A Bayesian Adaptive Trial in Duchenne Muscular Dystrophy

Stephen Lake, ScD

VP, Biometrics

Wave Life Sciences

15Jul2022

Outline

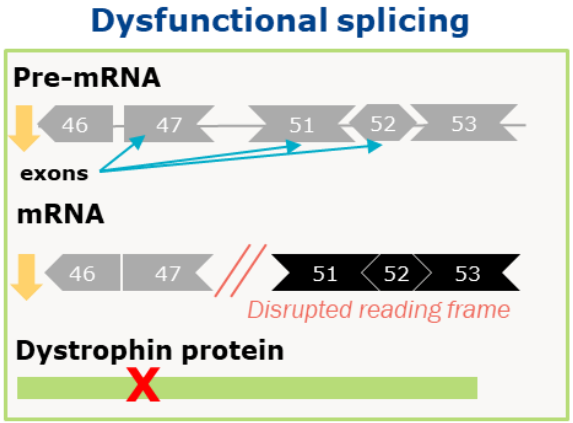
- Duchenne muscular dystrophy
- DYSTANCE 51: Pivotal study to assess safety and efficacy of suvodirsen in DMD
  - Study design
  - Statistical methods
  - Simulations
- CID experience and adaptive design considerations
- Summary

Duchenne Muscular Dystrophy

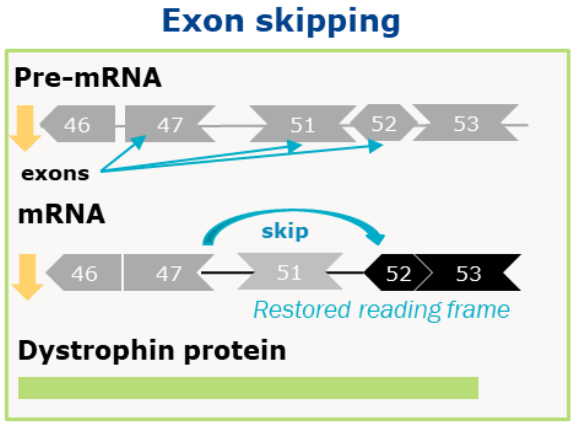
- Fatal, X-linked genetic neuromuscular disorder characterized by progressive, irreversible loss of muscle function, including heart and lung
  - >50% of DMD boys lose the ability walk before 15 years of age
- Genetic mutations in the dystrophin gene prevents the production of dystrophin protein, a critical component of healthy muscle function
- Impacts 1 in every 5,000 newborn boys each year;
  20,000 new cases annually worldwide
- Current disease modifying treatments have not demonstrably established clinical benefit

DYSTANCE 51

- DYSTANCE 51 was a phase 2/3 trial of suvodirsen, an investigational antisense oligonucleotide for DMD patients amenable to exon 51 skipping
- Strategy of exon skipping is to encourage the cellular machinery to "skip over" the exon(s) with the genetic mutation to create a shortened, but still functional, version of dystrophin protein
  - In the US, there are 3 approved therapies for DMD using this approach



Mutation can disrupt the code for making dystrophin protein



Exon skipping can bypass the mutation to restore expression of truncated but functional protein

DYSTANCE 51 objectives

- To evaluate the efficacy of suvodirsen by assessing changes in dystrophin levels
  - Change in dystrophin was the primary endpoint for US FDA with potential for accelerated approval
  - Dystrophin increase is the basis for accelerated approval for 3 therapies
- To evaluate the efficacy of suvodirsen by assessing changes in motor function by North Star Ambulatory Assessment (NSAA)
  - Demonstration of efficacy on clinical endpoints (e.g., 6MWT, NSAA) is required for EMA and Japan
  - NSAA
    - 17-items rated 0, 1, 2 with lower scores indicating less motor function
    - 8. Can you step down from the box using your right leg first?
    - 11. Get up from the floor using as little support as possible and as fast as you can (from supine)

