The Company of Biologists

RESEARCH ARTICLE

Natural history study and statistical modeling of disease progression in a preclinical model of myotubular myopathy

Suzie Buono[1,*], Arnaud Monseur[2,*], Alexia Menuet[1], Anne Robé[1], Catherine Koch[1], Jocelyn Laporte[3], Leen Thielemans[1,4], Marion Depla[1] and Belinda S. Cowling[1,‡]

ABSTRACT

Generating reliable preclinical data in animal models of disease is essential in therapy development. Here, we performed statistical analysis and joint longitudinal–survival modeling of the progressive phenotype observed in $Mtm1^{-/y}$ mice, a reliable model for myotubular myopathy. Analysis of historical data was used to generate a model for phenotype progression, which was then confirmed with phenotypic data from a new colony of mice derived via *in vitro* fertilization in an independent animal house, highlighting the reproducibility of disease phenotype in $Mtm1^{-/y}$ mice. These combined data were used to refine the phenotypic parameters analyzed in these mice and improve the model generated for expected disease progression. The disease progression model was then used to test the therapeutic efficacy of *Dnm2* targeting. *Dnm2* reduction by antisense oligonucleotides blocked or postponed disease development, and resulted in a significant dose-dependent improvement outside the expected disease progression in untreated $Mtm1^{-/y}$ mice. This provides an example of optimizing disease analysis and testing therapeutic efficacy in a preclinical model, which can be applied by scientists testing therapeutic approaches using neuromuscular disease models in different laboratories.

This article has an associated First Person interview with the joint first authors of the paper.

KEY WORDS: Centronuclear myopathy, Myotubular myopathy, Preclinical disease model, Dynamin, Antisense oligonucleotide, Therapy, Neuromuscular disorder

INTRODUCTION

In a research setting, mouse models of disease provide an excellent tool to investigate disease pathophysiology and test therapeutic approaches. Mouse lines that recapitulate the disease phenotype can be used to identify novel therapeutic targets, and map in a temporal and dose-dependent manner the response to treatment. Animal models

[1]Dynacure, Illkirch 67400, France. [2]Pharmalex, Mont-St-Guibert 1435, Belgium. [3]Institut de Génétique et de Biologie Moléculaire et Cellulaire, Institut National de la Santé et de la Recherche Médicale U1258, Centre National de la Recherche Scientifique UMR7104, Université de Strasbourg, 67404 Illkirch, France. [4]2 Bridge, Zoersel 2980, Belgium.
*These authors contributed equally to this work

‡Author for correspondence (belinda.cowling@dynacure.com)

S.B., 0000-0001-5694-2249; J.L., 0000-0001-8256-5862; M.D., 0000-0003-4651-9302; B.S.C., 0000-0001-6915-2596

may also be useful in identifying and testing diagnostic, prognostic or therapeutic biomarkers. Pathological phenotypes in animals may be compared with the human disease, to help gain insight into the underlying pathophysiology that is relevant for patients, and to test the potential of various therapeutic modalities targeting the mechanism of action, including small molecules (chemically derived) and various biologics (e.g. gene and cell therapies).

Many potential therapies for neuromuscular diseases have moved from proof-of-concept in animal studies towards clinical trials in the past decade (Cowling and Thielemans, 2019). Often, proof-of-concept is generated using preclinical animal models of neuromuscular disease; however, translating preclinical data to a clinical setting is often challenging. Mouse models of human disease have limitations, which may be linked to the relatively uniform genetic background of animal cohorts compared to humans, the potentially complex genetic involvement in inherited neuromuscular disorders (causative gene and epistatic mutations) and the targeting of pathogenic pathways, which might not be similar between species. Considering these limitations is important when interpreting animal data. Optimizing the generation of relevant, reliable and reproducible preclinical data on disease phenotype and therapeutic potential is of high importance. Natural history studies in mice are a useful way to understand disease progression and standardize phenotyping parameters across studies. This has been performed in recent years in mdx mice, a frequently used mouse line to investigate Duchenne muscular dystrophy (van Putten et al., 2019; Gordish-Dressman et al., 2018).

In this study, we focus on a preclinical animal model for myotubular myopathy. Myotubular myopathy is a non-dystrophic, debilitating rare congenital disease, associated with muscle weakness and abnormally located nuclei and other organelles in skeletal muscle (Ravenscroft et al., 2015). Myotubular myopathy [also called X-linked centronuclear myopathy (CNM), XLCNM, XLMTM; OMIM 310400] is due to mutations in the phosphoinositide phosphatase myotubularin (MTM1) (Laporte et al., 1996). There are an estimated 2650 living patients with myotubular myopathy, in the USA, EU, Japan and Australia (Vandersmissen et al., 2018). $Mtm1^{-/y}$ mice recapitulate myotubular myopathy; as observed in patients, mice display a severe myopathic phenotype and a reduced lifespan (Buj-Bello et al., 2002). This model has been used across laboratories from several continents, and has been used to provide the therapeutic proof-of-concept and supportive data contributing to the rationale for initiating several clinical trials (NCT04915846, NCT03199469, NCT04033159), thus highlighting the importance of data generated from this model. Previous studies suggested that increased DNM2 was largely responsible for the CNM phenotype observed in mice and patients (Cowling et al., 2014, 2011; Liu et al., 2011; Massana Munoz et al., 2019). Reduction of DNM2 was shown to rescue myotubular myopathy in mice ($Mtm1^{-/y}$ mice) by genetic cross (Cowling et al., 2014) or systemic delivery of antisense oligonucleotides (ASOs)

Disease Models & Mechanisms

(Tasfaout et al., 2017; Koch et al., 2020), or by reducing DNM2 using an adeno-associated virus-mediated shRNA approach targeting *Dnm2* (Tasfaout et al., 2018).

The goal of this study was to describe disease progression and the variability with which it can occur, and to inform which phenotypic parameters are sufficient to predict disease progression in $Mtm1^{-/y}$ mice. We performed statistical analysis and modeling of the progressive phenotype observed in $Mtm1^{-/y}$ mice from retrospective ('training') and prospective ('test') data generated in $Mtm1^{-/y}$ mice. These combined data were then used to improve both the phenotyping analysis performed in these mice and the model generated for normal disease progression. Finally, this model was tested and validated by performing a therapeutic dose–response study targeting *Dnm2*.

RESULTS
Natural history study analysis and model generation from training cohort

To generate a statistical model of disease progression in myotubular myopathy mice, we first focused on performing an analysis of previous natural history data generated from $Mtm1^{-/y}$ mice across several studies (Koch et al., 2020), referred to as the 'training cohort'. Analysis of 38 mice from weaning identified the normal survival curve for $Mtm1^{-/y}$ mice, with loss of survival starting from 3-4 weeks of age and no mice surviving past 8 weeks of age (Fig. 1A). Analysis of body weight progression in mice showed that the average weight of 3-week-old mice at the start of the study was 10.31±1.27 g (mean±s.d.). In the majority of mice, this increased from weeks 3-4, followed by a decline in body weight, with no mice reaching 20 g in this study (Fig. 1B). Next, a joint model was generated from the raw body weight data, considering both survival and evolution of body weight. Based on this model, the line of best fit and prediction intervals for the expected range of body weights for $Mtm1^{-/y}$ mice were modeled (Fig. 1B: black line, gray shadow, respectively).

Hanging time upside down from a cage lid with four paws was used as an indicator of whole-body strength in mice, and averaged ~50 s at 3 weeks of age in $Mtm1^{-/y}$ mice (Fig. 1C). From 5 weeks of age, almost no mice could perform the test for more than 5 s, compared to wild-type (WT) mice, which can normally hang for the maximum time tested, 60 s (Koch et al., 2020), at all timepoints analyzed. As statistical modeling for prediction to perform the test in this format was not feasible, we performed a time to event analysis, selecting hanging time ability cutoff times of 50 s (Fig. 1D) or 10 s (Fig. 1E). At 3 weeks of age, ~60% of $Mtm1^{-/y}$ mice could hang for more than 50 s, whereas by 5 weeks of age almost no mice could perform this test for more than 10 s (Fig. 1D,E), indicating rapid and severe progression of the myopathic phenotype during this age bracket.

To capture key disease elements and monitor severity and progression of disease in $Mtm1^{-/y}$ mice, previously a disease severity score (DSS) was created (Tasfaout et al., 2017) (Table S1). The DSS was designed to capture key phenotypic elements of the myopathy phenotype apparent in mice. This focused primarily on six key factors: body weight, hanging test (whole-body strength), kyphosis (curvature of the spine), walking ability (hindlimb muscle weakness), ptosis (eyelid muscle weakness) and difficulties in breathing (Table 1). DSS analysis was performed on training cohort data, and modeling was performed to identify the line of best fit and prediction interval of expected scores (Fig. 1F). Mice exhibited an average DSS of 0.88±0.55 (mean±s.d.) at 3 weeks of age, which progressed to 4.44±0.42 (mean±s.d.) by 8 weeks of age, suggestive
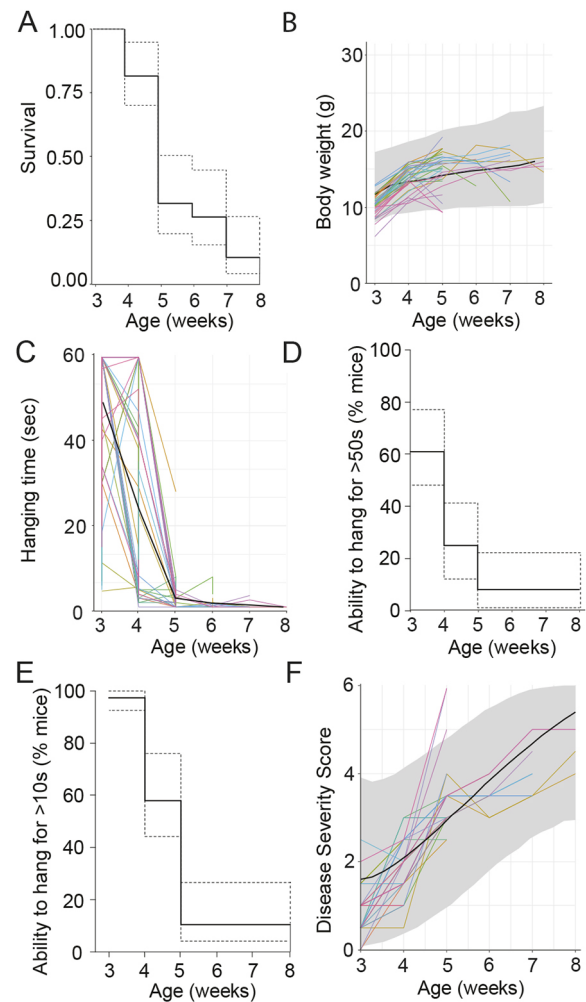

Fig. 1. Natural history study analysis and model generation from training cohort. (A) Survival of mice, line of best fit (solid line) and 95% confidence interval (CI) (dashed lines) are shown. (B) Body weight (g) progression weekly in $Mtm1^{-/y}$ mice. Individual mice (colored lines), line of best fit (black line) and prediction interval (shaded gray zone) are highlighted on the graph. (C) Hanging time (60 s maximum), with individual mouse progression (colored lines) and average (black line) shown. (D,E) Time to event analysis of hanging test, with cutoff times of 50 s (D) and 10 s (E) displayed. Graphs represent line of best fit (solid line) ±95% CI (dashed lines). (F) Disease severity score (DSS), with a score between 0 (unaffected) and 6 (most severely affected) awarded per mouse per week, based on six different phenotypes (see Table 1 for details). Modeling was performed to identify the line of best fit (black line) and prediction intervals (gray shaded area) for expected disease progression in $Mtm1^{-/y}$ mice. All data in this figure represent analysis of $Mtm1^{-/y}$ mouse phenotypes from the training cohort, starting from 3 weeks of age; individual mice represented as colored lines. n=38 male mice.

of a severe disease phenotype in $Mtm1^{-/y}$ mice surviving until 8 weeks of age. Overall, analysis of survival and phenotyping data from the training cohort confirm drastically reduced survival and a severe progressive myopathic phenotype in $Mtm1^{-/y}$ mice from weaning until 8 weeks of age.

