Case Studies in Designing Clinical Trials in Rare Disease

Melanie Quintana, PhD
Berry Consultants
CTMC Webinar 2019
July 16th 2019
Dominantly Inherited Alzheimer’s

- Rare genetic form of Alzheimer’s (<1% of total Alzheimer’s population)
- Early age of onset: 30-50
- Goal: Does the treatment slow cognitive progression?
GNE Myopathy

• Rare genetic muscle disease
• Slowly progressive muscle weakness and atrophy effecting different muscle groups at different stages of the disease
• Goal: Does the treatment slow decline of muscle strength?
Complexity in Rare Disease

• Heterogeneity in progression
• Large variability in key clinical endpoints
• Different endpoints are affected at different stages of the disease

Common Solutions:
• Enroll a more homogenous subset
• Enroll a large enough sample size to overcome heterogeneity

Both not ideal in a rare disease setting!
Solutions for Rare Disease

GNE Myopathy

• Natural History Study -> Disease Progression Model
• Joint Disease Modification Analysis incorporating all muscle groups

DIAN

• Natural History Study -> Disease Progression Model
• Disease Modification Analysis
• Adaptive Platform Trial with freq. interims and shared Controls

• Natural History Studies + Disease Progression Models -- Know what you are working with!
• Innovative Designs
  • More powerful analysis methods
  • Adaptive designs with frequent interims
  • Use all available data
DIAN: DOMINANTLY INHERITED ALZHEIMER'S
DIAN: SUMMARIZE NATURAL HISTORY
DIAN Observational Data

Cognitive performance by EYO

Non-Carriers (n=142)
Mutation Carriers (n=223)
Understand Natural Progression

Disease Progression Model

• Mathematical function quantitatively captures individual disease evolution in terms of a single or multiple disease-specific biomarkers and/or clinical outcome measures
  – Likely Monotonic
  – Ideally measures are modeled as a function of "disease age" / stage
DIAN Disease Progression

\[ Y_{ij} = \gamma_i + f(EYO_{ij} + \delta_i|\alpha) + \epsilon_{ij} \]

\[
f(x) = \begin{cases} 
0 & x \leq -15 \\
(1 + [x] - x)\alpha_{[x]} + (x - [x])\alpha_{[x]+1} & -15 < x \leq 15 \\
\alpha_{15} & x > 15 
\end{cases}
\]

- Expected progression as a function of EYO
  - Monotonically decreasing spline with knots at each integer value for EYO between -15 and +15
  - Subject-level random effect for the adjustment in the estimated age of onset (EYO_{ij})
  - Subject-level random effect for the cognitive score at the healthy stage EYO < -15
DIAN Disease Progression

\[ Y_{ij} = \gamma_i + f(EYO_{ij} + \delta_i|\alpha) + \epsilon_{ij} \]

\[ f(x) = \begin{cases} 
0 & x \leq -15 \\
(1 + [x] - x)\alpha_{[x]} + (x - [x])\alpha_{[x]+1} & -15 < x \leq 15 \\
\alpha_{15} & x > 15 
\end{cases} \]

- Expected progression as a function of EYO
  - Monotonically decreasing spline with knots at each integer value for EYO between -15 and +15
  - Subject-level random effect for the adjustment in the estimated age of onset (EYO_{ij})
  - Subject-level random effect for the cognitive score at the healthy stage EYO < -15
DIAN Disease Progression

\[ Y_{ij} = \gamma_i + f(EYO_{ij} + \delta_i | \alpha) + \epsilon_{ij} \]

\[
f(x) = \begin{cases} 
0 & x \leq -15 \\
(1 + [x] - x)\alpha_{[x]} + (x - [x])\alpha_{[x]+1} & -15 < x \leq 15 \\
\alpha_{15} & x > 15
\end{cases}
\]

• Expected progression as a function of EYO
  – Monotonically decreasing spline with knots at each integer value for EYO between -15 and +15
  – Subject-level random effect for the adjustment in the estimated age of onset (EYO\textsubscript{ij})
  – Subject-level random effect for the cognitive score at the healthy stage EYO < -15
DIAN Disease Progression
DIAN: VIRTUAL PATIENT SIMULATIONS
Virtual Patient Simulation

- Patient simulator for clinical trial design
Natural Cognitive Decline

Cog. Score

Years Since Onset
Subject 1:
Subject-level random effect: -.8
Years since onset at enrollment: 5
Enrolled to treatment group
DIAN: UNDERSTAND POWER
DIAN Initial Proposed Design

Proposed Design:

• 80 subjects per arm randomized 3:1 (Treatment: Control)
• Max length of follow-up: 4 years
• Primary Analysis Method: MMRM
Power DIAN Trial

- MMRM at 4 years

% Reduction Decline vs. Power
Common Primary Analysis: MMRM

• MMRM Issues:
  – Dilution of effect due to subjects not expected to progress (very early or very late disease)
  – Test effect at a single time point
DIAN: DESIGN INNOVATIONS
Common Primary Analysis: MMRM

• MMRM Issues:
  – Dilution of effect due to subjects not expected to progress (very early or very late disease)
  – Test effect at a single time point
Disease Progression Modification Analysis

• **DPMA: Assume proportional treatment effect at each EYO**

![Proportional Treatment Effect Graph](image)

- Proportional to the expected decline on control
- 50% Treatment effect
  - **EYO -5**: Abs. $\Delta = 0.125$
  - **EYO 5**: Abs. $\Delta = 0.4$
Disease Progression Modification Analysis

• **DPMA**: Assume proportional treatment effect at each EYO
  – Adjusts for expected decline given EYO
  – Uses all timepoints
    • Incorporate differential follow-up: Due to missing data; early interim analyses, extended follow-up
    • Extended follow-up = Greater Power
DPMA vs. MMRM

Power DIAN Trial

DPMA at 4 years
MMRM at 4 years

% Reduction Decline
DIAN Adaptive Platform Trial

Platform Trial w/ Multiple Drugs & Shared PBO

Adaptive Design w/ Frequent Interim Analyses for early success or futility
Borrowed Controls

Power DIA Trial

- Single PBO
- Shared PBO

% Reduction Decline

DPMA at 4 years
MMRM at 4 years

Berry Consultants
statistical innovation
Frequent Interim Analyses

Power DIAN Trial

- DPMA at 2 years
- DPMA at 3 years
- DPMA at 4 years

% Reduction Decline
Summary DIAN

• Natural History Studies + Clinical Trial Simulation = More Informed Trial Design!
  – Original DIAN Power = < 20%

• Need for better analysis methods that use all available data and adjust for expected progression
  – Innovative DPMA + Shared PBO leads to increase in DIAN power from <20% to > 80%!
GNE MYOPATHY
GNE: SUMMARIZE NATURAL HISTORY
GNE Natural History Data

- **Sample Size:** 38 Patients
- **Visits:** Every 3-6 months
  - Number of months from baseline per patient ranges from 0-32
- **Measurements taken on possible primary endpoints:**
  - Six minute walk
  - Quantitative Muscle Assessment (QMA) for multiple muscle groups
## Quantitative Muscle Assessment (QMA)

*Based on age, gender and BMI*

<table>
<thead>
<tr>
<th>Muscle Group</th>
<th>Measured Strength (kg)</th>
<th>Percent Predicted*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper Extremities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Grip</td>
<td>39.46</td>
<td>83.11%</td>
</tr>
<tr>
<td>R Grip</td>
<td>45.93</td>
<td>91.41%</td>
</tr>
<tr>
<td>L Wrist Ext</td>
<td>11.97</td>
<td>80.15%</td>
</tr>
<tr>
<td>R Wrist Ext</td>
<td>16.91</td>
<td>105.91%</td>
</tr>
<tr>
<td>L Shoulder Abd</td>
<td>21.50</td>
<td>88.53%</td>
</tr>
<tr>
<td>R Shoulder Abd</td>
<td>20.24</td>
<td>73.22%</td>
</tr>
<tr>
<td>L Elbow Flex</td>
<td>20.60</td>
<td>71.11%</td>
</tr>
<tr>
<td>R Elbow Flex</td>
<td>22.64</td>
<td>69.08%</td>
</tr>
<tr>
<td>L Elbow Ext</td>
<td>10.72</td>
<td>47.54%</td>
</tr>
<tr>
<td>R Elbow Ext</td>
<td>11.89</td>
<td>51.20%</td>
</tr>
<tr>
<td><strong>Sum Upper</strong></td>
<td><strong>221.84</strong></td>
<td><strong>77.02%</strong></td>
</tr>
<tr>
<td>L Dorsiflex</td>
<td>8.02</td>
<td>25.83%</td>
</tr>
<tr>
<td>R Dorsiflex</td>
<td>7.42</td>
<td>23.98%</td>
</tr>
<tr>
<td>L Hip Abd</td>
<td>27.48</td>
<td>77.38%</td>
</tr>
<tr>
<td>R Hip Abd</td>
<td>31.27</td>
<td>90.44%</td>
</tr>
<tr>
<td>L Hip Ext</td>
<td>56.40</td>
<td>117.54%</td>
</tr>
<tr>
<td>R Hip Ext</td>
<td>54.28</td>
<td>114.39%</td>
</tr>
<tr>
<td><strong>Sum Lower</strong></td>
<td><strong>300.83</strong></td>
<td><strong>74.15%</strong></td>
</tr>
<tr>
<td>Sum Strength</td>
<td>522.66</td>
<td>75.34%</td>
</tr>
</tbody>
</table>

* Harris-Love et al, Rehab Research Practice, 2014
Possible Primary Endpoints: 6 Min. Walk

Raw Scores

Change from Baseline + Model Fit

- Natural Progression +/- 1SD
- 50% Reduction in Decline
- 100% Reduction in Decline

Change (SD): Im
raw
- 1 Year
  -10.69 (32.69)
- 2 Year
  -21.38 (32.69)

N 80% Power:
- 1 Year: 1178
- 2 Year: 296

Power N=80:
- 1 Year: 0.11
- 2 Year: 0.3

Berry Consultants Statistical Innovation
Possible Primary Endpoints:
Upper Extremity Composite Subset*

Raw Scores

Change from Baseline + Model Fit

Change (SD): 
1 Year: -5.66 (10.9)
2 Year: -11.32 (10.9)

N 80% Power: 
1 Year: 468
2 Year: 120

Power N=80: 
1 Year: 0.21
2 Year: 0.63
Aug 22, 2017

Ultragenyx Announces Top-Line Results from Phase 3 Study of Ace-ER in GNE Myopathy

Study did not meet its primary endpoint

NOVATO, Calif., Aug. 22, 2017 (GLOBE NEWSWIRE) -- Ultragenyx Pharmaceutical Inc. (NASDAQ:RARE), a biopharmaceutical company focused on the development of novel products for rare and ultra-rare diseases, today announced that a Phase 3 study evaluating aceneuramic acid extended release (Ace-ER) in patients with GNE Myopathy (GNEM) did not achieve its primary endpoint of demonstrating a statistically significant difference in the upper extremity muscle strength composite score compared to placebo. The study also did not meet its key secondary endpoints. Adverse events were generally balanced between Ace-ER and placebo and safety was consistent with previously released Ace-ER data. Ultragenyx plans to discontinue further clinical development of Ace-ER.

"We are disappointed by these results, as we had hoped that Ace-ER would offer a new option for GNEM patients. We would like to thank the patients, caregivers, and investigators involved in the Ace-ER development program," said Emil D. Kakkis, M.D., Ph.D., Chief Executive Officer and President of Ultragenyx. "This outcome does not affect our overall strategy, as the company moves forward with multiple preclinical and clinical programs and regulatory filings."

The Phase 3 Ace-ER study enrolled 89 adults with GNEM able to walk > 200 meters in the six minute walk test. Patients were randomized 1:1 to Ace-ER at a dose of 6g/day or placebo for 48 weeks. The study did not meet the primary endpoint of demonstrating a statistically significant improvement in UEC score (+0.74 kg, p=0.5387) for Ace-ER treated patients (n=45, -2.25 kg) compared to placebo (n=43, -2.99 kg) patients for the change from baseline to 48 weeks. There were three pre-specified key secondary endpoints, including the lower extremity muscle strength composite score as measured by hand-held dynamometry (HHD), physical functioning using the Mobility domain of the GNE Myopathy-functional activity scale (GNEM-FAS), and a measure of muscle strength in knee extensors. The study did not meet any of these key secondary endpoints.
Model Muscle Decay

- **Goal:** Model the expected decay of each muscle over time

- Need to align patients based on an unknown “disease age”
Natural History Data

A) Subject A

B) Subject B

C) Subject C

D) Subject D

- Knee Flexion
- Ankle Dorsiflexion
- Knee Extension
- Grip
- Shoulder Abduction
- Elbow Flexion

Muscle Strength vs. Chronological Age
Decline vs. Age

A) Ankle Dorsiflexion

B) Knee Flexion

C) Knee Extension

D) Grip

E) Shoulder Abduction

F) Elbow Flexion
Decline vs. Latent Disease Age

\[ Y_{i,j,k} \sim N(\mu_{i,j,k},(\sigma^2 + \delta)) \]

\[ \mu_{i,j,k} = \text{logit}^{-1}[\theta_k + \beta_k(t_{i,j} - \alpha_i)] \cdot M_{i,k} \]

- Model proportion of muscle strength
  - normal distribution
  - variance is a function of mean and muscle-specific component

- Model mean based on logit decay function with components:
  - muscle specific location parameter (when the muscle begins to decay)
  - muscle specific slope parameter (rate of decay)
  - Subject-specific age adjustment parameter, determines “disease age”
  - Subject and muscle specific relative maximum, to account for variation in overall strength of the individual
Decline vs. Latent Disease Age

\[ Y_{i,j,k} \sim N(\mu_{i,j,k}, (\sigma_{k} \mu_{i,j,k})^2 + \delta) \]

\[ \mu_{i,j,k} = \text{logit}^{-1}[\theta_{k} + \beta_{k} (t_{i,j} - \alpha_{i})] \times M_{i,k} \]

- Model proportion of muscle strength
  - normal distribution
  - variance is a function of mean and muscle-specific component

- Model mean based on logit decay function with components:
  - muscle specific location parameter (when the muscle begins to decay)
  - muscle specific slope parameter (rate of decay)
  - Subject-specific age adjustment parameter, determines “disease age”
  - Subject and muscle specific relative maximum, to account for variation in overall strength of the individual
Decline vs. Latent Disease Age

\[ Y_{i,j,k} \sim N(\mu_{i,j,k}, (\sigma_{k} \mu_{i,j,k})^2 + \delta) \]

\[
\mu_{i,j,k} = \logit^{-1}[\theta_k + \beta_k (t_{i,j} - \alpha_i)] * M_{i,k}
\]

• Model proportion of muscle strength
  – normal distribution
  – variance is a function of mean and muscle-specific component

• Model mean based on logit decay function with components:
  – muscle specific location parameter (when the muscle begins to decay)
  – muscle specific slope parameter (rate of decay)
  – Subject-specific age adjustment parameter, determines “disease age”
  – Subject and muscle specific relative maximum, to account for variation in overall strength of the individual
Decline vs. Latent Disease Age

\[ Y_{i,j,k} \sim N(\mu_{i,j,k}, (\sigma_k \mu_{i,j,k})^2 + \delta) \]

\[ \mu_{i,j,k} = \text{logit}^{-1}[\theta_k + \beta_k (t_{i,j} - \alpha_i)] * M_{i,k} \]

- Model proportion of muscle strength
  - normal distribution
  - variance is a function of mean and muscle-specific component

- Model mean based on logit decay function with components:
  - muscle specific location parameter (when the muscle begins to decay)
  - muscle specific slope parameter (rate of decay)
  - Subject-specific age adjustment parameter, determines “disease age”
  - Subject and muscle specific relative maximum, to account for variation in overall strength of the individual
Decline vs. Latent Disease Age

\[ Y_{i,j,k} \sim N(\mu_{i,j,k}, (\sigma^2_{k}\mu_{i,j,k})^2 + \delta) \]

\[ \mu_{i,j,k} = \logit^{-1}(\theta_k + \beta_k (t_{i,j} - \alpha_i)) \cdot M_{i,k} \]

• Model proportion of muscle strength
  – normal distribution
  – variance is a function of mean and muscle-specific component

• Model mean based on logit decay function with components:
  – muscle specific location parameter (when the muscle begins to decay)
  – muscle specific slope parameter (rate of decay)
  – Subject-specific age adjustment parameter, determines “disease age”
  – Subject and muscle specific relative maximum, to account for variation in overall strength of the individual
Decline vs. Latent Disease Age

\[ Y_{i,j,k} \sim N(\mu_{i,j,k},(\sigma_{k}\mu_{i,j,k})^2 + \delta) \]

\[ \mu_{i,j,k} = \text{logit}^{-1}[\theta_k + \beta_k(t_{i,j} - \alpha_i)] \times M_{i,k} \]

