Analytical framework for assessment of long-term efficacy of therapies for Duchenne muscular dystrophy using external controls

Eugenio M. Mercuri[1]*; Damon Asher[2]; Kai Ding[2]; Matthew Furgerson[2]; Jianhua Jin[2]; Jing Li[2]; Xiao Ni[2]; James Signorovitch[3]; Linda P. Lowes[2]; Jacob S. Elkins[2]

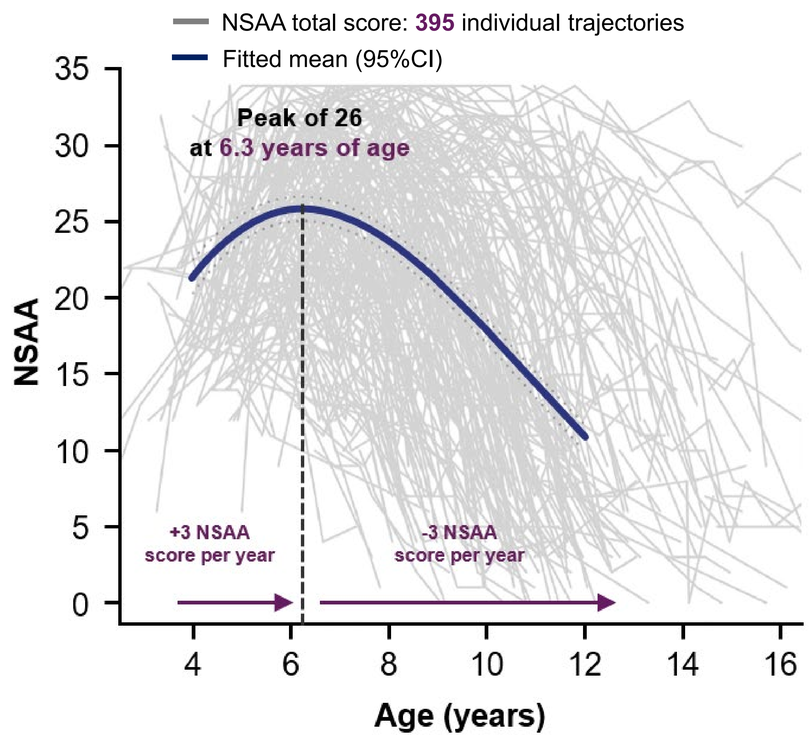
[1]Pediatric Neurology Institute, Catholic University and Nemo Pediatrico, Fondazione Policlinico Gemelli IRCCS, Rome, Italy; [2]Sarepta Therapeutics, Inc., Cambridge, MA, USA; [3]Analysis Group, Boston, MA, USA

*Presenting author

Background

Natural history and clinical trial placebo arms provide a potentially rich data source for predicting the course of untreated DMD. External controls are becoming increasingly important to elucidate longer-term outcomes[1-3]

- Due to the irreversible, progressive nature of DMD, it is against patients' interest to maintain placebo arms in clinical trials for extended periods of time
- This necessitates the use of external controls to assess long-term efficacy of treatments
- While extensive data on DMD progression in untreated patients are available, various sources of variability in these datasets may hinder their full application
- In DMD trials, use of external controls alongside appropriate statistical methods can help reduce bias, increase power, and facilitate more rigorous scientific inference[1-3]



Challenges with external control cohorts

- **Real-world studies comprise a large and diverse patient population**
- **Heterogeneous disease results in a variety of non-linear decline trajectories**
- **Standard of care may change over time or may differ between regions**
- **With time, nearly all untreated patients with DMD will lose function, resulting in a dominance of minimal scores ("floor effect")**
- **Data are often missing as patients age due to loss of function or missed visits**
- **With disease progression, score distribution is increasingly non-normal**

Objective

Here, we outline an analytical framework for comparing clinical trial data with external control data

