Infant and Toddler Development; CHAQ = Child Health Assessment Questionnaire; DEXA = dual energy x ray absorptiometry; HDD = hand-held dynamometry; LEFS = Lower Extremity Functional Scale; PODCI = Pediatric Outcomes Data Collection Instrument; PLP = pyridoxal 5' phosphate; PPI = inorganic pyrophosphate; RGI C = Radiographic Global Impression of Change; RSS = rickets severity scale
Clinic

Juvenile-onset

---

Whyte et al. JCI Insight. 2016
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

Clinic

Juvenile-onset
Clinic

Juvenile-onset

BOT2: Shuttle Run

Baseline
22.2 sec

6 Months
12.3 sec

36 Months
8.6 sec
Strensiq® (asfotase alfa) [hypophosphatasia; HPP]

Juvenile-onset

Biology
- TNSALP
- Substrates (PPI, PLP)

Pathology
- Skeletal System
  - Bone Mineralization (Bone Biopsy; DEXA)
  - Rickets Severity (RGI-C; RSS)
  - Growth

Clinic
- Physical Function
  - Ambulation (6MWT)
  - Gait (mPOMA-G)
  - Development Milestones (BSID-III; BOT-2)
  - Strength (HDD)

QoL / Survival
- Survival/Respiratory Status
- Activities of Daily Living/Pain (CHAQ; PODCI; LEFS; BPI-SF)

6MWT = 6 minute walk test; BOT 2 = Bruininks-Oseretsky Test of Motor Proficiency BPI SF = Brief Pain Inventory-Short Form; BSID III = Bayley Scales of Infant and Toddler Development; CHAQ = Child Health Assessment Questionnaire; DEXA = dual energy x ray absorptiometry; HDD = hand-held dynamometry; LEFS = Lower Extremity Functional Scale; PODCI = Pediatric Outcomes Data Collection Instrument; PLP = pyridoxal 5′ phosphate; PPI = inorganic pyrophosphate; RGI C = Radiographic Global Impression of Change; RSS = rickets severity scale
Clinic

Juvenile-onset

Performance-Oriented Assessment of Mobility Problems in Elderly Patients

Mary E. Tinetti, MD
Development and validation of a modified performance-oriented mobility assessment tool for assessing mobility in children with hypophosphatasia

Dawn Phillipsa,*, Donna Griffimb, Tracy Przybylskib, Erica Morrisonb, Amy L. Reevesb, Marc Vallee, Kenji P. Fujitad and Katherine L. Madsonb,1

aDivision of Physical Therapy, University of North Carolina, Chapel Hill, NC, USA
bShriners Hospitals for Children, St. Louis, MO, USA
cBiostatistics, Alexion Pharmaceuticals, Inc., Boston, MA, USA
dClinical Development, Alexion Pharmaceuticals, Inc., Boston, MA, USA
mPOMA-G Review and Adaptation

- An expert panel of physicians, physical therapists, and statisticians evaluated the suitability of the POMAG for assessing gait in children with HPP using observational, non-instrumented video footage
- Most POMA-G components were relevant and could be used
- Several modifications were recommended to adapt it for use in children with HPP resulting in the modified POMA-G (mPOMA-G)
- Modifications included:
  1. removing the rating of initiation of gait;
  2. expanding the assessment of step length and step continuity;
  3. removing the rating of path;
  4. adding new items within observations for step length and height;
  5. clarifying descriptions of specific items to increase sensitivity and consistency among raters; and
  6. Creating a scoring key that provides detailed instructions and illustrations
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

**mPOMA-G Validation**

- Concurrent validation of mPOMA-G scores was made to other outcome measures assessing functional impairments
- Pearson correlation coefficients demonstrated strong concurrent validity between mPOMA-G scores and
  - Childhood Health Assessment Questionnaire (CHAQ) Disability Index,
  - Pediatric Outcomes Data Collection Instrument (PODCI), and
  - 6-Minute Walk Test.

---

**Figure A**

(A) CHAQ Disability Index

Infantile

Childhood

![Graphs showing correlation between mPOMA-G and CHAQ scores](image-url)
Ryder: Use of Imported Clinical Assessment Tools in Rare Disease: A Case Study

mPOMA-G Application

• Conducted in accordance with GCP and after IRB review and approval. Parents or legal guardians of the patients provided written informed consent and patients provided written assent. Visible faces in videos were permanently blurred, and all videos (n = 64) were assigned a new masking code and randomized before each scoring.

• 3 trained physical therapists applied the mPOMAG to score videos of 14 children with HPP while walking.

• Patients (age range: 5–15 years) were enrolled in an open-label asfotase alfa clinical study (NCT00952484) with extension (NCT01203826) or a natural history study (NCT02235493).

• Videos of children in the treated group (n = 8) were taken before and after treatment; videos of children in the natural history group (n = 6) were taken at routine follow-up visits.
mPOMA-G Application

The median (range) rate of change per year was 2.51/year (0.0, 4.6) in asfotase alfa-treated patients compared with 0.33/year (0.0, 0.9) for untreated historical controls (p=0.0303, Wilcoxon rank-sum test)

Figure 7: MPOMA-G Results for Historical Controls vs. Treated Patients
Forward Recommendation

- In the development of rare/ultra-rare disease, build in a forward review of assessment tools in alternative disease areas with relevant morbidity/functional disability
- Consider its application in the development program and review/modify the clinimetric characteristics when applied to the specific disease under study
- Conduct rater training and assessment tool validation using established scales
- Apply to relevant natural history and study drug datasets
Henrik Zetterberg
Professor of Neurochemistry
University of Gothenburg and University College London
Development and validation of cerebrospinal fluid and blood biomarkers for neurodegenerative diseases

Henrik Zetterberg, MD, PhD
Department of Psychiatry and Neurochemistry,
University of Gothenburg, Sweden;
Institute of Neurology and UK Dementia Research Institute, UCL, UK
Fluid biomarker candidates of potential relevance to neurodegenerative disease

- Neurogranin
- NSE, SBPDs and UCHL1
- SNAP25, SYT1, SV2A
- P-tau
- Axon terminals
- Dendrites
- Astrogial cell
- Neuronal soma
- Brain capillary
- Neuroglia
- Amyloid plaques

Potential markers:
- S100B
- GFAP
- Interleukins/cytokines
- TREM2
- Chitotriosidase
- MCP-1
- Complement proteins
- APP and Aβ
- CSF/serum albumin ratio
- CSF PDGFRβ

Missing: TDP-43 α-syn (?)
A = amyloid pathology
CSF Aβ42 is decreased in AD

Aβ42 AD vs Control

… 90 more studies...
CSF Aβ42 is a marker of amyloid plaque pathology

Study design: 155 autopsy cases
Plaque counts – neocortex and hippocampus
Post-mortem CSF samples

CSF Aβ42 correlates with amyloid cortical amyloid plaque load

Strozyk et al, Neurology 2003;60:652-656.
CSF Aβ42 concentration correlates with amyloid PET

Study design: 118 patients with cognitive complaints examined for both CSF biomarkers - as part of clinical routine – 2 years and amyloid 18F-flutemetamol PET

Cut-offs: CSF Aβ42 < 647 pg/mL 18F-flutemetamol PET > 1.42

Original cohort n= 118

Validation cohort n= 38

Positive PET+CSF or Negative PET+CSF 92 %

Positive PET+CSF or Negative PET+CSF 97 %

Palmquist S, et al, JAMA Neurol 2014
CSF Aβ42 concentration may be decreased in neuroinflammatory conditions

Augutis et al., Multiple Sclerosis 2013
CSF Aβ42 concentration may be decreased in normal pressure hydrocephalus