- Model muscle strength
  - normal distribution
  - variance is a function of mean and muscle-specific component

- Model mean based on logit decay function with components:
  - muscle specific location parameter (when the muscle begins to decay)
  - muscle specific slope parameter (rate of decay)
  - Subject-specific age adjustment parameter, determines “disease age”
  - Subject and muscle specific relative maximum, to account for variation in overall strength of the individual
Disease Age Per Subject

Patient: A
Baseline DA: 15.58

Patient: B
Baseline DA: −8.12

Patient: C
Baseline DA: 10.38

Patient: D
Baseline DA: −9.88

Patient: A
Baseline DA: 15.58

Patient: B
Baseline DA: −8.12

Patient: C
Baseline DA: 10.38

Patient: D
Baseline DA: −9.88

Disease Age Per Subject

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion

Proportion Max Strength

Proportion Predicted Strength

Chronological Age

Disease Age

Knee Flexion
Ankle Dorsiflexion
Knee Extension
Grip
Shoulder Abduction
Elbow Flexion
Expected Muscle Decline vs. Disease Age
Expected Muscle Decline vs. Disease Age

G) Ankle Dorsiflexion

H) Knee Flexion

I) Knee Extension

J) Grip

K) Shoulder Abduction

L) Elbow Flexion
Expected Muscle Decline vs. Disease Age
Patient-Specific Prediction

A) Subject A

B) Subject B

C) Subject C

D) Subject D

E) Disease Age: 17.75

F) Disease Age: ~8.02

G) Disease Age: 11.04

H) Disease Age: ~11.33
Disease Age vs. Clinical Manifestation

Climbing Stairs
- Normal
- Slow, intermittent use of handrail
- Dependent on handrail
- Handrail and additional support
- Unable

Walking
- Normal
- Slow or mild unsteadiness
- Intermittent use of device
- Dependent on device
- Wheelchair dependent

Turning in Bed
- Normal
- Slow
- Great difficulty
- Can initiate but not turn alone
- Total assistance

Cutting Food & Handling Utensils
- Normal
- Slow
- Some help needed
- Food cut by someone
- Needs to be fed

Sit to Stand
- Independent
- Performs with substitute motions
- Requires use of arms
- Requires assistance
- Unable

Fine Motor Tasks
- Independent
- Slow
- Modified technique or assistive device
- Frequent assistance from caregiver
- Unable

Hygiene (Bathing & Toileting)
- Normal
- Increased effort
- Modified technique or assistive device
- Occasional assistance from caregiver
- Dependent

Dressing
- Normal
- Increased effort
- Modified technique or assistive device
- Occasional assistance from caregiver
- Dependent
**Disease Age 0:**
- Intermittent use of handrail when climbing stairs
- Mild unsteadiness when walking
- Use substitute motions when standing from sitting
- 6MWT: ~450 meters
- AMAT high functional group

**Disease Age 10:**
- Dependent on a handrail when climbing stairs
- Intermittent use of a device for walking
- Slow turning in bed, cutting food and handling eating utensils
- Increased effort with hygiene and dressing
- 6MWT: ~340 meters
- AMAT moderate functional group

**Disease Age 20:**
- Requires use of arms when standing from sitting
- Slow fine motor tasks
- 6MWT: ~230 meters, but several patients unable to complete
- AMAT low functional group

**Disease Age 30:**
- Requires handrail and additional support to climb stairs
- Dependent on device to walk
- Great difficulty turning in bed
- Needs help with fine motor tasks, cutting food & handling utensils
- Modified technique for hygiene and dressing
- 6MWT: Not able to complete
GNE: DESIGN + ANALYSIS INNOVATIONS
Incorporation of Treatment Effect

\[ Y_{i,j,k} \sim N(\mu_{i,j,k}, (\sigma_k \mu_{i,j,k})^2 + \delta) \]

\[
\mu_{i,j,k} = \begin{cases} 
\text{logit}^{-1}\left[ \theta_k + \beta_k (t_{i,j} - \alpha_i) \right] \ast M_{i,k} & \text{pre-treatment} \\
\text{logit}^{-1}\left[ \theta_k + \beta_k (t_{i,j} - T_i) \right] \ast M_{i,k} & \text{post-treatment}
\end{cases}
\]

\[
\text{Time of treatment subject } i
\]

**Treatment effect**: Constant % slowing in the rate of decline across all muscles under treatment compared to the rate of decline in the placebos

- Alt. interpretation: Slowing in number of years it will take to reach milestones
  - Ex: 50% slowing in rate of decline = will take subject 2x’s as many years to reach milestone under treatment
Incorporation of Treatment Effect

- Ability to detect treatment effect depends are where the patient is on the decline
- Muscle where we can best detect treatment effect is subject-specific
  - Given the patients disease age, which muscle is actively decaying
- Incorporate all muscles in the estimation of treatment effect
50% Reduction in Decline

<table>
<thead>
<tr>
<th>Natural Progression</th>
<th>Treatment*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of knee flexion strength (0.60-&gt;0.20)</td>
<td>42 years old</td>
</tr>
<tr>
<td>Ability to climb stairs: Dependent on a handrail</td>
<td>36 years old</td>
</tr>
<tr>
<td>Ability to walk: Intermittent use of an ambulatory device</td>
<td>38 years old</td>
</tr>
</tbody>
</table>

*Slow Decline by 50% with \( \gamma = 0.5 \)
GNE Proposed Design + Power

- N=50 Enrolled 2:1 Treatment vs. Control
- Follow all subjects for 2 years vs. Extended follow-up
- Primary Analysis: Disease Modification Analysis
- Treatment effects: 0% (Null) vs. 40-50% slowing in progression

<table>
<thead>
<tr>
<th>% Reduction</th>
<th>Scenario</th>
<th>No LTFU</th>
<th>10% Annual Rate LTFU</th>
<th>10% Annual Rate LTFU + Extended FU</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td></td>
<td>0.886</td>
<td>0.823</td>
<td>0.892</td>
</tr>
<tr>
<td>45%</td>
<td></td>
<td>0.797</td>
<td>0.722</td>
<td>0.817</td>
</tr>
<tr>
<td>40%</td>
<td></td>
<td>0.676</td>
<td>0.597</td>
<td>0.715</td>
</tr>
<tr>
<td>0%</td>
<td></td>
<td>0.012</td>
<td>0.011</td>
<td>0.011</td>
</tr>
</tbody>
</table>
Summary GNE

• Natural History Studies + Clinical Trial Simulation = More Informed Trial Design!
  – Traditional Design GNE= 21%
  – Need

• Need for better analysis methods that use all available data and adjust for expected progression
  – Innovative DPMA 21% to > 80%! 
References


RESEARCH ARTICLE

A novel cognitive disease progression model for clinical trials in autosomal-dominant Alzheimer’s disease

Guoqiao Wang1 | Scott Berry2 | Chengjie Xiong1 | Jason Hassenstab3 | Melanie Quintana2 | Eric M. McDade3 | Paul Delmar4 | Matteo Vestruci4,5 | Gopalan Sethuraman6 | Randall J. Bateman3
For the Dominantly Inherited Alzheimer Network Trials Unit

Clinical trial outcomes for Alzheimer’s disease are typically analyzed by using the mixed model for repeated measures (MMRM) or similar models that compare an efficacy scale change from baseline between treatment arms with or without participants’ disease stage as a covariate. The MMRM focuses on a single-point fixed follow-up duration regardless of the exposure for each participant. In contrast to these typical models, we have developed a novel semiparametric cognitive disease progression model (DPM) for autosomal dominant Alzheimer’s disease based on the Dominantly Inherited Alzheimer Network (DIAN) observational study. This model includes 3 novel features, in which the DPM (1) aligns and compares participants by disease stage, (2) uses a proportional treatment effect similar to the concept of the Cox proportional hazard ratio, and (3) incorporates extended follow-up data from participants with different follow-up durations using all data until last participant visit. We present the DPM model developed by using the DIAN observational study data and demonstrate through simulation that the cognitive DPM used in hypothetical intervention clinical trials produces substantial gains in power compared with the MMRM.

KEYWORDS
Alzheimer's disease, disease progression model, mixed effects model for repeated measures, proportional treatment effect

1 | INTRODUCTION

The Alzheimer’s disease (AD) field has progressively moved to studying potential disease-modifying therapies in earlier stages of disease, including in asymptomatic stages before dementia onset. The failure to develop effective disease-modifying treatments in the later dementia stages of disease1,2 has supported the notion that prevention is likely to be more effective. However, AD prevention trials require long periods of follow-up because of the slowly progressive nature of cognitive decline over many years, creating a major challenge in the implementation of prevention trials. For example, participant attrition and long enrollment durations combined with long treatment periods make implementation of large complex prevention trials impractical and less efficient. Several strategies are being pursued to mitigate
these challenges, including the development of surrogate biomarkers,\textsuperscript{3} more sensitive measures of cognitive disease progression,\textsuperscript{4} and platform trials to test several targets in parallel.\textsuperscript{5} In addition, the advancement of statistical methods provides an opportunity to greatly increase the power, speed, and efficiency of AD prevention trials.

Several strategies may be incorporated to increase power, including increasing overall sample sizes,\textsuperscript{6,7} adjusting the sample size based on results from interim analysis,\textsuperscript{8} using a targeted trial that enrolls only a very specific population,\textsuperscript{9} or using enrollment enrichment strategies to reduce heterogeneity of trial participants.\textsuperscript{10} Although these trial designs may improve the ability to detect effective treatments, they have drawbacks including increased time, expense, and limited generalizability. Because trial outcomes are typically determined with the mixed model for repeated measures (MMRM) using time since baseline as a categorical variable or even cross-sectional models that focus on comparing the absolute cognitive change from baseline to a fixed post baseline time point,\textsuperscript{6,7,11,12} typical clinical trials collect clinical assessments for a fixed duration for each participant. And the participants who were enrolled early and had completed the fixed follow-up duration were no longer active, while the late enrollees were still fulfilling the follow-up. Thus, the early enrollees with potentially the longest and most valuable exposure if continuously followed do not contribute to a stronger analysis. This lost opportunity is even more significant in prevention trials with extended enrollment times. Additionally, participants typically enter trials at different stages of disease; thus, statistical models that ignore disease stage and look at change from trial baseline introduce additional heterogeneity due to variability of disease stage. Overall, these shortcomings have resulted in trial designs that require large sample sizes and long exposure to achieve acceptable statistical power.\textsuperscript{2,6,7} We demonstrate that a disease progression model (DPM) built from the Dominantly Inherited Alzheimer Network (DIAN) observational cohort avoids each of these shortcomings to greatly increase power.

Autosomal-dominant AD (ADAD) is a rare genetic disorder caused by a mutation in 1 of 3 genes: amyloid precursor protein (\textit{APP}), presenilin 1 (\textit{PSEN1}), or presenilin 2 (\textit{PSEN2}). Mutation carriers are destined to develop dementia of the Alzheimer's type, generally at an early age, typically with the age of onset between 30 and 50 years.\textsuperscript{13} Age of onset for asymptomatic mutation carriers can be estimated from other carriers of the same mutation. The estimated years from symptom onset (EYO) for an individual is their current age minus their estimated age of onset.\textsuperscript{14} Within the DIAN observational study, we have shown that EYO is a reliable predictor of disease stage.\textsuperscript{14,15} Additionally, it also provides a remarkably consistent rate of cognitive disease progression across participants. We have developed a cognitive disease progression model (DPM) based on the DIAN observational study that models the rate of cognitive progression as a function of EYO. The consistency of the rate of decline as a function of EYO is striking. We demonstrate how this cognitive DPM is used in a hypothetical phase III interventional trial and compare the power of the cognitive DPM with the more conventional MMRM model typically used in AD clinical trials.

2 | MATERIALS AND METHODS

2.1 | DIAN observational study

The DIAN observational study is an international, multisite, longitudinal study of individuals from families with an established history of ADAD. Participants must have confirmation of a causal ADAD mutation in their family, with a 50% chance of inheriting the mutation. From the disease modeling aspects of this paper, we analyze only confirmed mutation carriers. The details of participants' demographics and the clinical, cognitive, imaging, and biochemical measures have been reported in previous publications.\textsuperscript{14} The data used in the DPM development include DIAN quality-controlled data from July 2008 to January 2015 consisting of 225 mutation carriers.

2.2 | Estimated years from symptom onset (EYO)

Each participant in the study has an estimated age of onset. The assignment of an age of onset for each participant is based on the mean age of onset for that person's specific matching mutation established through systematic review and meta-analysis.\textsuperscript{15} Each clinical assessment occurs at a time differential from the participant's estimated age of onset—we refer to the timing as the EYO. Specifically, the EYO is calculated as the age of the participant at the time of the clinical assessment minus this participant's estimated age of onset. For example, at an assessment, if a participant's age is 45.3 years, and the estimated age of onset for this participant is 50.2, then EYO for that assessment is $-4.9$, meaning that this participant is about 4.9 years to his/her symptom onset.
2.3 | Cognitive composite

The cognitive composite used in these analyses combines measures of episodic memory, executive functioning, processing speed, and mental status and was chosen to sensitively measure the cognitive decline, which occurs before the first symptom onset in preclinical AD. Three separate approaches using a mathematically optimized approach,\(^\text{16}\) basic principles of neuropsychology,\(^\text{17,18}\) and evaluating prior demonstrated domains in sporadic AD were compared and found to converge on the 4 domains included in the cognitive composite for this study. This composite is similar to other composites.\(^\text{19}\) Episodic memory is assessed with the DIAN Word List test delayed recall and the delayed recall score from the Wechsler Memory Scale-Revised Logical Memory IIA subtest.\(^\text{20}\) Executive functioning and processing speed is assessed with the Wechsler Adult Intelligence Scale-Revised Digit-Symbol Substitution test, and mental status with the Mini Mental State Examination (MMSE). This cognitive composite was developed by normalizing each individual test to a z-score before averaging. All components except the MMSE are normalized using the mean and standard deviation (SD) of each component score from mutation carriers well before symptom onset (EYO ≤ −15). However, the MMSE has a ceiling effect, the SD among those with EYO ≤ −15 is small, and using this SD will overweight MMSE. A simple smoothing spline model for the rate of decline of MMSE was fit, and the estimated SD from the model is used for the normalization. The details for the normalization are provided in the Supporting Information.

Next, the 4 z-scores are equally weighted to construct a single composite. The construction of the cognitive composite creates a single score with mean zero and SD near 1 for participants in a healthy state (EYO ≤ −15).

2.4 | The cognitive DPM based on the DIAN observational study

Let \(Y_{ij}\) be the \(j\)th cognitive composite measured at EYO\(_{ij}\) for participant \(i, i = 1, ..., k\). Let \(n_i\) be the number of observations for participant \(i\). The underlying rationale of the cognitive DPM is to represent a participant’s composite score at any given EYO as a function of this participant’s score in a healthy stage (defined as EYO ≤ −15 in this study) plus a decline from the relatively healthy stage to this particular EYO. The cognitive composite is modeled as a function of EYO\(_{ij}\) by using a mixed-effects model,

\[
Y_{ij} = \gamma_i + f(EYO_{ij} + \delta_i|x) + \epsilon_{ij} \text{ for } i = 1, ..., k; j = 1, ..., n_i.
\]

The random effect parameter \(\gamma_i\) is the individual cognitive composite at the relatively healthy stage defined as (EYO ≤ −15) in this study. The function \(f(x)\) represents the mean decline from the relatively healthy stage to a given EYO (the semiparametric model presented below). The random effect \(\delta_i\) represents a participant-level adjustment to incorporate the uncertainty in the estimate for the covariate age of onset.

Function \(f(x)\) is modeled as a monotonically decreasing spline with knots at each integer value for EYO between (inclusive) −15 and +15, represented by \(\alpha_{-15}, \alpha_{-14}, ..., \alpha_{15}\). The modeling puts no restriction on the shape of the decline curve—it does not enforce linearity, but only a decline in cognitive mean as the participant ages. The decline is defined for all continuous EYOs by using linear interpolation between the values of \(\alpha\)

\[
f(x) = \begin{cases} 
0 & x \leq -15 \\
(1 + [x] - x)\alpha_{[x]} + (x - [x])\alpha_{[x]+1} & -15 < x \leq 15, \\
\alpha_{15} & x > 15 
\end{cases}
\]

where \([x]\) is the floor function and it represents the largest integer less than \(x\). The parameters for this model are the values of \(f\) at each knot point \(\alpha_x\). The errors, \(\epsilon_{ij}\), are assumed to be independent and identically distributed with a normal distributions \(\mathcal{N}(0, \sigma^2)\). The prior distributions for the 2 random effects across participants are modeled as: \(\gamma_i \sim \mathcal{N}(0, 1), \delta_i \sim \mathcal{N}(0, 2)\); the variance, \(\sigma^2\), is assumed to be a weak (nearly non-informative) prior inverse-gamma distribution: \(\sigma^2 \sim IG(0.01, 0.01)\), and the \(\alpha\) are assumed to be proportional to a weak normal distribution restricted to decreasing values:

\[
\alpha_x \sim \mathcal{N}(\alpha_{x-1}, 100^2)I_{[\alpha_x > \alpha_{x-1}]}, \text{ for } x = -14, ..., 15,
\]

For identifiability, we assume that the mean cognitive score at EYO −15 (considered as healthy) is 0; thus, participants with EYO ≤ −15 or less are represented as \(Y_{ij} = \gamma_i + \epsilon_{ij}\). This preserves the interpretation of the random effect \(\gamma_i\) to represent the mean cognitive score in a healthy state of the individual. This monotonicity assumption in \(\alpha\) plays 2 very
important roles. The first is that it forces the estimates of the mean cognitive composite to decline over time—accounting for the scientifically expected result as part of the model. The second is that this restriction creates inherent smoothing as well, creating a smooth estimate of the mean decline over time.