Natural history study in test cohort of myotubular myopathy mice

Following on from analysis of historical data in the training cohort, we next performed a natural history study in 20 $Mtm1^{-/y}$ mice after weaning, designated the 'test cohort'. Of note, this colony was

Table 1. Disease severity scoring system factors

Disease scoring factor	Previous DSS – description (from Tasfaout et al., 2017)	Score	Optimized DSS – description	Score
Body weight	Difference in body weight between $Mtm1^{-/y}$ mouse and wild-type littermate	0: 0-1 g 0.5: >1-2 g 1: >2 g	Difference in body weight in $Mtm1^{-/y}$ mouse between week n and n+1	0: ≥0.25 g gain 0.5: ±0-0.25 g* 1: ≥0.25 g loss
Breathing difficulties	Frequency and amplitude qualitative evaluation based on observations	0: Normal 1: Altered	Not included	
Hanging test ability	Ability to hang from the lid of a cage for 60 s	0: 60 s 0.5: 5-60 s 1: <5	As before, but with a graduated score of 0-2, representing 0-60 s hanging time	[60−time (s)]/60 s×2=value (0-2) Examples: 1 s→(60−1)/60×2=1.97 15 s→(60−15)/60×2=1.5 30 s→(60−30)/60×2=1 60 s→(60−60)/60×2=0
Kyphosis	Curvature of the spine	0: No 0.5: Mild 1: Severe	Unchanged	
Ptosis	Drooping of the eyelid	0: No 1: Yes	Not included	
Walking difficulties	Ability to use hindlimbs	0: Normal 0.5: Splayed 1: Loss of use	Unchanged	
Maximum score		6		5

Previous disease severity scoring system (Tasfaout et al., 2017) and optimized disease severity scoring system based on National Health Service (NHS) data analysis and modeling from training and test cohorts of $Mtm1^{-/y}$ mice. The full standard operating procedure can be found in the Supplementary Materials and Methods. *Only relevant during growth phase of a mouse, from 0 to 12 weeks. DSS, disease severity score.

generated by *in vitro* fertilization (IVF) from samples taken from $Mtm1^{-/y}$ mice from the colony used to generate the training cohort. IVF and colony generation were performed in a separate animal housing facility at a different location, thus more accurately representing the comparison of preclinical research experiments by multiple research laboratories in different locations.

Analysis of survival in the test cohort of $Mtm1^{-/y}$ mice was performed. Loss of survival was observed from 3 to 4 weeks of age, and no mice survived past 12 weeks of age (Fig. 2A), suggesting more variation in survival in the test cohort compared to the training cohort (Fig. 1A). Based on the body weight data generated in the training cohort (based on raw data up to 8 weeks of age), a model was generated to display the line of best fit (median profile) and prediction interval expected for body weight progression in $Mtm1^{-/y}$ mice (Fig. 1B). Body weight was then analyzed weekly in the test cohort and overlaid on the model generated from training data, including extrapolation of expected prediction interval for 8-12 weeks of age. All mice fell within the expected weight range for $Mtm1^{-/y}$ mice over the period analyzed (3-12 weeks of age), confirming the reproducibility of body weight progression in this mouse line (Fig. 2B). Consistent with the training colony, no mice reached 20 g body weight.

To analyze whole-body strength, the hanging test was then performed in the test cohort of $Mtm1^{-/y}$ mice. As observed in the training cohort (Fig. 1C), hanging time rapidly declined in mice from 3 to 5 weeks of age (Fig. 2C). A statistical comparison between training and test cohorts was performed using a time to response analysis (Fig. 2D,E), confirming the reliability of this test as an indicator of whole-body strength in $Mtm1^{-/y}$ mice.

Finally, the combined DSS was analyzed. Based on the model generated from the training cohort (Fig. 1F), which included extrapolation of expected prediction interval for 8-12 weeks of age, all mice fell within the expected DSS over the period analyzed (Fig. 2F, 3-12 weeks of age), confirming the reproducibility and reliability of DSS progression in this mouse line, and the analytical validity of the model generated.

Optimized disease severity analysis based on combined natural history study data from training and test cohorts

Based on the large volume of data generated above, we next aimed to optimize the disease severity analysis performed in $Mtm1^{-/y}$ mice, by focusing on a statistical modeling approach. First, combined analysis of the survival of $Mtm1^{-/y}$ mice highlighted that deaths could occur from the age of weaning at 21 days (youngest age analyzed in this study), until 12 weeks of age, with less than 50% of $Mtm1^{-/y}$ mice surviving past 5 weeks of age, compared to 100% of WT mice (Fig. 3A).

Next we investigated the disease severity in the combined training and test cohorts. The DSS focuses on six factors – body weight, hanging test, kyphosis, walking ability, ptosis and difficulties breathing – as described in Table 1. The progression of each of the individual six factors of the combined training and test mouse cohorts is shown in Fig. 3B (average values shown). Although body weight, hanging, walking and kyphosis reflected early disease progression in $Mtm1^{-/y}$ mice from 3 to 7 weeks of age, breathing and ptosis did not reflect disease progression, with elevated scores only occurring after 7 weeks of age, when 20% of mice were still alive. Therefore, we decided to focus on the four factors best reflecting early disease progression. Body weight progression in $Mtm1^{-/y}$ mice (Fig. 3C, red) clearly identified a pattern of weight gain, not reaching a maximum of 20 g, before a sharp decline in body weight prior to death, compared to WT mice (black), which continued to increase in body weight. On average, body weight decline started 2.3 weeks before death, with an average body weight loss of 2.16±1.07 g. For this reason, the parameters measuring body weight were optimized to reflect body weight gain (0, expected in healthy juvenile mice), stabilization (0.5) or decline (1) relative to the prior week in the same mouse (Table 1), rather than relative to WT mice. Hanging test, reflecting whole-body strength, was identified as the best representative of the myopathic phenotype in $Mtm1^{-/y}$ mice, with a rapid decline in performance identified from 3 to 5 weeks of age, compared to WT mice, which could perform the test for the duration of the experiment (Fig. 3D). Consequently, hanging ability was given a higher weight of 2 points. It was also constructed as a continuous score between 0 and 2 instead
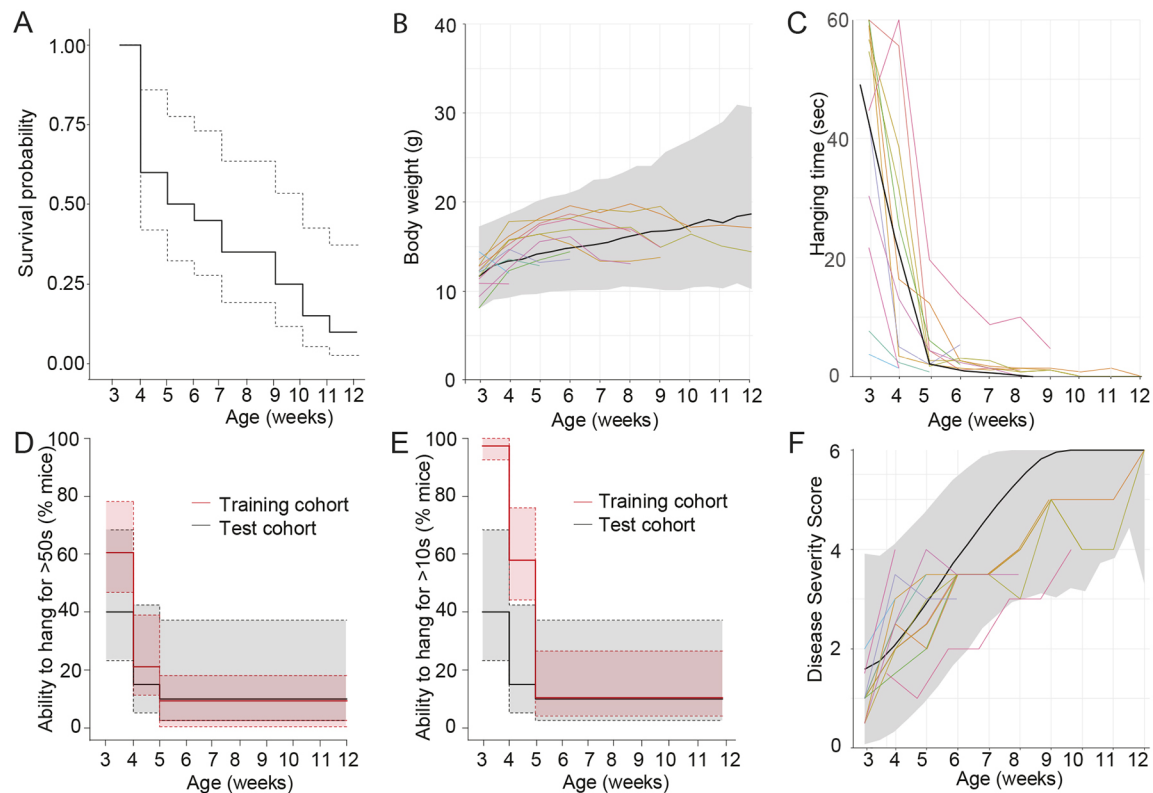
Fig. 2. Natural history study in test cohort of *Mtm1$^{-/y}$* mice. (A) Survival of mice, line of best fit (solid line) and 95% CI (dashed lines) are shown. (B) Body weight (g) progression weekly in *Mtm1$^{-/y}$* mice. Individual mice (colored lines), line of best fit (black line) and prediction interval (shaded gray zone), from the training cohort (Fig. 1B), are highlighted on the graph. (C) Hanging time (60 s maximum), with individual mouse progression shown (colored lines). The line of best fit from the training cohort (Fig. 1C) is highlighted in black. (D,E) Time to event analysis of hanging test, with cutoff times of 50 s (D) and 10 s (E) displayed. Graphs represent line of best fit (solid line) ±95% CI (dashed lines) of training (red, from Fig. 1D,E) and test (black) cohorts for comparison. (F) DSS, with a score between 0 (unaffected) and 6 (most severely affected) awarded per mouse per week, based on six different phenotypes (see Table 1 for details). Line of best fit (black) and prediction intervals (gray shaded area) are shown from the training cohort for reference (Fig. 1F). All data in this figure represent analysis of *Mtm1$^{-/y}$* mouse phenotypes from the test cohort (or training cohort where indicated), from 3 weeks of age; individual mice represented as colored lines. *n*=20 male mice.

of the previously used categorical scores of 0, 0.5 and 1 (Table 1; Table S1). The scores for kyphosis and walking were deemed appropriate, and no optimization was required. The line of best fit for each DSS parameter is shown in Fig. 3E, updated based on statistical analysis and modeling of natural history study data from training and test cohorts in mice. This reflected a progressive and severe development of all disease phenotypes analyzed in *Mtm1$^{-/y}$* mice from 3 to 12 weeks of age. Of note, the two functional tests, hanging ability (whole-body strength) and walking ability, were the first to decline in *Mtm1$^{-/y}$* mice, followed by physical parameters, kyphosis and finally body weight.