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>iNPH (n = 28)</th>
<th>HI (n = 20)</th>
<th>iNPH/HI ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFL</td>
<td>1,260 (640-2,290) †</td>
<td>825 (653-1,243)</td>
<td>1.53^b</td>
</tr>
<tr>
<td>MBP</td>
<td>1.5 (1.1-1.9) ↔</td>
<td>1.3 (1.0-1.5)</td>
<td>1.12 NS</td>
</tr>
<tr>
<td>Aβ38</td>
<td>637 (438-894) †</td>
<td>1,641 (1,231-2,173)</td>
<td>0.39^c</td>
</tr>
<tr>
<td>Aβ40</td>
<td>5,067 (3,634-6,573) †</td>
<td>10,083 (7,626-12,794)</td>
<td>0.50^c</td>
</tr>
<tr>
<td>Aβ42</td>
<td>221 (156-325) †</td>
<td>498 (391-669)</td>
<td>0.44^c</td>
</tr>
<tr>
<td>sAPPα</td>
<td>505 (338-739) †</td>
<td>1,110 (727-1,244)</td>
<td>0.46^c</td>
</tr>
<tr>
<td>sAPPβ</td>
<td>176 (110-258) †</td>
<td>414 (250-545)</td>
<td>0.43^c</td>
</tr>
<tr>
<td>t-tau</td>
<td>39 (34-50) †</td>
<td>84 (64-107)</td>
<td>0.47^c</td>
</tr>
<tr>
<td>p-tau</td>
<td>39 (33-50) †</td>
<td>59 (47-75)</td>
<td>0.67^d</td>
</tr>
<tr>
<td>IL-8</td>
<td>34 (26-38) ↔</td>
<td>31 (26-40)</td>
<td>1.10 NS</td>
</tr>
<tr>
<td>IL-10</td>
<td>0.66 (0-0.9) ↔</td>
<td>0.67 (0-0.8)</td>
<td>0.99 NS</td>
</tr>
<tr>
<td>MCP1</td>
<td>746 (602-874) †</td>
<td>628 (564-686)</td>
<td>1.19^b</td>
</tr>
<tr>
<td>Albumin CSF</td>
<td>287 (188-408) ↔</td>
<td>232 (203-280)</td>
<td>1.24 NS</td>
</tr>
<tr>
<td>Albumin ratio</td>
<td>6.8 (5.0-10) ↔</td>
<td>5.6 (4.5-6.4)</td>
<td>1.22 NS</td>
</tr>
</tbody>
</table>

Abbreviations: Aβ = amyloid β; HI = healthy elderly individuals; IL = interleukin; iNPH = idiopathic normal-pressure hydrocephalus; LCSF = lumbar CSF; MBP = myelin basic protein; MCP1 = monocyte chemoattractant protein 1; NFL = neurofilament light protein; NS = nonsignificant; p-tau = phosphorylated tau; sAPP = soluble amyloid precursor protein; t-tau = total tau.

^a Arrows indicate levels in iNPH in comparison with HI. Values are given as median (Q1–Q3 range).

^b p ≤ 0.05.

^c p ≤ 0.001.

^d p ≤ 0.01.

Jeppsson et al., Neurology 2013
...and there may be constitutively low Aβ producers who are close to the Aβ42 cutpoint for positivity

The CSF Aβ42/Aβ40 ratio corrects for this
CSF Aβ42/40 (or Aβ38) and PET Aβ

Cohort: Swedish BioFINDER
215 SCD/MCI (108 PET+ and 107 PET-)
PET: flutemetamol

The CSF Aβ42/Aβ40 ratio in clinical practice
CSF Aβ42/Aβ40 ratio – longitudinal data

Betthauer T et al., AAIC 2021
Neuroimaging: Multimodal Biomarkers
July 27, 2021
CSF $\text{A}\beta_{42}/\text{A}\beta_{40}$ ratio – longitudinal data

Betthauser T et al., AAIC 2021
Neuroimaging: Multimodal Biomarkers
July 27, 2021
How about plasma Aβ?

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect Size (95% CI)</th>
<th>AD</th>
<th>CTRL</th>
<th>Effect Size</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>% Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayeux, 1999</td>
<td>0.600 1.600</td>
<td>1.553</td>
<td>1.648</td>
<td>5.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arvanitakis, 2002</td>
<td>1.085 1.035</td>
<td>0.800</td>
<td>1.471</td>
<td>2.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fukumoto, 2003</td>
<td>1.057 1.090</td>
<td>0.911</td>
<td>1.227</td>
<td>4.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mayeux, 2003</td>
<td>1.168 1.168</td>
<td>1.061</td>
<td>1.287</td>
<td>4.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fagan, 2007</td>
<td>1.139 1.139</td>
<td>0.779</td>
<td>1.665</td>
<td>2.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fagan, 2007</td>
<td>1.000 1.000</td>
<td>0.580</td>
<td>1.723</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giedraitis, 2007</td>
<td>0.850 0.850</td>
<td>0.479</td>
<td>1.509</td>
<td>1.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posasrel, 2007</td>
<td>0.765 0.765</td>
<td>0.665</td>
<td>0.879</td>
<td>4.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schupf, 2008</td>
<td>0.994 0.994</td>
<td>0.986</td>
<td>1.002</td>
<td>5.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xu, 2008</td>
<td>0.848 0.848</td>
<td>0.632</td>
<td>1.139</td>
<td>2.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fagan, 2009</td>
<td>0.625 0.625</td>
<td>0.344</td>
<td>1.135</td>
<td>1.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roher, 2009</td>
<td>1.202 1.202</td>
<td>0.993</td>
<td>1.601</td>
<td>2.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedaghhat, 2009</td>
<td>1.313 1.313</td>
<td>1.219</td>
<td>1.415</td>
<td>4.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lui, 2010</td>
<td>0.926 0.926</td>
<td>0.878</td>
<td>0.977</td>
<td>5.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sundelöf, 2010</td>
<td>1.295 1.295</td>
<td>1.091</td>
<td>1.538</td>
<td>3.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head, 2011</td>
<td>0.645 0.645</td>
<td>0.604</td>
<td>0.689</td>
<td>5.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiu, 2013</td>
<td>1.841 1.841</td>
<td>1.328</td>
<td>2.552</td>
<td>2.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiu, 2013</td>
<td>1.217 1.217</td>
<td>1.175</td>
<td>1.261</td>
<td>5.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruiz, 2013</td>
<td>0.831 0.831</td>
<td>0.607</td>
<td>1.136</td>
<td>2.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yamamoto, 2013</td>
<td>1.012 1.012</td>
<td>0.531</td>
<td>1.927</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang, 2013</td>
<td>0.865 0.865</td>
<td>0.811</td>
<td>0.922</td>
<td>5.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bu, 2014</td>
<td>1.228 1.228</td>
<td>1.079</td>
<td>1.399</td>
<td>4.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rembach, 2014</td>
<td>1.013 1.013</td>
<td>0.953</td>
<td>1.077</td>
<td>5.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, 2014</td>
<td>0.991 0.991</td>
<td>0.979</td>
<td>1.103</td>
<td>5.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jiao, 2015</td>
<td>1.387 1.387</td>
<td>1.168</td>
<td>1.648</td>
<td>3.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim, 2015</td>
<td>0.865 0.865</td>
<td>0.759</td>
<td>0.986</td>
<td>4.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Janellidze, 2016</td>
<td>0.673 0.673</td>
<td>0.581</td>
<td>0.780</td>
<td>4.25</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All Studies

p=0.38718
Highly sensitive and precise mass spec methods work

Ovod et al. A&D, 2017

Nakamura et al., Nature, 2018
Plasma Aβ in the Insight46 cohort

Study design: Insight46 - epidemiological study people born 1946 (n= 414 cognitively unimpaired) APOE genotype, neuropsych testing, amyloid PET
Plasma Aβ42, Aβ42/40 using immunoassay (Simoa) and IP LC-MS/MS

Plasma Aβ42 and Aβ40/42 ratio by IP-MS/MS show high concordance with brain amyloidosis

Keshavan A et al., Brain 2021
Plasma Aβ42/Aβ40 ratio using a fully automated Cobas assay

- Purple = CU
- Yellow = MCI

AUC (CU) = 0.84 (0.80, 0.88)
AUC (MCI) = 0.79 (0.72, 0.85)

Palmqvist *et al.*, unpublished
The challenge

The fold reduction in CSF Aβ ratio is much greater than in plasma because of peripheral Aβ

Schindler et al., Neurology, 2019
The challenge, continued…

Plasma Aβ1–42/Aβ1–40

CSF Aβ1–42/Aβ1–40

CSF pTau/Aβ1–42

No “wiggle room” around cutoff

Plasma Aβ1–40 (pg/mL)

CSF Aβ1–42 (pg/mL)

CSF pTau (pg/mL)

Rabe C et al., under review
Group level enrichment/screening: Yes

Individual diagnostics: No, or maybe, but with great caution
T = tau pathology
<table>
<thead>
<tr>
<th>Study</th>
<th>Log2Ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF P-tau is increased in AD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CSF P-tau increase only in AD, not in (most) other neurodegenerative diseases

Itoh et al., Ann Neurol 2001
Differential detection of AD measuring different phospho-forms of tau in CSF

[Graphs showing box plots for different groups (Young, CU-, CU+, MCI+, AD, Non-AD, FTD) for pTau-181, pTau-199, pTau-202, pTau-205, pTau-217, pTau-231, and pTau-398]

Benedet, Gobom et al., unpublished
Different phospho-forms of tau can be measured in plasma

Ashton et al., Acta Neuropathol. 2021
Plasma tests as clinical tools to predict AD-type dementia in patients with subjective or mild cognitive impairment