To calculate the posterior distribution, an MCMC algorithm is used with a single chain after a burn-in of 10 000 observations and a chain length of 100 000. The algorithm uses adaptively updated Metropolis-Hastings steps for improved convergence and mixing. See the Supporting Information for further details.

2.5 | Estimated natural decline of ADAD population using DIAN observational study

We apply the mixed effects cognitive DPM to the DIAN observational data. The posterior mean and SD of the natural decline at each EYO point—the $\alpha$ in Equation (1)—are presented in Table 1. In Table S2, the posterior mean and SD of the natural decline is provided without the monotonicity assumption. Each decline parameter, $\alpha$, represents the mean decline in the cognitive composite from the healthy state. The model estimated decline from an EYO of $-15$ to an EYO of 0 (onset) is $-1.06$, meaning that the mean decline to the point of estimated symptom onset is approximately a 1 point $z$-score decline in the cognitive endpoint. In the 4 years following onset, there is an estimated additional 1.09 decline, meaning that the rate of decline is estimated to increase 3 to 4 times after onset compared with the 15 years before onset. The posterior mean of the model standard error of the cognitive test around the true mean, $\sigma$, is 0.333 with a SD of 0.019.

2.6 | Modeling therapeutic treatment effect

With the progression of cognitive decline estimated as a function of EYO, we are interested in modeling a potential treatment effect on the rate of cognition decline in the ADAD population. We extend the above-described mixed effects model of cognitive decline across natural history participants to incorporate a treatment effect under an experimental treatment with an assumption of a proportional treatment effect to the rate of natural cognitive decline. In particular, if a treatment is provided to subject $i$ at time $T_i$, measured on the time scale of EYO, then the model for the cognitive endpoint incorporates the effect of the treatment from the point of intervention forward. The multiplicative effect to the

**TABLE 1** The posterior mean (SD) of the mean cognitive decline for each EYO estimated by the cognitive disease progression model using Dominantly Inherited Alzheimer Network observational study

<table>
<thead>
<tr>
<th>EYO</th>
<th>Posterior Mean (SD) of $\alpha_{EYO}$</th>
<th>EYO</th>
<th>Posterior Mean (SD) of $\alpha_{EYO}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-15$</td>
<td>0 (0)</td>
<td>$1$</td>
<td>$-1.20$ (0.18)</td>
</tr>
<tr>
<td>$-14$</td>
<td>$-0.07$ (0.06)</td>
<td>$2$</td>
<td>$-1.40$ (0.29)</td>
</tr>
<tr>
<td>$-13$</td>
<td>$-0.14$ (0.08)</td>
<td>$3$</td>
<td>$-1.70$ (0.41)</td>
</tr>
<tr>
<td>$-12$</td>
<td>$-0.21$ (0.09)</td>
<td>$4$</td>
<td>$-2.15$ (0.44)</td>
</tr>
<tr>
<td>$-11$</td>
<td>$-0.27$ (0.09)</td>
<td>$5$</td>
<td>$-2.66$ (0.38)</td>
</tr>
<tr>
<td>$-10$</td>
<td>$-0.33$ (0.10)</td>
<td>$6$</td>
<td>$-2.93$ (0.34)</td>
</tr>
<tr>
<td>$-9$</td>
<td>$-0.39$ (0.10)</td>
<td>$7$</td>
<td>$-3.11$ (0.38)</td>
</tr>
<tr>
<td>$-8$</td>
<td>$-0.46$ (0.11)</td>
<td>$8$</td>
<td>$-3.37$ (0.37)</td>
</tr>
<tr>
<td>$-7$</td>
<td>$-0.53$ (0.11)</td>
<td>$9$</td>
<td>$-3.71$ (0.24)</td>
</tr>
<tr>
<td>$-6$</td>
<td>$-0.61$ (0.12)</td>
<td>$10$</td>
<td>$-3.86$ (0.24)</td>
</tr>
<tr>
<td>$-5$</td>
<td>$-0.68$ (0.12)</td>
<td>$11$</td>
<td>$-4.07$ (0.27)</td>
</tr>
<tr>
<td>$-4$</td>
<td>$-0.76$ (0.12)</td>
<td>$12$</td>
<td>$-4.29$ (0.37)$^a$</td>
</tr>
<tr>
<td>$-3$</td>
<td>$-0.83$ (0.13)</td>
<td>$13$</td>
<td>$-6.10$ (0.94)$^a$</td>
</tr>
<tr>
<td>$-2$</td>
<td>$-0.90$ (0.13)</td>
<td>$14$</td>
<td>$-7.77$ (1.51)$^a$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$-0.98$ (0.14)</td>
<td>$15$</td>
<td>$-9.22$ (1.73)$^a$</td>
</tr>
<tr>
<td>0</td>
<td>$-1.06$ (0.14)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$The minimum of this composite score is $-4.11$, which is achieved around estimated years from symptom onset (EYO) 11 or when all the components have scores of 0. These $z$-scores are less than the minimum and are model estimated decline that had the monotonic decline continued beyond EYO 11. However, autosomal-dominant Alzheimer’s disease patients rarely if any survive past EYO 10.
future rate of progression after intervention is modeled as $e^\theta$, with log-progression rate $\theta$. The cognitive composite $Y_{ij}$ with corresponding timing of the observation as $EYO_{ij}$, is modeled as

$$Y_{ij} = \gamma_i + g(EYO_{ij}|\delta_i, T_i, \theta) + \varepsilon_{ij} \text{ for } i = 1, \ldots, k; j = 1, \ldots, n_i.$$  

The decline function $g$ is a combination of the natural decline $f$ and the effect of the intervention $\delta_i$:

$$g(EYO_{ij}|\delta_i, T_i, \theta) = \begin{cases} 
   f(EYO_{ij} + \delta_i|\alpha) & EYO_{ij} \leq T_i \\
   f(T_i + \delta_i|\alpha) + e^{\theta} [f(EYO_{ij} + \delta_i|\alpha) - f(T_i + \delta_i|\alpha)] & EYO_{ij} > T_i
\end{cases}$$

The random effects $\gamma_i$ and $\delta_i$, the variance $\sigma^2$, and the mean decline at each knot ($\alpha$) are modeled in the same way as in Equation (1). The prior distribution for the decline parameters is varied from the original model to have an SD of 1.5 between yearly deviations. The choice was based on calibrating the model to have good type I error properties while still being relatively uninformative. The priors of $\alpha$’s are set to be

$$\alpha_x \sim N(\alpha_{x-1}, 1.5^2)I_{[\alpha_x < \alpha_{x-1}]}, \text{ for } x = -14, \ldots, 15.$$  

The effect of a treatment is captured by the parameter $\theta$. The value of $e^\theta$ is the proportional effect of the treatment and is referred as the cognitive progression ratio (CPR). If the CPR is equal to 1, then the rate of decline on the treatment is identical to the control (natural history or placebo) after the treatment intervention. A value of the CPR larger than 1 would indicate an increased rate of decline for the treatment compared with the control. If the CPR is less than 1, then the rate of decline is slower for the treatment than for the control. The value of the CPR is interpreted as the ratio of the rate of decline under the treatment to the control; thus, this quantity is directly interpretable as the size of the treatment effect and its clinical relevance. The CPR can be thought of much like a hazard ratio in a time to event analysis. For example, if the CPR is 0.70, then the rate of decline on the treatment is slowed by 30% compared with the control.

We explore the ramifications of different treatment effects by showing how a modeled treatment effect slows the cognitive decline. Figure 1 shows the estimated natural decline and the decline if a treatment is started at EYO −15 with a 30% slowing (blue) or 70% slowing (green). A mutation carrier is expected to naturally decline to a z-score of −1 approximately at EYO −1. A 30% treatment effect would lead to a delay of 3 years (from EYO −1 to EYO 2) to reach the same −1 z-score, whereas a 70% treatment effect would lead to a delay of 9 years.

### 2.7 Simulation of ADAD clinical trials

To demonstrate the behavior of the cognitive DPM, we simulated virtual ADAD clinical trials by using the DPM as the primary analysis. We selected the virtual patient simulation parameters to closely mimic the original DIAN-TU trial.
design. We simulated trials with 3:1 treatment to placebo randomization ratio of 80 patients. We include 2 interim analyses for efficacy. Overall, we make the following assumptions for our simulated trials:

- **Sample size**: 80 (60:20)
- **Duration**: 4-year follow up after the last enrolled participant for any group. If cognitive measures beyond 4-year follow-up are available for participants enrolled earlier in the trial, the data will be incorporated into the cognitive endpoint model to estimate the treatment effect.
- **Accrual rate**: a mean of 5 per month simulated from a Poisson distribution
- **Frequency of cognitive assessments**: every 6 months
- **Dropout rate**: 5% annually, simulated independent of cognitive value, meaning missing completely at random
- **Simulation of an individual participant**
  - An EYO at enrollment is simulated as a uniform value over the integers from −15 to +10 (inclusive).
  - Expected natural cognitive progression: All new placebo participants behave like the cognitive model (with assumed monotonically decreasing cognitive mean values) estimates from the DIAN observational data, with respective variability of new measurements. In particular, posterior mean estimates given in Table 1 that were estimated using all DIAN observational participants are used to simulate the cognitive measure (including the SD of 0.333). The postbaseline scores are simulated assuming a treatment CPR value (1 for placebo, and different values of CPR for the experimental treatment).
  - A simulation of CDR global based on the baseline cognitive composite is conducted to determine if the additional entry criterion of CDR global ≤ 1 is met. If the CDR global criterion is not met, a substitute participant is simulated until the participant meets the CDR global condition.

- **Interim analysis**: 2 interim analyses are conducted: when the last participant in the treatment cohort reaches 2 years and 3 years of follow-up. At each interim, the treatment will be stopped for efficacy if it demonstrates a statistically superior slowing of cognitive decline. If efficacy is not demonstrated at any interim analyses, then the experimental treatment will continue to the next interim analysis or the final analysis. The rules defined for efficacy success are as follows: a treatment will be stopped early for efficacy if the posterior probability that the treatment slows the rate of cognitive decline is greater than or equal to 0.9952 (see the Supporting Information for the calculation of this threshold to control type I error).
- **Primary analysis**: The primary analysis is performed when the last participant enrolled to the trial has been followed for 4 years. The null hypothesis of the primary analysis is $H_0: \text{CPR} = 1$ against the alternative hypothesis: $H_A: \text{CPR} < 1$. The treatment will be declared successful, and the null hypothesis rejected in favor of the alternative, and a slowing of the rate of cognitive decline concluded if the posterior probability of a CPR < 1 is greater than or equal to 0.9952. The threshold for final success has been determined to control the one-sided type I error at 0.025 or less taking into account the interim analyses.

Assuming hypothetical values of CPR ranging from 0% to 80%, we simulated 5000 trials for each reduction, and 50 000 draws from each MCMC algorithm to calculate posterior probabilities. Power is estimated as the proportion of the 5000 trials that meet a primary analysis of superiority. Using the same simulated data, analyses based on MMRM were also conducted. The MMRM analysis model included participant-level random effects, a fixed effect for the cognitive score at baseline for placebo and treated participants, fixed effects for time-varying rates of decline for placebo participants (where time is included as years since baseline and is used as categorical), and fixed effects for time-varying treatment effects for treatment participants. The treatment effect at year 4 is tested by contrasting the group difference. The power using MMRM is estimated as the proportion of the 5000 trials with $P$-value less than .05 for a 2-sided test. Simulations of the DPM were conducted by using Fortran and the MMRM analysis conducted using R.

### 2.8 Trial simulation results

Figure 2 presents the power using the cognitive DPM as a function of the assumed treatment effect with analyses at the second, third, and fourth years. Additionally, the power of the MMRM at the fourth year for the exact same assumptions is shown. The cognitive DPM yields a substantial increase in power compared with the traditional MMRM (Figure 2). For an assumed 40% reduction in the cognitive decline, the cognitive DPM provides 91.1% power whereas the MMRM only yield 26.7% power at the fourth year analysis. Importantly, we show that even for therapies with greater than 50%
effect on slowing disease progression, using an MMRM approach still yields less than 50% power to detect this effect at the fourth year. Even the second and third year analyses using cognitive DPM yield more power than MMRM when the reduction in cognitive decline is more than 30%.

The 3 main reasons for this substantial increase are as follows:

1. Positioning a participant in the model based on their EYO rather than change from baseline
2. Allowing the extended follow-up to contribute to the primary analysis in a very strong way
3. The assumption of a common proportional effect across time for the treatment arm rather than fitting the effect for a single visit

To understand the role that the first of these plays consider a participant coming in at an EYO of −14. The expected decline is very different from a participant coming in at an EYO of +1. The MMRM analysis characterizes the observations as time from baseline, while the cognitive DPM treats them as the progression relative to their stage of disease (EYO). The estimated SD for a single cognitive test using the cognitive DPM is 0.333. If we use the exact same simulation of virtual participants and calculate the SD in the 4-year change from baseline values, it is 0.85. As a heuristic argument, to create the same standard error in the change from baseline at 4 years (0.85/√n), one would need to enroll 6.52 participants for the MMRM relative to 1 for the cognitive DPM (a SD of 0.33).

We have simulated a wide range of different scenarios around the assumption of the natural decline, variability, accrual rates, and drop-out rates and displayed the robustness of the model to these changes in assumptions. In all the scenarios we investigated, the DPM led to large increases in power compared with the typical MMRM model (results not shown).

3 | DISCUSSION

Having research participants that are before onset, yet are destined to develop AD, and at a very predictable time, while tragic, provides an incredible scientific opportunity for developing disease-modifying therapies. While the ability to characterize those that will get AD allows studying the disease before onset, the ability to characterize the time of onset and the decline rate before and after onset makes an enormous difference in the ability to learn from prevention trials. We present a flexible cognitive DPM that provides a strong characterization of the cognitive decline of ADAD mutation carriers in the DIAN observational cohort. The model demonstrates the highly consistent behavior of cognitive decline and has been extended to estimate effects of potential disease modifying treatment effects. The model provides tremendous power increases—for example, it allows detection of a potential treatment effect with 80 ADAD patients compared with what normally would take more than 400 ADAD patients. We show that the cognitive DPM accounts for the heterogeneity between trial participants and efficiently uses the outcome assessment in the extended follow-up, and
thus reduces the required sample size to achieve sufficient power for clinical trials compared with common alternative approaches. In addition, the DPM allows for characterizing a clinically relevant estimate of effect size.

The cognitive DPM is developed based on a particular cognitive composite, but it is not restricted to this unique composite. Any composite or any single cognitive measurement can potentially be used in a similar model. A similar DPM could be used on a variety of endpoints as long as their behavior exhibits progressive decline. Furthermore, the DPM could be extended to jointly model multiple outcomes by assuming individual random effects for each outcome to share a multivariate normal distribution with a common age of onset and by using the same proportion for the treatment effect.

The model provides additional inferential strength based on the extended length of follow-up of some participants beyond 4 years based on the assumption of a proportional treatment effect. This is a desirable aspect of analyzing a progressive disease—the most valuable information is from those participants with the longest exposure. Many alternative analysis models, like the MMRM, do not increase their inferential strength when participants have extended follow-up. Using a proportional treatment effect to the rate of decline and using extended follow-up can be used in sporadic AD trials. For example, the change since baseline in ADAS-cog 11 and in ADAS-cog 14 in mild AD in EXPEDITION1 and EXPEDITION2 trials and the decline over time in standardized MMSE and in Bristol Activities of Daily Living Scale in the trial for donepezil and memantine indicated approximately proportional treatment effects from baseline to the end of study. But, without a reliable estimate of the age of onset in the preclinical stage, the ability to use a DPM with time based on EYO remains elusive in sporadic AD trials. Methods that can be potentially used to find an estimated age of onset for sporadic AD have been proposed.