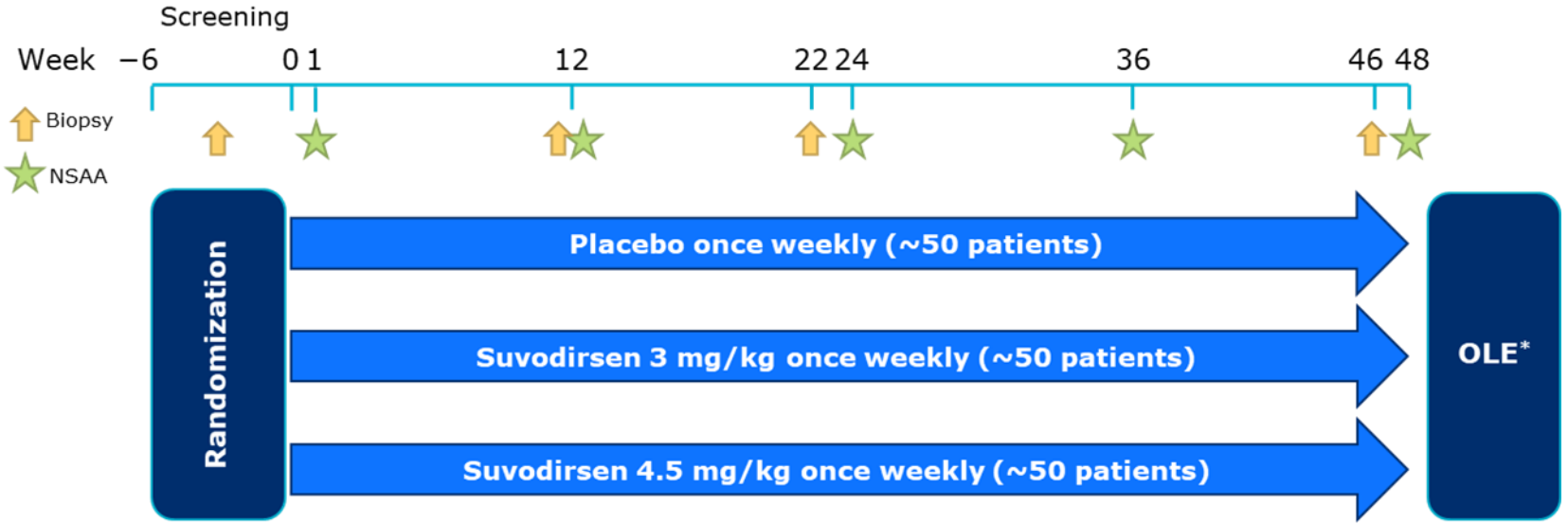
DYSTANCE 51 study design goals

- Through interim analyses, provide an opportunity to identify a suvodirsen treatment effect on dystrophin prior to study conclusion
- Maximize the probability of a definitive NSAA result
  - Select statistical methods that efficiently use data
  - Perform interim analysis to determine whether current sample size is sufficient
  - Incorporate historical control borrowing
- Adaptive design elements
  - Early demonstration of efficacy on dystrophin
  - Stop enrollment to low dose arm due to insufficient efficacy, do not pool doses for NSAA comparison
  - Stop enrollment due to high probability of NSAA success with current patient numbers

DYSTANCE 51 and CID program

- FDA announced CID program in August 2018
- Berry Consultants
- CID application submitted by Wave in September 2018
- Wave announced acceptance in January 2019
- First CID meeting end of January 2019
- Development of suvodirsen stopped in December 2019
  - Important to describe the trial design and CID experience to inform future trials in rare diseases

Overview of DYSTANCE 51 study design



- Primary endpoint
  - FDA: Dystrophin protein
  - EMA, PMDA: NSAA
- Secondary endpoint
  - FDA: NSAA
  - EMA, PMDA: Dystrophin protein

Study design goal:
Maximize the probability of a definitive NSAA result

- Select statistical methods that efficiently use data
  - Bayesian progression model (Wang et al. 2018; Quintana et al. 2019)
  - Flexible, non-parametric rate of progression of the NSAA for placebo patients
  - Proportional treatment effect estimated using available longitudinal data
  - Baseline covariates can be included
  - Meta-analytic approach dynamically determines the amount of borrowing (random data source effect)
- Perform interim analysis to determine whether current sample size is sufficient
  - Goldilocks approach (Broglio, Connor, Berry 2014)

Study design goal:
Maximize the probability of a definitive NSAA result

- Incorporate the capability to augment the placebo arm with historical data
  - Gain access to historical control datasets (placebo arm data from clinical trials)
  - Apply key inclusion criteria to increase similarity between the historical control and placebo datasets
    - Age, ambulation, steroid use
  - Adjust for covariates in the statistical model to account for differences in baseline covariates
  - Use method that appropriately accounts for borrowing from external datasets

Bayesian progression model for NSAA

$$Y_{ij} = \gamma_i + \exp\left(\theta_{t(i)} + \eta_i + \alpha X_i + \delta_{s(i)}\right) \sum_{k=0:j} \beta_k + \epsilon_{ij};$$

$$\epsilon_{ij} \sim N\left(0, \sigma^2_{t(i),s(i)}\right).$$

- Assume piecewise linear model for the placebo decline in NSAA
  - Each 12-week period is modelled as β_k with the sum of all β_k equal to the total decline over 48 weeks
- $\exp(\theta_T), T = 0, 1, 2$ quantifies the proportional NSAA slowing by treatment group
  - E.g, $\exp(\theta_2) = 0.75$ → 25% slowing with high dose suvodirsen, pool treatment arms
- $\delta_{s(i)}$ is a random data source effect to account for between-data source variability in NSAA progression rate with

$$\exp\left(\delta_s\right) \sim Gamma\left(1/\sigma^2_\delta, 1/\sigma^2_\delta\right) \text{ for } s = 1, \ldots, S;$$

$$\sigma_\delta \sim Unif\,(0,1).$$

NSAA comparisons

- Comparisons for enrollment stop

$$\Pr\left(\exp(\theta_2) < \mathrm{DRR_N} | y\right) > 0.90$$

  - $\mathrm{DRR_N}$ is derived through simulation at the design stage. It is the posterior mean of the treatment effect necessary to achieve to ensure that efficacy success is declared given the interim analysis sample size (N).

- Comparisons for treatment efficacy

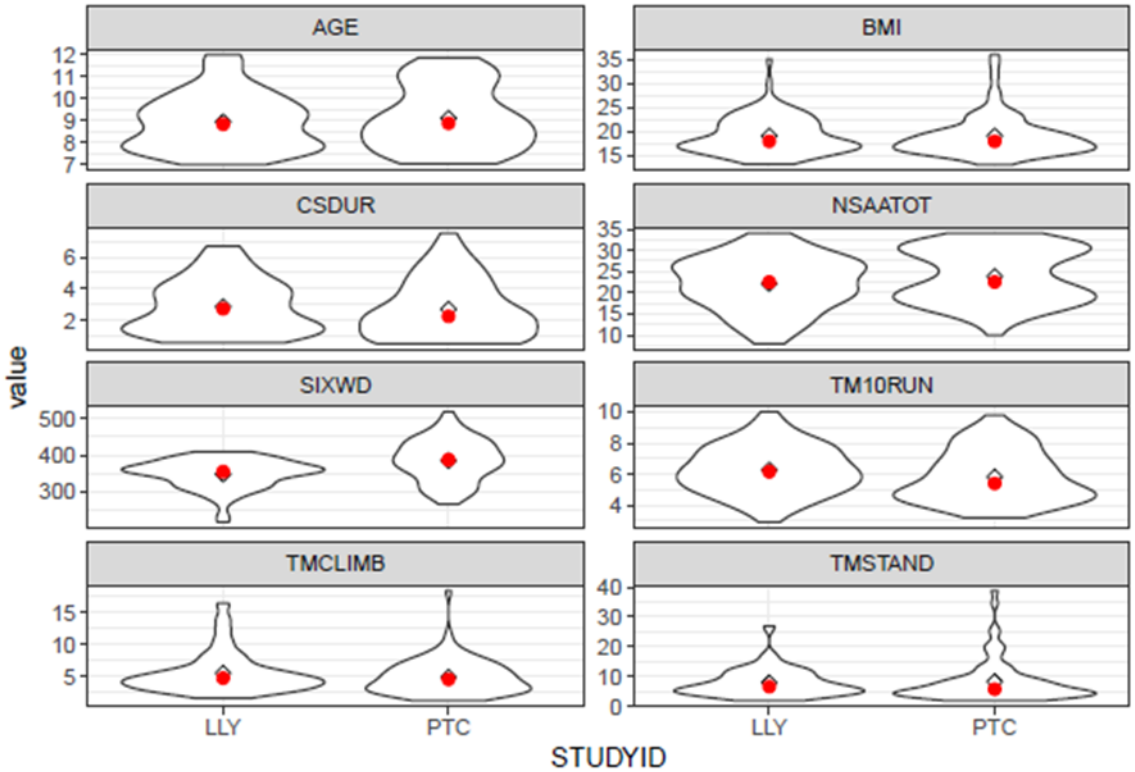
$$\Pr\left(\exp(\theta_2) < 1 | y\right) > 0.975$$