<table>
<thead>
<tr>
<th>Best model fit</th>
<th>Plasma P-tau217</th>
<th>APOE</th>
<th>Plasma NfL</th>
<th>MRI</th>
<th>Memory</th>
<th>Exec. function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parsimonious model</td>
<td>Plasma P-tau217</td>
<td>APOE</td>
<td>MRI</td>
<td>Memory</td>
<td>Exec. function</td>
<td></td>
</tr>
<tr>
<td>P-tau217, cognition and APOE</td>
<td>Plasma P-tau217</td>
<td>APOE</td>
<td>Memory</td>
<td>Exec. function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-tau217 and cognition</td>
<td>Plasma P-tau217</td>
<td>Memory</td>
<td>Exec. function</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-tau217 only</td>
<td>Plasma P-tau217</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic prediction by the memory clinic physicians</td>
<td>Clinical prediction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>AUC (95% CI)</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best model fit</td>
<td>0.92 (0.89–0.95)**</td>
<td>159</td>
</tr>
<tr>
<td>Parsimonious model</td>
<td>0.92 (0.88–0.95)**</td>
<td>161</td>
</tr>
<tr>
<td>P-tau217, cognition and APOE</td>
<td>0.90 (0.86–0.94)**</td>
<td>166</td>
</tr>
<tr>
<td>P-tau217 and cognition</td>
<td>0.89 (0.84–0.93)**</td>
<td>171</td>
</tr>
<tr>
<td>P-tau217 only</td>
<td>0.81 (0.75–0.87)*</td>
<td>207</td>
</tr>
<tr>
<td>Diagnostic prediction by the memory clinic physicians</td>
<td>0.72 (0.65–0.78)</td>
<td>228</td>
</tr>
</tbody>
</table>

*P < 0.05; **P < 0.001 vs the clinical prediction

Palmqvist et al., Nature Med. 2021
Establishing a cross-validated model

**Establishing the logistic regression model in BioFINDER**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma P-tau217 z-score</td>
<td>0.71</td>
</tr>
<tr>
<td>Memory</td>
<td>0.53</td>
</tr>
<tr>
<td>Executive function</td>
<td>0.95</td>
</tr>
<tr>
<td>One APOE ε4 allele</td>
<td>1.02</td>
</tr>
<tr>
<td>Two APOE ε4 alleles</td>
<td>0.99</td>
</tr>
</tbody>
</table>

AUC [95% CI] → 0.90 [0.87-0.94]

**Cross-validating the logistic regression model in ADNI**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma P-tau181 z-score</td>
<td>0.71</td>
</tr>
<tr>
<td>Memory</td>
<td>0.53</td>
</tr>
<tr>
<td>Executive function</td>
<td>0.95</td>
</tr>
<tr>
<td>One APOE ε4 allele</td>
<td>1.02</td>
</tr>
<tr>
<td>Two APOE ε4 alleles</td>
<td>0.99</td>
</tr>
</tbody>
</table>

AUC [95% CI] → 0.89 [0.85-0.93]

**http://predictAD.app**

Palmqvist et al., Nature Med. 2021
Donanemab lowers plasma P-tau217

Eli Lilly, unpublished
### TABLE 1. REDUCTIONS IN PHOSPHORYLATED TAU-181 WITH ADUCANUMAB TREATMENT VS PLACEBO

<table>
<thead>
<tr>
<th>Study</th>
<th>Aducanumab</th>
<th>Placebo</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMERGE (NCT02484547)</td>
<td>-13%</td>
<td>+8%</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>ENGAGE (NCT02477800)</td>
<td>-16%</td>
<td>+9%</td>
<td>P&lt;0.001</td>
</tr>
</tbody>
</table>

*a Values are with the higher of 2 doses used in the EMERGE and ENGAGE trials.*

Hansson O et al., unpublished
Diagnosing AD-type tau pathophysiology with a blood test: are we there yet?

Group level enrichment/screening: Yes

Individual diagnostics: Yes, at least we are getting there
Thanks!!

henrik.zetterberg@gu.se
h.zetterberg@ucl.ac.uk
Lesley Inker
Professor of Medicine
Tufts University School of Medicine
GFR Decline as a Surrogate Endpoint for Progression of CKD

Clinical Validation and Regulatory Acceptance of Biomarkers as Surrogate Endpoint
Translational Science in Drug Development: Surrogate Endpoints, Biomarkers, and More

May 25 2022

Lesley A Inker MD, MS
Co-Director, Chronic Kidney Disease-Epidemiology Collaboration
Tufts Medical Center & Tufts University School of Medicine, Boston MA
Background

- Kidney disease is slowly progressive
- Clinical trials to evaluate treatments to prevent or slow the progression to kidney failure require long duration of follow-up, leading to expensive and complex trials, or highly selected subset of participants
- Doubting of serum creatinine (57% decline in GFR) is accepted by regulators but still occurs late in disease course
- These challenges have likely contributed to the paucity of therapies to treat CKD

Number of RCT in kidney related domains compared to other medical fields

GFR slope and albuminuria are the two central biomarkers in CKD.
CKD-EPI Investigations of Surrogate Endpoints for Trials in CKD Progression

NIH U01 CKD-EPI includes evaluation of urine protein as surrogate

NKF-FDA Workshop May 2008 on UP

NKF-FDA Workshop December 2012 Lesser Decline in GFR

NKF-FDA-EMA Workshop March 2018 GFR Slope and UACR

CKD-EPI CT Funding in partnership with NKF and sponsors

Data identification, acquisition and cleaning; analyses; method development

Updated literature search; refined methods

Continual literature updates; Enhanced method development


UP, urine protein; GFR, glomerular filtration rate; UACR urine albumin to creatinine ratio; NKF, National Kidney Foundation
Use of GFR slope as surrogate endpoint

Inker et al JASN 2019
Levey et al AJKD 2014
Greene et al JASN 2019

Inker et al AJKD 2019
Levey et al AJKD 2019
Coresh et al JAMA 2014
Greene et al AJKD 2019

GFR slope

30% GFR decline
40% GFR decline
Doubling of SCr (57% GFR decline)
Kidney Failure

GFR (ml/min/1.73m²)

0 50 100 75 100

0 2 4 6 8 10 12 14

Years
Use of GFR slope as surrogate endpoint

Advantages

- Regardless of cause
  - Decreased GFR defines CKD
  - Level of GFR indicates severity
  - GFR decline is the definition of progression, for all causes
- Compared to time to event
  - Increased power
  - Includes fast and slow progressors
  - Includes patients who have GFR decline that might lead to endpoint even after the end of the trial

Limitations/complications

- eGFR can reflect GFR as well as non GFR determinants
- Nonlinearity
- Heterogeneity
- Informative censoring
- Acute effects
Challenge of acute effects in GFR slope

GFR = N \times \text{SNGFR}

**Control arm**
Declining N (number of nephrons)
Stable SNGFR (single-nephron GFR)

**Treatment arm**
short-term - ↓ SNGFR, no change in N
long-term - stable SNGFR, slower decline in N

T, Time
SNGFR, single nephron GFR

Chaudhari and Inker, Current Opinion in Nephrology 2020
Models for computation of GFR Slope

- **Goal:** Provide a set of models that accommodate the range of circumstances expected in trials of CKD progression

- **Linearity:** In general, reasonable assumption that moderate deviations from linearity in the chronic phase do not effect overall slope estimates in trials that are relatively short in duration

- **2-slope model** to allow for acute treatment effect on GFR that differs from chronic slope

Vonesh E, Tighiouart H, Ying J et al Mixed-effects models for slope-based endpoints in clinical trials of chronic kidney disease. Stats in Medicine 2019
Informative censoring: For studies with > 15 ESRD/Death events, used shared parameter models with Weibull survival times

Heterogeneity
- **Between subject:** Random slopes and intercepts
- **Within subject:** Power of the means model to allow greater variability at higher GFR
- **Treatment effect:** Allowed different slope variance in each group to accommodate non-uniform treatment effects

Model Selection
- Automated algorithm used to select first the most complicated model (shared parameter and all of heterogeneity components), followed by models that did not have one or more of the parameters
Trial Level Analyses: evaluate the association between treatment effects on GFR slope to that of the clinical endpoint across range of RCT’s

- Individual patient meta-regression
  - Consistent definitions
  - Correlation between errors in the estimated treatment effects

- Within study analyses:
  - Estimated treatment effects on GFR slope: GFRslopeTreatment - GFR SlopeControl
  - Estimated treatment effects on the clinical endpoints – Cox models, expressed as HR

- Bayesian meta-regression to obtain
  - Estimate of regression line as summarized by slope, intercept, RMSE and R²
  - Prediction intervals for HR on the clinical endpoints for future trial over a range of the treatment effect on the mean difference in GFR slopes

Inker L, Heerspink H, Tighioart H et al. GFR Slope as a Surrogate End Point for CKD Progression. JASN 2019
Trial-level analyses for the association of treatment effects on 3 year-total slope and chronic slope vs treatment effects on the clinical endpoint.