The delayed-start design (also referred to as the staggering-start design) has been considered in designs to demonstrate disease-modifying treatment effects. The cognitive DPM can be easily modified to analyze a treatment effect in a delayed start design by incorporating the starting time of each treatment as a function of the delayed start.

Compared with other well-established mixed effects models like the MMRM using time since baseline as a categorical variable, our model uses a stronger assumption of the proportional treatment effect, but this assumption can be relaxed by partitioning participants into different stages and then using multiple proportional treatment effects. Another assumption in our model is that we assumed a monotonic decline in the cognition over time. Although some individuals may violate this assumption in short random fluctuations, the monotonic decline has clear face validity in this population as there is essentially complete penetrance of these mutations leading to an inevitable cognitive decline. We have extended the model to allow parameters of “a learning curve” on the early visits after treatment. This allows the model to estimate a bump that may occur based on learning how to do well on the cognitive tests through practice.

ACKNOWLEDGEMENT

The DIAN observational study is supported by NIA grant U19 AG032438. We gratefully acknowledge the altruism of the participants and their families. The DIAN-TU trial is supported by NIH grants U01AG042791, R01AG146179, and R56AG053267. DIAN-TU Pharma Consortium has supports from AbbVie, Amgen, AstraZeneca, Biogen, Eisai, Janssen, Eli Lilly & Company/Avid Radiopharmaceuticals, Hoffman La-Roche/Genentech, Pfizer, and Sanofi. We are also grateful to GHR Foundation and Alzheimer’s Association for their continuous support and advocacy of the DIAN study and the DIAN-TU trial.

ORCID

Guoqiao Wang  http://orcid.org/0000-0002-6808-8737
Melanie Quintana  http://orcid.org/0000-0002-1763-2704

REFERENCES

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

The DIAN-TU Next Generation Alzheimer’s prevention trial: Adaptive design and disease progression model


*Department of Neurology, Washington University in St Louis, St Louis, MO, USA
bDepartment of Radiology, Washington University in St Louis, St Louis, MO, USA
cBerry Consultants, Austin, TX, USA
dIndiana Alzheimer Disease Center, Indiana University, Indianapolis, IN, USA
eMemory and Aging Program, Butler Hospital, Brown Medical School, Providence, RI, USA
fAlzheimer’s Disease Research Center, Keck School of Medicine of the University of Southern California, Los Angeles, CA, USA
gDepartment of Biostatistics, Washington University in St Louis, St Louis, MO, USA

Abstract

Introduction: The Dominantly Inherited Alzheimer Network Trials Unit (DIAN-TU) trial is an adaptive platform trial testing multiple drugs to slow or prevent the progression of Alzheimer’s disease in autosomal dominant Alzheimer’s disease (ADAD) families. With completion of enrollment of the first two drug arms, the DIAN-TU now plans to add new drugs to the platform, designated as the Next Generation (NexGen) prevention trial.

Methods: In collaboration with ADAD families, philanthropic organizations, academic leaders, the DIAN-TU Pharma Consortium, the National Institutes of Health, and regulatory colleagues, the DIAN-TU developed innovative clinical study designs for the DIAN-TU NexGen prevention trial.

Results: Our expanded trial toolbox consists of a disease progression model for ADAD, primary end point DIAN-TU cognitive performance composite, biomarker development, self-administered cognitive assessments, adaptive dose adjustments, and blinded data collection through the last participant completion.

Conclusion: These steps represent elements to improve efficacy of the adaptive platform trial and a continued effort to optimize prevention and treatment trials in ADAD.

© 2016 the Alzheimer’s Association. Published by Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Alzheimer’s prevention trial; Adaptive clinical trial; Biomarkers; Disease progression model; Cognitive composite; Dose adjustment; DIAN-TU; Amyloid; Tau; Autosomal dominant Alzheimer’s disease

1. Introduction

Alzheimer’s disease (AD) is a growing public and financial healthcare crisis. AD afflicts >5 million people in the United States (US), with an expected increase to 13.8 million by the year 2050 [1]. The costs for care of patients with AD and other dementias in 2015 was $226 billion, which is
predicted to increase beyond a trillion dollar annual cost by 2050 unless disease-modifying treatments are developed. Because of the severity and increasing prevalence of the disease, better treatments and prevention are urgently needed.

Development of highly effective AD treatments has been hampered by the lack of surrogate biomarkers, slow course of cognitive and clinical decline, variability in clinical phenotype, and variability when measuring cognition and functional impairments in AD. These challenges are especially difficult because validated diagnosis of AD in the absence of symptoms has not yet been achieved. Prevention trials thus must be long and large because of inability to predict if and when cognitive symptoms will start, lack of functional impairment, and difficulties identifying an at-risk study population before symptomatic cognitive decline.

To address key AD prevention trial design challenges in a population almost certain to develop AD—the autosomal dominant AD (ADAD) population—a public-private partnership was formed in 2010 to facilitate collaborative discussions among industry (the Dominantly Inherited Alzheimer Network Trials Unit [DIAN-TU] Pharma Consortium, dian-tu.wustl.edu/en/pharma-consortium-members), ADAD families, the Alzheimer’s Association, US National Institute on Aging (NIA), regulators including US Food and Drug Administration and European Medicines Agency, and researchers (DIAN observational study, Alzheimer’s Prevention Initiative, Anti-Amyloid in Asymptomatic Alzheimer’s disease [A4], Collaboration for Alzheimer’s Prevention, and others). Through these discussions, the DIAN-TU designed and executed a pioneering prevention trial for the ADAD population. The DIAN-TU (DIAN-TU-001, www.clinicaltrials.gov) is also a public-private partnership supported by Eli Lilly, Roche, Alzheimer’s Association, NIA, Avid Radiopharmaceuticals, GHR Foundation, Cogstate, Bracket, and Anonymous Foundation. Because the ADAD population is almost certain to develop cognitive impairment at a younger and predictable age, several of the challenges inherent in conducting prevention trials of late-onset, sporadic AD (SAD) are mitigated.

1.1. Autosomal dominant Alzheimer’s disease

ADAD is caused by mutations in the APP, PSEN1, or PSEN2 genes, which lead to early-onset AD. Because of the almost certain risk of developing AD and the predictability of the age at symptom onset [2,3], ADAD represents a uniquely informative population for clinical trials. Unlike SAD, the clinical onset of symptoms can be predicted throughout the lifespan, allowing drug trials to start years or even decades before symptoms occur [4]. Potential therapeutic drugs were developed using cellular and animal models of identified ADAD mutations [5–8]. These preclinical studies led to the development of agents used in anti-Aβ drug trials reaching clinical phase studies [9–15]. Whereas some remain in phase III trials, several have been discontinued because of adverse events (AEs) [16], worsened outcomes [17,18], or lack of benefit [16,19]. Of concern, many of these trials failed to show target engagement of the proposed drug, largely because of lack of studies that adequately incorporated biomarkers [14,15]. Results from recent trials involving anti-Aβ antibodies (e.g., gantenerumab, crenezumab, solanezumab, and aducanumab) have suggested that a greater magnitude of target engagement is critical and that treatment instituted earlier in the disease is possible and necessary for cognitive benefit [14,20–23].

For families known to carry ADAD mutations, prevention and treatment trials using novel therapies offer the potential to delay or even prevent dementia in asymptomatic individuals and improve mild symptoms or mitigate symptom progression. Because of the younger age, higher capability of central nervous system (CNS) repair, earlier stage of disease, and Aβ as a likely initiator of ADAD, the ADAD population is expected to provide an efficient and, compared with SAD, relatively rapid test for anti-Aβ therapeutics. The relative homogeneity of the DIAN-TU population, fewer comorbidities, low attrition to date, and the success in recruitment and completion of trial measures make the DIAN-TU and this patient population ideal for testing disease-modifying therapies.

1.2. Relationship to SAD

Changes in cognitive, clinical, biochemical, and structural measures in ADAD are similar to SAD [2,24,25]. This suggests a common pathophysiology and provides rationale that treatment trials in ADAD are likely to provide insight into common treatments with the much more prevalent SAD. Prior studies of statins in familial hyperlipidemia [26] demonstrated a remarkable normalization of cholesterol deposits, which predicted improvements of lifespan by treatment of hypercholesterolemia in the general population by several decades [27]. This is a historical precedent to what may be possible in ADAD and its relationship to SAD.

Three decades of accumulated research evidence supports the hypothesis that alterations in Aβ metabolism are necessary to cause AD [28,29]. In both SAD and ADAD, the presence of β-amyloidosis has been associated with cognitive decline in symptomatic individuals [30–32]. In ADAD, cerebral amyloidosis correlates with worse baseline performance on multiple cognitive composites and predicts greater decline over time in global cognition, working memory, and Mini-Mental State Examination (MMSE) in symptomatic mutation carriers [32]. Recent data from the DIAN observational study suggest that even in asymptomatic mutation carriers, the ratio of total tau to CSF Aβ1–42 is associated with longitudinal declines in episodic memory, language, and global cognition [33]. Similar results have been seen outside ADAD populations, in which Aβ-positive healthy older adults show decline in episodic memory compared with Aβ-negative healthy older
adults [30,31]. In contrast, an Icelandic mutation that reduces amyloidogenic processing of APP and also mildly decreases Aβ aggregation is protective against AD and age-related cognitive decline [34,35]. Together these data strengthen support for the amyloid hypothesis and the value of finding therapeutics targeting Aβ to treat both ADAD and SAD.

Regarding tau pathology, there is evidence in preclinical and clinical studies supporting a temporal relationship between Aβ plaque deposition and neurofibrillary tangle (NFT) development specifically in AD such that the development of Aβ plaques contributes to the subsequent spread of NFTs outside the entorhinal cortex and hippocampus [36–38]. In the aging process, NFTs can be seen starting in the fourth decade and appear to primarily develop in the entorhinal cortex and hippocampal areas [39] before the development of significant Aβ plaque. In the aging brain, the spread of NFT pathology appears to be limited in the absence of coexisting Aβ pathology [40–42]. Based on autopsy series of patients, the distribution of NFT pathology appears similar between ADAD and SAD [43]. Furthermore, recent studies in the DIAN have demonstrated similar distributions of tau positron emission tomography (PET) signals in ADAD and SAD [44]. Thus, the distribution and type of tau pathology among ADAD and SAD are analogous. Ongoing studies will clarify the role of aging in NFT development and further characterize tau pathology associated with cognitive and clinical impairment.

In SAD, a dynamic sequence of biomarker changes has been proposed [24,45] with cerebral amyloidosis being one of the first detectable changes, followed by measures of neurodegeneration and lastly cognitive and functional decline. Longitudinal and cross-sectional studies in ADAD [4,46,47] have supported the same relationship. Although further studies are required in both populations, the temporal order and progression of pathophysiological, cognitive, and clinical changes are shared in ADAD and SAD [3].

2. The DIAN-TU Alzheimer’s disease prevention trial platform

The DIAN-TU platform (NCT01760005, see Fig. 1) was launched to accelerate identification of effective drugs for prevention and treatment of ADAD [48,49]. In 2012, a phase II/III double-blind, randomized, pooled placebo-controlled 2-year biomarker trial began by testing two drugs—solanezumab (an antisoluble Aβ antibody) and gantenerumab (an antifibrillar Aβ antibody). The study transitioned within the DIAN-TU platform to the Adaptive Prevention Trial in 2014 as a 4-year phase III cognitive end point trial to determine whether Aβ antibody administration demonstrating CNS biomarker target engagement is able to prevent cognitive decline in cognitively normal individuals. The DIAN-TU is now operational in 7 countries (Australia, Canada, France, Italy, Spain, UK, and USA), 26 sites, and 4 languages (English, French, Italian, and Spanish) (see Fig. 2). Full enrollment of the first two drug arms has been achieved with a high assessment completion rate and low attrition. The DIAN-TU is one of five ongoing trials aiming to prevent or delay the onset and progression of AD [50].

![Fig. 1. Existing and proposed structure of the Dominantly Inherited Alzheimer Network Trials Unit (DIAN-TU) trial platform including the Next Generation (NexGen) prevention trial. Pink indicates enrollment of the first two arms, pink and green the biomarker phase, and green graduation to a cognitive end point. In this article, we describe design considerations for the two NexGen arms including proposed interim analyses based on biomarkers and performance on the DIAN-TU cognitive composite.](image-url)
2.1. DIAN-TU Next Generation

As enrollment for the first two arms was completed in late 2015, the DIAN-TU is preparing to launch two new drug arms as the Next Generation (NexGen) prevention trial (see Fig. 1). The NexGen prevention trial will be a multi-center, double-blind, randomized, pooled placebo-controlled, 4-year cognitive composite end point registration study of two potential disease-modifying therapies in 160 mutation carriers at risk for or with mild symptomatic ADAD. The NexGen prevention trial proposes to test the ability of a beta-secretase inhibitor (BACEi), gamma-secretase modulator, and other novel Aβ or tau-based therapies to slow or prevent cognitive decline in asymptomatic (clinical dementia rating [CDR] score 0) to mildly symptomatic (CDR 0.5 or 1) ADAD mutation carriers in the range of -15 to +10 years with respect to estimated years from symptom onset (EYO). EYO was calculated as the age of the participants at the time of the study assessment minus the mean mutation age of onset. For example, if a participant’s age was 45 years, the mean mutation age of onset for this participant was 50, then EYO would be -5. The DIAN-TU anticipates future development of AD-specific pathophysiology interventions that may include combination treatments. Each drug program will include at least 80 mutation carrier subjects who receive either drug or placebo randomized 3:1. Cognitive benefit will be determined over 4 plus years with the primary outcome being the DIAN-TU cognitive composite and secondary outcomes including multiple cognitive and clinical measures and cerebrospinal fluid and imaging biomarkers.

The NexGen prevention trial goals are to enable rapid testing of novel approaches, inform if dementia can be prevented or delayed in ADAD subjects, provide publicly accessible data and biological samples for research, and develop surrogate biomarkers to accelerate future AD trials. NexGen working groups evaluated additional designs that allow the platform to test drugs with diverse mechanisms of action more quickly. Significant innovations that may be implemented include an ADAD disease progression model (DPM) to detect changes in cognition with fewer participants compared with traditional mixed-effects model for repeated measures (MMRM), self-administered cognitive testing, a predefined dose escalation algorithm to safely maximize target engagement, adaptive trial design that includes both early biomarker and later cognitive interim analyses to inform early efficacy or futility, and novel imaging (e.g., tau PET and diffusion basis spectrum imaging [DBSI] magnetic resonance imaging [MRI]).

3. Methods and results

3.1. Primary end point—DIAN-TU cognitive composite

Utilizing the latest DIAN observational data limited to participants meeting DIAN-TU eligibility criteria, we developed a cognitive composite using sensitive and reliable tests that together represent cognitive domains affected in AD—particularly in the very early stages—including episodic memory, executive functioning, processing speed, and mental status. The DIAN-TU cognitive composite consists of the delayed recall score from the International Shopping List Test, the Logical Memory delayed recall score from the Wechsler Memory Scale—Revised, the Digit Symbol Coding test total score from the Wechsler Adult Intelligence Scale—Revised, and the MMSE total score. These measures
were selected based on their psychometric characteristics (reduced ceiling and floor effects and relatively low variability), sensitivity to subtle declines before clinical diagnosis, and face validity as indicators of the cognitive phenotype of AD. Because the DIAN-TU trial is a multisite international study, translations were necessary. Translation of the test stimuli for the International Shopping List Test followed procedures described in Lim et al. [51], which selects stimulus words representing commonly available items in the language and culture of interest. A similar procedure was followed for translation of the paragraph used in the Wechsler Memory Scale (Revised) Logical Memory test, wherein translations represent an adaptation of the content rather than a word-by-word translation [52]. This process relies on multiple reviews and revisions during translations with native speakers with the ultimate goal of achieving conceptual and cross-cultural equivalence. Power analyses indicate that the DIAN-TU composite is sensitive to decline and produces feasible sample size requirements to detect prescribed effect sizes.

We have vetted our composite by taking advantage of the longitudinal database from the DIAN observational study cohort (see Fig. 3). These analyses have shown that the proposed DIAN-TU cognitive composite can detect declines up to 10 years before estimated age of symptom onset. The composite has good distributional characteristics across the trial range of EYO and good separation of mutation carrier and noncarrier years before symptom onset. The DIAN-TU cognitive composite is a valid indicator of the cognitive phenotype of AD and is directly comparable with cognitive screen (MMSE) [21].

### 3.2. Primary analysis—ADAD DPM

Traditional trial designs in AD have focused on comparing the absolute change from baseline in a cognitive measure at a fixed time post randomization. This creates considerable variability because patients vary in their stage of disease at study entry and in their rates of decline. Historically, staging disease in SAD populations has been challenging because of heterogeneity at baseline and in the longitudinal trajectories of patients enrolled in therapeutic trials. Additionally, analyses focused on absolute changes at a predefined time point only weakly incorporate the longitudinal changes in progression. This results in trials with low power to demonstrate clinically important effects and increases the likelihood of type II error. Moreover, whereas clinically significant change may be clearly measurable in mild-moderate AD, preclinical or very mild stages will require mechanisms to detect small changes in outcomes that have meaningful downstream impacts.