Methods for external control cohort selection and adjustment of baseline characteristics

- **DMD disease heterogeneity** underscores the necessity for rigorous balancing of the external control population with the study population[1]
- **Pre-specified entry criteria** for the external control cohort are chosen to closely align with the treatment group's baseline characteristics
- **Logistic regression** can be used to estimate propensity scores used for the weighting procedure (eg, inverse probability treatment weighting[4]); patients in the external control cohort who are more like those in the treatment group will receive larger weights
- **Propensity score weighting** is implemented to ensure that key baseline characteristics of patients in the external control cohort are more comparable with those of the treatment group based on prognostic factors important to the specific population of the study (scan the QR code to access the **Supplement**)
  - To better contextualize outcomes of ambulatory patients with DMD, **baseline characteristics** that are important prognostic factors for loss of ambulation (eg, age and key functional outcomes) are utilized in the weighting procedure (see **Example 1**, as well as poster **P167**)

Example 1. Application of propensity score weighting: EMBARK PART 1[5]

Baseline characteristic, Mean (range)	EMBARK Part 1 delandistrogene moxeparvovec treatment n=64	EC cohort (before PSW) n=155	EC cohort (after PSW) n=143	Standardized mean difference EMBARK Part 1 vs EC cohort (after PSW)
Age, years	6.0 (4.1, 7.9)	6.4 (4.2, 8.0)	6.2 (4.2, 8.0)	-0.281
Weight, kg	21.2 (13.5, 37.4)	21.2 (14.0, 36.0)	21.2 (14.0, 36.0)	-0.198
Height, cm	108.6 (93.5, 127.0)	110.9 (94.9, 131.1)	110.6 (94.9, 131.1)	-0.285
NSAA total score	23.3 (14, 32)	26.0 (15, 32)	23.5 (15, 32)	-0.045
Time to rise from floor, s	3.5 (1.8, 5.8)	3.7 (1.9, 5.7)	3.5 (1.9, 5.7)	-0.011
Time of 10MWR, s	4.8 (3.2, 6.8)	4.8 (3.0, 6.7)	4.8 (3.0, 6.7)	0.034

Baseline characteristics of external controls who met entry criteria, before and after propensity score weighting, compared with baseline characteristics of delandistrogene moxeparvovec–treated patients in Part 1 of the EMBARK trial[5]

10MWR, 10-meter walk/run; EC, external control; NSAA, North Star Ambulatory Assessment; PSW, propensity score weighting.

Methods for use of various statistical models for the evaluation of efficacy; improved comparison through sensitivity analyses and additional external controls

- **Mixed-effects models with repeated measures** can be used for data with normal or close-to-normal distribution, provided that relatively few patients experienced a significant functional loss
- **Median regression models** can be used for non-normally distributed or skewed data (eg, if the follow-up includes ages at which functional decline and loss of ambulation are anticipated)
- A broad set of studies as sources of external control cohorts may be better than a single study; cross-validation and replication through an **independent additional external control** can provide an important test of the strength of evidence
  - Comparison of the external control utilized in the pooled 3-year functional analysis of delandistrogene moxeparvovec clinical trials with an independent natural-history dataset composed of patients receiving current standard of care showed that the external control used in the pooled analysis was representative of the DMD natural history population; change from baseline to Year 3 in NSAA total score of the two external control groups shows comparable disease courses (see **Example 2** in **Supplement**)
- **Sensitivity analyses,** in which the original protocol for adjustment of baseline characteristics is varied by adding or removing prognostic factors, can result in models with an improved convergence

Minimization of the impact of "missing" data due to loss of function and the "floor effect"