Inker L, Heerspink H, Tighioart H et al. GFR Slope as a Surrogate End Point for CKD Progression. JASN 2019.
Applying Trial Level Analyses to a New RCT

Trial Level Analysis in Previous RCTs
➢ Characterizes “causal association” between ITT-based estimates of treatment effects on surrogate & clinical endpoints

Application in New RCT
➢ Convert estimated treatment effect on surrogate(s) to probability of clinical benefit for newly tested intervention

Converting Treatment Effect on 3-Yr Total Slope to Probability of Clinical Benefit

<table>
<thead>
<tr>
<th>Estimated Effect on GFR Slope (ml/min/1.73m²/yr)</th>
<th>Large RCT</th>
<th>Moderate RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median HR and 95% Prediction Interval</td>
<td>PPV</td>
<td>Median HR and 95% Prediction Interval</td>
</tr>
<tr>
<td>0.5</td>
<td>0.77 (0.59, 0.99)</td>
<td>0.98</td>
</tr>
<tr>
<td>0.75</td>
<td>0.69 (0.52, 0.89)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>1.0</td>
<td>0.62 (0.47, 0.80)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>Threshold for effect on GFR slope to confer PPV ≥ 97.5</td>
<td>0.48</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Inker L, Heerspink H, Tighioart H et al
GFR Slope as a Surrogate End Point for CKD Progression JASN 2019
- Use of total slope instead of the clinical endpoint allows reduction in follow-up from 4-6 years to 2 years while improving efficiency by 17% to 64% (~sample size savings of 14% to 39%)

- Relative gains in power for slope analysis increase when baseline GFR is higher.

- Acute effect is critical consideration in selection of total vs chronic slope vs endpoint
Next steps/current work

- Update set of studies to account for well powered studies across more interventions

- Methods work on
  - Acute effects
  - Subgroups/interactions

- Joint models to combine slope with albuminuria as can be used in Phase II studies with shorter follow-up
Empirical data supports use of GFR decline as surrogate endpoints in RCTs evaluating therapies in CKD

When applying these data to the design of a future trial, the most appropriate endpoint for the new trial needs to be considered in the context of the trial phase, specific population, treatment, and design.
CKD-EPI CT Analytical Team

Andy Levey, Tom Greene and Josef Coresh

March 2018 CKD-EPI, CKD-PC, EMA, FDA and NKF Teams

Co Directors: Tom Greene, Hiddo Heerspink
Tufts: Juhi Chaudhari Hocine Tighiouart Jonathan Miao
Utah: Ben Haaland, Jian Ying, Willem Hardie
Chicago: Ed Vonesh
Groningen: Neils Jong
Nicole Gormley
Acting Division Director
Division of Hematologic Malignancies
US Food and Drug Administration
USE OF SURROGATE ENDPOINTS IN ONCOLOGY
DUKE MARGOLIS WORKSHOP
MAY 25, 2022

Nicole Gormley, MD
Division Director
Division of Hematologic Malignancies II
U.S. Food and Drug Administration
Outline

• Regulatory Considerations for Biomarker Development
• pCR Example
• MRD in Multiple Myeloma
• Future Directions
Potential uses of Biomarkers

• Prognostic Biomarker
• Clinical Uses
  – Screening/Early Detection
  – Monitor for relapse
  – Guide therapeutic decisions
• Regulatory Uses
  – Patient Stratification
  – Patient Selection/Enrichment
  – Risk-based treatment assignment
  – Intermediate Endpoint or Surrogate Endpoint
Biomarker as an Endpoint

• Intermediate clinical endpoint
  – Can be measured earlier than morbidity or mortality, but reasonably likely to predict clinical benefit

• Surrogate endpoint reasonably likely to predict clinical benefit

• Surrogate Endpoint
  – Clinical validation that the marker predicts clinical benefit
Development of Endpoints for Regulatory Use: Validation as a Surrogate

• Prentice Criteria
  – The surrogate must be a correlate of the true clinical endpoint
  – The treatment effect on the surrogate should capture the full effect of treatment on the clinical endpoint

• Meta-analytical methods
  – Patient-level data
  – Allow for assessment of Individual Level and Trial Level Surrogacy
    • Individual Surrogacy - Correlation between candidate surrogate and true clinical endpoint on an individual level
    • Trial Level Surrogacy - Correlation between effect of treatment on the candidate surrogate and the effect of treatment on the true clinical endpoint
  – Surrogate Threshold Effect
    • Minimum treatment effect on the surrogate necessary to predict an effect on the true clinical endpoint
Evidentiary Criteria

• **Meta-analysis Considerations**
  – Inclusion of more trials increases the statistical rigor of the analysis and may allow for more interrogation of the data to address uncertainties.
  – Inclusion of trials with a range of treatment effects (positive and negative trials) increases the accuracy and precision of trial level surrogacy assessment.
  – When designing a meta-analysis, consideration of MRD timing of assessment, missing data is important.
  – The trial populations and treatments included in the meta-analysis inform future applicability of the surrogate biomarker.
pCR Example

- Collaborative Trials in Neoadjuvant Breast Cancer
  - Conducted a pooled analysis of mature trials that had both pathologic complete response (pCR) and long-term outcome data
  - Objectives
    - Determine the association between pCR and EFS and OS
    - Determine the definition of pCR which best correlated with long-term outcomes
    - Identify breast cancer subtypes in which pCR best correlated with long-term outcome
    - Determine what magnitude of pCR improvement predicts long-term clinical benefit
# pCR Example

## pCR Pooled Analysis Results

<table>
<thead>
<tr>
<th>pCR definition</th>
<th>Event-free survival HR (95 % CI)</th>
<th>Overall survival HR (95 % CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ypT0 ypN0</td>
<td>0.44 (0.39–0.51)</td>
<td>0.36 (0.30–0.44)</td>
</tr>
<tr>
<td>ypT0/is ypN0</td>
<td>0.48 (0.43–0.54)</td>
<td>0.36 (0.31–0.42)</td>
</tr>
<tr>
<td>ypT0/is</td>
<td>0.60 (0.55–0.66)</td>
<td>0.51 (0.45–0.58)</td>
</tr>
</tbody>
</table>

Cortazar Ann Surg Oncol 2015
pCR Example

• Individual-Level Surrogacy

Cortazar Lancet 2014
pCR Example

• Trial-Level Surrogacy

R² 0.03 (95%CI:0.00,0.25)  
R² 0.24 (95%CI:0.00,0.70)
pCR Example

- **CTNeoBC Summary**
  - No pCR association with long-term outcomes (EFS and OS) at a trial level, only on an individual level
  - A standard definition that includes assessment of the nodes (ypT0ypN0 or ypT0/isypN0) should be used in future trials
  - Magnitude of pCR improvement that predicts long-term clinical benefit could not be established
    - Possibly due to heterogeneity of population, low pCR rates, lack of targeted therapies
MRD in MM Meta-analyses

Progression-Free Survival

<table>
<thead>
<tr>
<th>Study</th>
<th>Hazard ratio</th>
<th>HR</th>
<th>95%-CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korde 2015</td>
<td></td>
<td>0.10</td>
<td>[0.02; 0.61]</td>
</tr>
<tr>
<td>Mateos 2014</td>
<td></td>
<td>0.40</td>
<td>[0.25; 0.65]</td>
</tr>
<tr>
<td>Paiva 2008</td>
<td></td>
<td>0.35</td>
<td>[0.25; 0.50]</td>
</tr>
<tr>
<td>Silvennoinen 2013</td>
<td></td>
<td>0.28</td>
<td>[0.09; 0.89]</td>
</tr>
<tr>
<td>Random effects model</td>
<td></td>
<td>0.35</td>
<td>[0.27; 0.46]</td>
</tr>
</tbody>
</table>

Overall Survival

<table>
<thead>
<tr>
<th>Study</th>
<th>Hazard ratio</th>
<th>HR</th>
<th>95%-CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mateos 2014</td>
<td></td>
<td>0.48</td>
<td>[0.27; 0.88]</td>
</tr>
<tr>
<td>Paiva 2008</td>
<td></td>
<td>0.48</td>
<td>[0.30; 0.77]</td>
</tr>
<tr>
<td>Random effects model</td>
<td></td>
<td>0.48</td>
<td>[0.33; 0.70]</td>
</tr>
</tbody>
</table>