The next step was to combine these factors for an overall DSS, with a maximum value of 5. Individual progression of the optimized DSS is shown for *Mtm1$^{-/y}$* mice, confirming a rapid disease progression from 3 weeks of age (Fig. 3F). These data were then used to optimize the DSS model, represented on the graph, with the line of best fit and prediction interval shown (Fig. 3F). Not surprisingly, all mice from training and test cohorts fell within the predicted range for almost all timepoints analyzed. A standard operating procedure to perform this optimized DSS can be found in the Supplementary Materials and Methods.

Investigation of post-mortem muscle phenotypes in *Mtm1$^{-/y}$* mice

We next wanted to assess whether the disease severity phenotype observed in *Mtm1$^{-/y}$* mice correlated with muscle pathology. To this

end, we examined the muscle mass of several skeletal muscles. Muscle atrophy was similarly observed across tibialis anterior (TA), gastrocnemius and quadriceps skeletal muscles (2- to 3-fold reduction in mass, Fig. 4A; 1.5- to 1.9-fold relative to body weight, Fig. S1A,B). Muscle mass correlated with disease severity in *Mtm1$^{-/y}$* mice from all skeletal muscles analyzed (Fig. 4B-D). As the name suggests, CNM results in the abnormal positioning of nuclei internally within muscle fibers. Focusing on the TA muscle, fiber diameter and nuclei position were significantly altered in *Mtm1$^{-/y}$* mice versus WT mice (Fig. 4E-I). Although a significant correlation was observed when analyzing DSS compared to fiber size and nuclei position in WT and *Mtm1$^{-/y}$* mice, no correlation was observed when analyzing only *Mtm1$^{-/y}$* mice at 5 weeks of age (Fig. 4H,J). This may be explained by the muscle function parameters of the DSS (walking and hanging ability), which are already nearly maximally affected by this age, whereas the variable components at this age (kyphosis and body weight) are less related to muscle structure and function (Fig. 3E).

Validation of the model in a dose–response study following DNM2 reduction

The final step to validate the disease severity model generated was to test a therapeutic intervention in *Mtm1$^{-/y}$* mice. Reduction of DNM2 by systemic delivery of ASOs was shown to rescue myotubular myopathy in 3-week-old *Mtm1$^{-/y}$* mice in a dose–response manner (Tasfaout et al., 2017; Koch et al., 2020). Here, we
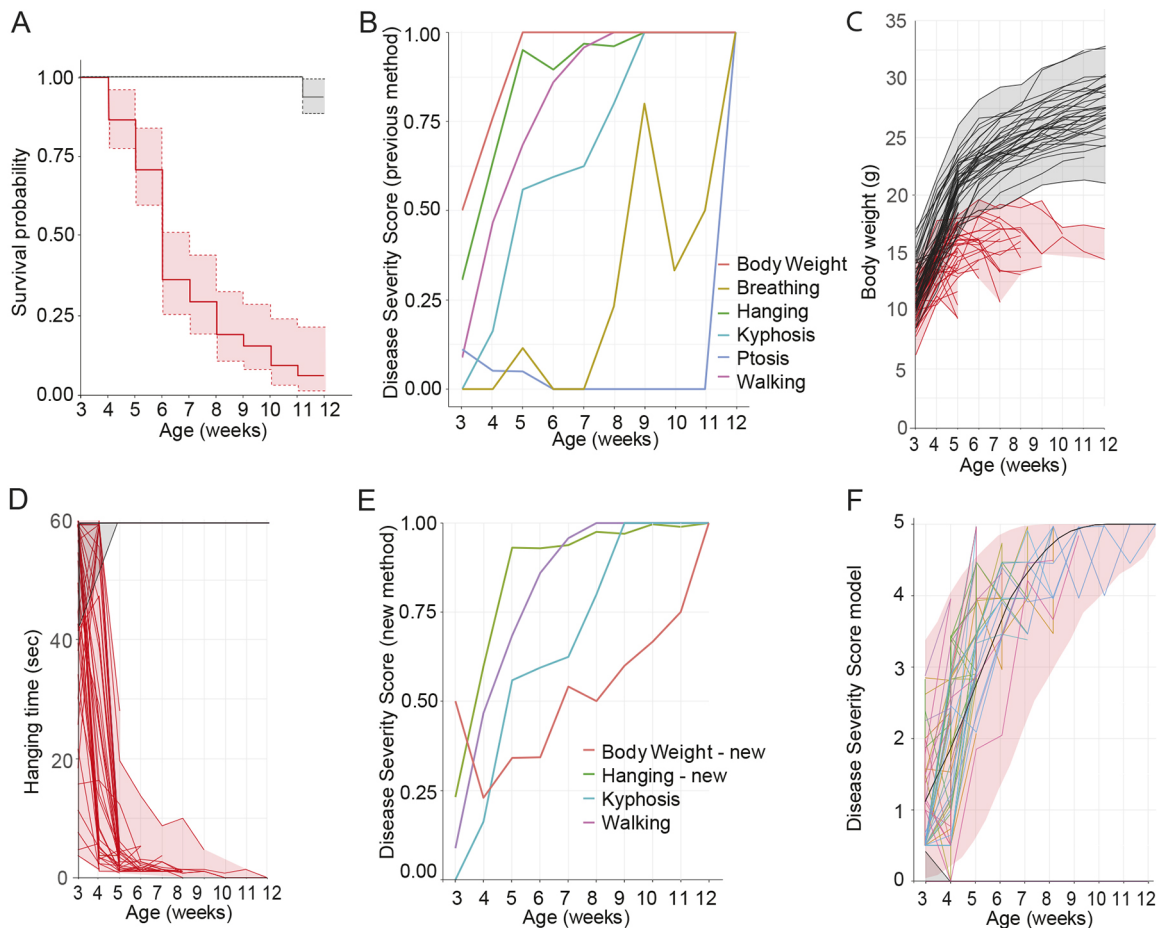
Fig. 3. Optimized disease severity analysis based on combined natural history study data from training and test cohorts. (A) Survival of mice, line of best fit (solid line) and 95% CI (dashed lines) shown for wild-type (WT; black) and *Mtm1⁻/y* (red) mice. (B) Line of best fit shown for progression of each of the six individual factors of the DSS for *Mtm1⁻/y* mice (see Table 1 for details). (C,D) Individual body weight (C) and hanging time (D) from combined natural history study data from training and test cohorts for WT (black) and *Mtm1⁻/y* (red) mice. (E) Line of best fit shown for progression of each of the selected four individual factors comprising the optimized DSS in *Mtm1⁻/y* mice (see Table 1 for details). (F) DSS progression of individual *Mtm1⁻/y* mice, with a score between 0 (unaffected) and 5 (most severely affected) awarded per mouse per week, based on four different phenotypes (see Table 1 for details, hanging test maximal value of 2 possible). Line of best fit (black), and prediction intervals (shaded area) modeling is shown. All WT mice had a score of 0 from week 4 onwards (black lines, shading). All data in this figure represent analysis of WT (*n*=64) and *Mtm1⁻/y* (*n*=58) mouse phenotypes from the training and test cohorts, from 3 weeks of age; individual mice represented as colored lines.

tested the dose–response effect of treatment of *Mtm1⁻/y* mice with ASOs targeting murine *Dnm2*, by weekly intraperitoneal dosing from 5 weeks of age, an age at which mice are already severely affected by the disease (Fig. 3). Low, mid and high (6.25, 12.5, 25 mg/kg, respectively) doses of the 16-nucleotide ASO candidate targeting murine *Dnm2* (DYN101-m), or ASO control (25 mg/kg) (Tasfaout et al., 2017), were delivered weekly by intraperitoneal injection from 5 to 12 weeks of age. A clear increase in survival was observed in *Mtm1⁻/y* mice at all doses tested (Fig. 5A), compared to untreated *Mtm1⁻/y* mice in this study. All WT mice survived for the duration of this study. As no dose response was identified for survival in this study, survival data across groups were combined and compared to those of untreated *Mtm1⁻/y* mice. Combined survival analysis, represented as mean±95% confidence interval (Fig. 5B), was clearly above the survival observed in untreated *Mtm1⁻/y* mice, indicating that survival significantly increased following DYN101-m administration in this trial.

To investigate the therapeutic effect of DNM2 reduction on disease severity in *Mtm1⁻/y* mice, we next analyzed each of the four DSS parameters in mice from the dose–response study. Regarding

physical parameters, body weight progression (Fig. 5C) and kyphosis (Fig. 5D), a clear improvement was observed in *Mtm1⁻/y* mice following reduction of DNM2 at all doses at the completion of the study (12 weeks of age). Of note, *Mtm1⁻/y* mice injected with the low dose (6.25 mg/kg) showed a delayed response for improvement in presentation of kyphosis, compared to mid and high doses (12.5 mg/kg and 15 mg/kg, respectively). A clear improvement was observed in functional parameters, walking (Fig. 5E) and hanging ability (Fig. 5F), at mid and high doses, whereas no improvement was observed in these parameters at low dose, despite the improved survival. Furthermore, a delayed improvement in hanging test ability was observed in mid compared to high dose, suggesting a dose–response effect.

Finally, the combined DSS was analyzed in *Mtm1⁻/y* mice treated with ASOs targeting *Dnm2*. An improvement in DSS was observed at all dose ranges tested (Fig. 5G), with all doses resulting in mice with reduced DSSs, outside of the predicted range for untreated *Mtm1⁻/y* mice of the same age (Fig. 5H). Importantly, a dose–response effect was observed, with higher doses resulting in a more rapid improvement of the disease phenotype, thus validating the
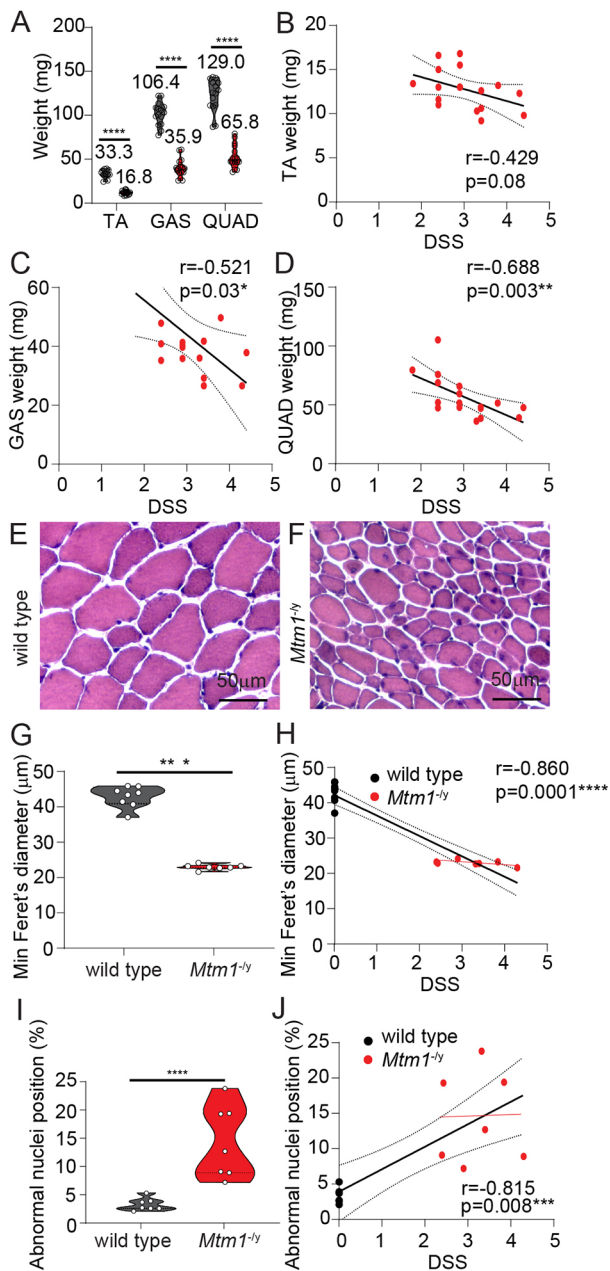
Fig. 4. Skeletal muscle analysis relative to disease severity in *Mtm1⁻/ʸ* mice. (A) Muscle mass of tibialis anterior (TA), gastrocnemius (GAS) and quadriceps (QUAD) muscles from WT and *Mtm1⁻/ʸ* mice. (B-D) Correlation analysis of TA (B), GAS (C) and QUAD (D) muscle mass relative to DSS for *Mtm1⁻/ʸ* mice. Line of best fit and 95% CI are highlighted. Representative Hematoxylin and Eosin (HE) staining of TA muscles from WT (E) and *Mtm1⁻/ʸ* (F) mice (scale bars: 50 µm). (G,H) Analysis of fiber size, represented as minimum feret's diameter presented alone (G) or relative to DSS (H). (I,J) Analysis of fibers with altered nuclei positioning presented alone (I) or relative to DSS (J). (H,J) Line of best fit and 95% CI highlighted for all mice (black), and line of best for *Mtm1⁻/ʸ* mice alone (red). (A,G,I) Represented as violin plots, individual mouse data shown. Spearman correlation analysis performed for C,D,H,J; Pearson correlation analysis for B; unpaired *t*-test for A (TA and QUAD),I; unpaired *t*-test with Welsh's correction for A (GAS),G. *$P<0.05$, **$P<0.01$, ***$P<0.001$, ****$P<0.0001$.

disease severity model generated, and confirming a reduction of the myopathy features in a dose–response manner following *Dnm2* reduction.