The DIAN observational study has identified highly predictable and consistent changes that occur decades before onset of clinical symptoms [4], and we have used this valuable information to create a DPM that dramatically improves our ability to detect drug effects in DIAN-TU trials. In subjects who are not yet symptomatic, we can estimate accurately and reliably the number of years until symptom onset [3]. With longitudinal follow-up, we are continuously able to gather a more comprehensive understanding of the rate of cognitive decline across the spectrum of EYO and how these declines manifest across different cognitive domains. One important finding from the analyses of the longitudinal data from the DIAN observational study is that the disease progression in the cognitive composite in mutation carriers is nonlinear (Fig. 4, left panel). Therefore, we have developed a cognitive end point nonlinear mixed-effects model with EYO as the measure of time for each subject, not time from randomization (Wang Q, Berry S, Quintana M, Xiong C, McDade E and Bateman R, manuscript in preparation). The model estimates the expected rate of cognitive decline for untreated mutation carriers. By incorporating two important random effects for each individual—a subject’s EYO and a subject’s baseline cognitive performance—the rate of decline becomes more homogeneous and with much less variability (Fig. 4, right panel). Within this DPM, we develop a “proportional” disease modification parameter for treatments relative to placebo subjects. This parameter is identified more precisely by the entire longitudinal data for treated and placebo subjects.

Whereas we assume that the expected proportion of the change in the cognitive measure under a treatment relative
to placebo is constant across EYO, the expected absolute difference in the cognitive readout for drug treatment relative to placebo will vary depending on the EYO of the subject and the duration of treatment. Fig. 5 provides an illustration of hypothetical proportional changes to the disease progression and the gradual increase in the absolute effect for a range of EYO entry points based on a treatment effect over the 4-year duration of the trial. The percent slowing of the progression rate in Fig. 5 ranges from 0% to 70% encompassing reported cognitive and clinical effects of Aβ immunotherapies; 34% reduction in cognitive decline for solanezumab [22] and 66%–75% reduction in MMSE and CDR for aducanumab [21]. Importantly, the estimated standard deviation of a single visit in the DIAN-TU cognitive composite is 0.33 (estimate of the DPM fit to the observational cohort from data freeze 9). Thus, a change in the DIAN-TU cognitive composite over time of 0.25 creates a statistical “effect” size of 0.25/0.33 ≈ 0.756. For a 1-unit change in the cognitive composite, this corresponds to a statistical effect size of 1.0/0.33 ≈ 3. Thus, effect sizes of 0.756 for asymptomatic and 3 for symptomatic are quite large in the DIAN population. For a drug that causes 50% slowing of the rate of decline, treatment effect sizes of 0.756/2 ≈ 0.378 and 3/2 = 1.5 are statistically achievable with sample sizes from 80 to 140 subjects with 4 plus years of follow-up.

Importantly, the DPM provides for very high power to determine whether a treatment slows the rate of decline in a randomized trial. The DPM shows dramatically increased power compared with an MMRM that has weak measures of the treatment effect over time and a weak allowance for capturing the influential covariate, EYO. For example, with 60 mutation carrier treatment subjects and 40 mutation carrier placebo participants enrolled, the power to determine superiority when there is a 30% slowing of disease progression is 0.24 for the MMRM and 0.90 for the proposed DPM. Additionally, the precision of this model allows more timely decisions of superiority and futility for treatments with clear benefit or minimal to no benefit.
3.3. DIAN-TU NexGen secondary outcome measures

A wide range of secondary analyses will be conducted, including efficacy analysis using MMRM to compare with the DPM and analyses of additional individual cognitive and clinical tests. Modeling for both the primary and secondary analyses will adjust for well-established risk factors and covariates such as baseline age, APOE e4 status, education, and gene mutation type. Analyses will be done on functional measures of activities of daily living, all biomarker measures (e.g., CSF Aβ1–40, Aβ1–42, tau and p-tau181, regional amyloid PET, volumetric MRI, and tau PET), and on the incidence of CDR conversion from baseline.

3.3.1. Exploratory self-administered Internet-based cognitive assessments

It is expected that the reliability of the observations of the subtle changes over time increases with the number of measurements used to estimate that change. An exploratory aim for secondary analysis in the NexGen prevention trial will be to introduce self-administered brief cognitive assessments using a validated Internet-based testing portal. The self-administered battery will include the computerized measures that comprise the 12-minute Cogstate Brief Battery, which is a component of the larger cognitive assessment battery currently in use in DIAN-TU. The overall goal is to collect more frequent cognitive assessments to test the hypothesis that this will increase reliability of the measurement of subtle cognitive change without substantially increasing study costs and burden for DIAN-TU participants. To our knowledge, this is the first trial design to incorporate self-administered assessments in a clinical trial investigating an AD disease-modifying agent.

3.3.2. Biomarker effects of DIAN-TU drugs

Convergence and consensus in the field has generated significant support for biomarkers in AD prevention trials [2,53]. The DIAN observational study was established as the first international registry and longitudinal study of ADAD to determine the temporal course of biomarker changes relative to AD symptom onset and to support clinical trials in this largely asymptomatic population. Discovery and validation of AD biomarkers for identifying individuals most likely to respond to treatments, measuring responses to treatments, and predicting clinical benefit of treatments (surrogate biomarkers) are a high priority to increase the efficiency of trials for AD.

3.3.2.1. Tracking target engagement

Drugs are developed to target pathologic proteins and processes (e.g., increased Aβ production and/or deposition) and the efficacy with which they engage the target can be assessed with biomarkers. The DIAN-TU NexGen prevention trial is testing two drugs in parallel to assess target engagement using primary biomarker readouts including amyloid PET imaging for anti-Aβ antibodies and CSF Aβ1–40 and Aβ1–42 for BACEi.

3.3.2.2. Finding a surrogate biomarker for AD clinical trials

From inception, the DIAN-TU has assessed many biomarkers. Should a drug demonstrate cognitive benefit, it will be possible to analyze which biomarkers best predicted therapeutic efficacy, findings of particular benefit for future AD trials. The need for predictive biomarkers is especially important in slowly progressive disorders like AD, without which prevention trials will necessarily be large, long, and expensive [54]. The inclusion of multiple biomarkers will allow for the analysis of multiple potential responses to therapies and also better reflect the array of biomarkers used to predict progression from the observational study.

As there is not yet an established predictive biomarker for AD, we have included a diverse panel of imaging and fluid AD biomarkers to determine drugs’ impact on different aspects of the pathobiology including atrophy (MRI), connectivity (functional MRI), metabolism (fluodeoxyglucose PET), amyloid pathology (amyloid PET), tau pathology (tau PET), vascular factors (MRI white matter), and biochemical changes (CSF). Novel CSF markers of synaptic dysfunction and neurodegeneration, imaging markers including tau PET, and MRI diffusion sequences to better understand amyloid-related imaging abnormalities (ARIAs) will be implemented. Exploratory biofluid and imaging analyses include measurement of CSF tau, phosphorylated tau, VILIP-1 (visinin-like protein 1, neuronal injury/degeneration) [55], and NGRN (neurogranin, [post-] synaptic dysfunction) [56–58]. Potential future exploratory biomarkers evaluated independently at trial completion include CSF synaptosomal-associated protein-25 (SNAP-25, [pre-] synaptic dysfunction) [59], neurofilament light chain (NFL, axonal injury) [60], and plasma tau (neuronal injury/degeneration) [61].

In particular, inclusion of tau PET imaging for the DIAN-TU NexGen drug arms will significantly enhance the trial and provide an opportunity to maximize understanding of the role of tau in AD. Inflection of CSF tau appears to predict the clinically critical cognitive decline in AD. Quantification of aggregated brain tau will be used as a secondary outcome measure to determine whether therapeutics lowering amyloid in presymptomatic and mild AD can decrease insoluble brain tau levels and whether treatment timing influences the outcome; quantify the rate of aggregated brain tau changes in placebo-treated ADAD participants in various stages of disease; test the hypothesis that aggregated tau levels can predict progression from asymptomatic to symptomatic AD; and analyze tau PET dynamics in the context of other imaging and CSF biomarkers in mutation carriers and noncarriers.
3.3.2.3. Biomarker analyses and understanding AEs

Extensive characterization of DIAN-TU participants may also elucidate mechanisms of ARIAs that have been increased in frequency by some amyloid therapeutics. The DIAN observational cohort has demonstrated normal brain MRIs until the time of transition to symptomatic AD, at which point atrophy, white matter disease, and microhemorrhages (MHEs) begin to manifest [62]. Prevalence of MHEs in the DIAN observational cohort is 15% overall with only 2.9% having greater than one MHE. The lack of other age- and disease-related comorbidities that are risk factors for MHEs in the DIAN population allows us to characterize ARIAs related to AD and to drug effects. To do so in greater detail, we will use DBSI, a novel form of diffusion-tensor imaging analysis, permitting separation of free water (edema) from cellular infiltrate more consistent with possible inflammatory processes [63,64].

3.3.2.4. A community resource

As the DIAN-TU is a public-private partnership with funding from the National Institutes of Health (NIH), data and samples will serve as a resource for researchers. The trial will generate a repository of CSF and blood samples, imaging data, and biomarker results providing a resource for qualified investigators to help develop novel and better-validated AD clinical trial biomarkers.

3.4. Dose-adjustment algorithm

As a critical component of the adaptive platform trial design, combined safety/biomarker-target analyses at years 1 and 2 will enable dose adjustment to maximize biomarker target engagement. Although dose optimization is conventionally performed during phase I/II studies, we propose an opportunity for dose adjustment in DIAN-TU NexGen for multiple reasons. First, doses identified in phase I/II studies of amyloid specific agents have not consistently translated to target engagement and efficacy in larger, later phase studies [23,65]. Second, with early evidence of biomarker target engagement, we ensure that decisions are made quickly and thereby avoid the mistake of identifying a subtherapeutic dose after 4 years, preventing the loss of precious time and resources. Third, the current DIAN-TU population has had few drug-related AEs indicating higher doses may be tolerated in this younger population. By starting with or progressing to higher doses in NexGen and then performing an early safety/biomarker analysis at year 1, we will be well positioned to assess the appropriate dose in the ADAD population.

A dose-escalation algorithm would account for safety signals and biomarker target engagement at interim analyses of both NexGen drugs, with precise cutoffs and safety profiles tailored to each compound based on preclinical and clinical data. After 1 year of drug and placebo exposure, a drug not meeting a minimal acceptable change in CSF Aβ or decrease in amyloid PET accumulation compared with baseline may be offered an opportunity to increase dose (provided there is also an acceptable safety profile). A clinical safety and biomarker interim analysis by an unblinded team will determine whether drug dosage should increase, decrease, continue at current dose, or be dropped because of futility of low biomarker engagement.

3.5. Interim analyses

Because of significant power gains from the ADAD DPM, it may be possible to declare early success in the NexGen prevention trial. Thus, interim analyses of cognitive outcomes are proposed when the last enrolled participant in a treatment cohort reaches 2 and 3 years of follow-up. At each interim, a regimen may be stopped for futility if the experimental therapy demonstrates a lack of efficacy (probability of at least a 20% slowing of cognitive decline is <5%) or stopped early because of demonstration of cognitive efficacy (the primary analysis using only the interim data showing statistically superior slowing of progression). If neither of these decisions is reached, then the regimen will continue to the next interim analysis or the final analysis when the last participant has completed 4 years of treatment. A detailed stopping boundary for the interim efficacy analyses will be developed.

Biomarker interim analysis at year 1 will examine change from baseline for a drug’s primary biomarker of target engagement. For drug programs that have a dose escalation at year 1, a year 2 evaluation of safety and biomarker engagement will be performed similar to year 1. However, dose-escalated drug programs that still have not safely demonstrated high impact on their biomarker of CNS target engagement at year 2 will be discontinued.

3.6. Extended randomized follow-up

Because of the extended periods of monitoring needed for cognitive change in prevention trials, any additional data on cognitive changes because of drug effects are highly valuable in determining the potential for drugs to slow or delay cognitive loss. We simulated collecting and analyzing all randomized data collected in the DIAN-TU trial from first patient in to last patient visit by having all participants continue randomized treatment until trial completion. Depending on the assumptions of trial enrollment duration and effect size, various improvements in power ranged from 1% to 30%. For example, a trial with a drug showing 30% slowing of disease progression (60 active to 40 placebo in mutation carrier subjects) has a 60% chance of demonstrating efficacy with exactly 4 years of follow-up, whereas the same scenario allowing randomized participants to continue until the last subject reaches 4 years has 90% power. The valuable data collected beyond the 4-year point provides highly informative measures to better estimate...
drug effects. These improvements in power take no additional time for the trial and can be implemented with modest costs of follow-up.

3.7. NexGen drug candidates and combination therapy

Diversifying the drug portfolio has been a longstanding DIAN-TU platform goal to increase the likelihood of finding an effective therapy and mitigating the inherent risk of individual drug failure because of unpredicted toxicity, lack of availability, or lack of efficacy. Two therapeutics will be chosen from the DIAN-TU Therapy Evaluation Committee imminent use category and are likely to include a class to target Aβ production. Future therapeutics tested in the DIAN-TU platform may include tau-based treatments and other non-Aβ targets.

An original goal of NexGen prevention trial design included testing of combination therapy to delay onset and progression of AD. Preclinical data suggest that combining therapeutics targeting the amyloid cascade (anti-Aβ and BACEi) significantly enhances reduction of amyloid load and plaque number beyond either monotherapy alone [66,67]. The existing and NexGen arms will be testing individual therapeutics that are a possible first combination [68–71], and the DIAN-TU platform’s multiple monotherapies and pooled placebo allow an opportunity to test combinations alongside monotherapy. Moreover, the platform provides a unique opportunity for companies to work together minimizing complicated licensing and contracts. However, as yet there has been no disease-modifying combination nominated for consideration by the DIAN-TU Therapy Evaluation Committee. Thus, combination therapy remains a desirable future aim of the DIAN-TU platform.

3.8. Engaging ADAD participants

There is a growing movement to include patient-centered outcomes in the drug development process. The DIAN-TU uses surveys [72], webinars, and an annual ADAD Family Conference (http://dian-tu.wustl.edu/en/adad-family-conference) to gather feedback on trial design and risk tolerance in the ADAD community. Pooled placebo and ability of individuals with undisclosed mutation status to participate in the trial—two central features of the original trial design—were incorporated as a direct result of patient engagement. Pooling data across placebo subjects from different drug arms provides participants with a 75% chance of receiving active drug rather than 50% in traditional 1:1 randomization. However, pooled placebo requires special attention to issues of equivalence and comparability. For example, sensitivity analyses are used to confirm that placebo groups across arms can be combined for the purpose of pooled placebo analyses. Further operational similarities across arms in the DIAN-TU support similarity in enrollment, assessments, and duration of the study. Inspired by ADAD families, the DIAN-TU NexGen design aims to test more drugs more quickly by decreasing required participant number using the ADAD DPM, declaring futility or success earlier, and maximizing dose and effect size via the dose-adjustment algorithm.

Because ADAD is a rare disease, we have implemented multiple outreach strategies to recruit and engage eligible participants through the DIAN observational study, DIAN Expanded Registry (DIAN EXR), and referrals from partnering clinicians and sites with >3500 potential participants identified. The DIAN-TU EXR offers exploratory genetic counseling and exome sequencing for families with early-onset AD but without confirmation of an ADAD mutation. It is possible that through the EXR, meaningful information such as remote cognitive testing could be captured. Similar programs such as the Brain Health Registry, www.brainhealthregistry.org, are underway in the aging population. However, no such registry exists in ADAD. Additionally, with the platform design of the DIAN-TU, there will be times enrollment is on hold. During these periods, the EXR could continue to capture cognitive assessments and other data that might be helpful at the start of new drug arms such as those proposed in the NexGen prevention trial.

4. Discussion

A therapy delaying the onset of AD dementia by 5 years that is introduced by 2025 would reduce the number of expected cases by 42% by 2050 and save the US $367 billion annually on costs of care [73]. Although a treatment advance in the ADAD population would have to be tested in SAD, we believe the data that support DIAN-TU outcomes may be utilized as supportive or as one of two pivotal trials to demonstrate effectiveness in SAD. A full evaluation of the comparability between ADAD and SAD is underway with large ongoing observational studies in Alzheimer’s Disease Neuroimaging Initiative, National Alzheimer’s Coordinating Center, and DIAN, which will help inform about the translatability between ADAD and SAD.