Landgren BMT 2016 Munshi Jama Oncol 2016
MRD in MM Meta-analyses

• Remaining Questions
  – Does MRD in MM have trial level surrogacy using individual patient level data?
  – What is the threshold that best correlates with clinical benefit?
  – What is the appropriate timing of assessment?
  – Does Sustained MRD better correlate with long-term outcomes?
  – Should MRD be assessed in those only in CR, VGPR, PR?
BELLINI Trial: A Cautionary Tale

- Phase 3, double-blind, randomized, placebo-controlled trial of bortezomib and dexamethasone with or without venetoclax in patients with relapsed/refractory, multiple myeloma who had received 1-3 prior lines of therapy

<table>
<thead>
<tr>
<th></th>
<th>Venetoclax Arm</th>
<th>Placebo Arm</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORR</td>
<td>82.0% (75.8, 87.1)</td>
<td>68.0% (57.8, 77.1)</td>
</tr>
<tr>
<td>MRD negativity rate (10^-5)</td>
<td>13.4% (8.9, 19.0)</td>
<td>1.0% (0.0, 5.6)</td>
</tr>
<tr>
<td>Median PFS (mos) (95% CI)</td>
<td>22.4 (15.3, NR)</td>
<td>11.5 (9.6, 15.0)</td>
</tr>
<tr>
<td>Hazard Ratio (95% CI)</td>
<td>0.63 (0.44, 0.90)</td>
<td></td>
</tr>
</tbody>
</table>

### BELLINI Trial: A Cautionary Tale

<table>
<thead>
<tr>
<th></th>
<th>PFS HR (95% CI)</th>
<th>OS HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients (N=291)</td>
<td>0.630 (0.443-0.897)</td>
<td>2.027 (1.042-3.945)</td>
</tr>
<tr>
<td>High-risk cytogenetics (N=49)</td>
<td>1.206 (0.577-2.520)</td>
<td>NE</td>
</tr>
<tr>
<td>Standard-risk cytogenetics (N=213)</td>
<td>0.544 (0.354-0.837)</td>
<td>1.505 (0.727-3.115)</td>
</tr>
<tr>
<td>t(11;14) (N=35)</td>
<td>0.110 (0.022-0.560)</td>
<td>0.343 (0.031-3.842)</td>
</tr>
<tr>
<td>BCL-2 high (N=140)</td>
<td>0.502 (0.294-0.856)</td>
<td>1.446 (0.568-3.678)</td>
</tr>
<tr>
<td>BCL-2 low (N=37)</td>
<td>1.387 (0.431-4.468)</td>
<td>NE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Ven (N=194)</th>
<th>Pbo (N=97)</th>
<th>Ven (N=20)</th>
<th>Pbo (N=15)</th>
<th>Ven (N=93)</th>
<th>Pbo (N=47)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORR</td>
<td>82%</td>
<td>68%</td>
<td>90%</td>
<td>47%</td>
<td>86%</td>
<td>68%</td>
</tr>
<tr>
<td>≥CR</td>
<td>26%</td>
<td>5%</td>
<td>45%</td>
<td>7%</td>
<td>32%</td>
<td>4%</td>
</tr>
<tr>
<td>≥VGPR</td>
<td>59%</td>
<td>36%</td>
<td>70%</td>
<td>27%</td>
<td>68%</td>
<td>34%</td>
</tr>
<tr>
<td>uMRD</td>
<td>13%</td>
<td>1%</td>
<td>25%</td>
<td>0%</td>
<td>17%</td>
<td>2%</td>
</tr>
</tbody>
</table>

CI, confidence interval; CR, complete response; HR, hazard ratio; NE, not estimable due to no events in placebo; ORR, overall response rate; OS, overall survival; Pbo, placebo; PFS, progression-free survival; uMRD, undetectable minimal residual disease ($10^{-5}$); VGPR, very good or better partial response.

a. t(4;14), t(14;16), or del(17p)

b. No high-risk cytogenetics
**Concerning OS results**

- Need evaluation of endpoints that can be assessed at Early timepoints \textit{and} Late timepoints that provide definitive evidence of clinical benefit
  - Bellini Trial showed divergent OS and ORR, PFS, MRD results
- Additional Information is needed on MRD as an endpoint in MM
MRD Today and Future Considerations

• MRD results used to support accelerated approval in ALL
  – Blinatumomab approval in MRD-positive B-cell Precursor ALL
    • Accelerated approval based on MRD response rate and hematological relapse-free survival
• MRD results have been included in Prescribing Information in CLL
  – Venetoclax, Obinutuzumab
• MRD results have been included in the Prescribing Information in MM
  – Daratumumab, Abecma
  – Currently recommended as a secondary endpoint
• Ongoing efforts in various diseases to formally evaluate MRD
Conclusions

• Validated Endpoints are needed for Regular Approval
• pCR and MRD are not validated surrogate endpoints
• Existing uncertainty and remaining questions regarding these endpoints for regulatory purposes
• MRD, pCR and other biomarker assessments in clinical trials should be discussed with the Agency
• FDA is committed to working with the community on the development of biomarkers.
Thanks...

• Laleh Amiri- Kordestani
• Marc Theoret
• Julia Beaver
Session 3: Clinical Validation and Regulatory Acceptance of Biomarkers as Surrogate Endpoints

Moderator:

- Norman Stockbridge, US Food and Drug Administration

Panelists:

- Steve Ryder, Rallybio Inc.
- Henrik Zetterberg, University of Gothenburg
- Lesley Inker, Tufts University
- Nicole Gormley, US Food and Drug Administration
- Aliza Thompson, US Food and Drug Administration
- Jeff Allen, Friends of Cancer Research
Session 3: Clinical Validation and Regulatory Acceptance of Biomarkers as Surrogate Endpoints

Discussion Questions:

1. What are the challenges associated with validating biomarkers, and what approaches may support efficient biomarker validation?

2. What characteristics and processes are shared by programs with a strong track record in evaluating candidate surrogates?

3. What more can be done to assist developers in validating candidate surrogates?

4. How can early involvement and communication with regulatory agencies support biomarker validation?
Break

We will be back momentarily.

The next panel will begin at 2:05 p.m. (U.S. Eastern Time)
Session 4: Beyond Surrogate Endpoints: Other Ways Translational Science Can Support Drug Development

2:05 pm – 3:30 pm EST
Leslie Gordon
Medical Director and Co-Founder
Progeria Research Foundation
Beyond Surrogate Endpoints: Other Ways Translational Science Can Support Drug Development

Translational Science in Drug Development: Surrogate Endpoints, Biomarkers, and More
May 24, 25, 2022
Duke Margolis Center for Health Policy

Leslie B. Gordon, MD, PhD
The Progeria Research Foundation
Hasbro Children’s Hospital & Alpert Medical School of Brown University
Boston Children’s Hospital Boston and Harvard Medical School
• Volunteer Medical Director, The Progeria Research Foundation

• In-kind donations: Receive medication for Progeria clinical trials from 3 drug companies (names not included at FDA’s request) at no cost

• Sources of Funding for Research: The Progeria Research Foundation; FDA
Progeria: An Ultrarare Fatal Premature Aging Disease

- Segmental “Premature Aging”
- Prevalence 1/20 million
- 19 children in US
- ~400 children worldwide

- Autosomal Dominant
- Lifespan Ave 14.5 yrs.
- Death due to premature atherosclerosis
Clinical Signs of HGPS
CV and Neurovascular Disease

• Global, Progressive
• Heart Failure, Strokes

MRI 5 year old with carotid obstruction

Human HGPS Vascular Disease

- Calcific Plaques
- Thick Fibrotic Adventitia
- Medial Cell Death with Extracellular Matrix Deposition

Olive et al, Hypertension, 2010
Assays Demonstrating Extremely Stiff Vessels In HGPS

Avg. PWV 3.5 x normal (40-60 y.o.)
Recurrent *de novo* point mutations in lamin A cause Hutchinson–Gilford progeria syndrome

Maria Eriksson*, W. Ted Brown†, Leslie B. Gordon‡, Michael W. Glynn§, Joel Singer∥, Laura Scott∥∥, Michael R. Erdos*, Christiane M. Robbins*, Tracy Y. Moses*, Peter Berglund¶, Amalia Dutra*, Evgenia Pak*, Sandra Durkin§, Antonei B. Csoka#, Michael Boehnke∥∥, Thomas W. Glover§ & Francis S. Collins*

We were catapulted into a new phase…
Mutation Optimizes $LMNA$ Internal Splice Site

Exon 11 \[\rightarrow\] Exon 12

Mutant Splicing
150 bp deletion (50 aa)
“progerin”

Lamin A: Inner Nuclear Membrane Protein
- Lines the inner nuclear membrane-Scaffolding
- Binds chromatin to effect transcription
- Structural and signaling effects
- Expressed by Differentiated Cell Types
- Undergoes post-translational processing that is defective in HGPS due to 50 aa deletion
- Thus, progerin is short, permanently farnesylated and toxic to cells