Histopathological analysis of dose–response study with DNM2 therapy

To support the improvement in survival and myopathic phenotype observed following ASO delivery to *Mtm1⁻/ʸ* mice, we next investigated post-mortem skeletal muscle specimens at the molecular level at 12 weeks of age. Comparison of treated *Mtm1⁻/ʸ* mice was only possible with WT mice, as untreated *Mtm1⁻/ʸ* mice do not survive until this age. Following weekly intraperitoneal administration of ASO targeting *Dnm2* into mice, a significant dose-dependent reduction of *Dnm2* mRNA was observed (39-69% reduction, Fig. 6A). Skeletal muscle mass was increased when analyzed alone or relative to body weight in a dose–response manner in both gastrocnemius and TA skeletal muscles (Fig. S1C-F), which correlated with the reduction in DSS observed (Fig. 6B, Fig. 5G; Fig. S1G).

We next investigated whether the reduced myopathic phenotype observed following *Dnm2* reduction correlated with an amelioration of the structural defects observed in *Mtm1⁻/ʸ* mice. Corresponding to the improved whole-body strength observed following *Dnm2* reduction (Fig. 5), an improvement in fiber size was also observed, with no statistical difference observed at the top dose tested compared to that of WT control mice (Fig. 6C,D). Furthermore, this correlated with the improvement in DSS (Fig. 6E). Although mice with the lowest dose still presented with 20.5±3.2% of centralized nuclei, consistent with previously published data in *Mtm1⁻/ʸ* mice (Tasfaout et al., 2017; Koch et al., 2020), versus 0-2% in WT mice, this was reduced to 9.4±4.6% in the high-dose group, suggesting a dose-dependent improvement in nuclei position in our study (Fig. 6D,F). This was further supported by a clear amelioration of the abnormal organelle accumulations observed by succinate dehydrogenase (SDH) staining of skeletal muscles (Fig. 6D). Importantly, a significant correlation was observed between the DSS and nuclei position across mice and across doses (Fig. 6G). Of note, although an improvement in survival and disease severity was observed in low-dose mice at 12 weeks of age, which was outside of the modeled expected values of *Mtm1⁻/ʸ* mice at 12 weeks of age (Fig. 5A,H), these mice displayed similar results to untreated *Mtm1⁻/ʸ* mice at 5 weeks of age (Fig. 6B,E,G, red overlay). Therefore ASO-mediated reduction of *Dnm2* resulted in a significant dose-dependent improvement in disease phenotype, with a clear improvement outside the expected disease progression in untreated *Mtm1⁻/ʸ* mice.

DISCUSSION

The purpose of this study was to validate statistically the disease phenotype in *Mtm1⁻/ʸ* mice and to use these data to generate the first joint longitudinal–survival model of disease progression in this mouse line, which is highly relevant for testing therapeutic approaches preclinically in mice. This study provides an example of optimizing analysis of disease progression and testing therapeutic efficacy in a preclinical model of myotubular myopathy.

Modeling disease progression in *Mtm1⁻/ʸ* mice

In this study, we performed disease progression modeling, to provide an optimized model of disease progression in *Mtm1⁻/ʸ* mice over time based on natural history survival and phenotyping data. The consistency of the disease phenotype between mice from the same colony, and between colonies, is important to understand preclinical disease models. Variability may occur in the phenotype between mice in the same colony, which may affect interpretation of the data generated. Here, we observed minor differences in survival between colonies, which were factored into the survival probability
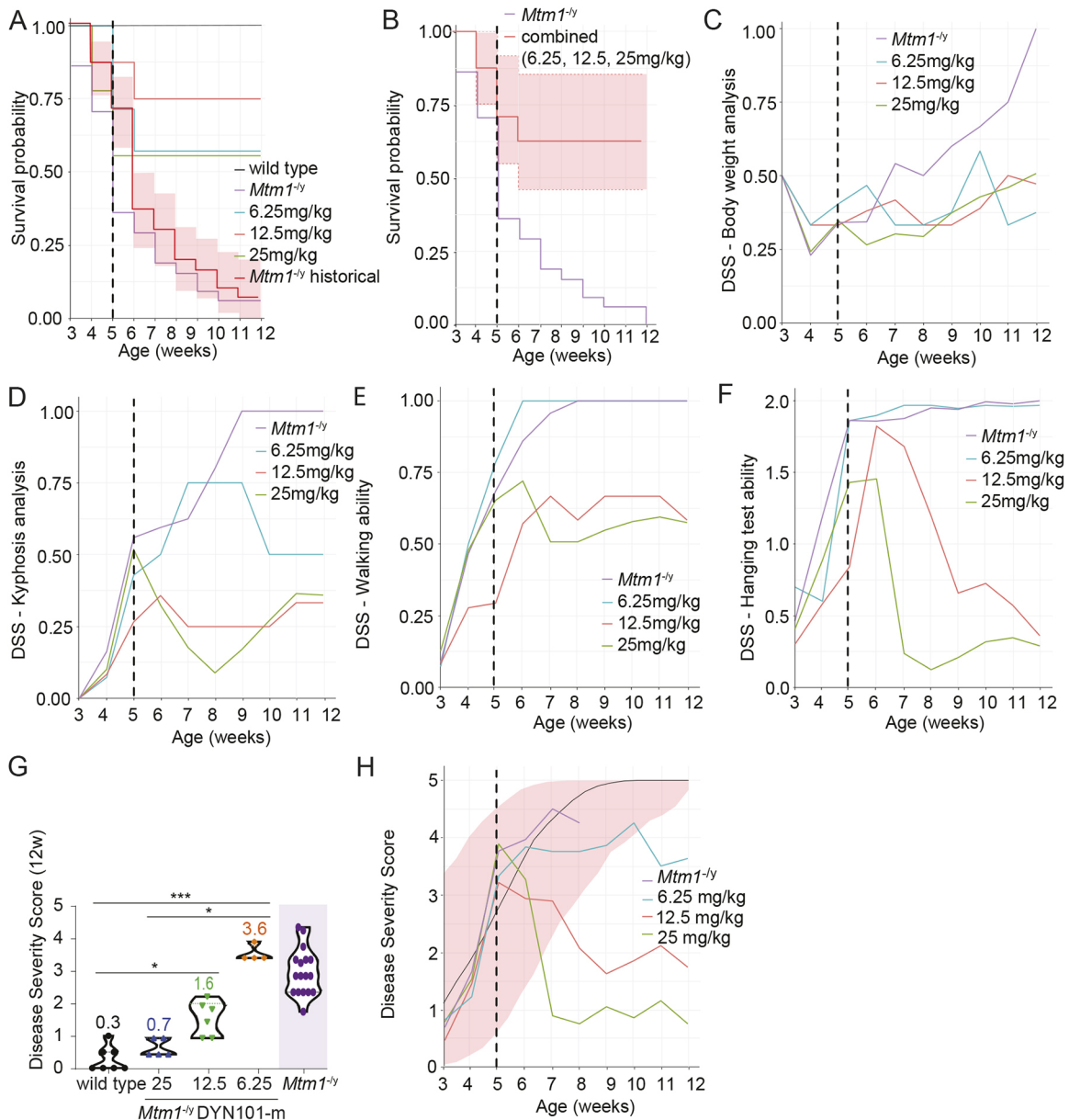
Fig. 5. Validation of the model in a dose–response study with DNM2 therapy. Three doses were selected: low, mid and high (6.25, 12.5, 25 mg/kg; blue, red, green lines, respectively) doses of DYN101-m targeting murine *Dnm2* were delivered weekly by intraperitoneal injection from 5 to 12 weeks of age to $Mtm1^{-/y}$ mice, and compared to WT (black) and $Mtm1^{-/y}$ control mice injected with antisense oligonucleotide (ASO) control (purple) (*n*=7 mice/group). (A,B) Survival of mice shown for all groups separately (A) or with DYN101-m-treated groups combined (B). Average (solid line) and 95% CI (dashed lines) are shown for grouped data (B). (C-F) Individual progression (average) of individual DSS factors: body weight analysis (C), kyphosis (D), walking ability (E) and hanging test performance (F). (G) DSS of WT and $Mtm1^{-/y}$ mice at 12 weeks of age, with a score between 0 (unaffected) and 5 (most severely affected) awarded per mouse, based on four different phenotypes (C-F). DYN101-m dosing indicated where relevant (mg/kg). DSSs from $Mtm1^{-/y}$ mice at 7 weeks of age from Fig. 4 shown for reference only (purple). (H) DSS model (Fig. 3D) represented by line of best fit (black) and prediction intervals (shaded area) modeling is displayed; line of best fit for joint longitudinal survival and phenotyping of $Mtm1^{-/y}$ mice (control, low, mid, high dose) in the dose–response study is indicated. Vertical dashed black lines indicate time of first injection. Kruskal–Wallis test performed for G. *$P<0.05$, ***$P<0.001$.

curve generated (Fig. 3A). Survival rates observed in training and test cohorts were consistent with recently published data from the same mouse line housed in three different laboratories (Maani et al., 2018; Gayi et al., 2018; Danièle et al., 2018). Furthermore, both body weight progression and overall disease severity parameters were consistent between colonies (Fig. 2B-F), and used to optimize the model generated for disease severity mapping in $Mtm1^{-/y}$ mice (Fig. 3). Of interest, muscle mass correlated with disease severity parameters in 5-week-old $Mtm1^{-/y}$ mice; however, no correlation was observed between overall disease severity and muscle histology

analysis (Fig. 4), suggesting that muscle histopathology does not reflect the full disease spectrum in this model.