Potential sources of historical placebo data

Study Name	Investigated Therapy	Data Manager or Investigator (Sponsor)	No. of Placebo Patients	Status
Tadalafil DMD (NCT01865084)	Tadalafil	C-Path D-RSC (Eli Lilly)	116	Available
ACT-DMD (NCT01826487)	Ataluren	C-Path D-RSC (PTC Therapeutics)	115	Available
B5161002 (NCT02310763)	Domagrozumab	Pfizer	40	Available*
DEMAND II (NCT01153932)	Drisapersen	Biomarin	18	Not Available
DEMAND III (NCT01254019)	Drisapersen	Biomarin	61	Not Available

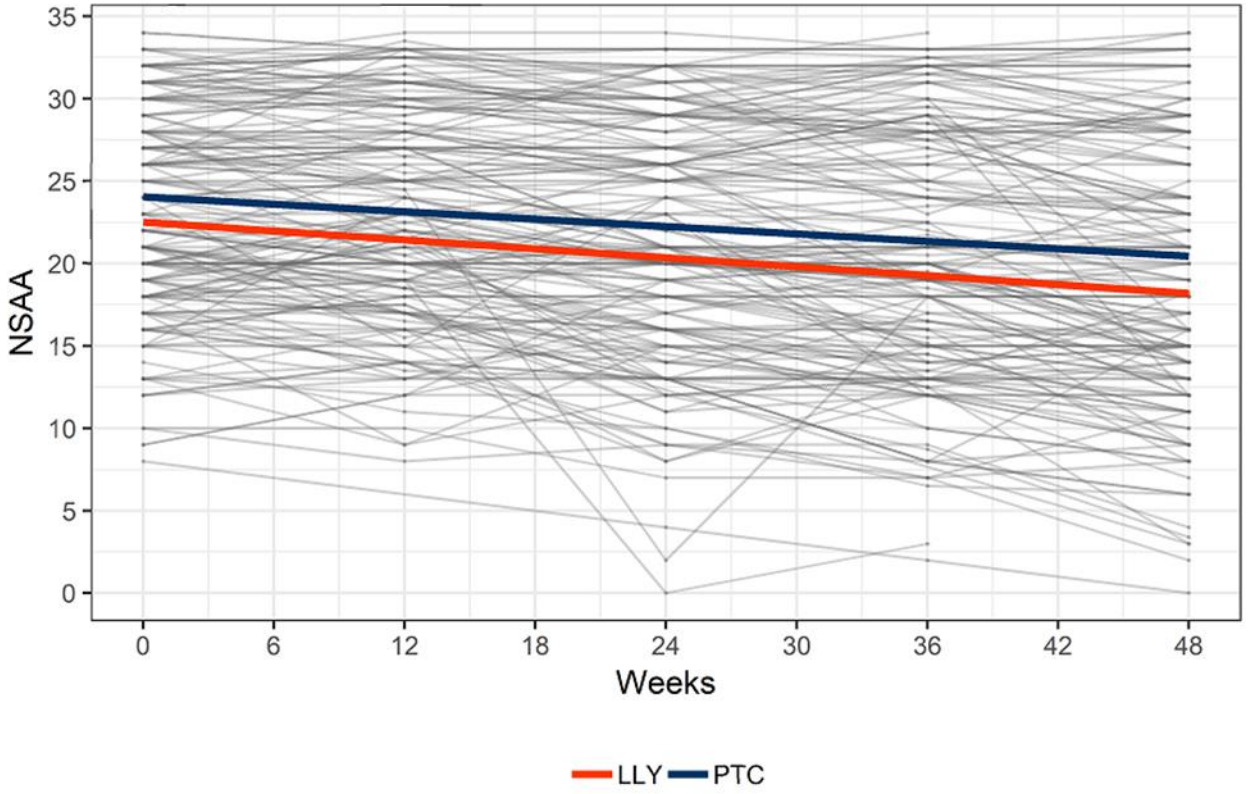
* Data access discussion underway at time of discontinuation of suvodirsen development. Pfizer team conducted analyses to assist with clinical trial simulations.

Baseline covariates in tadalafil and ataluren study



LLY: Eli Lilly tadalafil study, PTC: PTC Therapeutics ACT DMD study, BMI: body mass index, CSDUR: duration of steroid usage, NSAATOT: NSAA score, SIXWD: 6-minute walk distance, TM10RUN: timed 10-meter walk run, TMCLIMB: timed 4-stair climb, TMSTAND: time to stand from the floor

NSAA LME slope in tadalafil and ataluren studies



NSAA LME summaries from historical data

Dataset	N	Baseline NSAA	SD baseline NSAA	Mean change from baseline	SD change from baseline	Residual error
Tadalafil DMD	90	22.5	6.0	−4.3	3.7	2.4
ACT-DMD	76	24.0	5.8	−3.5	3.4	2.2
B5161002	26	20.6	5.9	−4.6	5.8	3.2

Abbreviations: DMD, Duchenne muscular dystrophy; NSAA, North Star Ambulatory Assessment.