HGPS is Caused by a Single Base Silent Mutation in the $LMNA$ Gene (c.1824 C>T, G608G)
Biology Leads The Way Towards Treatment Trials

Diagnostic Testing
HGPS Progerin-producing Cells and Mouse Models
Clinical Studies
Clinical Trials
Progerin Causes Nuclear Blebbing In Cultured Cells
% Blebbed Cells Increases with Passage Number

Normal Fibroblast Nuclei
Progeria Fibroblast Nuclei
Human Progerin-Producing Mouse Models Created

- Human BAC Transgenic G608G Mouse Model
  (Varga et al (Collins) PNAS 2006)
  - Mice Are Small,
  - Develop CVD but not plaques,
  - Die Early, cause of death unknown
  - Human Progerin Produced

- Mouse Knock-in G609G Mouse Model
  (Osorio et al (Lopez-Otin) Sci Transl Med 2011)
  - Mouse Progerin Produced
  - Mouse Are Small
  - Develop CVD but not plaques,
  - Die Early, cause of death unknown

- Additional endothelial-specific and VSMC-specific mouse models have also been developed
Biology Leads Us To Potential Treatment

A. Lamin A Generation

1. Farnesyltransferase → CAAX
2. Zmpste24 or RCE1 → C-44X
3. ICMT → C-OCH₃
4. Zmpste24 → Lamin A

B. Progerin Generation

- 50 amino acids deleted
- Farnesyltransferase
- Zmpste24 or RCE1
- ICMT
- Progerin
- no cleavage site

Lonafarnib FTI
Farnesyltransferase Inhibition as Treatment (not all using the FTI in our trials)

Normal  HGPS  HGPS with FTI, 72 hrs.

FTI Lonafarnib Normalized human HGPS Fibroblast Nuclear Shape

FTI ABT-100 Improved Disease in Zmpste24 Deficient Mouse Model, Including lifespan

When treated with FTI tipifarnib after birth, Cardiovascular disease did not develop

When allowed to develop cardiovascular disease for 9 months, then treated with FTI tipifarnib, Normal vasculature detected

Capell et al 2005; Glynn et al, 2005; Toth et al, 2005; Fong et al, 2006
Improvements With Lonafarnib Treatment in Children: Changes in the Arteries and Extended Survival

PNAS 2012: Clinical trial of a farnesyltransferase inhibitor in children with Hutchinson–Gilford progeria syndrome

1/27 deaths vs. 9/27 deaths
88% reduced risk of death during the 2 years of treatment

Current unpublished estimate of average lifespan extension is 4.3 years

Lonafarnib (Zokinvy) is our first FDA approved drug for Progeria
Biology Leads Us To Clinical Trials

Acetyl-CoA + acetoacetyl-CoA

3-Hydroxy-3-methylglutaryl-CoA

HMG-CoA Reductase

Statins

Mevalonic Acid

Isopentyl-PP

Geranyl-PP

Lonafarnib + Pravastatin + Zoledronate

Bisphosphonates

Farnesyl-PP Synthase

Farnesyl-PP

Squalene

Geranylgeranyl-PP

Geranylgeranylated proteins

Lonafarnib

Cholesterol

Farnesyl-preprogerin

Everolimus

FTis

Farnesyltransferase

Preprogerin

Disposal

Autophagy

Disposal

Autophagy

Everolimus

Lonafarnib
Zmpste24 Mutations do result in progeroid disease in humans, but not identical to HGPS and not progerin-producing (abnormal prelamin A causes disease)

This model is not progerin-producing, no CVD

Zmpste24 mice have spontaneous fracture and neuro. deficits, unlike HGPS

Human Clinical Trial of HGPS in Combination with Lonafarnib, Pravastatin and Zoledronic acid Showed No Benefit Over and Above Lonafarnib Monotherapy

A great animal model, but not optimal for drug development in HGPS
Animal Husbandry: G609G Homozygote:

- soft gel-based chow on the floor of cage +
- introduction of a caretaker mouse in each cage

- original 50% survival at 103 days (Osorio et al., 2011)
- new extended the mean lifespan = 168 days

- allowed the cardiovascular phenotype to worsen similar to that observed clinically in patients.

- cardiovascular function progressed to extreme stiffening and diffuse vascular calcification.
Extended Mouse Lifespan Potentiates Overlap with Human Cardiovascular Disease

Lonafarnib Therapy: PWV Improved at 168 Days of Age; Everolimus does not add benefit

Murtada Et al…Humphrey; Yale U.; Unpublished
Getting The Word Out for Maximal Success

Collection and Distribution of Best Practices and Guidance for Basic Scientists

❖ New Publications
❖ Investigator Surveys
❖ Email Blitz’s with new information
❖ Resource Center
❖ Posters at Scientific Meetings
Centralizing Disease-Specific Animal Testing To Optimize Outcomes and Comparability

Assess Candidate Intervention (i.e. supporting in vitro data and biologic plausibility)

Choose Most Appropriate HGPS Mouse Model

Implement Controlled Intervention Study

Survival Study with Pathology

Centralized Serial Phenotypic Assessment (weight, progerin levels, etc)

Send Mice and/or Samples to Investigator for Specialized Analyses

Gating For Human Trial
Potential New Treatments’ effects on Progeria Mouse Model Survival*

% Increase in Progeria mouse lifespan compared to controls

- **DNA base editing**: 140% (Koblan et al, 2021)
- **RNA therapeutics**: 61.6% (Erdos et al, 2021)
- **progerin**: 50% (Kang et al, 2021)
- **Ionafarnib**: 24.9%

*Note that mouse models in use were not the same across all studies*
Determination and Collaboration

Finding…
Diagnosing…
Studying…
Treating…
CURING

Together, we WILL find the cure!
www.progeriaresearch.org
Estelle Marrer-Berger
Senior Translational Safety Leader
Roche
Optimizing early clinical investigations by increasing the predictive value of non-clinical activities

Estelle Marrer-Berger, Antje Walz, and Imein Bousnina

Duke-Margolis Center for Health Policy / May 24-25, 2022
Targeting intracellular Wilms tumor 1 in AML with a TCR-like T-cell bispecific antibody

- WT1 oncoprotein is an intracellular, transcription factor, overexpressed in leukemias (AML, ALL) and solid cancers (ovarian cancer and mesothelioma)

*Quantification of the RMF peptide on AML blasts*
Augsberger et al., Blood, 2021

- In adults, WT1 expression is restricted to a few tissues: (kidney podocytes, Sertoli and granulosa cells in the testes and ovaries, few mesothelial cells and 1% of bone marrow cells)
A human/patient-centric non-clinical approach to bring WT1 TCB to patients?

- Lack of cross-reactive TOX animal species and the standard non-clinical toolbox not applicable
- Increased risks for off-target / off-tumor cross-reactivity

Reduce and manage the «Unknown»

Increase the predictive value: predicting from human to human
Non clinical strategy for Entry into Human based on NAMs

*In vitro / ex vivo derived therapeutic index, starting dose, and PAD*

- Ligandome elucidation with human vital organ lysates
- Study in HLA-A2 Tg mice using an hemi-surrogate
- In vitro testing in human 2D/3D models
- Safety risks
- PK/PD assessment
- Translation to safe starting dose and risk mitigations

Non-clinical safety strategy combining new «state of the art» activities to potentiate risk identification in the absence of a cross-reactive species. Integration of qualitative and quantitative assessments to define a therapeutic index. Define starting dose and risk management plan.
Risk identification using human 2D / 3D in vitro systems

Organs/Cells Tested
- Astrocytes
- iPS-Derived Astrocytes
- Lund Mesencephalic Cells
- Retinal Pigment Epithelial Cells
- Microvascular Endothelial Cells
- Bone Marrow Derived Mononuclear Cells
- Epidermal Keratinocytes
- Adult Melanocytes
- Renal Proximal Tubule Epithelial Cells
- Renal Mixed Epithelial Cells
- Pulmonary Artery Endothelial Cells
- Aortic Endothelial Cells
- Aortic Smooth Muscle Cells
- Beating Cardiomyocytes
- Bronchial Epithelial Cells
- Small Airway Epithelial (Alveolar Cells)
- Lung Chip
- Small Airway Epithelial (Alveolar Cells)
- Intestinal Organoids (iPS-derived Duodenum and Colon)
- Liver Spheroids
- Pancreatic Islet Spheroids
- Lung Squamous Cell Carcinoma (Tissue Positive Control)

Human primary cells (2D or 3D) → Allogeneic or autologous PBMCs → WT1 TCB (Dose response @ 24, 48, 72h) → WT1 TCB + INFγ mimic inflammation / infection

Cell Death
- LDH
- Caspase 3/7

Cytokines
- Granzyme B
- TNF-a
- IFN-g
- IL-2
- IL-6
- IL-8
- IL-10

Physiological Parameters
- AST (Liver)
- Cardiomyocytes: - Beat Rate - Base Impedance

Supernatant → Microscopy → Electrophysiology
WT1 TCB consistently induced minimal (?) lysis in liver spheroids (7 donors) co-cultured with allogenic PBMCs (3 donors)

Bright field pictures of liver spheroids in co-culture experiment 72 hours after treatment
And now what?