Understanding the natural history of the disease model is the first step in determining study parameters for testing a therapeutic approach, such as timing of the intervention and appropriate parameters to test for therapeutic effects. Here, the data generated were used to select the relevant parameters to analyze disease progression in $Mtm1^{-/y}$ mice and improve the disease severity scoring system using statistically powered data. The model developed here from phenotyping of over 50 $Mtm1^{-/y}$ mice from two
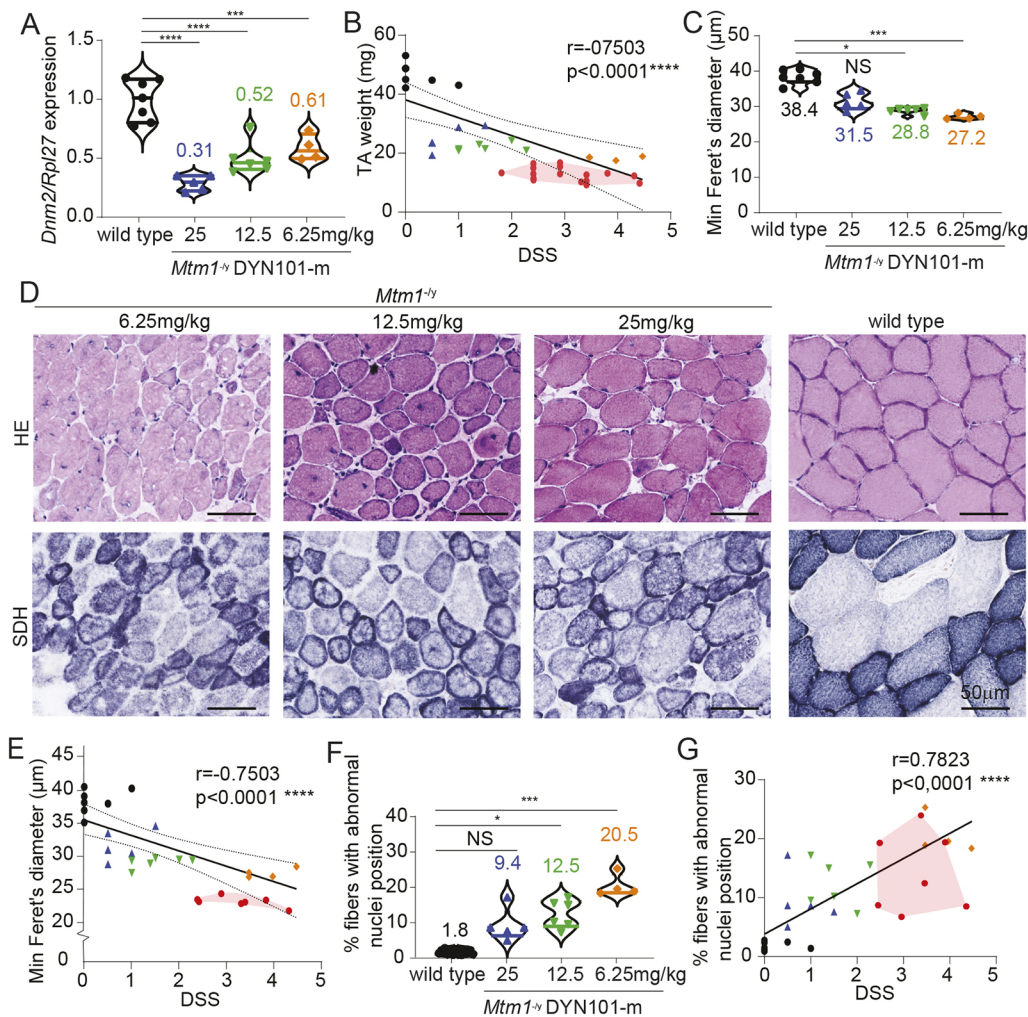
Fig. 6. Correlation analysis of post-mortem skeletal muscle specimens with disease severity following DNM2 therapy. (A) *Dnm2* mRNA expression quantified by quantitative RT-PCR analysis, relative to *Rpl27* expression, from TA muscles, represented as mean±s.d. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test performed. (B) TA muscle mass relative to DSS. (C,E) Analysis of fiber size, represented as minimum feret's diameter alone (C) or relative to DSS (E). (D) HE (top row) and succinate dehydrogenase (SDH; bottom row) staining of TA muscles (scale bars: 50 μm). (F,G) Analysis of fibers with altered nuclei positioning displayed alone (F) or relative to DSS (G). (A,C,F) Represented as violin plots; individual mouse data shown, *n*=4-7 mice per group. (C,F) Kruskal–Wallis (non-parametric) followed by Dunn's multiple comparisons test performed. (B,E,G) Spearman correlation tests performed. 12-week-old *Mtm1*[−/y] mice injected with 6.25, 12.5 or 25 mg/kg are represented by orange, green and blue points, respectively; WT mice are represented by black dots. B, E and G contain as an overlay *Mtm1*[−/y] mice at 5 weeks of age, reproduced from Fig. 4, purely for comparative purposes, and are not included in statistical analyses here. *P<0.05, ***P<0.001, ****P<0.0001. NS, not significant.

independent colonies suggests a progressive myopathic phenotype in mice, which provides a window for testing therapeutic potential post-weaning (Fig. 3). This statistical model will be made available to any researcher working with this mouse line, which can be of use to validate future therapeutic approaches tested on this mouse line in different laboratories. One strength of this approach is that the data used to generate this model came from two independent colonies housed in independent breeding facilities, thus more faithfully represent variation observed by different research laboratories. This modeling approach suggests a method to validate preclinical phenotyping data and therapeutic efficacy that may be applied by scientists testing therapeutic approaches using other neuromuscular disease models. Sharing of historical phenotyping data and modeling approaches across publicly accessible platforms is an important collaborative step that researchers can contribute to, which can help to increase the reliability of preclinical data generated. Supporting this notion, a second study (Sarikaya et al., 2022) published in this issue focuses on disease progression of an independent cohort of *Mtm1*[−/y] mice on the C57B6/J background. This comprehensive study supports the consistency of the progressive myopathic phenotype observed in this mouse line (motor performance and muscle mass – the latter supported by MRI imaging), whilst highlighting subtle differences likely due to the distinct mouse strains. In addition, the complementary study by Sarikaya and colleagues includes detailed omics analysis, which sheds light on the molecular mechanisms that accompany the phenotyping and

histopathological data presented here, and provides insight into the temporal pathogenic increase observed in DNM2 protein expression.

Use of preclinical models to test therapeutic efficacy
It is of high importance to statistically validate the model(s) used for neuromuscular disorders and analyze therapeutic efficacy in an unbiased manner. We used the model generated here of disease progression in *Mtm1*[−/y] mice to validate whether reducing *Dnm2* expression improves survival and disease phenotype in these mice. Targeting *Dnm2* by ASO-mediated reduction blocked disease progression (low dose) and resulted in a significant reduction in disease phenotype (mid/high dose), with a clear improvement outside the normal range of disease progression in *Mtm1*[−/y] mice in all dose groups (Fig. 5). This study followed on from the initial proof-of-concept study validating targeting *DNM2* with ASOs as a potential target for therapy (Tasfaout et al., 2017) and was carefully designed to test disease reversion in a dose–response study (Figs 5 and 6) (Koch et al., 2020).

Performing statistically powered blinded studies can help improve the correct interpretation of preclinical data, as can reproducing data from more than one colony. Initiatives such as the TREAT-NMD neuromuscular network offer support for preclinical research (https://treat-nmd.org/research-overview/), as well as guidance and advice to scientists on the different aspects of translational research related to therapy development programs in neuromuscular disorders, with the aim of improving translational efficacy for the benefit of patients

[TREAT-NMD advisory committee (TACT)] (Willmann et al., 2020). Therapeutic data in mouse models of disease often provide the main or sole data source supporting therapeutic improvement in disease presentation, with the majority of the supportive non-clinical data aimed at reducing any potential safety risk(s) to patients and understanding pharmacokinetics to translate the dose to the human context (Cowling and Thielemans, 2019). Currently three clinical trials (NCT04915846, NCT03199469, NCT04033159, the latter supported by data generated here) have been initiated following therapeutic proof-of-concept observed from three independent therapeutic approaches in this mouse model (Buj-Bello et al., 2008; Buono et al., 2018; Tasfaout et al., 2017; Maani et al., 2018; Gayi et al., 2018). Of note, preliminary therapeutic efficacy has been observed in myotubular myopathy patients following preliminary data generated with the most advanced of the therapeutic approaches (Shieh et al., 2020), supporting the potential utility of $Mtm1^{-/y}$ mice as a disease model for relevant myotubular myopathy disease phenotypes in patients.

In conclusion, we present here a statistical modeling approach for disease progression in myotubular myopathy mice and validate the reliability of this model by testing a therapeutic approach in this mouse line in a dose–response study. This approach can be applied by researchers across the neuromuscular field, to support the generation of reliable and reproducible preclinical data. Using this approach may thus improve confidence in preclinical therapeutic data generated from neuromuscular disease models across different laboratories or across different cohorts.

MATERIALS AND METHODS
Generation of $Mtm1^{-/y}$ mice
$Mtm1^{-/y}$ or WT 129SvPAS mice were previously generated and characterized by crossing $Mtm1$ heterozygous females obtained by homologous recombination with WT males (Buj-Bello et al., 2002). The training cohort was previously generated at the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC; Illkirch, France), and historical published data generated from this cohort were obtained and analyzed from 3 weeks of age [studies 3, 5, $n=38$ $Mtm1^{-/y}$ mice (Koch et al., 2020)]. To generate the test cohort, IVF was performed in the 129SvPAS strain at Janvier Laboratories (Saint-Berthevin, France), then offspring were transferred to Chronobiotron (Strasbourg, France) for colony generation of the test cohort. Animal experimentation was approved by the institutional ethical committee, the training cohort and ASO administration were approved by the Comité d'Ethique pour l'Expérimentation Animale IGBMC-Institut Clinique de la Souris (APAFIS#5453-2016052510176016 v5), and the test cohort was approved by the Comité Régional d'Ethique en Matière d'Expérimentation Animale de Strasbourg Chronobiotron (20183-2019040817583412 v5 and v6). Test cohort mice were analyzed from 3 weeks of age ($n=20$ $Mtm1^{-/y}$ mice). For both the training and test cohorts, daily observation was performed when necessary. In the case of severe phenotype, scoring was performed to determine whether action was needed to reduce pain, or whether humane endpoints were reached, in which case the mice were sacrificed. Mice were humanely sacrificed when required according to national and European legislations on animal experimentation. Male mice were analyzed in this study. Animals were housed in a temperature-controlled room (19-22°C) with a 12:12 h light/dark cycle, with free access to food. See Table S2 for additional information.

ASOs
ASOs were chemically modified with phosphorothioate in the backbone and cEt modifications on the wings with a deoxy gap (3-10-3 design). ASOs were synthesized by IONIS Pharmaceuticals as previously described (Tasfaout et al., 2017). Both the 16-nucleotide ASO candidate targeting murine $Dnm2$ (DYN101-m, GGCATAAGGTCACGGA) and the control sequence with no homology to the mouse genome (ASO-Ctrl, GGCCAATACGCCGTCA) were previously validated (Tasfaout et al., 2017; Buono et al., 2018). ASOs were dissolved in filtered and autoclaved

sterile D-PBS (Life Technologies, #14190-144). Intraperitoneal injections of 6.25, 12.5 or 25 mg/kg of ASO were performed in $Mtm1^{-/y}$ or WT male mice weekly, from 5 to 12 weeks of age ($n=7$ mice/group). Mice were sacrificed 2 days after the final injection (Koch et al., 2020).

Generation of mouse cohorts
Training cohort data were generated from historical data of $Mtm1^{-/y}$ mice located at the IGBMC animal facility (Koch et al., 2020). The test cohort was derived as follows: IVF was performed (Janvier Laboratories, Rennes, France) with samples taken from $Mtm1^{-/y}$ mice from the colony used to generate the training cohort, colony generation was performed, and then the animals were transferred to Chronobiotron animal facility for colony amplification and phenotyping.