Study design goal: assess suvodirsen treatment effect on dystrophin prior to study conclusion

- Dystrophin quantified from shoulder biopsy
- Biopsies collected at baseline and one post-baseline time point (week 12, 22 or 46)
  - Post-baseline time point determined by randomization order
- Dystrophin interim analyses occur when biopsy results available for a specific time point
  - These analyses are augmented with dystrophin data from the Phase 1 OLE
- Dystrophin analyses also used to determine whether to stop enrollment into suvodirsen low dose arm

$$D_{ij} = \tau_j + \vartheta_{t(i),j} + \epsilon_{ij};$$

$$\epsilon_{ij} \sim N(0, \sigma^2_{t(i),j}).$$

Dystrophin comparisons

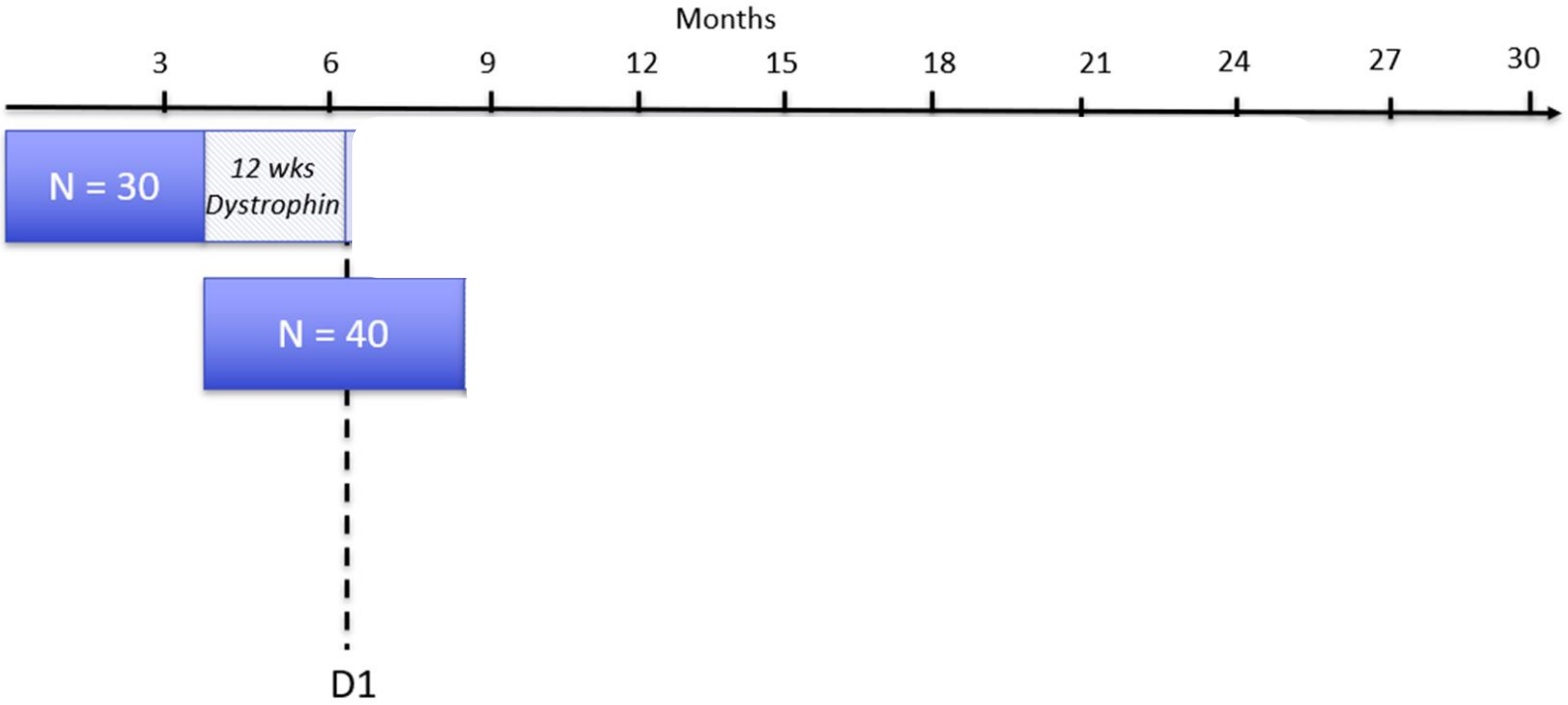
- Comparisons for stopping randomization to low-dose arm

$$\Pr(\vartheta_{2,j} > 0 \;\&\; \vartheta_{1,j} < 0.5\vartheta_{2,j} | d) > 0.95 \text{ for } j = 1, 2.$$

- Comparisons for treatment efficacy

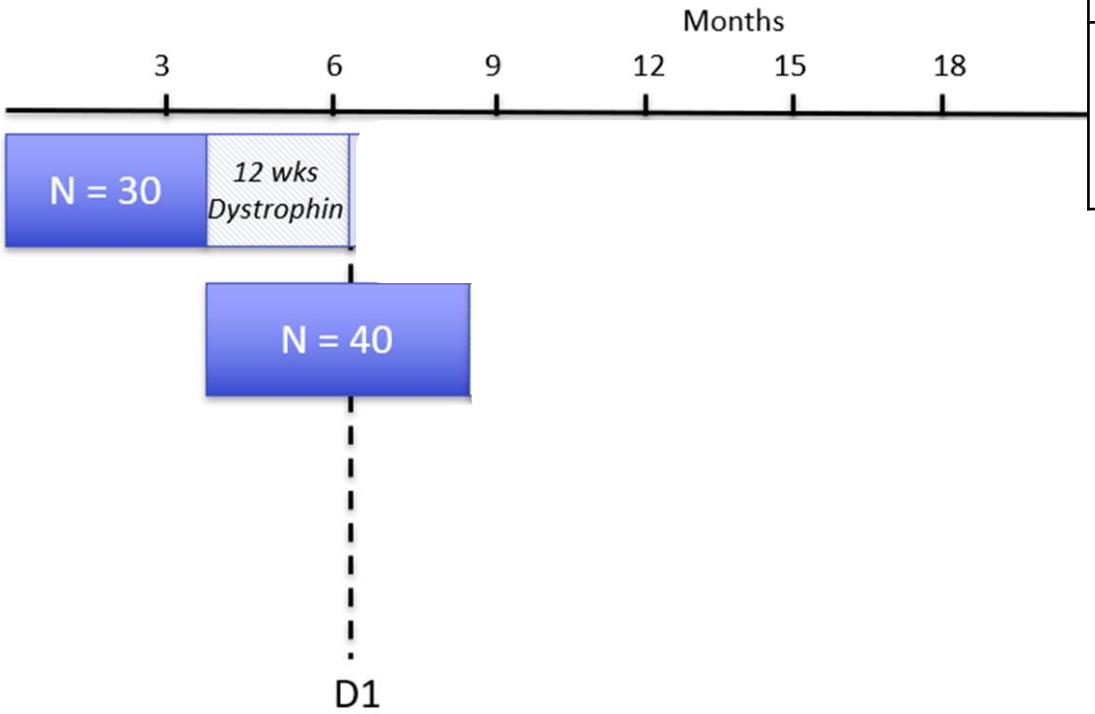
1. $\Pr(\vartheta_{2,1} > 0 | d) > 0.991667$ at the first dystrophin interim analysis,
2. $\Pr(\vartheta_{2,2} > 0 | d) > 0.991667$ at the second dystrophin interim analysis, or
3. $\Pr(\vartheta_{2,3} > 0 | d) > 0.991667$ at the final dystrophin analysis.