• The signal needs to be further assessed, qualitatively and quantitatively

• The in-vitro effect needs to be «translated» to a human body

• A threshold of significance needs to be established and be related to a drug concentration/exposure

• A therapeutic index needs to be defined

• Causality? Exacerbating contexts (pathologies, medications..)?

  ➢ Exclusion/Inclusion criteria, monitoring and mitigation need to be defined
Safety assessment based on dynamic in vitro killing assays using human liver spheroids (3D): *Mild but consistent signal at 1 µg/ml*

**Conclusions**

- The signal is **consistent** across donors and endpoints monitored.
- The **significance threshold was conservatively defined at 1 µg/mL** based on the increased LDH/caspase3/7, cytokines, and AST observed in the allogenic co-cultures (worst-case scenario).
- Alloreactivity amplifies the signal, though, compared to the negative control DP47, 1 µg/mL triggers a minimal effect.
  - **Causality? Do we have a therapeutic window?**
Causality of the signal observed: a CYP8B1 epitope?

WT1 TCB interactome elucidation to identify potential off-targets

Go after causality of the signal: technical, biological artefact or REALITY?

NO WT1 expression in liver spheroids

**Ligandome elucidation**

- **BsaB IMMUNOPRECIPITATION**
- Human primary hepatocyte lysates
- Off-target HLA-complexes
- WT1 TCB

**TARGET IDENTIFICATION**

- Decoy hits
- #PSMs

**Synapse stability evaluation**

Stable synapse formation only obtained with the targeted RMF peptide

Confirmation that the safety threshold at 1 μg/mL is conservatively defined
Deriving a TI from an harmonized in vitro / ex vivo dynamic testing using diseased and healthy human systems

**Safety**
- Primary liver spheroids
- PBMCs
- Hepatocytes
- Kupffer cells
- Liver sinusoidal endothelial cells

**Efficacy**
- Primary AML sample
- AML cells
- PBMCs
- MS-5 feeder cells

**Therapeutic Index**
- AML lysis (%)
- Liver LDH release

**With harmonized experimental design**
- Time course: 24, 48, 72h
- E:T (5:1)
- PBMCs (allogeneic vs autologous)
- Readouts: killing, cytokines, T cell activation

**PKPD with AUCE**

**Translational PKPD strategy**
- Patient-centric MABEL starting dose
- Safety margin prediction
- Prediction of efficacious exposure

*Van De Vyver A & Eigenmann M et al. AAPS J. 2021 Dec 3;24(1):7

Patient-centric starting dose prediction

Reducing the number of patients treated with sub-therapeutic doses

Higher starting dose: 5 µg (standard MABEL) to 150 µg (patient-centric MABEL)

Classical in vitro MABEL

- Uses the most sensitive tumor cell line, most sensitive readout (T cell activation) to derive a safe starting dose
- Starting dose safe BUT much lower than expected therapeutic dose

Patient centric MABEL

- Efficacy prediction using patient-derived material (ex vivo)
- Safety prediction based on primary healthy and diseases in vitro systems
- Starting dose is close to therapeutic dose and reflects a balanced risk/benefit

Increased relevance for patients

Saved 3 additional cohorts of patients with sub-clinical doses
WHAT WE LEARNT

Effect, thus therapeutic index varies across time-points.

Therapeutic index expressed as potency ratio (efficacy marker/safety marker).

- Therapeutic index ≥1 at 24 and 48 hours.
- Therapeutic index ≤1 at 48 hours.
- Therapeutic index >1000 at 72 hours.
Understand your system and your testing framework (1)
*Leverage the systems’ strengths, be aware of its limitations*

**Biological systems are highly dynamic and respond dynamically to stimuli**

It is critical to evaluate concentration / effect relationships throughout a time-course, for as long as the system allows.

---

**Experimental system routinely applied**

Co-culture of Tumor cells & PBMCs

**T-cell mediated drug response is highly dynamic**

Capture highly dynamic drug response

**PKPD analysis over full time course**

Relate tumor killing to PD readouts
Understand your system and your testing framework (2)
*Leverage the systems’ strengths, be aware of its limitations*

Consistency of the signal

The full cascade of events is observed with WT1 TCB: T cell activation, target cell killing, cytokine and AST increase; controls trigger the expected effects

### T cell activation and synapse formation

- **CD69 early activation marker**
- **CD107a degranulation marker**

### Target cell killing

- **PODO/SVRed + PBMC (8237): Target cell nuclei (red)**

### Physiological function
Patient-derived material enables the balancing act to achieve the highest safe starting dose for CD3 bispecifics

Most sensitive readout in most sensitive cancer cell line; starting dose targets EC20

Most relevant readout in most relevant test system: bone marrow from AML patients; starting dose targets EC20

<table>
<thead>
<tr>
<th></th>
<th>Blinatumomab</th>
<th>WT1 TCB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting dose to efficacious dose</td>
<td>10 000 fold</td>
<td>~100 fold</td>
</tr>
</tbody>
</table>

→ Patient-centric starting dose (2020) vs Standard MABEL starting dose (2013): 0.150 mg vs 0.005 mg
→ Patient-centric based efficacious dose range prediction: 10 to 15 mg
→ In the clinic, the starting dose was safe & CRs are observed at predicted exposures
Broad regulatory approval on a novel non-clinical approach & innovative IMP

- First patient dosed in November 2020; the EiH dose predicted from the patient-derived AML blasts + autologous PBMCs was SAFE
- So far, no evidence of liver toxicity up to the dose of 12 mg ($C_{\text{max}} > 1 \text{ ug/mL}$)
- The patient-centric framework is of high predictive value for the estimation of the pharmacologically active dose range
OUTLOOK
Our accomplishments and vision

*Embrace the uniqueness of the patient to match drug and dose*

- Understanding disease mechanisms
- Validate drug target & MoA
- Understanding dose / effect relationships

- Efficacy prediction based on tumor tissues
- Safety prediction based on healthy tissues
  → Safe starting dose, close to therapeutic dose
  → Safety assessment; risks and mitigation measures

- **Ex-vivo testing** on tumours from patients enrolled in the DE and expansion in parallel to in-vivo testing to:
  → identify responders and predict the efficacious dose
  → verify the predictive value of the EiH data and of power of the model+framework

Responder A & B
Dose A: 1 mg
Dose B: 2 mg
Schedule: Q3W

Preclinical → Ph1 DE → Ph1 Expansion → Ph2 → Ph3

Translational

- **Patient-derived material**
  - Normal tissue
  - Tumour tissue

Ex-vivo testing on tumours from patients enrolled in the DE and expansion in parallel to in-vivo testing to:

- Understand disease mechanisms
- Validate drug target & MoA
- Understand dose / effect relationships
Predict individual target exposure based on individual ex vivo EC$_{90}$

Assess individual ex vivo potency (EC$_{50}$)

Combine ex vivo EC$_{50}$ with individual PK
Academic collaborators
University Hospital, LMU Munich
• Prof Marion Subklewe
• Gerulf Hänel
Doing now what patients need next
Safety mitigation for off-tumor mediated killing: Dasatinib “switches off” the CD3 signaling and rapidly neutralizes TCR engagement.

Dasatinib prevents target cell killing from pre-activated PBMCs

- Absence of dasatinib [ON]
- Presence of dasatinib [OFF]

Dasatinib prevents cytokine release from pre-activated PBMCs

*In vitro “T cell dependent cellular cytotoxic assay”

Doing now what patients need next
Christine Garnett
Clinical Reviewer
Division of Cardiology and Nephrology
U.S. Food and Drug Administration
Clinical Translational Science: Leveraging Adult Efficacy Data for Pediatrics using Bridging Biomarkers

Christine Garnett, PharmD
Division of Cardiology and Nephrology, OND, CDER, FDA
Disclaimer and Acknowledgements

My presentation reflects my opinion and is not considered official FDA guidance.