DSS analysis
DSS was performed to monitor the clinical appearance of $Mtm1^{-/y}$ mice. The DSS was designed to evaluate the clinical evolution of six indicators of myopathy in mice: body weight difference, hanging test ability, kyphosis, hindlimb position whilst walking, breathing ability and ptosis, as described previously (Tasfaout et al., 2017). Details of DSS analysis, both previous and updated method based on analyses performed in this study, are represented in Table 1.

Hanging test
Mice were placed on a grid (cage lid, dimensions 410×270 mm), which was then inverted and held 40 cm above the cage litter; the latency to fall was measured three times for each mouse, with a minimum of 5 min interval between trials to allow a recovery period. The latency time measurements began from the point when the mouse was hanging free on the grid and ended with the animal falling to the cage underneath the grid. The maximum time measured was 60 s. Results are expressed as an average of three trials.

Statistics and modeling
For the 'training cohort', historical data were accessed, where the sample size (n) was established for each individual study (Koch et al., 2020). For the 'test cohort', we calculated the statistical power using R software function 'pwr2::ss.1way' (k, number of groups; alpha; beta; f, effect size; delta, smallest difference among k group; sigma, s.d., i.e. square root of variance; and B, iteration times). The statistical power was calculated as $n=15$ for the 'test cohort' to see a difference between groups; however, $n=20$ was used to optimally model the progression of the disease in $Mtm1^{-/y}$ mice. For the 'dose–response' ASO study, the prespecified effect size was established to be $n=7$ (Tasfaout et al., 2017).

To model body weight progression, a joint model considering survival and body weight progression was used. The evolution was not linear with time but rather changed, as the square root of weeks and the log weight are modeled in order to account for the natural lower bound at 0. Random slopes and intercept for each mouse are included. See Supplementary Materials and Methods for a more detailed description of the joint models, including for DSSs. Correlation analyses were performed with a Spearman or Pearson correlation test, as noted in the figure legends. Additional statistical analyses were performed as stated in the figure legend where appropriate, with $P<0.05$ considered significant. The normality of the residuals and the variance homogeneity were assessed to apply the appropriate statistical test with GraphPad Prism 9.

Exclusion criteria
All exclusion/inclusion criteria were pre-established. All male mice from 3 weeks of age were included. For the dose–response study, mice entered the study at 3 weeks of age; however, mice that died or were sacrificed for humane endpoints before starting the injection protocol at 5 weeks of age were excluded and replaced. For sample analysis, the exclusion criteria were (1) technical issues not allowing the data to be measured, or (2) sample(s) detected as an outlier with GraphPad Prism 9 (parameters: ROUT, Q=5%).

Randomization and blinding
For the dose–response experiment, $Mtm1^{-/y}$ or WT mice were randomly assigned to groups at 3 weeks of age ($n=7$ mice per group). $Mtm1^{-/y}$ or WT mice from the same litter were kept in the same cage. Where possible, mice

from the same litter were allocated to different dosing groups. For the histological analyses in the dose–response study, the ASO dose administered was blinded to avoid any bias.

Quantitative RT-PCR
RNA was isolated from organs using a NucleoSpin RNA kit (Macherey Nagel). For muscle tissue, the supplementary protocol available from the manufacturer entitled 'isolation of RNA from fibrous tissue' (https://www.mn-net.com/fr/nucleospin-rna-mini-kit-for-rna-purification-740955.50) was used. Reverse transcription was carried out on 250 ng aliquots using SuperScript™ IV Reverse Transcriptase (ThermoFisher Scientific). Quantitative PCR was done with PowerUp™ SYBR™ Green Master Mix (ThermoFisher Scientific) in a Quant Studio 3 Real-Time PCR System (ThermoFisher Scientific). The relative expression of $Dnm2$ mRNA was normalized to $Rpl27$. Primers used were as follows: $Rpl27$ Forward, 5′-AAGCCGTCATCGTGAAGAACA-3′; $Rpl27$ Reverse, 5′-CTTGATC-TTGGATCGCTTGGC-3′; $Dnm2$ Forward, 5′-ACCCCACACTTGCA-GAAAAC-3′; $Dnm2$ Reverse, 5′-CGCTTCTCAAAGTCCACTCC-3′. mRNA analysis followed most Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) standards.

Histological analysis of skeletal muscle
Air-dried transverse cryosections (8 µm) were fixed and stained with Hematoxylin and Eosin (HE) or SDH, and image acquisition performed with a slide scanner NanoZoomer 2 HT equipped with brightfield and the fluorescence module L11600-21 (Hamamatsu Photonics). Minimum feret's diameter was analyzed in wheat germ agglutinin (WGA) sections from TA mouse skeletal muscle, using a plugin developed in ImageJ – 'MyoMage'. Minimum feret's diameter was calculated in >500 fibers per mouse. The percentage of TA muscle fibers with centralized or internalized nuclei was counted in >500 fibers using the cell counter plugin in ImageJ (http://rsb.info.nih.gov/ij/) or FIJI analysis software. Qualitative SDH staining analysis was performed. Fiber to fiber variation in intensity in SDH staining is normal and is indicative of the oxidative state of the fiber. SDH staining is normally relatively homogeneous within each individual fiber. Accumulation within the center or the periphery of a fiber of SDH staining indicates an abnormal distribution.

References
Buj-Bello, A., Laugel, V., Messaddeq, N., Zahreddine, H., Laporte, J., Pellissier, J. F. and Mandel, J. L. (2002). The lipid phosphatase myotubularin is essential for skeletal muscle maintenance but not for myogenesis in mice. *Proc. Natl. Acad. Sci. USA* **99**, 15060-15065. doi:10.1073/pnas.212498399

Buj-Bello, A., Fougerousse, F., Schwab, Y., Messaddeq, N., Spehner, D., Pierson, C. R., Durand, M., Kretz, C., Danos, O., Douar, A.-M. et al. (2008). AAV-mediated intramuscular delivery of myotubularin corrects the myotubular myopathy phenotype in targeted murine muscle and suggests a function in plasma membrane homeostasis. *Hum. Mol. Genet.* **17**, 2132-2143. doi:10.1093/hmg/ddn112

Buono, S., Ross, J. A., Tasfaout, H., Levy, Y., Kretz, C., Tayefeh, L., Matson, J., Guo, S., Kessler, P., Monia, B. P. et al. (2018). Reducing dynamin 2 (DNM2)

rescues DNM2 -related dominant centronuclear myopathy. *Proc. Natl. Acad. Sci. USA* **115**, 11066-11071. doi:10.1073/pnas.1808170115

Cowling, B. S. and Thielemans, L. (2019). Translational medicine in neuromuscular disorders: from academia to industry. *Dis. Model Mech.* **13**, dmm041434. doi:10.1242/dmm.041434

Cowling, B. S., Toussaint, A., Amoasii, L., Koebel, P., Ferry, A., Davignon, L., Nishino, I., Mandel, J. L. and Laporte, J. (2011). Increased expression of wild-type or a centronuclear myopathy mutant of dynamin 2 in skeletal muscle of adult mice leads to structural defects and muscle weakness. *Am. J. Pathol.* **178**, 2224-2235. doi:10.1016/j.ajpath.2011.01.054

Cowling, B. S., Chevremont, T., Prokic, I., Kretz, C., Ferry, A., Coirault, C., Koutsopoulos, O., Laugel, V., Romero, N. B. and Laporte, J. (2014). Reducing dynamin 2 expression rescues X-linked centronuclear myopathy. *J. Clin. Invest.* **124**, 1350-1363. doi:10.1172/JCI71206

Danièle, N., Moal, C., Julien, L., Marinello, M., Jamet, T., Martin, S., Vignaud, A., Lawlor, M. W. and Buj-Bello, A. (2018). Intravenous administration of a MTMR2-encoding AAV vector ameliorates the phenotype of myotubular myopathy in mice. *J. Neuropathol. Exp. Neurol.* **77**, 282-295. doi:10.1093/jnen/nly002

Gayi, E., Neff, L. A., Massana Muñoz, X., Ismail, H. M., Sierra, M., Mercier, T., Décosterd, L. A., Laporte, J., Cowling, B. S., Dorchies, O. M. et al. (2018). Tamoxifen prolongs survival and alleviates symptoms in mice with fatal X-linked myotubular myopathy. *Nat. Commun.* **9**, 4848. doi:10.1038/s41467-018-07058-4

Gordish-Dressman, H., Willmann, R., Dalle Pazze, L., Kreibich, A., van Putten, M., Heydemann, A., Bogdanik, L., Lutz, C., Davies, K., Demonbreun, A. R. et al. (2018). "Of mice and measures": a project to improve how we advance duchenne muscular dystrophy therapies to the clinic. *J. Neuromuscul. Dis.* **5**, 407-417. doi:10.3233/JND-180324

Koch, C., Buono, S., Menuet, A., Robé, A., Djeddi, S., Kretz, C., Gomez-Oca, R., Depla, M., Monseur, A., Thielemans, L. et al. (2020). Myostatin: a circulating biomarker correlating with disease in myotubular myopathy mice and patients. *Mol. Ther. Methods Clin. Dev.* **17**, 1178-1189. doi:10.1016/j.omtm.2020.04.022

Laporte, J., Hu, L. J., Kretz, C., Mandel, J.-L., Kioschis, P., Coy, J. F., Klauck, S. M., Poustka, A. and Dahl, N. (1996). A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved in yeast. *Nat. Genet.* **13**, 175-182. doi:10.1038/ng0696-175

Liu, N., Bezprozvannaya, S., Shelton, J. M., Frisard, M. I., Hulver, M. W., McMillan, R. P., Wu, Y., Voelker, K. A., Grange, R. W., Richardson, J. A. et al. (2011). Mice lacking microRNA 133a develop dynamin 2-dependent centronuclear myopathy. *J. Clin. Invest.* **121**, 3258-3268. doi:10.1172/JCI46267

Maani, N., Sabha, N., Rezai, K., Ramani, A., Groom, L., Eltayeb, N., Mavandadnejad, F., Pang, A., Russo, G., Brudno, M. et al. (2018). Tamoxifen therapy in a murine model of myotubular myopathy. *Nat. Commun.* **9**, 4849. doi:10.1038/s41467-018-07057-5

Massana Munoz, X., Buono, S., Koebel, P., Laporte, J. and Cowling, B. S. (2019). Different in vivo impacts of dynamin 2 mutations implicated in Charcot-Marie-Tooth neuropathy or centronuclear myopathy. *Hum. Mol. Genet.* **28**, 4067-4077. doi:10.1093/hmg/ddz249

Ravenscroft, G., Laing, N. G. and Bönnemann, C. G. (2015). Pathophysiological concepts in the congenital myopathies: blurring the boundaries, sharpening the focus. *Brain* **138**, 246-268. doi:10.1093/brain/awu368

Sarikaya, E., Sabha, N., Volpatti, J., Pannia, E., Maani, N., Gonorazky, H. D., Celik, A., Liang, Y., Onofre-Oliveira, P. and Dowling, J. J. (2022). Natural history of a mouse model of X-linked myotubular myopathy. *Dis. Model. Mech.* **15**, dmm049342. doi:10.1242/dmm.049342

Shieh, P. B., Bonnemann, C. G., Müller-Felber, W., Blaschek, A., Dowling, J. J., Kuntz, N. L. and Seferian, A. M. (2020). Re: "moving forward after two deaths in a gene therapy trial of myotubular myopathy" by Wilson and Flotte. *Hum. Gene. Ther.* **31**, 787. doi:10.1089/hum.2020.217