Trial schematic (assuming 8 patients/month)



D1, D2, DF: Dystrophin Analyses
N1, N2, N3, N4, NF: NSAA Analyses

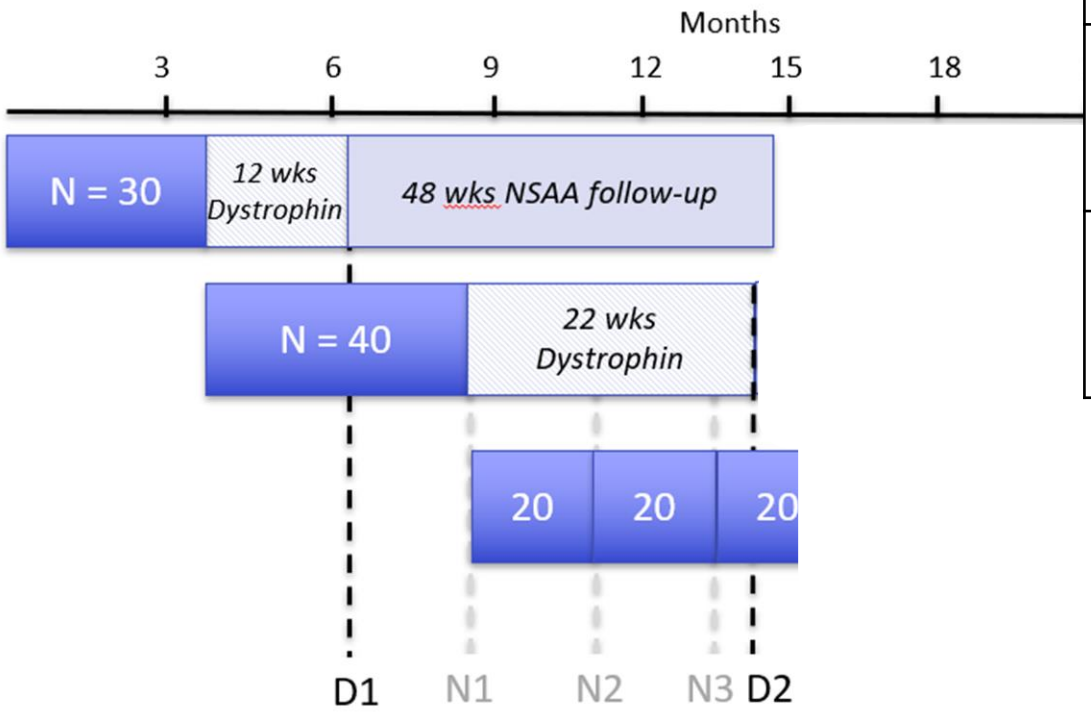
Trial schematic (assuming 8 patients/month)



Interim analyses	Potential outcomes
Dystrophin endpoint (D1)	Demonstration of efficacy Stop enrollment to low dose arm due to insufficient efficacy