- In patients with loss of function (NSAA total score of zero), **missing data** resulting from inability to complete functional assessments can be replaced (imputed) using values derived from other available information and a set of logical rules ("logic-based imputation")
  - For example, for patients unable to rise from the floor or walk/run 10 meters, a zero can be imputed for RFF velocity (rises/s) or 10MWR velocity (m/s) and 30 s can be imputed for RFF time or 10MWR time to ensure patients not able to perform a task can be retained in the analysis
- **"Floor effect,"** an important limitation in DMD, especially during the late ambulatory stage of disease, can be minimized by transforming the data (depending on the outcome chosen) or by converting continuous outcomes into discrete outcomes (eg, using responder analysis instead of score changes)
- **In DMD, missing data and "floor effect"** often coincide with disease progression: patients are more likely to "miss" assessments due to their inability to perform tasks (eg, 10MWR for non-ambulatory individuals)

Conclusion(s)

In Duchenne muscular dystrophy clinical trials, when long-term follow-up in a placebo arm is against patients' interest, comparing long-term functional outcomes from treated patients with those of a well-matched external control cohort can be an effective approach, provided that suitable patient populations and robust statistical methods are employed

Acknowledgments & Disclosures

Acknowledgments: Editorial support was provided by Vojislav Pejovic, PhD (HCG) in accordance with Good Publication Practice (GPP) 2022 guidelines (https://www.ismpp.org/gpp-2022) with funding from Sarepta Therapeutics, Inc., Cambridge, MA, USA. Studies 101, 102, and ENDEAVOR were sponsored and funded by Sarepta Therapeutics, Inc., Cambridge, MA, USA, and also funded by F. Hoffmann-La Roche Ltd, Basel, Switzerland. The authors would like to thank the patients, their families, the investigators, and the trial staff for their participation in the studies.

Disclosures: EMM has received fees from AveXis, Biogen, and F. Hoffmann-La Roche Ltd.; **JS** is a full-time employee of Analysis Group, Inc., a consulting firm that has received payment from Sarepta Therapeutics; **DA**, **KD**, **MF**, **JJ**, **JL**, **XN**, **LPL**, and **JSE** are employees of Sarepta Therapeutics and may have stock options.

Abbreviations: 10MWR, 10-meter walk/run; **DMD**, Duchenne muscular dystrophy; **NSAA**, North Star Ambulatory Assessment; **PSW**, propensity score weighting; **RFF**, rise from floor.

References

1. Muntoni F, et al. *Neuromuscul Disord*. 2022;32(4):271-283.
2. Goemans N, et al. *Neurology*. 2020;95(10):e1381-e1391.
3. FDA. Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products. Draft Guidance. February 2023. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-and-conduct-externally-controlled-trials-drug-and-biological-products. Accessed: February 2025.
4. Austin PC. *Multivariate Behav Res*. 2011;46(3):399-424.
5. Mendell JR et al. *Nat Med*. 2025;31(1):332-341.

SCAN THE QR CODE

To access the full poster on your mobile device, including any supplementary materials, please scan the code using your QR reader application. NB: There may be associated costs for downloading data. These costs may be high if you are using your smartphone abroad. Please check your mobile data tariff or contact your service provider for more details.

Presented at the 2025 MDA Clinical & Scientific Conference

March 16-19, 2025; Dallas, TX, USA

Analytical framework for assessment of long-term efficacy of therapies for Duchenne muscular dystrophy using external controls

Eugenio M. Mercuri[1]*; Damon Asher[2]; Kai Ding[2]; Matthew Furgerson[2]; Jianhua Jin[2]; Jing Li[2]; Xiao Ni[2]; James Signorovitch[3]; Linda P. Lowes[2]; Jacob S. Elkins[2]

[1]Pediatric Neurology Institute, Catholic University and Nemo Pediatrico, Fondazione Policlinico Gemelli IRCCS, Rome, Italy; [2]Sarepta Therapeutics, Inc., Cambridge, MA, USA; [3]Analysis Group, Boston, MA, USA

