

# Investigating survivorship bias in disease progression modelling for Duchenne muscular dystrophy

Authors: Stephen Duffull (1), Anna Largajolli (1), Francesco Bellanti (1), Federica Alessi (2), Sara Cazzaniga (2), Paolo Bettica (2)

<sup>1</sup>Certara, NJ, USA

<sup>2</sup>Italfarmaco SpA., Milan, Italy

## Background and Objective

- Duchenne muscular dystrophy (DMD) is a severe progressive muscular disease caused by mutations in the X-linked gene encoding for the dystrophin protein. DMD affects both skeletal and smooth muscle and is associated with degenerative loss of function.
- Onset of symptoms occurs at an early age and most children are unable to walk by 12 years, life expectancy 20-30 years.
- The NorthStar Ambulatory Assessment (NSAA) can be used to assess disease status, scores range from 0 (non-functional) to 34 (normal functional status)
- A disease progression (DP) model has been developed previously for DMD by the C-Path's Duchenne Regulatory Science Consortium (D-RSC) [1].
- DP was found to be delayed for patients with higher baseline score and for older patients which contrasts with the general heuristic understanding of DMD.
- Survivorship bias occurs when the sample is enriched with survivors (e.g. slow disease progression) as non-survivors do not meet inclusion criteria

## Objectives

- To investigate whether the apparent delay in onset of disease progression for DMD in older children could be a function of survivorship bias.

## Disease Progression model for DMD

- The D-RSC model has the general form

$$Score_i = Gmax \cdot \left(1 - \exp(-g_i \cdot age_i)\right) \cdot \left(1 - \frac{DPmax_i \cdot age_i^{\gamma_i}}{DP_{50_i}^{\gamma_i} + age_i^{\gamma_i}}\right) \quad (1)$$

$$DP_{50_i} = \theta_{DP50} \cdot \left(\frac{BAGE}{8.78}\right)^{\theta_{BAGE}} \cdot \left(\frac{BSCORE}{23}\right)^{\theta_{BSCORE}} \quad (2)$$

$$\theta_{BAGE} = 0.654, \theta_{BSCORE} = 0.625$$

- $BAGE$  = age at start of trial,  $BSCORE$  = score at start of trial.
- $\gamma = 6.55$ ,  $g = 0.375$ ,  $DPmax = 0.904$ ,  $Gmax=24$  (fixed)

## Understanding the model characteristics

- The purpose of a disease progression model is to provide a continuous description of the disease status of any patient after some index time (e.g. disease start)
  - Disease progression models provide a longitudinal solution
- Disease progression models are therefore ideally constructed from longitudinal data in which a patient is followed throughout their whole course of disease – which is generally impractical
  - Therefore these models are often constructed from short longitudinal or cross-sectional studies
- A delay in the disease progression with baseline age and score will translate to longer survival times as patients age – however the data may only speak to longer survival times based on those who were selected to be in the study
- The question therefore is whether the covariate effects of  $BAGE$  and  $BSCORE$  on  $DP_{50}$  are true or an artefact of the study design?

## Testing for artefactual covariate relationships

- A simulation study was conducted in which the true covariate relationships were known. The simulation study here does not necessarily reflect the settings from the original study.

## Simulation model

- The structural model and parameter values were the same as equation (1). No covariates were included against  $DP_{50}$  equation (2), hence  $DP_{50} = \theta_{DP50}$
- All structural and random effects model were based on [1]
- Simulations were performed in R without any covariate relationships. Estimation using SAEM + IMP in NONMEM 7.5.1 using the structural and covariate models in equations 1 and 2.