D1, D2, DF: Dystrophin Analyses
N1, N2, N3, N4, NF: NSAA Analyses

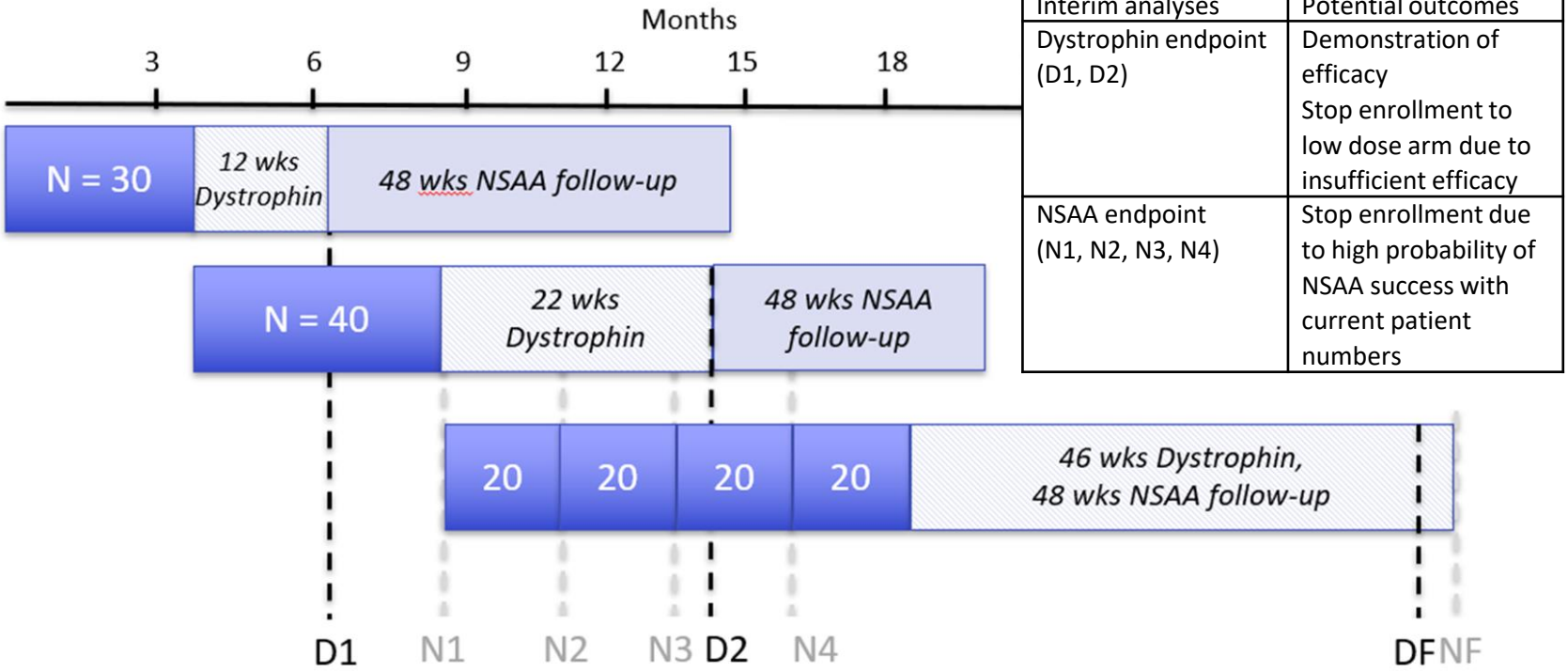
Trial schematic (assuming 8 patients/month)



Interim analyses	Potential outcomes
Dystrophin endpoint (D1, D2)	Demonstration of efficacy Stop enrollment to low dose arm due to insufficient efficacy
NSAA endpoint (N1, N2, N3)	Stop enrollment due to high probability of NSAA success with current patient numbers

D1, D2, DF: Dystrophin Analyses
N1, N2, N3, N4, NF: NSAA Analyses

Trial schematic (assuming 8 patients/month)



Interim analyses	Potential outcomes
Dystrophin endpoint (D1, D2)	Demonstration of efficacy Stop enrollment to low dose arm due to insufficient efficacy
NSAA endpoint (N1, N2, N3, N4)	Stop enrollment due to high probability of NSAA success with current patient numbers

D1, D2, DF: Dystrophin Analyses
N1, N2, N3, N4, NF: NSAA Analyses

Simulations

- Extensive simulations conducted to understand the operating characteristics of the trial design
- Performance metrics include probabilities of early and overall success, effective sample size with borrowing, mean months to dystrophin success, mean sample size, probability of pooling doses, mean treatment effect estimate, etc.
- 14 parameters defined a simulation scenario

Simulation parameters

- Base case and sensitivity analysis
  - Misspecification
- Range of treatment effects
- 231 different parameter combinations

Parameter	Base Case	Sensitivity Analysis Values	
		Lower	Higher
1. Accrual rate (patients/month)	8	4	10
2. Mean baseline dystrophin	0.5	0.15	1
3. Dystrophin measurement error	0.5	0.25	1
4. Dystrophin mean-variance assumption	Squared mean	Not squared mean	
5. Dystrophin placebo arm change from baseline	0	NA	NA
6. Low dose fraction of the high dose effect	0.75	0, 0.25, and 0.5	1
7. Week the treatment effect on dystrophin reaches full effect	46	12, 22	NA
8. Multiplicative difference in baseline dystrophin between OLE patients and DYSTANCE 51 patients	No difference	50% lower	50% higher
9. Additive difference in change from baseline dystrophin between OLE patients and DYSTANCE 51 patients	0	-0.5	+0.5
10. Mean baseline NSAA	23	20	26
11. SD baseline NSAA	6	3	10
12. Mean change from baseline to 48 weeks in NSAA for the placebo patients	-4.06	-4.56, -5.06	-3.56, -3.06
13. SD of the NSAA change from baseline (slope)	4	2	6
14. NSAA residual error	2.5	1	4

Dystrophin simulation results (base case)

Scenario effects		Mean months to flag	Cumulative probability of success flag			Probability of drop low dose		Overall estimated median effect
Dystrophin	NSAA		D1	D2	Total	D1	D2	High dose
0	0	29	0.004	0.013	0.018	0.000	0.000	0.00
1	0.2	22	0.03	0.48	0.98	0.00	0.01	0.98
1.5	0.3	16	0.09	0.83	0.99	0.00	0.00	1.51
2	0.4	14	0.16	0.94	1.00	0.00	0.00	2.02
3	0.5	12	0.26	1.00	1.00	0.00	0.00	3.07
4	0.6	11	0.40	1.00	1.00	0.00	0.00	4.14
10	0.7	10	0.62	1.00	1.00	0.00	0.00	9.99

Abbreviation: NSAA, North Star Ambulatory Assessment.

NSAA simulation results (base case)