*Presenting author

Supplementary information

Figure S1. Key considerations for using external control data

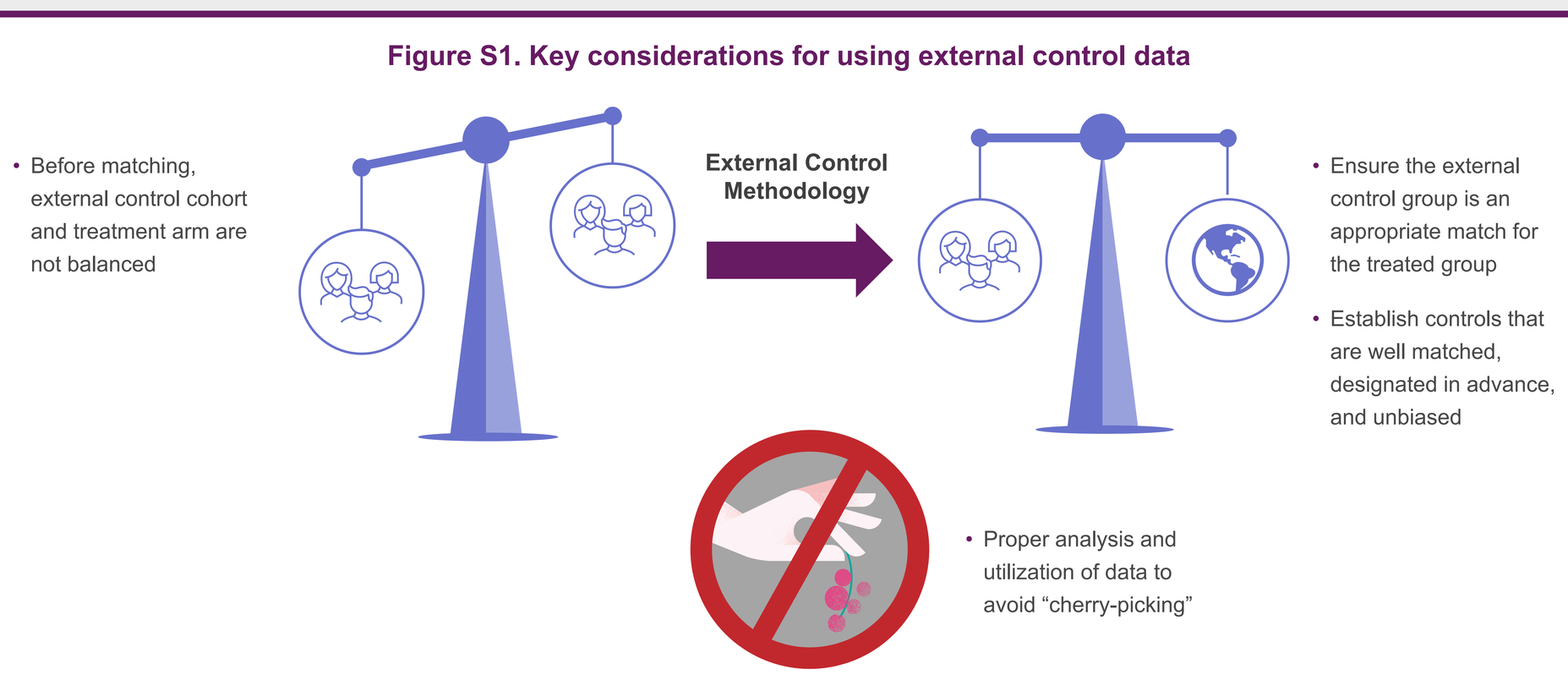
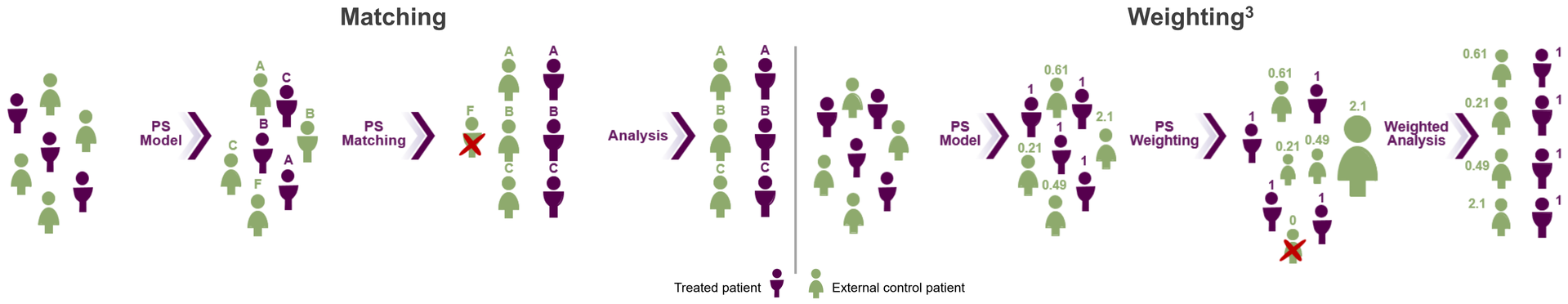