In support of the US National Alzheimer’s Project Act plan of finding an effective disease-modifying therapy by 2025, the DIAN-TU NexGen prevention trial design was developed through collaborations with DIAN-TU NexGen and Pharma Consortium members, DIAN and DIAN-TU subjects, patient groups (Alzheimer’s Association and ADAD Family Forum), and academic, clinical trial, and statistics experts. Central themes emerged including a focus on the adaptive trial platform to minimize cycle time, ensure maximal learning from the limited population of ADAD participants, and increase the chance of finding a treatment that provides cognitive benefit. The trial becomes stronger as more regimens are added, and regular interims with more stringent definitions of failure and success allow faster decision making. Substantial power gains are achieved by using the DPM. The NexGen prevention trial aims for a larger effect size by employing
a drug-specific dose-escalation algorithm to maximize target engagement, biomarker changes, and cognitive benefit over time. Finally, the trial will provide additional value to the Alzheimer’s field by investigating a comprehensive array of fluid and imaging biomarkers and cognitive tests. Inherent limitations include the possibility of small effects in a small sample size and as yet no clear relationship between biomarker changes and clinical efficacy.

The DIAN-TU NexGen prevention trial will benefit from time- and cost-saving efficiencies established by the existing DIAN-TU platform operations and infrastructure. We have accounted for the common risks of choosing a single unproven drug in a long-term prevention trial, subject enrollment and retention, selection of cognitive outcomes, and power to detect a clinically significant drug effect. If the DIAN-TU trial successfully identifies an effective therapeutic or surrogate biomarker, the impact on patients, societies, and economies may be substantial.

Acknowledgments

The NexGen prevention trial design work was supported by an Anonymous Foundation, the Alzheimer’s Association, GHR Foundation, and the DIAN-TU Pharma Consortium. The DIAN-TU is supported by the Alzheimer’s Association, Eli Lilly, Roche, NIH U01AG042791, NIH U01AG042791-S1, NIH R1A046179, Avid Radiopharmaceuticals, Cogstate, Bracket, and GHR Foundation. Data collection and sharing for this project was supported by The Dominantly Inherited Alzheimer Network (DIAN, U19AG032438) funded by the National Institute on Aging and the German Center for Neurodegenerative Diseases (DZNE). This manuscript has been reviewed by DIAN Study investigators for scientific content and consistency of data interpretation with previous DIAN Study publications. We acknowledge the altruism of the participants and their families and contributions of the DIAN and DIAN-TU research and support staff at each of the participating sites for their contributions to this study. The authors would also like to thank Jacki Mallmann and Amanda Fullbright for technical and administrative assistance.


RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed articles and books pertaining to clinical trial design and held numerous in-person and teleconference meetings with a variety of experts including clinicians, researchers, statisticians, patient advocates, and regulatory, clinical development, imaging, and informatics professionals.

2. Interpretation: This article outlines innovative trial design features that may be implemented in the Dominantly Inherited Alzheimer Network Trials Unit (DIAN-TU) Next Generation (NexGen) prevention trial and considered by others developing adaptive trials for autosomal dominant Alzheimer’s disease (ADAD) mutation carriers; the model fit provides a powerful way to simulate future subjects in the DIAN-TU trial and significant improvements in power over traditional analytical methods.

3. Future directions: Launch of the DIAN-TU NexGen prevention trial will enable testing of new therapeutics to determine if one or more is able to prevent or slow cognitive decline in the ADAD population.

References


Bayesian model of disease progression in GNE myopathy

M. Quintana | J. Shrader | C. Slota | G. Joe | J.C. McKew | M. Fitzgerald | W.A. Gahl | S. Berry | N. Carrillo

1 Berry Consultants, Austin, Texas
2 Department of Rehabilitation Medicine, Clinical Center, National Institutes of Health, Bethesda, Maryland
3 Therapeutics for Rare and Neglected Diseases Program, National Institutes of Health, Bethesda, Maryland
4 National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland
5 Kansas University Medical Center, Kansas City, Kansas
6 RTI Health Solutions, Durham, North Carolina

Correspondence
M. Quintana, Berry Consultants, Austin Texas.
Email: melanie@berryconsultants.net

Funding information
National Human Genome Research Institute (NHGRI); National Center for Advancing Translational Sciences (NCATS); NIH Clinical Center

One Sentence Summary: A Bayesian repeated measures model based on quantitative muscle strength data from a prospective Natural History Study was developed to determine disease progression and design clinical trials for GNE myopathy, a rare and slowly progressive muscle disease.

GNE myopathy is a rare muscle disease characterized by slowly progressive weakness and atrophy of skeletal muscles. To address the significant challenges of defining the natural history and designing clinical trials for GNE myopathy, we developed a Bayesian latent variable repeated measures model to determine disease progression. The model is based on longitudinal quantitative muscle strength data collected as part of a prospective Natural History Study. The GNE Myopathy Progression Model provides an understanding of disease progression that would have otherwise required a natural history of unfeasible duration. “Disease age,” the model-generated measure of disease progression, highly correlates with a variety of clinical, functional and patient-reported outcomes. With the incorporation of a treatment effect parameter to the GNE Disease Progression Model, we describe a novel GNE Myopathy Disease Modification Analysis that significantly increases power and reduces the number of subjects required to test the effectiveness of novel therapies when compared to more traditional analysis methods. The GNE Myopathy Disease Progression Model and Disease Modification Analysis can be applied to muscle diseases with prospectively collected muscle strength data, and a variety of rare and slowly progressive diseases.

KEYWORDS
Bayesian, clinical trial, disease progression model, GNE myopathy, muscle disease

1 | INTRODUCTION

GNE myopathy, a rare genetic muscle disease with an estimated prevalence of 6/1 000 000, is caused by biallelic mutations in GNE, the gene that encodes the rate-limiting enzyme of sialic acid biosynthesis. The underlying pathophysiology of GNE myopathy is thought to involve decreased sialylation of muscle glycoproteins. The disease affects skeletal muscle and typically presents in young adults (20-40 years of age) with foot drop, caused by weakness of the anterior tibialis muscle, a distal lower extremity muscle primarily responsible for ankle dorsiflexion. Over the course of subsequent years, the disease slowly progresses to cause atrophy and weakness of other skeletal muscle groups, with relative sparing of the...
quadriceps. It typically takes decades for patients to progress from early to late stages of the disease. This relentlessly progressive disease results in a significant loss of skeletal muscle leading to disability, wheelchair use, and eventually patients may become incapacitated and dependent on caregivers.

In general, the understanding of disease progression as measured by clinical outcome measures is critical for successful clinical trial development. Clinical trial simulators based on disease progression models have been utilized to make evidence-based decisions in clinical trial design in both Alzheimer's Disease and Parkinson's Disease. Furthermore, Bayesian disease progression models that measure the proportional slowing in the rate of progression of disease have also been shown to lead to substantial power gains of more traditional approaches and approved by regulatory agencies as the primary analysis method to study novel therapies in Alzheimer's disease.

The development of disease progression models is particularly relevant in rare disorders where no known therapy exists, patient resources are limited, and suitable outcome measures to quantify progression across the entire spectrum of the disease are difficult to identify. A lack of understanding of the rates of progression and poor selection of outcome measures can lead to trials that are underpowered and unable to provide definitive answers about the effectiveness of a novel therapy. In rare diseases, this wastes resources in an area of an unmet medical need burdened with limited patient resources and therapeutic development.

The understanding of the disease mechanism of GNE myopathy has led to the development of potential treatment strategies. However, because GNE myopathy is a rare and slowly progressive disease involving different muscles at different stages of the disease, the selection of appropriate primary endpoints applicable to an entire clinical cohort that could be used to determine the potential effectiveness of new therapies in clinical trials is challenging. A clinical endpoint, such as 6-minute walk test (6MWT), is only relevant for a very narrow range of subjects who have progressed enough to have difficulty walking but not to become nonambulatory.

Here, we present a disease progression model of GNE myopathy that will allow us to understand disease progression and make evidence-based decisions regarding the design of clinical trials and selection of key outcome measures. Specifically, we used quantitative muscle strength collected longitudinally on GNE myopathy patients (Section 2) to develop a Bayesian Disease Progression Model (DPM) (Section 3) that quantifies disease progression as it relates to observed muscle strength for six primary muscle groups. The DPM synthesizes the longitudinal decline of all subjects (each providing information on a brief segment of disease progression) by simultaneously estimating a latent “disease age” parameter for each subject, aligning the subjects based on their disease age, and jointly estimating the onset and rate of decline for each of the muscle groups. The subject-specific “disease age” alignment parameter and its correlation with other clinical and functional outcome measures, allows the characterization of the entire disease spectrum, as well as predictions of future progression for each subject, without conducting a Natural History Study (NHS) of unfeasible duration (Section 4). Furthermore, a single treatment effect parameter can be incorporated into the DPM to provide a powerful analysis tool to understand the possible disease progression effects of an experimental treatment, allowing the design of clinical trials with more power and fewer patients, compared to traditional endpoints and analyses (Section 5).

2 | MUSCLE STRENGTH DATA

2.1 | Subjects

We obtained muscle strength data from subjects who participated in a prospective single-center study “A Natural History Study of Patients with GNE Myopathy” (NIH study 11-HG-0218; ClinicalTrials.gov identifier: NCT01417533). In this paper, we report a progression model based on data collected before April 1, 2016, where there were 38 subjects ranging in age at baseline from 25 years old to 65 years old. The total follow-up across all 38 subjects is 44 years. Subjects range from having follow-up from 0 months (baseline only visit) to 35 months with an average of 14 months. Subjects range from having a single visit to a maximum of 6 visits with an average of 3 visits per subject.

2.2 | Quantitative muscle strength assessment

Muscle strength was evaluated in this study using the Quantitative Muscle Assessment (QMA—Aeverl Medical, Gainesville, GA). The QMA is a standardized system to measure peak muscle force during maximal voluntary isometric contraction and is a valid, reliable, reproducible and sensitive measure of muscle strength. Measurements included in the DPM were bilateral averaged ankle dorsiflexion, knee flexion, knee extension, grip, elbow flexion, and shoulder abduction. Since strength depends on several nondisease variables such as age, gender and size of the individual, muscle
strength was expressed as a proportion of the subject’s predicted normal values for their sex, age, and body mass index (BMI) based on regression equations of the National Isometric Muscle Strength Database Consortium and published literature. Proportion of predicted strength = (observed value in kg)/(predicted value in kg). A value of 1 indicates “normal” muscle strength based on sex, age, and BMI.

2.3 | Skeletal muscle decline in GNE myopathy

Figure 1A-F presents the proportion of predicted muscle strength for six different muscles as a function of the chronological age of subjects at the time of measurement. Different muscles are affected as a function of chronological age. Ankle dorsiflexion was below 50% of predicted in all patients and had reached full decline in the majority (Figure 1A). Knee extension was relatively preserved, with the majority of patients above 50% of predicted (Figure 1C). The decline of other muscles varied among subjects (Figure 1B, D-F), demonstrating that patients of a similar chronological age have different levels of disease progression. Disease duration, a patient-reported estimate to the number of years that a subject has been affected by the disease at the time of evaluation, showed a slightly better pattern of muscle decay (Figure S1). Based on this, we conclude that disease progression cannot be estimated based on chronological age or disease duration alone.

Figure 2A-D shows examples of longitudinal data for the proportion of strength of the primary muscle groups for 4 subjects. At baseline, Subject A was 30 years old and the proportion of dorsiflexion and knee flexion strength was already 0. In contrast, Subject D was 49 years at baseline, and while this subject’s dorsiflexion had completely declined, all other muscle groups were close to or above 1, meaning this subject had similar strength on those muscle groups as a healthy subject of same age, sex, and BMI. The comparison of Subject D to Subject A demonstrates that disease progression cannot be estimated based on chronological age alone. Subject D also shows that subject-level and muscle-level maximum strengths vary and need to be accounted for in our modeling.

3 | GNE MYOPATHY DISEASE PROGRESSION MODEL

A goal of the DPM is to align individuals using a single latent disease stage parameter, referred to as “disease age,” to characterize the systematic disease progression as it relates to muscle strength. To explore this concept, we addressed whether the estimation of this disease age parameter can be based on the observed strength of a single muscle. Throughout and without loss of generality, we assume that disease age is defined as 0 when the proportion of knee flexion strength is 0.50. Knee flexion was selected because subjects had a wide range of knee flexion muscle strength and it had the most defined decay pattern as a function of chronological age (Figure 1B). Figure S2B shows the evident decline of individual muscles as a function of knee-flexion–based disease age, suggesting that other muscles follow a particular sequence of involvement in GNE myopathy, with a systematic pattern of decline.

The above results demonstrate that muscle strength across the six primary muscle groups follows a systematic pattern of decline once aligned based on the observed strength of a single muscle (knee flexion). To provide more accurate alignment of subjects in terms of disease age, the DPM is a joint repeated measures mixed effects model that simultaneously models the observed muscle strength for six muscles as a function of a single subject-level latent disease age. The GNE DPM includes parameters for muscle-specific timing and rate of progression as a function of disease age as well as muscle-specific variability. The GNE DPM also includes subject-specific estimation of disease age as well as inherent muscle strength. Using the GNE NHS, we jointly analyze data on N subjects, with nᵢ visits/measurements for K total muscles. For each subject i, tᵢ is their chronological age in years at visit j and Yᵢ,j,k is the observed muscle strength for muscle k at visit j. The measured values are modeled jointly using the following repeated measures mixed effects model:

\[ Y_{i,j,k} \sim N \left( \mu_{i,j,k}; \sigma_{i,j,k}^2 + \delta^2 \right); \]

\[ \mu_{i,j,k} = \text{logit}^{-1} \left[ \theta_k + \beta_k \left( t_j - a_i \right) \right] \ast M_{i,k}; \]

\[ i = 1 \ldots N; j = 1 \ldots n_i; k = 1 \ldots K. \]

Within the GNE DPM, we assume that each observed subject and visit-specific muscle strength is distributed normally with a mean that is defined using a logistic decay function as well as a subject and muscle-specific random effect and a variance that is a function of the mean.
FIGURE 1 Muscle strength in subjects with GNE myopathy. (A-F) The proportion of predicted muscle strength for 6 muscle groups collected for all subjects and time points (y-axis) plotted as a function of the chronological age of subjects at the time of evaluation. Note the majority of subjects have completely absent ankle dorsiflexion, but there were subjects of both 30 and 60 years of age that had measurable ankle dorsiflexion strength. Example subjects A-D (see Figure 2) are labeled on each plot. (G-L) DPM fit plots for each muscle as a function of their estimated disease age at the time of each visit. For each subject, the proportion of predicted muscle strength is adjusted by the estimated subject- and muscle-specific inherent muscle strength, M. The posterior mean (solid black line) is shown. Red: proportion of inherent strength at disease age 0. Blue: disease the age at which the model-estimated proportion of predicted muscle strength is 0.50. Example subjects A-D (see Figure 2) are labeled on each plot.
FIGURE 2  Longitudinal muscle strength data and estimated disease progression for 4 GNE myopathy subjects. The muscle strength for 6 different muscle groups, ankle dorsiflexion (red), knee flexion (yellow), knee extension (gray), grip (green), shoulder abduction (blue), and elbow flexion (purple), is plotted for the 4 example subjects labeled on Figure 1. (A-D) Longitudinal strength is plotted as a function of age at each time point (A-D, circle). Subjects A-C have similar chronological ages but different degrees of disease progression. At baseline, Subject A (age 29) and Subject C (age 35) have completely declined on ankle dorsiflexion and knee flexion; in contrast, Subject B (age 29) is actively losing ankle dorsiflexion and knee flexion strength. (E-H) Estimated disease progression for the subjects above and the subject's disease age at baseline is noted. The posterior mean (dotted line), and the 95% predictive intervals (shaded region) are shown for a period of 10 years.

The logit of the mean divided by the subject and muscle-specific random effect, $M_{i,k}$, is assumed to be linear as a function of the latent subject-specific "disease age," $t_{ij} - a_i$, with a muscle-specific intercept/timing of decline, $\theta_k$, and muscle-specific rate of decline, $\beta_k$. The subject and muscle-specific random effect, $M_{i,k}$, is introduced to account for subject-level variation in individual inherent strengths. Without loss of generality, to constrain the model for identifiability, we assume $\theta_1 = 0$. This constraint implies that at chronological age $a_i$, a subject has a mean of 50% of their inherent strength for knee flexion (where $k = 1$ implies knee flexion). Thus, $a_i$ is interpreted as the chronological age for subject $i$ when their knee flexion is at 50% of their inherent strength.