Table 1: Study designs

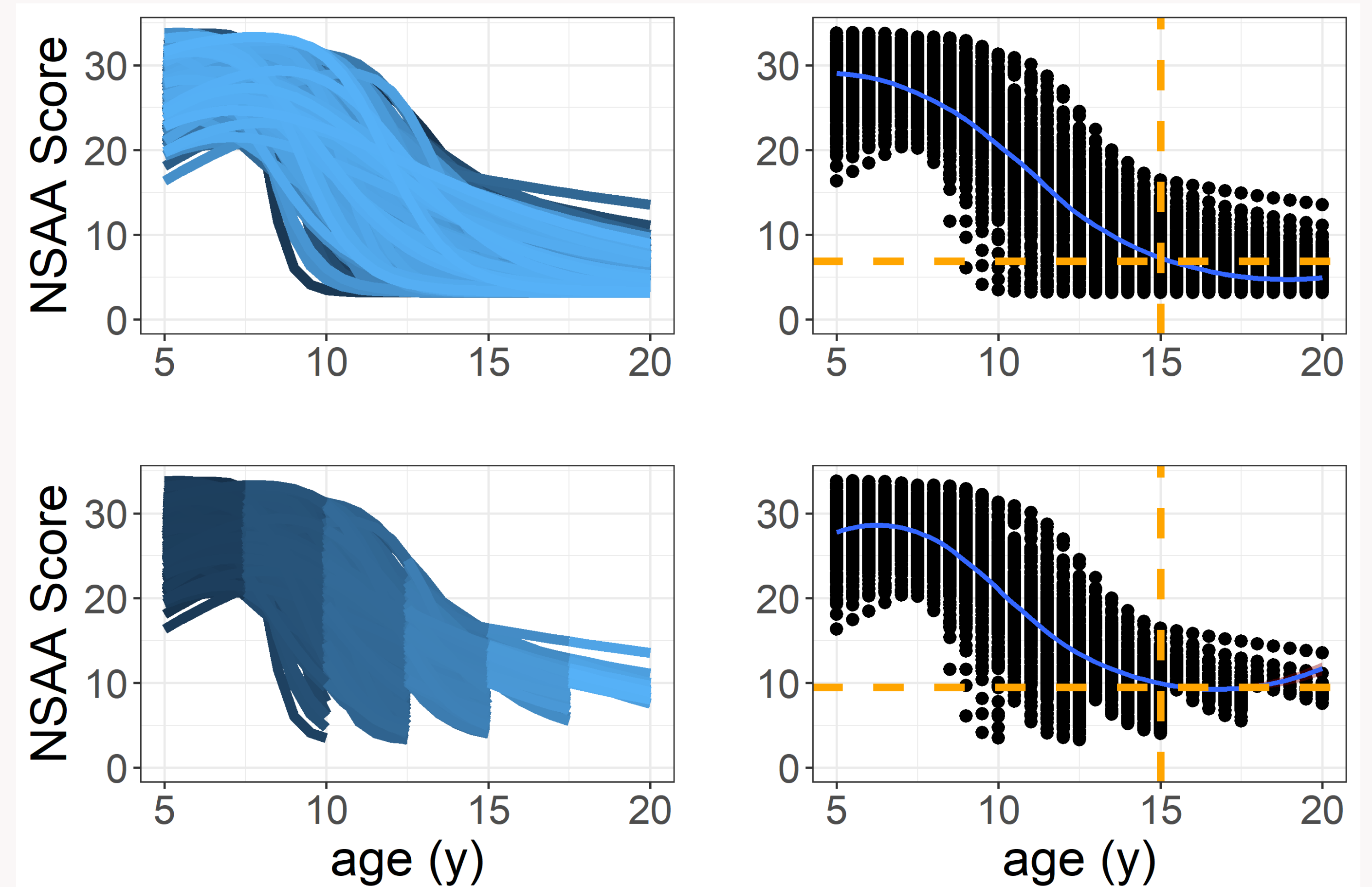
Study A = Long Duration Study	Study B = Short Duration Study
All patients were enrolled at 5 years of age and followed to 20 years or until NSAA < 3	Patients were enrolled in 6 cohort groups with baseline ages at 5, 7.5, 10, 12.5, 15 and 17.5 years of age and followed for 2.5 years or until NSAA < 3.
The single cohort included 250 patients with 31 observations per patient	Each cohort included 250 patients with 6 observations per patient
Minimum baseline score at start of study = 10	Minimum baseline score at start of study = 10
NSAA measurements were taken every 6 months	NSAA measurements were taken every 6 months

## Acknowledgements

The research team acknowledges the use of the [Duchenne Muscular Dystrophy \(DMD\) Clinical Trial Simulation tool](#), developed by C-Path's Duchenne Regulatory Science Consortium (D-RSC), as a key resource in supporting our clinical trial design and/or analysis

## Results

Figure 1: Simulated data



Notes. The top row is a simulation from Study A and the bottom row from Study B. The left-hand panels are the individual data as spaghetti plot and right-hand as individual data points. The orange dashed line shows the average prediction of NSAA score at 15 years. For the longitudinal study the y-intercept is 6.9 years and for study B the y-intercept is 9.5 years

## Data Simulation

- Study A consists of a single long-cohort study. Study B consists of 6 short term cohort studies.
- The expected NSAA score at 15 years for Study B was 9.5 years compared to 6.9 years for Study A
- Study B predicts a better outcome at 15 years of age of almost 2.5 years

## Data Estimation

- The  $BAGE$  and  $BSCORE$  covariate effects were not able to be estimated for Study A as the model was unstable
- The estimated parameter values from the original study [1] and from study B are shown in Table 2 (and were associated with a significant drop in the objective function value. .

Table 2: Covariate parameter estimates on  $DP_{50}$

	True values	Study B estimates (RSE%)
BSCORE	0	0.339 (5.3%)
BAGE	0	0.493 (3.2%)

## Limitations / Discussion / Inference

### Limitations

- This work does not replicate the cohorts present in the original analysis of [1]. Therefore the parameter estimates for the covariate effects and any other parameters are not necessarily representative of the true disease progression.
  - A full review of the DMD model with validation is provided in Largajolli [2]
- We considered only 1 set of designs for Study B which are not necessarily those that may be conducted in future studies nor are they suggested to be appropriate designs for DMD studies.
- The inclusion/exclusion details used here are not necessarily representative of those for a clinical study.

### Discussion

- This work highlights that covariate relationships may be found that are not necessarily explicable based on known disease progression (e.g. older patients have slower disease progression).
- The purpose of this work is to illustrate that covariate effects may be seen even if they are not included in the true model and that these covariate effects may reflect the study design.
- In this case the finding that baseline AGE and SCORE were positively correlated with better outcomes is a feature that aligns with survivorship bias (i.e. that the patients were eligible to enrol in the study indicates they are "survivors" – perhaps those with slow disease progression).

### Inference

- It is possible but not explored here that the shorter the longitudinal study (e.g. the extrema case of a cross-sectional study) and more strict the inclusion/exclusion criteria may amplify these findings.

## References

- Lingineni et al. CPT Pharmacometrics Syst Pharmacol. 2022;11:318–332
- Largajolli et al. Validating the D-RSC disease progression modelling framework for Duchenne muscular dystrophy and evaluating the long-term impact of Givinstat therapy in pediatric patients. PAGE Poster 2025



Want to learn more?  
 << Scan Here