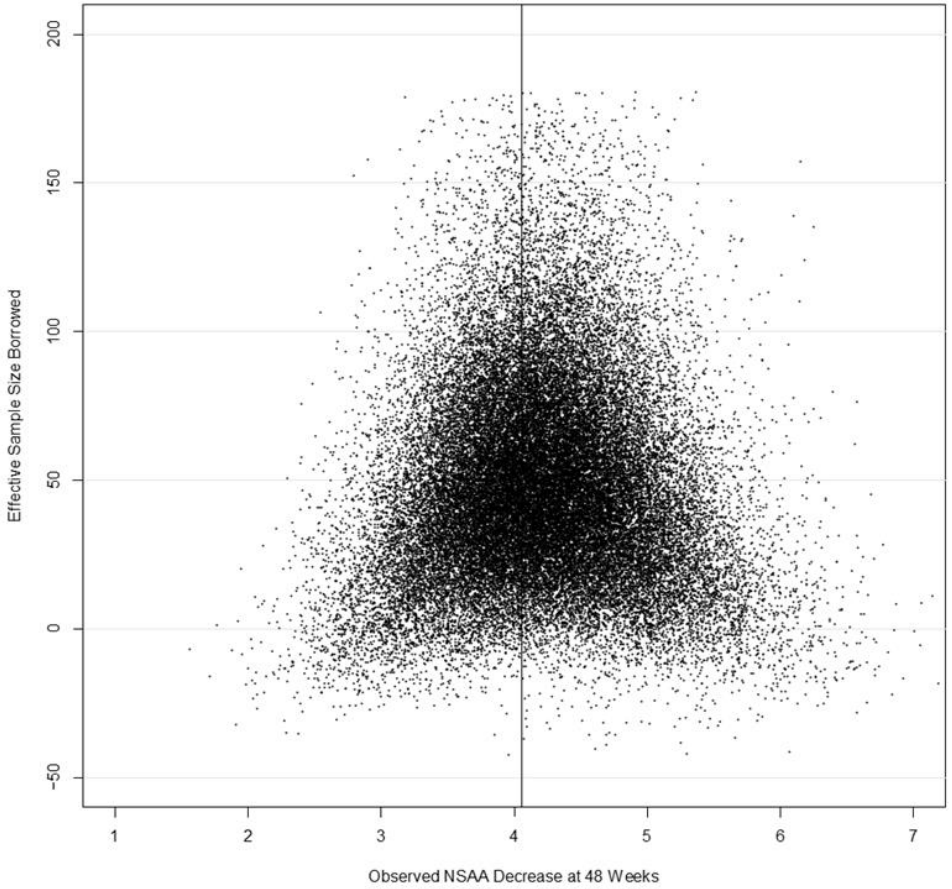
Scenario effects			Mean effective sample size	Cumulative probability of success					Overall estimated median effect
Dystrophin	NSAA	Mean N	Borrowed	70	90	110	130	Total	High dose
0	0	150	45	0.000	0.000	0.001	0.001	0.017	0.02
1	0.2	149	46	0.00	0.00	0.00	0.01	0.32	0.23
1.5	0.3	148	45	0.00	0.01	0.02	0.02	0.64	0.34
2	0.4	147	46	0.01	0.03	0.05	0.07	0.92	0.44
3	0.5	142	42	0.02	0.06	0.12	0.20	0.98	0.55
4	0.6	133	44	0.04	0.12	0.25	0.44	1.00	0.63
10	0.7	120	43	0.07	0.25	0.50	0.70	1.00	0.73

Abbreviation: NSAA, North Star Ambulatory Assessment.

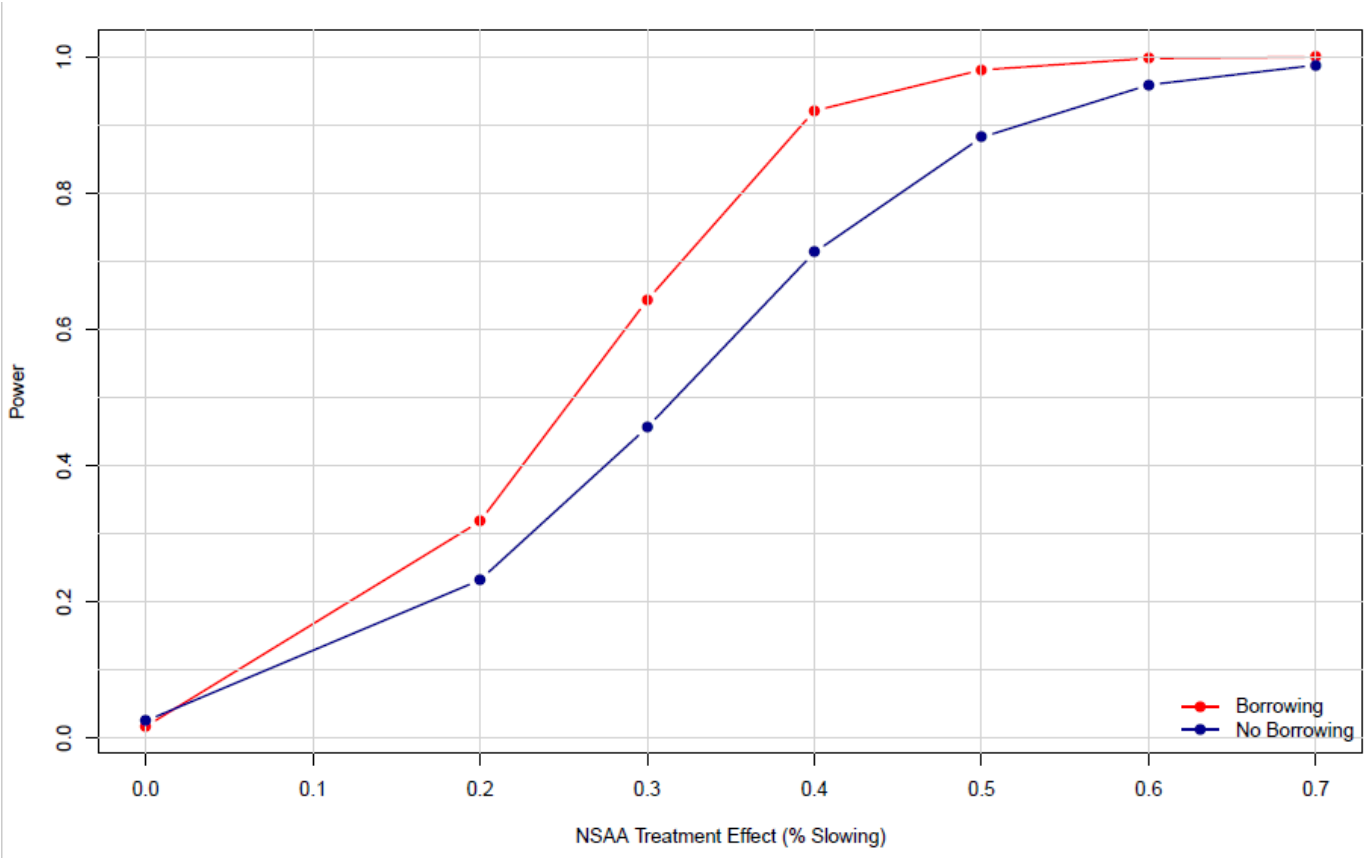
Effective sample size borrowed from historical data

$$N_h \approx N\left(\frac{\mathrm{se}(\sum_k \beta_k)}{\mathrm{se}_h(\sum_k \beta_k)}\right)^2$$