The variance of the observed muscle strengths is assumed to be a function of an overall muscle-specific variance parameter, $\sigma^2_k$, to measure the within subject variability of muscle strength at a normal state as well as the subject, muscle and visit-specific mean. The mean/variance relationship is expected as muscle strength is lost and is documented in studies evaluating muscle strength as a function of normal aging. To allow for additional variance in the proportion of muscle strength once a muscle is completely declined ($\mu_{i,j,k} = 0$), a dispersion parameter is added, $\delta$, that is assumed to be constant across all muscle groups.
3.1 Prior specification

To complete the model specification, the following noninformative prior distributions are placed on the muscle-specific location and rate or decline parameters parameters:

\[ \theta_k \sim N(0, 2^2) ; \]
\[ \beta_k \sim N(-\infty, 0) \left( 0, .4^2 \right) . \]

For the muscle-specific location and slope parameters, normal prior distributions centered around zero for the location and truncated at zero for the slope (to allow for decay only) with standard deviations of 2 and .4 were assigned, respectively. The inverse logit of the muscle-specific location parameter, \( \logit^{-1}[\theta_k] \), corresponds to the expected muscle strength at disease age zero. A value of zero, as is assumed for knee flexion, would correspond to an expectation of 50% muscle strength at disease age zero. With a prior standard deviation of 2, there is 95% prior probability that the muscle strengths range from 2% to 98% at disease age zero and as such the prior encompasses a wide range of values and is noninformative.

Similarly, the prior for the muscle-specific slope parameter \( \beta_k \) results in 95% prior probability that muscle strength could decline anywhere from 0% to 19% within one year after muscle strength is at 50%. As such, in a slowly progressive disease where each muscle is likely to take over 10 years to fully decline, the prior encompasses a wide range of values and is noninformative.

The muscle-specific variance, \( \sigma_k^2 \), is assumed to have the following inverse gamma prior:

\[ \frac{1}{\sigma_k^2} \sim \text{Gamma}(11, 0.225) . \]

This prior has an expectation of .0225 and a weight of 10 subjects worth of information. The induced prior on the standard deviation has a mean of .15 and there is a 95% prior probability that the standard deviation is between .10 and .20.

The constant dispersion parameter, \( \delta \), is assumed to come from a gamma distribution:

\[ \delta \sim \text{Gamma}(2,100) \]

with prior expectation of approximately 0.02. The prior is assumed to be informative to only allow a small amount of variation when muscle strength has fully declined.

The latent subject-specific chronological age when knee flexion is 50% of inherent strength, \( a_i \), is assumed to be distributed normally (truncated at 0) with a mean of 30 (the expected age when knee flexion is 50%) and an inverse gamma hyperprior on the variance as follows:

\[ a_i \sim N(0,\infty)(30, \sigma_a^2) ; 1/\sigma_a^2 \sim \text{Gamma}(11,100) . \]

A hierarchical model for \( a_i \) was created to allow the variability of the chronological age when knee flexion is 50% across all subjects to be estimated from the data. Disease progression is assumed to be a function of the latent subject-specific disease age, \( t_{i,j} - a_i \), that is the chronological age at the time of the visit minus the latent model-estimated \( a_i \). If the variability of \( a_i \) across subjects is estimated to be vary low, this is an indication that chronological age is a good predictor of disease progression. The inverse gamma hyperprior distribution on the variance is noninformative with prior expectation of 10 and a weight of 10 subjects worth of information.

The muscle and subject-specific maximum inherent strength parameter, \( M_{i,k} \), is also assumed to have a hierarchical distribution that is centered on a common subject-specific mean, \( \omega_i \), as follows:

\[ M_{i,k} \sim N(0,\infty)(\omega_i, \sigma_M^2) ; 1/\sigma_M^2 \sim \text{Gamma}(101, 2.25) ; \]
\[ \omega_i \sim N(0,\infty)(1, \sigma_\omega^2) ; 1/\sigma_\omega^2 \sim \text{Gamma}(1001, 1) . \]

This hierarchical distribution allows the inherent muscle strength across muscles to be centered on a model-estimated subject-specific mean inherent strength and implicitly introduces correlation of the inherent strengths across muscles.
within a single subject. Furthermore, the hierarchical model allows the amount of variability of inherent strength across muscles within a single subject, $\sigma^2_{M,i}$, as well as the amount of variability of mean subject-specific inherent strengths across subjects, $\sigma^2_{\mu,i}$, to be estimated from the data. In particular, the within subject variance (across muscles), $\sigma^2_{M,i}$, is assumed to have an inverse gamma distribution with a prior expectation on the variance of approximately 0.02. If this variance is estimated to be small there will be strong correlation in inherent strength within a subject across muscle groups. The hierarchical subject-specific mean, $\omega_i$, is assumed to come from a normal distribution with a mean of 1 and an across subject variance, $\sigma^2_{\mu,i}$, that has an inverse gamma distribution with prior expectation on the variance of approximately 0.001. Both prior distributions on the variance parameters for the hierarchical model on the muscle and subject-specific proportion of inherent strength are informative and carry expectations that the values will be small. This assures that the subject-specific mean will not vary much from 1 and the muscle-specific inherent values will be highly correlated within a subject and will not vary much from each other.

### 3.2 Posterior estimation

The joint posterior distributions of all model parameters were fit using Markov chain Monte Carlo. The values of the model parameters were sampled from the joint posterior distributions using Gibbs sampling. When closed forms of the full conditionals are not available, Metropolis-Hastings was used. The posterior means and 95% equal-tailed credible intervals are provided as estimates of the muscle-specific location and shape decay parameters. We provide posterior means of the subject-specific disease age, and subject- and muscle-specific proportion of inherent muscle strength. Posterior mean and 95% equal-tailed prediction intervals are presented for the muscle strengths for each subject and muscle. Example code and convergence diagnostics are included as supplementary material in a computation appendix.

### 3.3 Predictive distributions

Predictive distributions for current and future muscle decline for subjects form the NHS are obtained using samples from the posterior distribution of all model parameters described in Section 3.2. For example, samples from the predictive distribution of muscle decline for subject $i$ at any past, current or future age $j^*$ can be obtained using the following algorithm:

1. Sample one set of values from the posterior distributions for all model parameters: $\hat{\theta}_k, \hat{p}_k, \hat{\sigma}_k, \hat{\delta}^2, \hat{\alpha}_i, \hat{M}_{t,k}$;
2. Sample the predicted muscle strength, $Y^*_{i,j^*,k}$ for each muscle $k = 1 \cdots K$ from the following normal distribution given the set of sampled model parameters specified above:

\[
Y^*_{i,j^*,k} \sim \mathcal{N}\left(\hat{\mu}_{i,j^*,k}, (\sigma_k \hat{\mu}_{i,j^*,k})^2 + \hat{\delta}^2 \right);
\]

\[
\hat{\mu}_{i,j^*,k} = \logit^{-1}\left[\hat{\theta}_k + \hat{p}_k (j^* - \hat{\alpha}_i) \right] \ast \hat{M}_{t,k};
\]

\[
i = 1 \cdots N; k = 1 \cdots K.
\]

For each subject in the NHS, we summarize the predictive distribution of muscle decline using the mean and 95% prediction intervals from the predictive distribution.

### 4 CHARACTERIZATION OF THE NATURAL PROGRESSION OF GNE MYOPATHY

Progressive muscle weakness is evident systemically when the strength of individual muscle groups, adjusted by the subject-specific proportion of inherent strength ($P/M$), was modeled jointly using the GNE DPM and plotted as a function of estimated disease age (Figure 1G-L). This suggests that the strength decrement in individual muscles was related to a common latent disease process, as opposed to sporadic factors such as age or disease duration (Figure 1A-F).

#### 4.1 Sequential muscle involvement in GNE myopathy

The onset and rate of progression varied among different muscle groups in GNE myopathy. Figure 3 presents the posterior mean proportion of inherent strength for each muscle as a function of disease age, over the course of disease progression.
The shaded region presents the pointwise 95% credible intervals for the mean proportion of inherent strength. Systematic, sequential muscle decay is evident with advancing disease age. The anterior tibialis muscle, which is primarily responsible for ankle dorsiflexion, is affected early in the disease progression and exhibits a relatively fast rate of progression. The ankle dorsiflexion strength was below 0.50 in all subjects and close to 0 in the majority of them (Figure 1G). Knee flexion and elbow flexion progressed at a rate similar to that of ankle dorsiflexion, but start declining approximately 10 and 20 years later, respectively (Figure 1H and L). Shoulder abduction was estimated to be involved early in the disease, but with a slower rate of progression, reaching full decline at an approximate disease age of 40 (Figure 1K). Conversely, the quadriceps, responsible for knee extension, progressed at the slowest rate (Figure 1I). The model-estimated proportions of inherent strength at a disease age 0 are ankle dorsiflexion (0.64), shoulder abduction (0.56), knee extension (0.70), grip (0.77), and elbow flexion (0.84) (Figure 1, red line). Therefore, when knee flexion is at 50% strength, we expected ankle dorsiflexion to be at 4%, shoulder abduction at 56%, knee extension at 70%, grip at 77%, and elbow flexion at 84%. The model-estimated disease ages at which each muscle has a muscle strength of 0.50 of inherent strength are ankle dorsiflexion (−7.7), shoulder abduction (2.6), elbow flexion (9.2), and grip (10.6) (Figure 1, blue line). Therefore, we expect ankle dorsiflexion to be at 50%, 7.7 years before knee flexion is at 50%, shoulder abduction 2.6 years after, elbow flexion 9.2 years after, and grip 10.6 years after. Knee extension is not expected to reach a strength of 0.50 over the course of the disease under consideration (disease ages −40 to 40).
### TABLE 1  Correlation of disease age with clinical endpoints and PROs as compared to chronological age and disease duration

<table>
<thead>
<tr>
<th></th>
<th>Chronological Age</th>
<th>Disease Age</th>
<th>Disease Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite total strength</td>
<td>0.11</td>
<td>−0.85 ****</td>
<td>−0.08</td>
</tr>
<tr>
<td>Composite UE strength</td>
<td>0.20 *</td>
<td>−0.84 ****</td>
<td>0.01</td>
</tr>
<tr>
<td>Composite LE strength</td>
<td>0.00</td>
<td>−0.77 ****</td>
<td>−0.16</td>
</tr>
<tr>
<td>6MWT</td>
<td>−0.19</td>
<td>−0.81 ****</td>
<td>−0.30 **</td>
</tr>
<tr>
<td>AMAT endurance</td>
<td>−0.08</td>
<td>−0.86 ****</td>
<td>−0.21 *</td>
</tr>
<tr>
<td>AMAT functional</td>
<td>−0.26 **</td>
<td>−0.78 ****</td>
<td>−0.34 ***</td>
</tr>
<tr>
<td>AMAT total</td>
<td>−0.18</td>
<td>−0.84 ****</td>
<td>−0.29 **</td>
</tr>
<tr>
<td>HAP adjusted activity score</td>
<td>−0.32 **</td>
<td>−0.70 ****</td>
<td>−0.35 **</td>
</tr>
<tr>
<td>HAP maximum activity score</td>
<td>−0.28 **</td>
<td>−0.58 ****</td>
<td>−0.22 *</td>
</tr>
<tr>
<td>IBMFRS</td>
<td>−0.40 ***</td>
<td>−0.70 ****</td>
<td>−0.49 ****</td>
</tr>
</tbody>
</table>

Clinical endpoints and PROs obtained contemporaneously to muscle strength were analyzed to determine whether disease age accurately represents disease progression as measured by endpoints other than strength. Composite muscle strength is expressed as percent of predicted. Pearson correlation coefficient and P values (two-tailed): *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001 for the model-generated disease age, compared to patient’s age and disease duration (age at baseline minus age of onset) at baseline.

**Abbreviations:** AMAT: Adult Myopathy Assessment Tool; HAP: Human Activity Profile; IBMFRS: Inclusion Body Myositis Functional Rating Scale; LE: lower extremity; UE: upper extremity; 6MWT: 6-minute walk test.

### 4.2 Estimates of future progression in individual subjects

The GNE DPM provides predictions of current and future muscle decline given the overall posterior results described in Section 4.1 as well as the posterior distribution for a subject’s estimated disease age and inherent strength as described in Section 3.3. Figure 2A-C illustrates 4 subjects of similar chronological age, but different mean model-estimated disease ages of 17.75, −8.02, 11.04, and −11.33. The posterior means and 95% predictive intervals of muscle strengths for these 4 subjects are shown in Figure 2E-H. Subject A (Figure 2E) was 32 years old at the last visit and declined below 0.30 in all muscles other than knee extension. The model predicts that, within the next 10 years, the most actively decaying muscles in this subject will be elbow flexion, grip, and shoulder abduction, with elbow flexion reaching full decline by age 40. Subject B (Figure 2F) is close in age but is at an earlier stage of the disease with a mean estimated disease age of −8.02 years. Ankle dorsiflexion is predicted to have the highest rate of progression, with a full decline expected at age 38; knee flexion is predicted to decrease to be less than 0.30 within the next 10 years, and both shoulder abduction and elbow flexion are predicted to decline at or below 0.50 by age 38. The lower extremity muscles in Subject C (Figure 2G) have almost fully declined except knee extension at age 38; in the next decade, the upper extremity muscles will have the fastest progression, with shoulder abduction and elbow flexion expected to decline below 0.20 by age 42.

### 4.3 Disease age correlation with clinical and functional endpoints

To assess whether the model-estimated “disease age” accurately represents disease progression as measured by endpoints other than muscle strength, we evaluated a variety of clinical and patient-reported outcomes (PROs) related to muscle function collected contemporaneously with the muscle strength data (Section 2.2). These included the 6MWT, the Adult Myopathy Assessment Tool (AMAT), and the Inclusion Body Myositis Functional Rating Scale (IBMFRS). These evaluations provide information regarding physical function, the ability to perform activities of daily living (ADLs) and the patient’s perspective on disease severity.

We determined that the model-based disease age, based entirely on muscle scores, has a strongly significant correlation (p < 0.0001) with each of the endpoints tested. In comparison, chronological age and disease duration showed weak to moderate correlations with the other endpoints (Table 1).

The impact of advanced disease age on specific ADLs was evaluated using a validated PRO, the IBMFRS. For GNE myopathy patients, the IBMFRS total score decreased by approximately 5 points with an advancing disease age of 10 years. Furthermore, individual items from the IBMFRS, including climbing stairs, walking, turning in bed and adjusting covers, cutting food and handling utensils, sit to stand, fine motor tasks, hygiene, and dressing show decreased scores with advancing disease age (Figure 4). Upright mobility tasks such as walking, climbing stairs, and performing sit to stand
FIGURE 4  Activities of daily living as reported by subjects of advancing “disease age.” Data obtained as part of NIH Protocol 11-HG-0218 “A Natural History Study of GNE Myopathy.” Only data from same patient visits as muscle strength data used for generation of the GNE DPM are used. Scores for individual items from the Inclusion Body Myositis Functional Rating Scale (IBMFRS) on each subject (circles) are plotted as a function of “disease age.” The moving average (+/− 10 years; dotted line) was calculated as the average of all data points within a 10-year window. A progression bar (bottom) is shown based on the moving average. Possible scores are 4 (normal; dark green), 3 (light green), 2 (yellow), 1 (orange), or 0 (unable or dependent; red). IBMFRS items include climbing stairs, walking, turning in bed and adjusting covers, cutting food and handling utensils, sit to stand, fine motor tasks, bathing and toileting, and dressing. Possible scores are 4 (normal; dark green), 3 (light green), 2 (yellow), 1 (orange), or 0 (unable; red)

are affected before a disease age of negative 10 years, but ultimately all ADLs are affected with advancing disease age in subjects with GNE myopathy. The “quadriiceps-sparing” feature of GNE myopathy likely contributes to the preservation of upright mobility task performance (walking, sit to stand, and climbing stairs) with gait aid and physical assistance required past disease age of 30 years.

5  |  DESIGN OF EVIDENCE-BASED CLINICAL TRIALS FOR GNE MYOPATHY

5.1  |  Clinical trial performance and power using traditional endpoints

We performed power calculations to explore the effect of choosing a single primary endpoint for a clinical trial of GNE myopathy. Using the NHS data collected concurrently as part of the NHS, we characterized the expected mean rate of
relative to natural history. The treatment effect, the rate of progression of the disease. In the DPM, we estimate the natural rate of decay for each muscle, outcome parameter in several rare diseases, as a primary endpoint would require 296 patients, using a composite of upper least 96 patients are needed if a single primary endpoint is used. Specifically, using the 6MWT, a widely accepted primary ment compared to placebo, trial length of 2 years, 1:1 randomization, and power of 80% using a two-sample t-test, at decline for several endpoints, including composite upper and lower extremity muscle strength, functional endpoints (6MWT, AMAT), and PROs (IBMFRS and HAP) (Table 2). If we assume a 50% reduction in the rate of decline for treatment compared to placebo, trial length of 2 years, 1:1 randomization, and power of 80% using a two-sample t-test, at least 96 patients are needed if a single primary endpoint is used. Specifically, using the 6MWT, a widely accepted primary outcome parameter in several rare diseases, as a primary endpoint would require 296 patients, using a composite of upper or lower extremity strength would require 164 and 150 patients, respectively.