Figure S2. Propensity score matching vs propensity score weighting[1,2]



Matching

- External control patient(s) matched with a corresponding treated patient
- Utilized with a large sample size
- Average treatment effects calculated using data from included patients only

Weighting[3]

- Each external control patient assigned a weight based on the degree of similarity to treated group
- Can be used with smaller sample sizes
- Data from all external-control and treated patients are retained
- Matching efficiency is enhanced by eliminating external control patients whose propensity scores fall outside the range of the treated patients’ propensity scores (this range is referred to as the common support region)
- Average treatment effects calculated using data from all patients in the common support region[4]

Example 2: Application of an independent natural history cohort to cross-validate the external control used in the pooled 3-year functional analysis of delandistrogene moxeparvovec trials

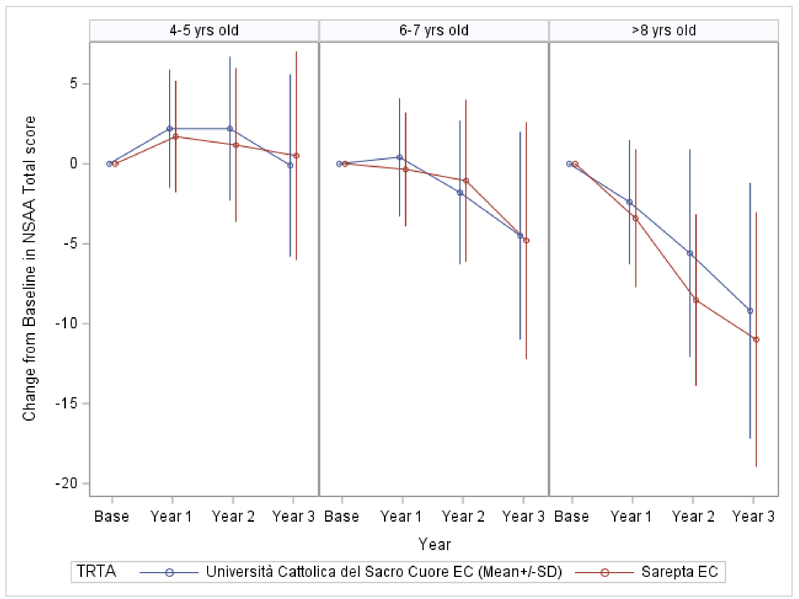
- Comparison of baseline characteristics (**Table S1**) and change from baseline in NSAA Total score (**Figure S3**) between the external control cohort data utilized in the pooled analyses 3-year vs EC (after pre-entry inclusion criteria and before propensity score weighting) and an independent dataset from the Università Cattolica del Sacro Cuore composed of patients receiving current standard-of-care