Tasfaout, H., Buono, S., Guo, S., Kretz, C., Messaddeq, N., Booten, S., Greenlee, S., Monia, B. P., Cowling, B. S. and Laporte, J. (2017). Antisense oligonucleotide-mediated Dnm2 knockdown prevents and reverts myotubular myopathy in mice. *Nat. Commun.* **8**, 15661. doi:10.1038/ncomms15661

Tasfaout, H., Lionello, V. M., Kretz, C., Koebel, P., Messaddeq, N., Bitz, D., Laporte, J. and Cowling, B. S. (2018). Single intramuscular injection of AAV-shRNA reduces DNM2 and prevents myotubular myopathy in mice. *Mol. Ther.* **26**, 1082-1092. doi:10.1016/j.ymthe.2018.02.008

van Putten, M., Putker, K., Overzier, M., Adamzek, W. A., Pasteuning-Vuhman, S., Plomp, J. J. and Aartsma-Rus, A. (2019). Natural disease history of the D2-mdx mouse model for Duchenne muscular dystrophy. *FASEB J.* **33**, 8110-8124. doi:10.1096/fj.201802488R

Vandersmissen, I., Biancalana, V., Servais, L., Dowling, J. J., Vander Stichele, G., Van Rooijen, S. and Thielemans, L. (2018). An integrated modelling methodology for estimating the prevalence of centronuclear myopathy. *Neuromuscul. Disord.* **28**, 766-777. doi:10.1016/j.nmd.2018.06.012

Willmann, R., Lee, J., Turner, C., Nagaraju, K., Aartsma-Rus, A., Wells, D. J., Wagner, K. R., Csimma, C., Straub, V., Grounds, M. D. et al. (2020). Improving translatability of preclinical studies for neuromuscular disorders: lessons from the TREAT-NMD Advisory Committee for Therapeutics (TACT). *Dis. Model Mech.* **13**, dmm042903. doi:10.1242/dmm.042903
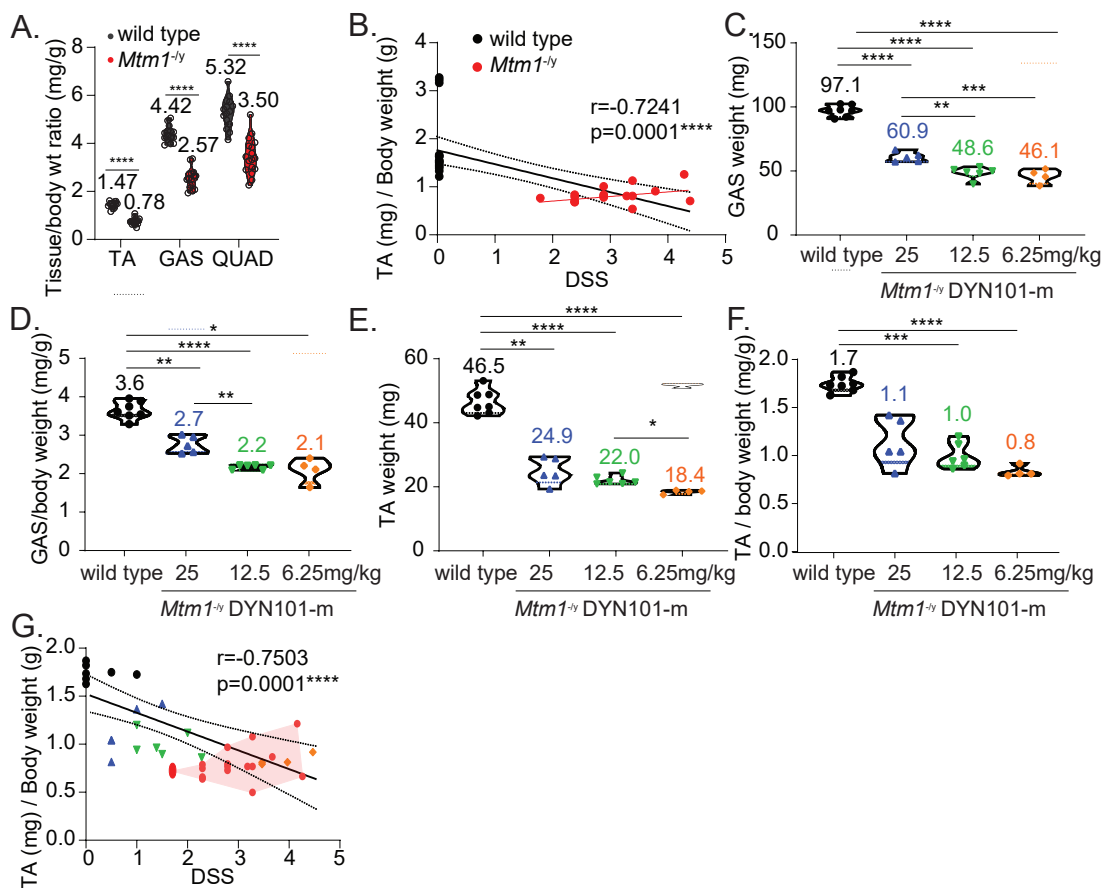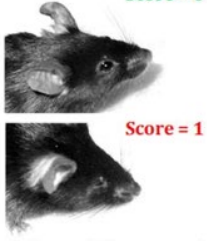
Disease Models & Mechanisms

Fig. S1. Analysis of post mortem skeletal muscle speci-mens and correlation with disease severity

(A) Muscle mass of tibialis anterior (TA), gastrocnemius (GAS) and quadriceps (QUAD) muscles relative to body weight, from 5 week old wildtype and $Mtm1^{-/y}$ mice. Unpaired t-test with Welsh's correction performed. (B) Correlation analysis of TA muscle mass/body weight ratio, relative to disease severity score (DSS) for 5 week old $Mtm1^{-/y}$ (red) and wild type mice (black). Line of best fit and 95%CI highlighted in black for all mice (r= -0.7241; p < 0.0001), line of best fit for $Mtm1^{-/y}$ alone (red line). Muscle mass presented alone (GAS-C; TA-E) or relative to body weight (GAS-D; TA-F), for wildtype and $Mtm1^{-/y}$ mice injected with 6.25, 12.5 or 25mg/kg DYN101-m targeting murine $Dnm2$ reduction. (G) TA/ body weight ratio represented relative to DSS for wild type (black), $Mtm1^{-/y}$ mice injected with 6.25 (orange), 12.5 (green), 25mg/kg (blue) DYN101-m. 5 week old $Mtm1^{-/y}$ mice from (B) reproduced here for comparison purposes only (red dots, shading), and are not included in statistical analysis. Data represented as a violin plots (A,C-F), individual mouse data shown. Statistical analysis: (A) Unpaired t-test with Welsh's correction; Spearman correlation tests performed for (B), (G) Ordinary 1-way ANOVA followed by Dunnett's multiple comparisons test performed for (C)-(F). *p<0.05, **p<0.01, ***p<0.001,, ****p<0.0001.

Table S1. Previous disease severity scoring system (Tasfaout, Buono et al. 2017). A scoring system was set up to evaluate the clinical evolution of six centronuclear myopathy features. Difference of body weight between $Mtm1^{-/y}$ versus WT littermate, ability to perform the hanging test, walking manner, presence or absence of ptosis and kyphosis and breathing difficulties (frequency and amplitude evaluation based on clinical observations) are recorded and a score of 0, 0.5 or 1 is given to each clinical readout. The sum represents the DSS. The higher the DSS, the more severe the phenotype, minimum 0 (healthy mouse), maximum 6 (severely affected mouse) (Tasfaout, Buono et al., 2017).

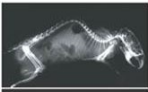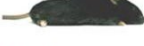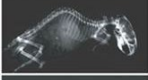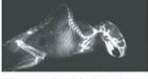| Disease scoring category	DSS Description	Score
Body weight	Score 0-1 Difference in body weight between $Mtm1^{-/y}$ mouse and wildtype littermate	0: 0-1g 0.5: >1-2g 1: >2g
Ptosis	Score 0 or 1 Drooping of the eyelid	0 : No 1 : Yes Ptosis: Score = 0 Score = 1 Adapted Yang et al. Immunology and Microbiology (2007)
Hanging test ability	Score of 0-1, representing 0-60 seconds hanging time.	0: 60s 0.5: 5-60s 1: <5

Kyphosis	Score 0-1 Curvature of the spine	Score 0: no curvature of the spine Score 0.5: mild curvature of the spine **Kyphosis:** Score = 0 Score = 0.5 Score = 1 Adapted from Gabellini et *al.* Nature (2006) Score 1: severe curvature of the spine
Breathing difficulties	Score 0 or 1 Frequency and amplitude evaluation based on clinical observations	0: no breathing alternation 1: breathing alternation
Walking difficulties	Score 0-1 Ability to use hindlimbs	Score 0: normal use of hindlimbs Score 0.5: splayed use of hindlimbs Score 1: loss of use of hindlimbs **Walking & gait:** Score= 0 Score= 0.5 Score=1
Maximum score	6	

Table S2. Husbandry and housing conditions for training and test cohorts

		TRAINING COHORT	TEST COHORT
		Training cohort data was generated from historical data of *Mtm1*$^{-/y}$ mice located at the IGBMC animal facility, Illkirch, France (Koch, Buono et al. 2020). N=38 mice	Test cohort data: In vitro fertilization (IVF) was performed (Janvier Laboratories, Rennes, France) with samples taken from *Mtm1*$^{-/y}$ mice from the colony used to generate the training cohort. Following IVF and colony generation, mice were transferred to Chronobiotron animal facility (Strasbourg, France), for colony amplification and phenotyping (test cohort). N=20 mice
Husbandry conditions		Light cycle: 12hr/12hr Temperature: 19-22°C Enrichment: nests to help nesting.	Janvier Laboratories : Light cycle: 12hr/12hr Temperature: +21°C ± 2°C Enrichment: all cages contain sticks to gnaw (Aspenbrick). Breeding cages were systematically enriched with nests.
Housing conditions		Light cycle: 12hr/12hr Temperature: 19-22°C Enrichment: nests	Chronobiotron animal facility: Light cycle: 12hr/12hr Temperature: +21°C ± 2°C Enrichment: nests and cardboard tunnel
Weaning age		3 weeks of age	3 weeks of age
Description of cage		*Cage type 2:* Individually ventilated cages: no Floor surface: 370 cm² *Cage type 3:* For housing during experimental procedure, cage type 3 was rarely used because of the larger size and increased distance to access food *Access to food and water:* Gel diet was added inside the cage to facilitate access. Food pellets were added at multiple places inside the cages.	Chronobiotron animal facility: *Cage type 2:* Individually ventilated cages: no Brand: Tecniplast (#1264C Eurostandard) or Erhet Floor surface: 370 cm² Dimensions: L x W x H : 268 x 215 x 141 mm *Cage type 3:* For housing during experimental procedure, cage type 3 was rarely used because of the larger size and increased distance to access to the food. Individually ventilated cages: no Brand: Zoonlab Floor surface: 830 cm² Dimensions: L x W x H : 425 x 265 x 150 mm *Access to food and water:* Gel diet was added inside the cage to facilitate access. Food pellets were added at multiple places inside the cages. Longer nozzles on water bottles were used to reduce the distance between floor and bottle nozzle.
Animal density		4 animals per cage: type II cage 5 animal per cages: type III cage	4 animals per cage: type II cage 5 animal per cages: type III cage
Genotype ratio		Typically 1 or 2 wildtype mice were housed in each cage with *Mtm1*$^{-/y}$ mice from the same litter, whenever feasible.	
Change frequency		Once per week	
Diet	**Food composition:**	SAFE® gel diet breeding	SAFE® gel diet water
	Cereals	17.3%	
	Proteins	6.6%	
	Vitamins, minerals	1.2%	
	Fibers	1.4%	1.8%
	Water	72.9%	98.2%