I am grateful to Drs. Norman Stockbridge, Lynne Yao and Tom Fleming for their insights and contributions to this presentation.
Pediatric Extrapolation

An approach to providing evidence in support of effective and safe use of drugs in the pediatric population when it can be assumed that the course of the disease and the expected response to a medicinal product would be sufficiently similar in the pediatric and reference (adult or other pediatric) population.
Factors Influencing Extrapolation Approaches

- **Disease Similarity**
  - Common pathophysiology, disease definition, course of disease

- **Response Similarity**
  - Similar pharmacology, response endpoints
  - Exposure-response relationship

- **Existing Data**
  - Quantity and quality of existing data
  - Sources: clinical, nonclinical, real world, registries, experience with similar drugs
Pediatric Extrapolation Approaches

**Same disease and response**
- Controlled trial using bridging biomarkers
  *confidence in similarity of disease
  *less confidence in similarity of exposure-response in children

**High confidence**
- Pharmacokinetic and safety study using exposure matching

**Similarity of Disease and Response to Treatment**

**Different disease or response**
- Adequate and well-controlled trial(s) using clinical or surrogate endpoints

**Evidence to Support Similarity**

**Large gaps in knowledge**
Use of Biomarkers in Pediatric Extrapolation

<table>
<thead>
<tr>
<th>Pharmacodynamic</th>
<th>Bridging</th>
<th>Surrogate</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Used for dose selection, disease similarity, response similarity</td>
<td>• Extrapolate efficacy from adults to children for drugs that are effective in adults with similar disease</td>
<td>• Substitute for a direct measure of how a patient feels, functions, or survives</td>
</tr>
</tbody>
</table>
Disease processes in pediatric and adult settings are closely related biologically.

In adults, intervention is safe and has substantial effects on FFS measures and biomarker.

Effects on the bridging biomarker capture effects on the principal causal pathway through which the disease process meaningfully influences FFS measures.

Intervention does not have important unintended effects on FFS measures that are not captured by the bridging biomarker.

In adults, intervention’s net effect on FFS measures is consistent with what would be predicted by the level of intervention’s effect on the bridging biomarker.

FFS = Feels, functions, survives
FDA Uses Bridging Analyses of Pediatric Hemodynamic Data to Adult Exercise Capacity in the Approval of Tracleer® (Bosentan) for Pediatric Pulmonary Arterial Hypertension Patients 3 Years of Age and Older

On September 5, 2017, the US Food & Drug Administration (FDA) approved Tracleer (bosentan) for the treatment of pulmonary arterial hypertension (PAH) (WHO Group 1) in pediatric patients aged 3 years and older. This is the first approval of a drug for the treatment of pediatric PAH with idiopathic or congenital PAH to improve pulmonary vascular resistance (PVR), which is expected to result in an improvement in exercise ability. FDA’s efficacy evaluation relied on the findings from one of the trials – BREATHE-3, an open-label, uncontrolled study in 19 pediatric patients with PAH aged 3 to 15 years which measured PVR, a cardio-pulmonary hemodynamic variable. FDA conducted analyses using data from previously approved programs in adults that established the relationship between improvements in the 6-minute walk distance (6MWD) and PVR in adults and showed that the relationship was consistent across different approved drug classes (e.g., endothelin receptor antagonist, prostanooids, PDE5 inhibitor, and soluble guanylate cyclase stimulator). The observed reduction in PVR in pediatrics from the BREATHE-3 study was used to bridge the bosentan efficacy findings in adults.

–American College of Clinical Pharmacology, 2017
PVR as Bridging Biomarker for Pulmonary Arterial Hypertension

- Adult and pediatric PAH subtypes of idiopathic, heritable and associated with congenital heart disease are similar in pathophysiology.
- PVR is a hemodynamic measure of pulmonary arterial pressure and cardiac output. PVR is on the causal pathway through which the disease process impacts how patients feel, function and survive.

PVR = Pulmonary vascular resistance
Improvement in $\Delta 6\text{MWD}$ Corresponds to Decrease in $\Delta \text{PVR}$ in Adults

Shown are the observed data by treatment assignment overlaid with regression slope and 95% confidence interval. Black error bars represent mean and standard deviation $\Delta 6\text{MWD}$ within each decile of $\Delta \text{PVR}$. 

Population Slope:
$-0.055$ (95% CI: $-0.62, -0.047$); $p < 0.0001$
Consistent Relationship Across Drug Classes and Drugs in Adults

Forest plot of mean (95% CI) regression slopes shown by drug class (left) and individual drugs (right). The dashed line is the mean slope of pooled data.
PVR explains the treatment effect on 6 min walk distance in adults

- Bosentan had significant effects on Δ6MWD and ΔPVR:
  - Clinical endpoint, Δ6MWD: +35 m
  - Biomarker, ΔPVR: -250 dyne*sec/cm^5

- 50% treatment effect on Δ6MWD explained by ΔPVR in the data analytical model with and without treatment

- No imbalance of deaths or serious adverse events in both adults and children
Bosentan significantly reduced $\Delta PVR$ in children and adults

Box plots show the mean (white circles), median (notch); 95% CI of median (width of notch); 25th and 75th percentile (width of box); 1.5* interquartile range (whiskers); and outliers (filled circles).
Bosentan Indication

• Tracleer® is indicated for the treatment of pulmonary arterial hypertension (PAH) (WHO Group 1):
  • in adults to improve exercise ability and to decrease clinical worsening. Studies establishing effectiveness included predominantly patients with WHO Functional Class II-IV symptoms and etiologies of idiopathic or heritable PAH (60%), PAH associated with connective tissue diseases (21%), and PAH associated with congenital heart disease with left-to-right shunts (18%).
  
• in pediatric patients aged 3 years and older with idiopathic or congenital PAH to improve pulmonary vascular resistance (PVR), which is expected to result in an improvement in exercise ability.
Conclusions

• Use of bridging biomarkers in pediatric extrapolation is distinct from other roles for biomarkers:
  • Not PD marker that is used to support dose selection
  • Not validated surrogate endpoint that can reliably predict the net effect of the intervention on feels, functions, or survives outcomes.

• To establish a bridging biomarker in registrational decision-making, the biomarker should satisfy the 5 core criteria

• Pediatric extrapolation using a bridging biomarker has been used to approve drugs for pediatrics
  • Bosentan for pediatric PAH
References


• ICH E11A: Pediatric Extrapolation Guideline (draft, currently under public consultation)

• ADEPT 7 workshop (September 1, 2021) https://cersi.umd.edu/2017-drug-development-pediatric-heart-failure-workshop


Session 4: Beyond Surrogate Endpoints: Other Ways Translational Science Can Support Drug Development

Moderator:
• David Strauss, US Food and Drug Administration

Panelists:
• Leslie Gordon, Brown University
• Estelle Marrer-Berger, Roche
• Christine Garnett, US Food and Drug Administration
• Anthony Durmowicz, Cystic Fibrosis Foundation
• Lynne Yao, US Food and Drug Administration
Session 4: Beyond Surrogate Endpoints: Other Ways Translational Science Can Support Drug Development

Discussion Questions:

1. What translational approaches assist in drug development programs beyond use of surrogate endpoints?

2. What benefits and challenges exist in using these translational approaches to support drug development?

3. How can translational science approaches support regulatory submissions for accelerated approval or traditional approval?

4. Is there more that can be done to encourage use of these approaches?
Session 5: Opportunities and Challenges for Incorporation of Translational Science in Clinical Development Programs

3:30 pm – 4:15 pm EST
Session 5: Opportunities and Challenges for Incorporation of Translational Science in Clinical Development Programs

Moderator:

• Michael Pacanowski, US Food and Drug Administration

Panelists:

• Jeffrey Siegel, US Food and Drug Administration
• David Reese, Amgen
• Jen Farmer, Friedrich’s Ataxia Research Alliance
• Steve Hoffmann, Foundation for the National Institutes of Health
Session 5: Opportunities and Challenges for Incorporation of Translational Science in Clinical Development Programs

Discussion Questions:

1. Reflecting on the meeting, what are key strategies for optimizing the use of surrogate endpoints and other translational approaches for drug development?

2. What are the challenges to taking a biomarker from discovery to validation?

3. Is there more that can be done to facilitate the process? What mechanisms might be able to increase the use of translational research studies?

4. What are key strategies for facilitating collaboration between stakeholders, with the overall goal of improving therapeutic development and approval?

5. What are future considerations and next steps for advancing translational science studies and increasing the use and acceptability of these approaches?
Closing Remarks | Day 2

Michael Pacanowski
Director of the Division of Translational and Precision Medicine
U.S. Food and Drug Administration
Thank You!

Contact Us

healthpolicy.duke.edu

Subscribe to our monthly newsletter at dukemargolis@duke.edu

1201 Pennsylvania Avenue, NW, Suite 500
Washington, DC 20004

DC office: 202-621-2800
Durham office: 919-419-2504

Follow Us

DukeMargolis

@DukeMargolis

@DukeMargolis

Duke Margolis