Table S1. Comparison of baseline characteristics

Baseline characteristic	Università Cattolica del Sacro Cuore (N=106)	Sarepta EC (N=174)
Age, years		
Mean (SD)	6.7 (1.2)	6.5 (1.0)
Median [Q1, Q3]	6.5 [5.8, 8.0]	6.42 [5.8, 7.3]
NSAA total score		
Mean (SD)	24.2 (4.8)	24.3 (4.3)
Median [Q1, Q3]	25.5 [21.0, 28.0]	25 [21.0, 28.0]
Time to rise from floor, s		
Mean (SD)	4.59 (1.84)	4.54 (1.82)
Median [Q1, Q3]	4.27 [3.20, 5.30]	4.16 [3.20, 5.30]
Rise from floor velocity, 1/s		
Mean (SD)	0.25 (0.10)	0.25 (0.09)
Median [Q1, Q3]	0.23 [0.19, 0.31]	0.24 [0.19, 0.31]
Time of 10MWR, s		
Mean (SD)	5.64 (1.54)	5.13 (1.01)
Median [Q1, Q3]	5.54 [4.59, 6.75]	5 [4.30, 5.90]
10MWR velocity, m/s		
Mean (SD)	2.12 (0.24)	2.02 (0.38)
Median [Q1, Q3]	1.96 [1.71, 2.20]	2 [1.69, 2.33]

10MWR, 10-meter walk/run; EC, external control; NSAA, North Star Ambulatory Assessment.

Figure S3. Comparison of the change from baseline in NSAA total score



EC, external control; NSAA, North Star Ambulatory Assessment; SD, standard deviation.

Conclusion(s)

In Duchenne muscular dystrophy clinical trials, when long-term follow-up in a placebo arm is against patients’ interest, comparing long-term functional outcomes from treated patients with those of a well-matched external control cohort can be an effective approach, provided that suitable patient populations and robust statistical methods are employed

Acknowledgments & Disclosures

Acknowledgments: Editorial support was provided by Vojislav Pejovic, PhD (HCG) in accordance with Good Publication Practice (GPP) 2022 guidelines (https://www.ismpp.org/gpp-2022) with funding from Sarepta Therapeutics, Inc., Cambridge, MA, USA. Studies 101, 102, and ENDEAVOR were sponsored and funded by Sarepta Therapeutics, Inc., Cambridge, MA, USA, and also funded by F. Hoffmann-La Roche Ltd, Basel, Switzerland. The authors would like to thank the patients, their families, the investigators, and the trial staff for their participation in the studies.

Disclosures: EMM has received fees from AveXis, Biogen, and F. Hoffmann-La Roche Ltd.; **JS** is a full-time employee of Analysis Group, Inc., a consulting firm that has received payment from Sarepta Therapeutics; **DA**, **KD**, **MF**, **JJ**, **JL**, **XN**, **LPL**, and **JSE** are employees of Sarepta Therapeutics and may have stock options.

Abbreviations: 10MWR, 10-meter walk/run; **DMD**, Duchenne muscular dystrophy; **NSAA**, North Star Ambulatory Assessment; **PSW**, propensity score weighting.

References

1. Nimmo A, et al. *Transpl Int*. 2022;35:10105.
2. Kuss O, et al. *Dtsch Arztebl Int*. 2016;113(35-36):597-603.
3. Austin PC. *Multivariate Behav Res*. 2011;46(3):399-424.
4. Rosenbaum PR and Rubin DB. *Biometrika*. 1983;70(1):41-55.

SCAN THE QR CODE

To access the full poster on your mobile device, including any supplementary materials, please scan the code using your QR reader application. NB: There may be associated costs for downloading data. These costs may be high if you are using your smartphone abroad. Please check your mobile data tariff or contact your service provider for more details.

Presented at the 2025 MDA Clinical & Scientific Conference

March 16-19, 2025; Dallas, TX, USA