### 5.2 GNE disease modification analysis

Since GNE myopathy is characterized by muscle atrophy, a disease-modifying treatment would be expected to slow or stop the rate of progression of the disease. In the DPM, we estimate the natural rate of decay for each muscle, $\beta_k$. To determine whether a therapy leads to a modification in disease progression, we incorporated a single treatment effect parameter into the DPM to provide a GNE Myopathy Disease Modification Analysis (DMA). Since different muscles progress at different stages of the disease, it is crucial to incorporate all muscles in the estimation of the treatment effect. The single treatment effect parameter is referred to as the disease progression ratio, ie, $\gamma$, that measures the proportional change in the rate/timing of progression relative to placebo (or equivalently change in the rate of disease progression), after the initiation of treatment, $T_i$.

$$Y_{i,j,k} \sim N\left(\mu_{i,j,k}, \left(\sigma_k \mu_{i,j,k}\right)^2 + \delta^2\right);$$

$$\mu_{i,j,k} = \begin{cases} 
\logit^{-1} \left[\hat{\theta}_k + \beta_k \left(t_{ij} - a_i\right)\right] \ast M_{i,k} & \text{if } t_{ij} \leq T_i \\
\logit^{-1} \left[\hat{\theta}_k + \beta_k \left(t_{i} - a_i\right) + \beta_k \gamma_0 \left(t_{ij} - T_i\right)\right] \ast M_{i,k} & \text{if } t_{ij} > T_i \text{ and } \text{treat}_i = 0 \\
\logit^{-1} \left[\hat{\theta}_k + \beta_k \left(t_{i} - a_i\right) + \beta_k \gamma_0 \left(t_{ij} - T_i\right)\right] \ast M_{i,k} & \text{if } t_{ij} > T_i \text{ and } \text{treat}_i = 1.
\end{cases}$$

This DMA model is identical to the DPM specified above before intervention. In particular, a placebo effect parameter, $\gamma_0$, is introduced to differentiate how the placebo group progresses from natural progression. If $\gamma_0 = 1$, disease progression is the same under natural progression and once a placebo is given. If $\gamma_0 < 1$, a placebo has a slowing in the rate of progression relative to natural history. The treatment effect, $\gamma$, for those in the treatment group ($\text{treat}_i = 1$) is then the change in the rate of decline across all muscle groups relative to placebo or the change in the rate of disease progression. In particular, if $\gamma = 1$ there is no additional effect of the treatment on the rate of disease progression relative to placebo. If $\gamma < 1$, the

<table>
<thead>
<tr>
<th>Primary Endpoint</th>
<th>1-year Decline Mean (SD)*</th>
<th>Sample Sizeb (80% Power; 50% Reduction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite total strength (%predicted)</td>
<td>–2.01 (5.66)</td>
<td>252</td>
</tr>
<tr>
<td>Composite UE strength (% predicted)*</td>
<td>–1.67 (5.28)</td>
<td>318</td>
</tr>
<tr>
<td>Composite LE strength (% predicted)*</td>
<td>–2.28 (6.85)</td>
<td>288</td>
</tr>
<tr>
<td>Composite UE strength (kg)*</td>
<td>–4.52 (10.2)</td>
<td>164</td>
</tr>
<tr>
<td>Composite LE strength (kg)*</td>
<td>–9.55 (20.6)</td>
<td>150</td>
</tr>
<tr>
<td>subset 6MWT &gt; 200 m²</td>
<td>–5.66 (10.9)</td>
<td>120</td>
</tr>
<tr>
<td>subset 6MWT &gt; 200 m²</td>
<td>–11.4 (20.72)</td>
<td>106</td>
</tr>
<tr>
<td>6MWT (meters)</td>
<td>–10.7 (32.7)</td>
<td>296</td>
</tr>
<tr>
<td>AMAT total score</td>
<td>–1.47 (2.52)</td>
<td>96</td>
</tr>
<tr>
<td>HAP adjusted activity score</td>
<td>–1.11 (5.52)</td>
<td>778</td>
</tr>
<tr>
<td>HAP maximum activity score</td>
<td>–3.24 (7.96)</td>
<td>192</td>
</tr>
<tr>
<td>IBMFRS total score</td>
<td>–0.49 (1.65)</td>
<td>352</td>
</tr>
</tbody>
</table>

*The estimated yearly change (a negative value means decline) for each endpoint is estimated from the Natural History Study assuming a linear rate of decline.

bEstimated number of subjects that would be needed for a 2-year trial assuming 1:1 randomization with other available primary endpoints assuming treatment effect ($\gamma$) of 50% to reach 80% power.

cThe subset of subjects able to walk >200 meters on 6MWT at baseline.

**Abbreviations.** AMAT: Adult Myopathy Assessment Tool; HAP: Human Activity Profile; IBMFRS: Inclusion Body Myositis Functional Rating Scale; LE: lower extremities; UE: upper extremities; SD, standard deviation; 6MWT: 6-minute walk test.
FIGURE 5  Potential Effect of treatment on muscle progression. (A-C) The potential effect of treatment on muscle progression is
represented at different stages of muscle decline. The solid line corresponds to the progression of muscle decline without intervention. The
dashed lines represent different treatment effects, $\gamma$. If $\gamma = 1$, there is no treatment effect. If $\gamma < 1$, the treatment slows the rate of progression.
If $\gamma = 0.5$, the treatment slows the rate of muscle progression by 50%. If $\gamma = 0$, the treatment has stopped the rate of progression
treatment slows the rate of decay of each muscle group relative to placebo, to the extreme of a $\gamma = 0$, where the treatment
has stopped the decay of the muscles altogether. For example, if $\gamma = 0.5$, the treatment slows the rate of decay of all
muscles by 50% relative to placebo and, equivalently, the rate of progression by 50% (Figure 5). Therefore, if under the
natural progression of the disease, it would be expected to take 5 years for a subject to become nonambulatory, under a
treatment that slows the progression by 50%, it would now take 10 years to become nonambulatory. If treatment occurs
at early stages of disease progression for a particular muscle (Figure 5A), the absolute magnitude of the treatment effect
is greater than if the treatment occurs when muscles have almost fully declined (Figure 5C).
For the placebo effect parameter, $\gamma_0$, we assume that the parameter has a beta prior that is spread to be defined between
0 and 2, has a mean of 1 and a standard deviation of 0.15. We assume a uniform prior between 0 and 2 for $\gamma$.
In the DMA, we test the hypothesis that the disease progression ratio is greater than or equal to 1 (no or worsening effect)
against the hypothesis that the disease progression ratio is less than 1 (benefit of the treatment), comparing subjects
receiving drug vs placebo.

$$H_0: \gamma \geq 1$$
$$H_A: \gamma < 1$$
If the posterior probability that the disease progression ratio is less than 1 ($H_A$) is greater than 0.975 at the final analysis,
then the primary analysis is successful at showing a statistically significant reduction in the rate of disease progression.
The posterior probability threshold of .975 is chosen to maintain overall control of a type I error at 2.5%. Within the
Bayesian setting, given the Unif(0, 2) prior distribution on $\gamma$ and thus equal prior probability of $H_0$ and $H_A$, the posterior
probability of $H_A$ threshold of .975 is akin to requiring a Bayes Factor threshold of 39 (i.e., very strong evidence) to claim a
statistically significant reduction in the rate of disease progression.

Given the correlations of other clinical outcome measures with disease age provided in Section 4.3, we can relate the
slowing of the rate of disease progression in terms of the disease progression ratio to changes in other clinical measures.
For example, if an intervention that slows the rate of disease progression by 50% is given to a patient starting at a disease
age of $-10$, it would take 20 years (instead of 10 years estimated under natural progression) for the subject to go from
>90% strength on knee flexion and nearly normal functioning on all ADLs to being at 50% strength on knee flexion,
dependent on a handrail when climbing stairs, intermittently using an ambulatory device when walking and starting to
have difficulties with ADLs (e.g., cutting food, bathing, toileting, and getting dressed) (Figure 3 and Figure 4).

5.3  Clinical trial power and performance using GNE DMA
We characterized the performance of a clinical trial in which the primary efficacy analysis is the GNE DMA. To inform
the design of the clinical trial to determine the efficacy of a potential therapy for GNE myopathy, simulations of the design
under multiple scenarios were performed. Scenarios included different assumed treatment effects ($\gamma$), numbers of future
randomized patients, the duration of treatment, randomization ratio, and the frequency of visits. For each scenario, 1000 trials were simulated. For each simulation, we assume that the newly simulated subjects would be similar to the natural history subjects in terms of the distribution of disease ages (mean disease age of 10) and the mean location and rate of decline for each muscle group (as shown in Figure 3). It was determined that a trial enrolling 50 subjects randomized 2:1 (treatment: placebo) followed for a duration of 2 years and follow-up visits every 6 months would have a power of 87.5% if the treatment slowed the rate of decline by at least 50%. Thus, using the DMA, we are able to conduct a clinical trial that allows more patients to receive the investigational therapy, with higher power and fewer patient resources than a clinical trial that uses a single primary endpoint and a more traditional analysis method.

Under the assumption that the treatment does not slow the rate of progression, a conservative type I error rate of less than 1% is achieved. To explore the sensitivity of the type I error rate to the assumptions made in terms of the distribution of disease ages enrolled in the trial and in terms of the effect of the placebo, we investigate four additional scenarios.

- Distribution of disease age: randomized subjects are at an earlier stage of the disease (mean disease age of 5) or later stage of the disease (mean disease age of 15) on average.
- Placebo effect: subjects have a faster (10% faster; $\gamma_0 = 1.10$) or slower (10% slower; $\gamma_0 = 0.90$) rate of decline on placebo compared to their natural history.

In particular, type I error rates are relatively insensitive to enrolling earlier or later disease staged subjects (1.4%-1.5%) compared to the subjects in the NHS. As expected, type I error increases slightly to 2% if the randomized subjects have a slower disease progression once they are given placebo but is still controlled at the 2.5% level.

### 6 | DISCUSSION

Recently, the FDA has initiated a Model-Informed Drug Development Pilot Program to, among other things, incorporate statistical models based on clinical data to facilitate drug development. Drug development for rare and slowly progressive diseases, such as GNE myopathy, can be facilitated significantly by the use of disease progression modeling based on clinical data. In the case of GNE myopathy, characterizing the spectrum of disease progression for this rare and slowly progressive disease would require a prospective NHS of unfeasible duration and size. Here, we describe the development of a DPM to define overall long-term disease progression in GNE myopathy based on aligning and consolidating short-term observations of disease progression across a diverse set of subjects at different stages of the disease. The model is based on longitudinal muscle strength data collected from a diverse cohort of GNE myopathy subjects evaluated in a prospective NHS. Muscle strength was selected as the basis for the model since skeletal muscle atrophy leading to weakness is the sole complication of GNE myopathy; the clinical manifestations of the disease are direct consequences of the reduced muscle strength.

The DPM aligns subjects based on a latent disease age that is a better measure of disease progression than chronological age or disease duration. With advancing disease age, there is progressive muscle weakness, affecting subjects in a systematic pattern of muscle decline and a decrease in physical function and ability to perform ADLs. The GNE DPM allows the description of long-term sequential muscle involvement in GNE myopathy and provides estimates for the onset of muscle involvement and the rate of decay of individual muscle groups across the entire disease course. To account for the variable onset and the rate of progression among different muscle groups in GNE myopathy, the DPM uses a logistic decay function with a separate onset and rate of progression parameter for each muscle. The onset and rate of decline of muscle strength for each muscle are modeled jointly and are connected through the subject-level disease age parameter. As depicted in Figure 1G-L and Figure 3, once aligned based on disease age, the NHS subjects provide an overall picture of the systematic progression of each muscle group across the a disease course that spans 50 years of the disease.

The GNE DPM provides prediction of future long-term disease progression for each subject conditional upon the subject-level disease age and inherent muscle strength. The prediction of future long-term disease progression is made possible by aligning and consolidating a diverse set (in terms of disease age) of short-term progressions from the NHS.

Finally, the DPM can be extended to provide a primary analysis tool, the DMA, which leads to a substantial increase in power over standard endpoints and traditional analysis methods, allowing for the determination of treatment effect with a feasible sample size and trial duration, the ability to enroll a much wider range of patient severities, and the capacity to detect clinically meaningful changes. The DPM is the primary analysis method of a multicenter pivotal randomized control trial currently being developed with the NIH to study the efficacy of a novel therapy for GNE myopathy. The use
of the DPM as a primary analysis tool within this trial provides 87.5% power with 50 patients when nearly double the sample size would be needed using more traditional primary endpoints and analysis methods.

The accuracy of the GNE DPM and the long-term subject-level predictions of future muscle decline rely on the degree to which the modeling assumptions are meet: (1) muscle decline and disease progression follows a logistic decay shape as a function of disease age; (2) conditional upon the disease age/timing of initial progression of a participant and the inherent strength of the participant, the rates of progression across all muscle groups will be similar and subjects are not expected to have different rates of progression; and (3) long-term predictions span disease ages that are represented in the NHS. The DPM assumes that muscle decline and disease progression follows a logistic decay shape as a function of disease age and that conditional upon their disease age, subjects will have similar onset and rate of progression on each of the muscle groups. However, it may be the case that the rate of decline for each muscle group is subject-dependent. To address this assumption, the accuracy of predictions from the DPM can be tested as subjects return for follow-up visits by comparing their actual measurements to the model predictions. Despite the limitations of the DPM, several findings support the accuracy of the estimates. The DPM accurately predicts the rate of decline of the quadriceps (knee extension) to be considerably slower than that of other muscles groups, which is consistent with the description of GNE myopathy as a “quadriceps-sparing myopathy.” The DPM also estimates the anterior tibialis muscle as the first muscle to be affected, as previously described for this disease. Finally, the model-estimated disease age correlates very strongly with clinical and patient-centered outcomes not included in the modeling.

While the loss of muscle strength is directly related to functional ability, the relationship is not always linear. As seen in Duchenne Muscular Dystrophy, a substantial decrease in strength early on may be associated with a small change in function, while a small change in strength late in the disease can have a significant impact on a subject’s functioning. Thus, it is essential that the DPM takes into account not only changes in muscle strength but that also characterizes changes in patients’ functioning, the need for an ambulatory device, and the ability to perform ADLs. We specifically evaluated the DPM to determine if it reflects clinically-relevant and patient-reported changes. We found that as model estimated disease age increased, the ability to perform ADLs decreased; subjects went from independently functioning to requiring ambulatory devices to walk, requiring intermittent assistance or performing modified techniques, to eventually complete dependence on a caregiver and wheelchair use. Furthermore, the DPM can capture the potential treatment effect of prolonging function and the ability to perform ADLs in GNE myopathy subjects by a slowing of the disease's natural progression. As described in the results, slowing the disease progression by 1 or 2 decades could prolong the time GNE myopathy subjects have to independently walk, climb stairs, or perform ADLs.

The DPM was developed to characterize disease progression in GNE myopathy, but it can also be used to strengthen clinical trial design and provide solutions to the challenging drug development aspects of GNE myopathy. Since its implementation in 1983, the US Orphan Drug Act has aimed at stimulating the development of therapies for rare diseases. However, the challenges of drug development for slowly progressive rare diseases remain unresolved. In a recent publication, Hay et al reported that only 10.4% of drugs that enter phase 1 are approved, and the likelihood is lower for neurodegenerative diseases at 9.8%. Half of the failures during Phase 3 are attributed to lack of efficacy, but it is unclear how many of these failures were, in fact, attributable to poor study design and inappropriate endpoints, rather than lack of drug efficacy, since a poor study design would not accurately discern efficacy. The GNE DPM provides an innovative approach to drug development. GNE myopathy represents many of the difficulties in developing novel therapies, including rareness of the disease, slow progression, no functional or clinical measures that are relevant for a wide range of disease severity. In this case, the understanding of progression, provided by the GNE DPM has allowed the simulation and optimization of different trial designs for this rare and slowly progressive disease.

ACKNOWLEDGEMENTS

This study was supported by the National Human Genome Research Institute (NHGRI), by the National Center for Advancing Translational Sciences (NCATS), and by the NIH Clinical Center, all under the National Institutes of Health, Bethesda, Maryland, USA.

AUTHOR CONTRIBUTIONS

COMPETING INTERESTS
Include any financial interests of the authors that could be perceived as being a conflict of interest.

ORCID

M. Quintana https://orcid.org/0000-0002-1763-2704
N. Carrillo https://orcid.org/0000-0003-0374-0808

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Quintana M, Shrader J, Slota C, et al. Bayesian model of disease progression in GNE myopathy. *Statistics in Medicine.* 2019;38:1459–1474. [https://doi.org/10.1002/sim.8050](https://doi.org/10.1002/sim.8050)