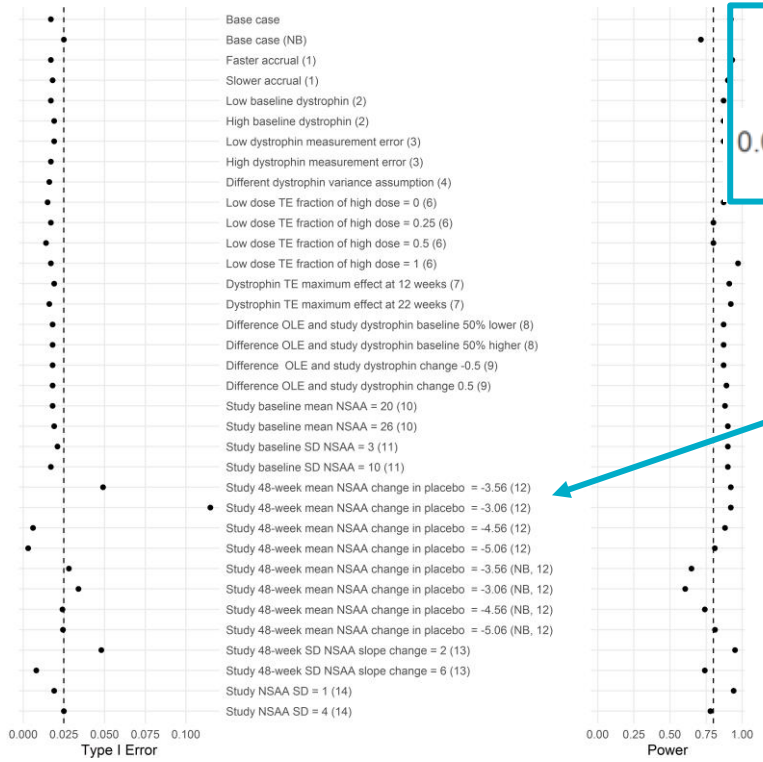
Effective sample size borrowed:
$N_h - N$

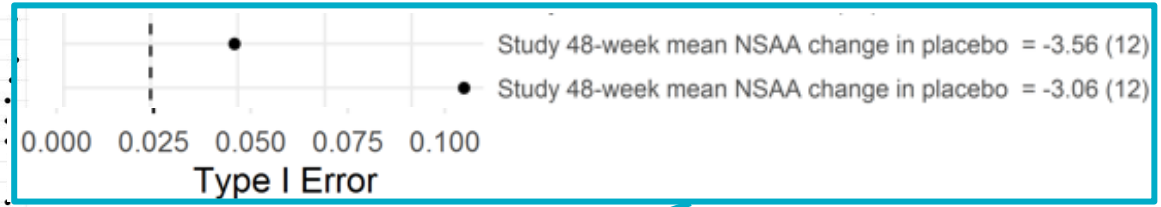


NSAA power


Power
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
NSAA Treatment Effect (% Slowing)
Borrowing
No Borrowing

Type I error and power for NSAA sensitivity analyses





Sensitivity analysis pertaining to prior distributions

- Data source random effect dictates the amount of historical data borrowing. If the variability between datasets is high, then there is less borrowing.

$$\exp(\delta_s) \sim Gamma\left(1/\sigma_\delta^2, 1/\sigma_\delta^2\right) \text{ for } s = 1, \ldots, S;$$
$$\sigma_\delta \sim Unif(0,1)$$
$$\mathrm{E}(\exp(\delta_s)) = 1$$
$$\mathrm{Var}(\exp(\delta_s)) = \sigma_\delta^2$$

- Sensitivity analysis explored the impact of changing the range of the uniform prior of σ_δ

Upper bound of U(0, b)	Type I error	Power
1	0.0160	0.896
2.0	0.0161	0.887
5	0.0169	0.882

CID experience

- Excellent opportunity for Wave to collaborate with FDA statistical leadership and the review division
  - Well attended meetings with statistical support across Offices/Divisions
  - Insights from reviewers with experience in different therapeutic areas
- CID meetings are not brainstorming sessions
- Timeline is tight, especially if also engaging other regulatory agencies
  - Meeting 1 at the end of January 2019
  - Briefing document for Meeting 2 was initially due two weeks after the first meeting
    - Meeting 2 delayed until June 2019 to provide time for simulations

Innovative trial design considerations

- Gaining access to historical data can be tough (and is getting tougher).
  - Critical Path Institute
- Endpoint instruments and training evolve over time, which can complicate use of historical data.
- Plenty of statistical alternatives exist. Choose a method and agree to sensitivity analyses.
  - Inverse probability weighting
- This is not BIO201; Complicated methods, simulations and presentations
- Extensive documentation required.
  - DMC charter, ISC charter, SAP, DAP
- Logistical challenges

Summary

- Adaptive design supported different trial objectives
- Placebo arm augmentation with historical data can be challenging
- CID program provided a productive sounding board for innovative trial design
  - Facilitated interactive exchange
  - Flexibility in response to emerging needs (2 teleconferences)
- Next time
  - Futility dystrophin analysis
  - More aggressive enrollment stopping
  - Group sequential option for NSAA
  - Study size constraints
- Innovative trial designs are the way forward for efficient drug development
  - 17 other trials recruiting DMD patients when DYSTANCE 51 started
- Lake, Quintana, et. al (2021, Statistics in Medicine)

Thank You

Stephen Lake

VP, Biometrics

Wave Life Sciences

Melanie Quintana

Scott Berry

Kristin Broglio

Jennifer Panagoulias

Michael Panzara