Supplementary Materials and Methods

Joint Models for Disease Severity Scores. Joint models for disease severity scores are presented in the associated manuscript. The mice where followed over a period $[0; \tau\,[$. For each subject i, we observe the following:

- Longitudinal measures: $\{Y_{ij}; j = 1, \dots, n_i\}$ at times $\{t_{ij}; j = 1, \dots, n_i\}$
- Survival time and indicator:

$$T = \min(ST, C) \text{ and } \delta = I(ST \leq C) = \begin{cases} 1 \text{ if uncensored observation} \\ 0 \text{ if censored observation} \end{cases}$$

The longitudinal measurements and time to event were jointly measured via a latent bivariate process, which was realized independently in each subject $W_i = \{W_{1i_i}, W_{2i}\}$

The longitudinal sub-model is defined as: The observed response y_{ij} scaled between 0 and 1 ($i = 1, \dots, M \text{ and } j = 1, \dots, n_i$), is defined as : $y_{ij} \sim Beta(a_{ij}, b_{ij})$; where the beta distribution is defined by the parameters a_{ij} and b_{ij} , which are defined, respectively, by the mean μ_{ij} and "the sample size" ν of the distribution as follows:

$$a_{ij} = \mu_{ij} * \nu$$
$$b_{ij} = (1 - \mu_{ij}) * \nu$$

The parameter ν is estimated from the data, and the mean μ_{ij} is defined as a mixed model with logit-link function, where *T* is a constant to center time:

$$\chi_{ij} = \alpha_i + \beta * (T + t_{ij})$$
$$\mu_{ij} = \frac{1}{1 + exp(-\chi_{ij})}$$

The equation of the mean contains a random effect on the intercept $W_{1i} = \alpha_i \sim N(0, \sigma_\alpha^2)$. In turn, the survival sub-model is then defined via a Weibull survival model:

$$S(t_{ij}) = exp(-(\zeta_i * t_{ij})^k)$$

Where: $\quad \zeta_i = exp(-W_{2i})$

$$W_{2i} = \gamma * \alpha_i$$

Where γ expresses the induced association. A negative value of gamma implies that the larger value of the DSS Score the smaller the probability to stay in the study. A positive value would however imply the opposite. For simplicity in estimation, k = 1 in all estimations of the model indicating a constant failure rate. The model for the evolution of weight over time is similar to the one proposed here over.

The main difference lies in the equation for the mean. Indeed, the equation is assumed to be a linear gaussian model that evolves a the square root of time rather than a beta model.

The disease severity data from individual mice are included only until the point of death. The death of the mouse is then factored into the joint longitudinal survival model for disease severity scores. The above model used here allows the expected range to be presented following the correction for limited sample size, variability, and survival differences. The output is presented as the line of best fit (black line) and prediction intervals (shaded zone). A correct model fit occurs when 95% of the observed individual data is contained, for each time point, in the intervals depicted. Individual data from all living mice is shown as an overlay of the model to confirm this point on each graph, and thus the validity of the model.

Standard Operating Procedure for Disease Severity Score

This SOP describes the disease severity scoring (DSS) system, based on the original (DSS) published in Tasfaout, Buono et al 2017, and subsequently optimized in Buono et al (associated manuscript).

Objective

The goal of this SOP is to detail the procedure of allocating a disease severity score (DSS) in mice. The DSS is a parameter which is used for the evaluation of the disease severity of the myopathic phenotype in $Mtm1^{-/y}$ mice. The calculation is based on 4 parameters: difference of body weight, hanging test ability, kyphosis and walking difficulties. The minimal score is 0 (no myopathic phenotype) and the maximal score is 5 (characterizing a very severe myopathic phenotype).

1. Abbreviations and Definitions

DSS	Disease Severity Score
NHS	Natural History Study
SOP	Standard Operating Protocol
NA	Not applicable

2. Risk Assessment / Related documents

3. Ethics

Mice should be handled according to national legislation on animal care and experimentation. This SOP is in alignment with the French and European legislation. Procedures and Protocols should be approved by the institutional Ethics Committee.

4. Materials and Equipment

- Type 3 cage: 425 x 265 x 150 mm
- Metal Grid : 410 x 270 mm
- Timer
- Scale
- Lid of transport box (optional)

5. Reagents

NA

6. Protocol

1) Body weight

Measure body weight, to 2 decimal places. The score for this parameter (score between 0 and 1) is the difference of a *Mtm1$^{-/y}$* mouse body weight from week n to n+1:

Table 1: Body weight to score conversion

Body weight	SCORE
X ≥ 0.25g	0
-0.25g > X > 0.25g	0.5
X ≤ -0.25g	1

X(g) = (body weight week n+1(g)) – (body weight week n(g))

2) Hanging test ability

This test must be done <u>one mouse at a time</u>:

a. Take one type 3 cage and one grid.
b. Place mouse in the middle of the grid
c. Turn the grid upside down. The suspending animal should hold on to the grid in order to avoid falling.
   - Test must be set up at a certain height, around 40 cm, for mouse to not being influenced to jump.
d. Prevent the mouse from turning over to the other side of the grid or at pellets food and bottle place by barring with the hand (without touching the mouse) or by using the lid of a transport box.
e. The latency to fall will be measured <u>three times</u> (<u>60 seconds each</u>) for each mouse, with a minimum interval of 10 minutes between trials.
   - The latency time measurements begin when the mouse is hanging free on the wire and end with the animal falling to the cage underneath the wire or grid.
   - If performing the whole body hanging test for the first time, mouse can fall as soon as grip is turned upside down. If this is the case, the first assay can be considered as familiarization of the mouse with the testing conditions and will not be considered as one of the three assays of the test. Only time (seconds) for the three next trials will be reported in the dedicated table (Table 2).
   - If a mouse falls for any other reason that muscle strength default (eg if a mouse is not willing to do the test, voluntary jump from the grip, or falls because of your hands…), this will not be considered as a trial. This must be recorded in the note section in the dedicated table (Table 2)
   - A mouse should normally explore the grid. If a mouse stays in place and does not explore the grip. This must be recorded in the note section in the dedicated table (Table 2)
f. Time (seconds) when mouse falls should be reported in the dedicated table (Table 2).
   - 60 seconds is the maximum time allowed

Table 2: Data collection table for hanging test

Mice ID number	Mice genotype	Age (weeks)	Hanging time (seconds)			Note
			Test 1	Test 2	Test 3	

Note: Mouse genotype should be blinded until experiment is complete

The score for this parameter (score between 0 and 2) represents the mean hanging time according to the formula: Time (s) > [(60-Time (s))/60s]*2 = value (0-2)

Examples:
1s > (60-1)/60*2=1.97
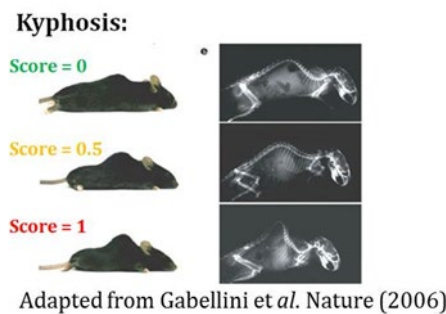15s > (60-15)/60*2=1.5
30s > (60-30)/60*2=1
60s > (60-60)/60*2=0

3) Kyphosis

This parameter reflects the curvature of the spine. The score is noted as:

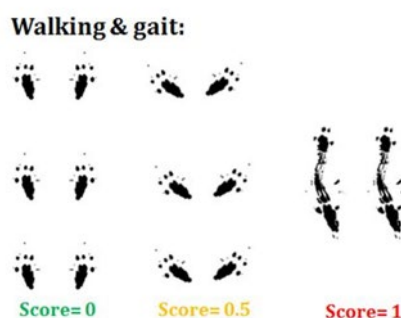Adapted from Gabellini *et al.* Nature (2006)

Score 0: no curvature of the spine
Score 0.5: mild curvature of the spine
Score 1: severe curvature of the spine

4) Walking difficulties

This parameter reflects the ability to use hindlimbs and the ability of mice for walking. The score is noted as:

Score 0: normal use of hindlimbs
Score 0.5: splayed use of hindlimbs
Score 1: loss of use of hindlimbs

Note: if score 1:
- indicate if the hindlimbs are paralyzed or not
- indicate if the mouse can still move

7. Calculation of DSS

Disease Severity Score (DSS) is the sum of the scores of the 4 parameters: body weight, hanging test ability, kyphosis and walking difficulties. The maximal DSS is 5.

Note: always refer to ethics application for humane endpoints.

8. Data Collection

All the data concerning the DSS will be reported in the dedicated table (table 3).

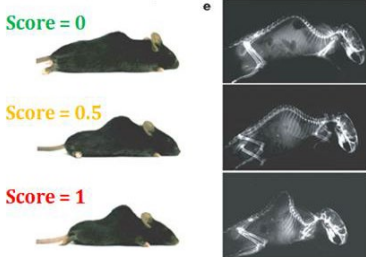**Table 3: Data collection table for Disease Severity Score**

Mice ID	Mice genotype	Age (weeks)	Body weight	Score body weight	Score hanging	Score walking	Score kyphosis	TOTAL DSS

Note: Mouse genotype should be blinded until experiment is complete

9. Data Analysis

The data will be expressed as a sum of the 4 different parameters: body weight, hanging test ability, kyphosis and walking difficulties

Table 4. Disease severity score (DSS). Tabular summary of DSS procedure, based on natural history study data analysis and modeling from training and test cohorts of *Mtm1^{-/y}* mice (Buono et al, associated manuscript).

Disease severity score	DSS Description	Score
Body weight	Score 0-1 Difference in body weight between *Mtm1^{-/y}* mouse from week n to n+1.	Score 0: $x \geq 0.25g$ Score 0.5: $-0.25 > x > 0.25$ Score 1: $x \leq -0.25g$ With x being the weight difference
Hanging time	Score of 0-2, representing 0-60 seconds hanging time.	Time (secs) > [(60-Time (s))/60s]*2 = value (0-2) <u>Examples:</u> 1s > (60-1)/60*2=1.967 15s > (60-15)/60*2=1.5 30s > (60-30)/60*2=1 60s > (60-60)/60*2=0
Kyphosis	Score 0-1 Curvature of the spine	Score 0: no curvature of the spine Score 0.5: mild curvature of the spine Score 1: severe curvature of the spine **Kyphosis:** Score = 0 Score = 0.5 Score = 1 Adapted from Gabellini *et al.* Nature (2006)
Walking difficulties	Score 0-1 Ability to use hindlimbs	Score 0: normal use of hindlimbs Score 0.5: splayed use of hindlimbs Score 1: loss of use of hindlimbs **Walking & gait:** Score= 0 Score= 0.5 Score=1
Maximum score	5	

References

Koch, C., S. Buono, A. Menuet, A. Robe, S. Djeddi, C. Kretz, R. Gomez-Oca, M. Depla, A. Monseur, L. Thielemans, L. Servais, C. N. M. S. G. NatHis, J. Laporte and B. S. Cowling (2020). "Myostatin: a Circulating Biomarker Correlating with Disease in Myotubular Myopathy Mice and Patients." Mol Ther Methods Clin Dev **17**: 1178-1189.

Tasfaout, H., S. Buono, S. Guo, C. Kretz, N. Messaddeq, S. Booten, S. Greenlee, B. P. Monia, B. S. Cowling and J. Laporte (2017). "Antisense oligonucleotide-mediated Dnm2 knockdown prevents and reverts myotubular myopathy in mice." Nat Commun **8**: 15661.